УДК 004.056.5
DOI: 10.25559/SITITO.17.202104.922-931

Научная статья

Анализ подходов к обнаружению атак в зашифрованном трафике

М. С. Полянская
ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова», г. Москва, Российская Федерация
Адрес: 119991, Российская Федерация, г. Москва, ГСП-1, Ленинские горы, д. 1
m.s.polyanskaya@mail.ru

Аннотация
Задача автоматического обнаружения сетевых вторжений активно изучается с 1980-ых годов. Отдельный интерес представляет обнаружение атак в зашифрованном трафике, доля которого в Интернете возрастает. Целью данной работы было проанализировать возможные подходы к обнаружению атак в зашифрованном трафике. В разделе 3 проанализированы подходы, основанные на незашифрованных метаданных, а также на альтернативных криптосистемах. Основными методами контроля трафика являются сигнатурный (на основе правил) и поведенческий (на основе обнаружения аномалий). Задача анализа зашифрованного трафика нетривиальна и будет рассматриваться в контексте второго подхода. В данной статье рассмотрены методы машинного обучения, подходящие для решения задачи анализа зашифрованного трафика, с учётом существующей практики обнаружения атак на основе аномалий. Несмотря на большой потенциал криптографических методов, наиболее практичным подходом, на данный момент, признан анализ метаданных. Также весьма перспективны многосторонние вычисления, которые позволяют анализировать полезную нагрузку пакетов, но не требуют перехода на альтернативное шифрование.

Ключевые слова: система обнаружения вторжений, машинное обучение, криптосистемы

Автор заявляет об отсутствии конфликта интересов.

Для цитирования: Полянская М. С. Анализ подходов к обнаружению атак в зашифрованном трафике // Современные информационные технологии и ИТ-образование. 2021. Т. 17, № 4. С. 922-931. doi: https://doi.org/10.25559/SITITO.17.202104.922-931

Original article

Analysis of Approaches to Detecting Attacks in Encrypted Traffic

M. S. Polyanskaya
Lomonosov Moscow State University, Moscow, Russian Federation
Address: 1 Leninskie gory, Moscow 119991, GSP-1, Russian Federation
m.s.polyanskaya@mail.ru

Abstract

The automatic detection of network intrusions has been under active study since the 1980s. Of particular interest is the detection of attacks in encrypted web traffic, the percentage of which on the Internet is increasing. The purpose of this article was to analyze possible approaches to detecting attacks in encrypted web traffic. Section 3 analyzes approaches based on unencrypted metadata as well as alternative cryptosystems. The main methods for controlling web traffic are the signature method (based on rules) and the behavioral method (based on anomaly detection). The task of analyzing encrypted traffic is not trivial, and it will be considered in the context of the second approach. This article discusses machine learning methods suitable for solving the problem of encrypted traffic analysis, taking into account the existing practice of detecting attacks based on anomalies. Despite the great potential of cryptographic methods, the most practical approach, at the moment, is the analysis of metadata. Multilateral computing, which allows analysis of the payload of packets, but does not require conversion to alternative encryption, is also very promising.

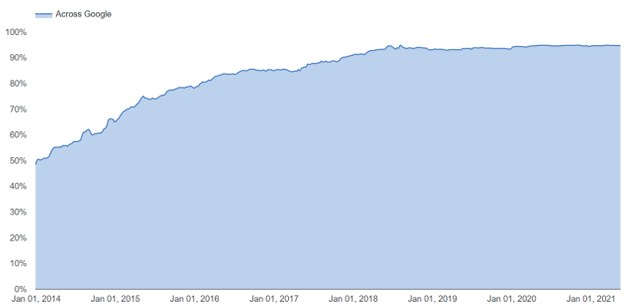
Keywords: intrusion detection system, machine learning, cryptosystems

The author declares no conflict of interest.

For citation: Polyanskaya M.S. Analysis of Approaches to Detecting Attacks in Encrypted Traffic. *Sovremennye informacionnye tehnologii i IT-obrazovanie = Modern Information Technologies and IT-Education*. 2021; 17(4):922-931. doi: https://doi.org/10.25559/SITITO.17.202104.922-931

1. Введение

Согласно данным Google Transparency Report, доля зашифрованного трафика в Интернете постоянно возрастает: от приблизительно 50% в 2014 году до более 90% по состоянию на январь 2021 года; 96% наиболее посещаемых сайтов по умолчанию используют протокол HTTPS (Рис. 1) [1].



Р и с. 1. Зашифрованный трафик через Google
F i g. 1. Trend Encrypted traffic across Google [1]

Большинство систем защиты – intrusion detection systems (IDS) – выступают своеобразным man-in-the-middle: перед анализом трафика он дешифруется, а после шифруется снова и отправляется адресату. Возникает проблема доверия к IDS, тем более когда она аппаратно отделена от обслуживаемого сервера. В свете вышесказанного, перспективна задача обнаружения атак в зашифрованном трафике.

Основными методами контроля трафика являются сигнатурный (на основе правил) и поведенческий (на основе обнаружения аномалий). Задача анализа зашифрованного трафика нетривиальна и будет рассматриваться в контексте второго подхода.

Целью данной работы было проанализировать возможные подходы к обнаружению атак в зашифрованном трафике. В разделе 3 проанализированы подходы, основанные на незашифрованных метаданных, а также на альтернативных криптосистемах.

2. Обзор подходящих методов машинного обучения

В данном разделе рассмотрены методы машинного обучения, подходящие для решения задачи анализа зашифрованного трафика, с учётом существующей практики обнаружения атак на основе аномалий (вообще говоря, в незашифрованном трафике).

Перечислим традиционные методы машинного обучения:

- Логистическая регрессия
- Наивный байесовский классификатор
- Дерево принятия решений и случайный лес
- Метод k ближайших соседей
- Метод опорных векторов

Все они встречаются в тематической литературе, но, как правило, в сравнении с тестируемым нейросетевым подходом.

Как эффективные инструменты IDS, в литературе упоминаются нейросети: Deep Neural Network (DNN), а также её разновидности Deep Belief Network (DFN), Recurrent Neural Network (RNN) [2], Convolutional Neural Network (CNN) [3].

Deep Neural Network (DNN) – нейронная сеть с несколькими слоями между входным и выходным слоями; вообще говоря, превосходит простые методы при наличии достаточного обучающего набора данных.

Deep Belief Network (DBN) – разновидность DNN, в которой процесс обучения оптимизирован с помощью послойного жадного алгоритма. DBN не требуют так много размеченных обучающих данных, которых как раз не хватает в области обнаружения атак.

IDS, основанная на DBN, описана в статье [4].

Recurrent Neural Network (RNN) позволяет учитывать предыдущее состояние нейронов, то есть обрабатывать серии событий во времени, что полезно при обнаружении атак, представляющих собой серию последовательных запросов.

Отдельно следует отметить LSTM (Long short-term memory), которые более приспособлены, чем обычные RNN, к ситуации, когда важные события разделены временными лагами с неопределённой продолжительностью и границами.

IDS, основанная на RNN, описана в статье [5].

Convolutional Neural Network (CNN) имеет слои, которые выделяют высокоуровневые признаки входа, таким образом понижая количество обучаемых параметров; CNN традиционно используется для распознавания изображений. IDS, основанная на CNN, описана в статье [3]: pcap-файлам трафика сопоставлены битовые изображения.

3. Подходы к обнаружению атак на веб-приложения в зашифрованном трафике

3.1. Анализ метаданных

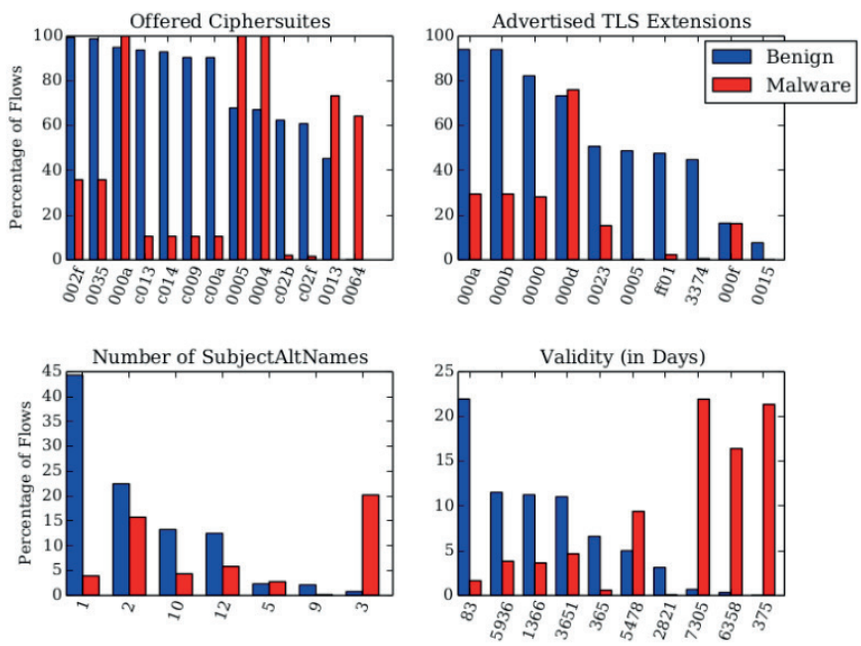
В работе [6] инженеры Cisco предлагают подход, который лёг в основу набора продуктов для обнаружения вредоносной активности в корпоративной сети Cisco Encrypted Traffic Analytics (Cisco ETA). Cisco ETA обнаруживает в сети не атаки как таковые, а вредоносные программы. Для рассмотрения выбран именно пример Cisco ETA, так как авторами подробно описаны отобранные признаки. Не все эти признаки релевантны для искусственно смоделированных атак.

Экспериментальным путём авторы остановились на наборе признаков TLS+DNS+HTTP+BD+SPLT, где:

- TLS – информация из незашифрованного TLS-рукопожатия, в особенности clientHello и clientKeyExchange, которая содержит ценную информацию о клиентской библиотеке TLS: предлагаемые наборы шифров (ciphersuites), расширения TLS (TLS extensions), количество альтернативных имён (SubjectAltNames), срок действия сертификата сервера (validity), длина открытого ключа клиента. Подразумевается, что авторы вредоносных программ используют

[1] HTTPS encryption on the web [Электронный ресурс] // Google Transparency Report, 2021. URL: https://transparencyreport.google.com/https/overview?hl=en (дата обращения: 15.08.2021).

отдельный набор библиотек и конфигураций TLS, а также часто используют устаревшие параметры (Рис. 2).



Р и с. 2. Различие признаков в TLS-рукопожатии у честных пользователей (Benign) и вредоносных программ (Malware) [6, С. 37]
F i g. 2. Significance difference in TLS handshake between honest users (Benign) and malware (Malware) [6, p. 37]

- DNS – информация о DNS-запросах клиента: длина доменного имени, количество IP-адресов в DNS-ответе, поле TTL, наличие запрошенного доменного имени в рейтинге самых популярных сайтов Alexa. Доменные имена менее динамичны, чем IP, поэтому позволяют создавать более надёжные чёрные списки. Этот контекст может быть полезен: к примеру, боты используют DNS для связи с управляющим сервером. Вредоносные доменные имена часто сгенерированы алгоритмами Domain Generation Algorithms (DGA) и имеют особенности.

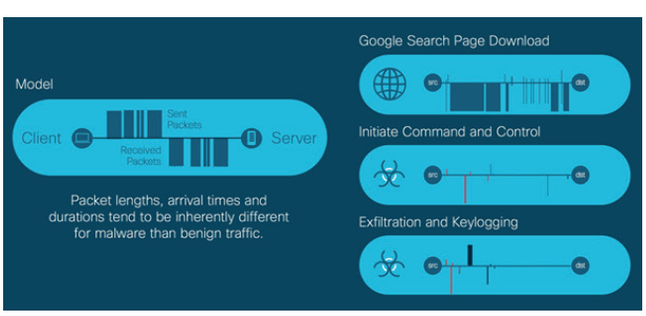
Признак нерелевантен для искусственно смоделированных атак.

- HTTP – заголовки из незашифрованных HTTP-запросов клиента: наличие определённых полей, значения Content-Type, User-Agent, Server и код возврата в ответах.

Признак нерелевантен для искусственно смоделированных атак.

- BD – распределение байтов в пакете (для различения зашифрованного и незашифрованного содержимого: в зашифрованном распределение байтов равномерное; также для определения типов незашифрованных файлов).
- SPLT – последовательность длин пакетов и временных промежутков между ними (Рис. 3).

Признак нерелевантен для искусственно смоделированных атак.



Р и с. 3. Иллюстрация признака SPLT для обычных запросов и вредоносной активности (установление контроля; отправка данных кейлоггера)[2]
F i g. 3. Illustration of the SPLT feature for normal requests and malicious activity (instituting control; sending keylogger data)

Авторы применяли к данным l1-логистическую регрессию с 10-кратной кросс-валидацией, получив высокую точность и интерпретируемые результаты. Авторы также опробовали метод SVM и не обнаружили существенных различий в точности. Метод SVM требует больше вычислений и не обладает

[2] Rehak M., Anderson B. Securing Encrypted Traffic on a Global Scale [Электронный ресурс] // Cisco. Jan 26, 2018. URL: https://blogs.cisco.com/security/securing-encrypted-traffic-on-worldwide-scale (дата обращения: 15.08.2021).

интерпретируемостью, потому рекомендуется логистическая регрессия.

Тем же исследователям [7] по данным TLS удалось обнаруживать и различать в сети 18 семейств вредоносных программ (Bergat, Deshacop, Dridex и т.д.).

Признаки, непосредственно относящиеся к TLS-потоку: TLS+BD+SPLT – также обеспечивают высокую точность.

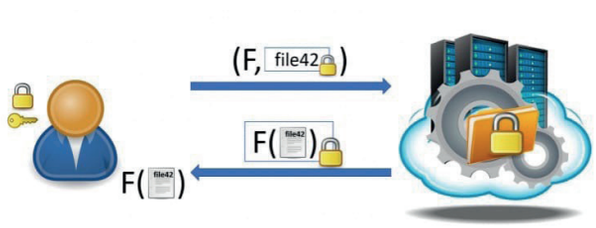
3.1.1. Оценка подхода

Подход к анализу зашифрованного трафика, основанный на анализе метаданных, уже имеет успешную практическую реализацию. Он хорошо встраивается в существующую инфраструктуру. Результаты интерпретируемы и предоставляют исследователям информацию о вредоносных программах.

Отбор релевантных признаков – ключевой момент подхода, и может представлять сложность.

Заметим, что инженеры Cisco фокусировались на обнаружении в сети вредоносных программ, а не атак в чистом виде.

3.2. Полностью гомоморфное шифрование



Р и с. 4. Иллюстрация полностью гомоморфного шифрования[3]

F i g. 4. An illustration of a fully homomorphic encryption

Рассмотрим подход к анализу зашифрованного веб-трафика, основанный на замене алгоритма шифрования веб-трафика на полностью гомоморфный. Использование данного криптопримитива в сочетании с машинным обучением подробно описано в работе[4].

Гомоморфное шифрование – криптографический примитив, представляющий собой функцию шифрования, удовлетворяющую дополнительному требованию гомоморфности относительно каких-либо алгебраических операций над открытыми текстами.

То есть, для гомоморфной схемы шифрования ∃ класс функций F такой, что для ∀ f ∈ F ∃ функция g такая, что:

$Enc(f(a_1, a_2, \ldots, a_n)) = g(Enc(a_1), Enc(a_2), \ldots, Enc(a_n))$

для любых открытых текстов $a_1,\ldots,a_n$. При этом для вычисления g не требуется знать открытых текстов и секретных параметров схемы.

Будем говорить, что приведенная криптосистема гомоморфна по функциям F.

Особый интерес представляют *полностью гомоморфные* криптосистемы (FHE, fully homomorphic encryption): гомоморфные одновременно по сложению и умножению (без ограничений на количество операций) [8]. Если криптосистема с такими свойствами может зашифровать два бита, то, поскольку операции сложения и умножения формируют над битами полный по Тьюрингу базис, становится возможным безопасно вычислить любую вычислимую функцию. В данном случае предлагается применить это свойство для решения задач машинного обучения [22].

FHE используются в удалённых безопасных вычислениях по следующему алгоритму:

1. генерация ключей – генерирование клиентом открытого ключа pk и секретного ключа sk;
2. шифрование – шифрование клиентом открытого текста m с использованием секретного ключа sk;
3. отправка клиентом шифрованного текста c и открытого ключа pk на сервер;
4. вычисление – получение сервером функции F, использование F и открытого ключа pk для выполнения вычислений над шифрованным текстом c;
5. отправка сервером результата клиенту;
6. расшифрование – расшифрование клиентом полученного от сервера значения с использованием sk.

Задача создания криптостойкой FHE впервые была сформулирована Рональдом Ривесом, Леонардом Адлеманом и Майклом Дертоузосом в 1978 году, и оставалась нерешённой до создания криптосистемы Джентри[5] (см. Приложение) в 2009 году [9], [10], которая имеет множество «последователей».

Схема называется *почти гомоморфной (SHE, somewhat homomorphic encryption)*, если поддерживает ограниченное количество выполнений гомоморфных операций; в FHE такая схема дополняется бутстрэппингом – техникой регулярного сброса накопившегося шума – которая понижает производительность.

Первой попыткой решить задачу слепой классификации на стороне сервера методом был проект Cryptonets. Его основная идея – шифрование входных данных заданной нейросети с помощью SHE-схемы BGV (Brakerski-Gentry-Vaikuntanathan [11]) и гомоморфное распространение сигналов по нейросети.

Заметим, что существующие FHE поддерживают только операции над целыми числами. В работе[6] предлагается решение этой проблемы путём перехода к дискретизованным нейросетям (DiNN); описан переход от традиционной нейросети к эквивалентной DiNN.

[3] The Three Musketeers of Secure Computation: MPC, FHE and FE [Электронный ресурс] // KU Leuven, 2021. URL: https://www.esat.kuleuven.be/cosic/blog/the-three-musketeers-of-secure-computation-mpc-fhe-and-fe (дата обращения: 15.08.2021).

[4] Minelli M. Fully Homomorphic Encryption for Machine Learning : Doctoral Thesis. Université Paris sciences et lettres, 2018 [Электронный ресурс]. URL: https://tel.archives-ouvertes.fr/tel-01918263/document (дата обращения: 15.08.2021).

[5] Gentry C. A Fully Homomorphic Encryption Scheme : Thesis of Doctor of Philosophy. Stanford University, 2009 [Электронный ресурс]. URL: https://crypto.stanford.edu/craig/craig-thesis.pdf (дата обращения: 15.08.2021).

[6] Minelli M. Fully Homomorphic Encryption for Machine Learning : Doctoral Thesis. Université Paris sciences et lettres, 2018 [Электронный ресурс]. URL: https://tel.archives-ouvertes.fr/tel-01918263/document (дата обращения: 15.08.2021).

3.2.1. PPIDS

Использование подхода, основанного на FHE, в контексте IDS встречается в работе [12]. В ней рассматривается IDS на основе правил в сочетании с гомоморфной схемой Domingo-Ferrer [13], [14], которая появилась раньше криптосистемы Джентри и не является криптостойкой, зато проста в использовании и не имеет проблем с производительностью.

3.2.2. Оценка подхода

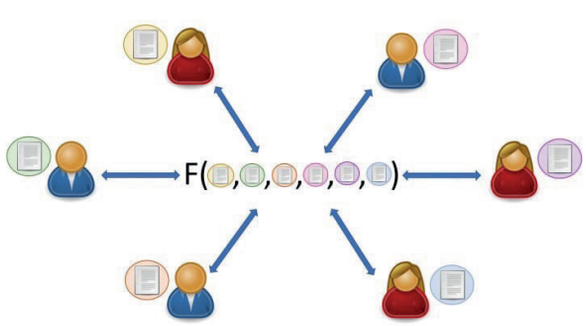
По сравнению с анализом метаданных достоинство подхода, основанного на FHE, в том, что он позволяет анализировать полезную нагрузку пакетов, которая наиболее релевантна.

Основной недостаток подхода – это требование перехода на новый, малоизученный способ шифрования, который имеет определённые ограничения. Одной из существенных проблем известных FHE является их крайне низкая производительность, которую возможно повысить, накладывая ограничения на количество операций (SHE). К. Джентри даёт следующую оценку своей FHE: выполнение одного поискового запроса Google с использованием зашифрованных на FHE слов потребует увеличения времени вычисления примерно в 10^{12} раз.

При использовании подхода приходится модифицировать архитектуру нейросетей: переходить к DiNN; менять функции активации, чтобы они расходовали меньше ресурсов SHE (levels). Это понижает качество классификации.

По причине низкой производительности на сегодняшний день данный подход является неудачным выбором для применения в таких динамичных системах, как сетевой трафик.

3.3. Многосторонние вычисления



Р и с. 5. Иллюстрация конфиденциальных многосторонних вычислений[7]

F i g. 5. An illustration of confidential multilateral computing

Конфиденциальные многосторонние вычисления (Secure Multi-Party Computation, MPC) – множество протоколов, позволяющих нескольким участникам произвести вычисление, зависящее от тайных входных данных каждого из них, таким образом, чтобы ни один участник не смог получить никакой информации о чужих тайных входных данных.

То есть, у каждого участника p_1, p_2, …, p_N есть тайные входные данные d_1, d_2, …, d_N соответственно. Участники хотят найти значение $F(d_1, d_2, …, d_N)$, где F – известная всем участникам вычислимая функция.

Впервые задача MPC была поставлена Э. Яо в 1982 году [15] на примере задачи миллионеров (сравнение двух чисел).

Важным частным случаем конфиденциальных двусторонних вычислений является забывчивая передача (oblivious transfer, OT).

В протоколах забывчивой передачи фигурируют отправитель и получатель. Отправитель хранит N секретов $\{x_1,…,x_N\}$ и по запросу передаёт получателю x_i так, что: 1) получатель не может узнать других секретов; 2) отправитель не может узнать i.

Протокол забывчивой передачи был впервые рассмотрен М. Рабиным в 1981 году[8].

В случае IDS в протоколе фигурируют две стороны: клиент (в данном контексте – охраняемый сервер) и IDS. Входные данные IDS – обученная модель классификации; входные данные клиента – подлежащий классификации веб-запрос. IDS может быть основана на том, что охраняемый сервер выполняет расшифрование запроса и перед выполнением запроса обращается с ним к «забывчивому» серверу для проверки.

В протоколе Garbled Circuit, предложенном Э. Яо (см. Приложение), рассматриваются два абонента, вычисляющих логическую схему, в виде которой может быть представлена любая программа, поэтому допустимо основываться на нём. Существует множество альтернативных подходов для различных задач (особых случаев вычислительных схем, большего числа абонентов и т.д.). Следует отметить протокол Goldreich-Micali-Widgerson [16], который хорошо применим к булевым и арифметическим схемам и выполняется за количество раундов, пропорциональное глубине схемы.

Пример решения задачи конфиденциального машинного обучения с помощью MPC встречается в работе [17]. В ней рассматриваются линейная регрессия, логистическая регрессия и простая нейросеть, которые используют протокол OT из статьи [18] на этапе обучения, протокол двухстороннего вычисления Garbled Circuit на этапе предсказания.

3.3.1. ZIDS

Использование подхода, основанного на MPC, в контексте IDS встречается в работе [19]. В ней рассматривается сигнатурная IDS, которая работает на протоколе забывчивых вычислений, разработанном авторами.

3.3.2. Оценка подхода

Как и в предыдущем подходе, достоинством подхода к анализу веб-трафика, основанного на многосторонних вычислениях, является анализ полезной нагрузки пакета. В отличие от FHE, MPC совместимы с текущими стандартами шифрования.

Основной недостаток подхода – требование интерактивности, то есть в несколько моментов времени клиент вынужден взаимодействовать с сервером и выполнять вычисления, в отличие от предыдущих случаев, в которых клиент лишь посылает

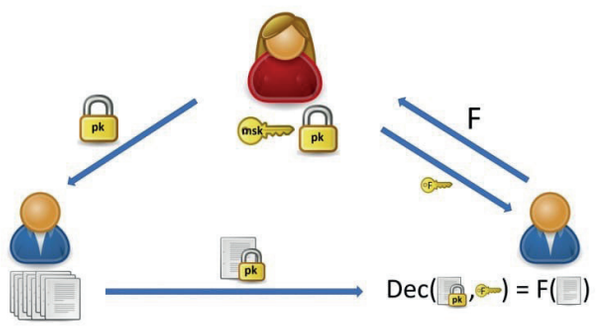
[7] The Three Musketeers of Secure Computation: MPC, FHE and FE [Электронный ресурс] // KU Leuven, 2021. URL: https://www.esat.kuleuven.be/cosic/blog/the-three-musketeers-of-secure-computation-mpc-fhe-and-fe (дата обращения: 15.08.2021).

[8] Rabin M. O. How To Exchange Secrets with Oblivious Transfer. Harvard University Technical Report 81, 1981 // IACR Eprint archive. Article number: 187. URL: https://eprint.iacr.org/2005/187 (дата обращения: 15.08.2021).

вход и получает выход. Хотя производительность у MPC-решений намного выше, чем у FHE, они влекут задержки в сети и высокую нагрузку на сеть.
На практике для построения MPC-классификатора так же, как для FHE, приходится модифицировать архитектуру целевой нейросети, решая аналогичные проблемы.

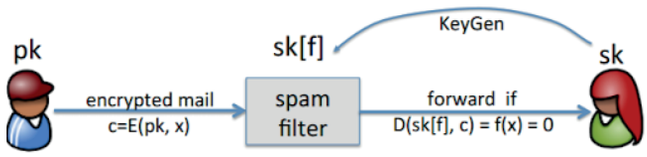
3.4. Функциональное шифрование

Функциональное шифрование (Functional Encryption, FE) – шифрование, позволяющее безопасно вычислить некоторую (одностороннюю) функцию от открытого текста.



Р и с. 6. Иллюстрация функционального шифрования[9]
F i g. 6. Functional encryption illustration

То есть, для схемы функционального шифрования ∃ функция f и функция g такие, что:
$f(a_1, a_2, ..., a_n) = g(Enc(a_1), Enc(a_2), ..., Enc(a_n))$
для любых открытых текстов $a_1,...,a_n$. При этом для вычисления g не требуется знать открытых текстов и секретных параметров схемы.
Пусть $F: K \times X \to \{0,1\}$ (k – некоторый параметр функции). Модель FE состоит из следующих этапов:
1. (pk, mk) – генерируется пара открытого и главного секретного ключа;
2. sk ← (mk, k) – генерируется секретный ключ для k;
3. c ← Enc(pk, x) – зашифровывается сообщение x;
4. y = Dec(sk, c) = F(k, x) – используется sk для вычисления F(k, x) из c.
В случае IDS в качестве x выступает анализируемый веб-запрос, в качестве k – параметры обученной модели машинного обучения на стороне IDS, F(k, x) – предсказание метки запроса. В отличие от FHE, эта схема позволяет вычисляющей стороне узнать непосредственно значение функции. Это преимущество для анализатора (Рис. 7), так как он, осведомлённый о результате проверки, может самостоятельно заблокировать вредоносную активность.



Р и с. 7. Иллюстрация конфиденциального фильтра спама на основе функционального шифрования
(здесь: sk[f] := sk, sk := mk) [8, С. 58]
F i g. 7. An illustration of a confidential spam filter based on functional encryption (here: sk[f] := sk, sk := mk) [8, p. 58]

Подход носит теоретический характер. Криптопримитив FE в данный момент малоизучен. Одна из его реализаций для отдельной задачи (шифрование на основе идентификаторов) – схема Сахаи-Уотерса [20].

4. Классификация атак

Методы обнаружения атак, основанные на альтернативных криптосистемах, позволяют анализировать пакеты на прикладном уровне и различать множество классов атак прикладного уровня, таких как XSS, SQL-инъекции, cookie poisoning, HTTP flood и т. д.
Метод, основанный на метаданных, – с учётом релевантных метаданных и имеющихся датасетов, смоделированных искусственно, – вероятно, позволит обнаруживать и различать атаки классов:

- Denial of Service (DoS)
- Попытки неавторизованного доступа
- Сканирование портов (probing)

5. Заключение

Несмотря на большой потенциал криптографических методов, наиболее практичным подходом, на данный момент, признан анализ метаданных. Также весьма перспективны многосторонние вычисления, которые позволяют анализировать полезную нагрузку пакетов, но не требуют перехода на альтернативное шифрование [21], [23-25].
Авторами запланированы исследования в области анализа метаданных, с учётом современной практики машинного обучения и особенностей поставленной задачи (например, актуальна оптимизация скорости обнаружения атак).

Приложение

1. Криптосистема Джентри (Gentry's Fully Homomorphic Encryption Scheme)

Для схемы используются идеальные решётки J в цепочках многочленов по модулю $f_n(x) = x^n+1$. Эрмитова нормальная форма идеальной решётки имеет вид:

[9] The Three Musketeers of Secure Computation: MPC, FHE and FE [Электронный ресурс] // KU Leuven, 2021. URL: https://www.esat.kuleuven.be/cosic/blog/the-three-musketeers-of-secure-computation-mpc-fhe-and-fe (дата обращения: 15.08.2021).

$$HNF(J) = \begin{bmatrix} d & 0 & 0 & 0 & 0 \\ -r & 1 & 0 & 0 & 0 \\ -[r^2]_d & 0 & 1 & 0 & 0 \\ -[r^2]_d & 0 & 0 & 0 & 0 \\ \vdots & & & \ddots & \\ -[r^{n-1}]_d & 0 & 0 & 0 & 1 \end{bmatrix}$$, где $d = det(J)$ и r — корень для $f_n(x)$ по модулю d.

Генерация ключей

1. Выбирается произвольная *n*-мерная целочисленная решётка v, где каждый вход v_i выбирается наугад, как *t*-разрядное число. С помощью этого вектора v формально определяется многочлен $v(x)= \Sigma v_i x^i$, а также соответствующая матрица поворота (составленная из входов v_i) V.
2. Вычисляется инверсия для v(x) по модулю $f_n(x)$, то есть многочлен w(x) степени не более n-1, что $v(x)\times w(x)=const \bmod f_n(x)$. Тогда матрица поворота W, соответствующая w(x), является инверсией для V, то есть V×W = I (I – единичная матрица).
3. Проверяется, что эрмитова нормальная форма для V имеет вид, указанный выше, а именно все столбцы, кроме левого, образуют единичную матрицу. В таком случае вся матрица V может быть получена с помощью элементов r и d.

Открытым ключом будет являться матрица V, заданная числами r и d. *Закрытым ключом* будут являться два многочлена (v,w).

Шифрование

Шифрование выполняется побитово. Пусть требуется зашифровать бит b ∈ (0,1).

На входе имеется бит b и открытый ключ V. Выбирается шумовой вектор u, компоненты которого принимают значения 0,1,-1. Затем вычисляется вектор $a = 2\times u+b\times e_1 = <2u_o+b, 2u_1,...,2u_{n-1}>$. Шифротекст вычисляется по формуле $c = a \bmod V = [a\times V^{-1}]\times V$.

Расшифрование

На входе имеется вектор c и матрицы V и W. Исходный бит b получается в результате операций:

$a = [c \times W/d] \times V$

$b = a_0 \bmod 2$

Бутстрэппинг

Вышеописанная схема является почти гомоморфной (SHE) из-за накопления ошибки *u*. Перейти к полностью гомоморфной схеме (FHE) позволяет бутстрэппинг, задача которого – построить функцию Recrypt, которая принимает шифротекст E(a) с шумом $N_0 < N$, где N – порог, и вычисляет новый шифротекст E(a), с шумом менее $\sqrt{N}$.

Пусть π – открытый текст. ψ_1 = Encrypt (pk_1, π) – подразумевается результат выполнения гомоморфной операции.

Вводится новая пара ключей (pk_2,sk_2). Recrypt – это гомоморфное выполнение над шифротекстом ψ_1 процедуры расшифрования Dε с использованием зашифрованного на pk_2 секретного ключа

$\mathsf{Recrypt}(\mathrm{pk}_2, D_\mathcal{E}, \langle \overline{\mathrm{sk}_{1j}} \rangle, \psi_1)$.

Set $\overline{\psi_{1j}} \xleftarrow{R} \mathsf{Encrypt}_\mathcal{E}(\mathrm{pk}_2, \psi_{1j})$ where ψ_{1j} is the j-th bit of ψ_1

Set $\psi_2 \leftarrow \mathsf{Evaluate}_\mathcal{E}(\mathrm{pk}_2, D_\mathcal{E}, \langle \langle \overline{\mathrm{sk}_{1j}} \rangle, \langle \overline{\psi_{1j}} \rangle \rangle)$

Output ψ_2

Таким образом, новый шифротекст ψ_2 – это зашифрованный на pk_2 результат Decrypt (sk_1, ψ_1) = π (π – исходный открытый текст). ψ_2 = Encrypt (pk_2, π).

2. Криптосистема Яо (Yao's Garbled Circuit)

Два абонента Алиса и Боб хотят вычислить известную обоим булеву схему, не выдавая друг другу своих входов.

1. Алиса искажает цепь – создаёт Garbled Circuits.

В таблице истинности цепи Алиса заменяет все нули и единицы на соответствующие метки: случайно сгенерированные строки из k бит. Затем Алиса зашифровывает метки выходных значений в таблице, используя соответствующие входные метки, и случайно переставляет строки.

a	b	c
0	0	0
0	1	0
1	0	0
1	1	1

a	b	c
X_0^a	X_0^b	X_0^c
X_0^a	X_1^b	X_0^c
X_1^a	X_0^b	X_0^c
X_1^a	X_1^b	X_1^c

Искаженная таблица
$Enc_{X_0^a, X_0^b}(X_0^c)$
$Enc_{X_0^a, X_1^b}(X_0^c)$
$Enc_{X_1^a, X_0^b}(X_0^c)$
$Enc_{X_1^a, X_1^b}(X_1^c)$

2. Алиса отправляет Бобу искажённую цепь и метку своего входа.
3. Боб получает метку для своего входа с помощью Алисы через забывчивую передачу.
4. Боб вычисляет схему, получая шифрованный выход.

Боб имеет искажённую таблицу и метки обоих входов. Он расшифровывает строки таблицы метками, при этом шифрование проводилось таким образом, чтобы расшифровать строку в таблице было возможно, только имея две корректные входные метки, в противном случае запись при расшифровке должна давать случайный мусор (об этом ниже). Таким образом Боб узнаёт верную метку X^c_x. Он отправляет её Алисе, чтобы та извлекла x по своей таблице меток.

Как Боб узнаёт, какое из расшифрованных значений верно? Один из способов реализовать это – при шифровании, основанном на псевдослучайной функции, предварительно добавить к открытому тексту n нулей. Вероятность того, что Боб случайно получит такое значение, расшифровывая неверные шифротексты, пренебрежимо мала.

References

[1] Wang Z., Fok K.W., Thing V.L.L. Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study. *Computers & Security*. 2022; 113:102542. (In Eng.) doi: https://doi.org/10.1016/j.cose.2021.102542

[2] Lakshmanarao A., Shashi M. A Survey On Machine Learning For Cyber Security. *International Journal of Scientific & Technology Research*. 2020; 9(01):499-502. Available at: https://www.ijstr.org/final-print/jan2020/-A-Survey-On-Machine-Learning-For-Cyber-Security.pdf (accessed 15.08.2021). (In Eng.)

[3] Wang W., Zhu M., Zeng X., Ye X., Sheng Y. Malware traffic classification using convolutional neural network for representation learning. *2017 International Conference on Information Networking (ICOIN).* IEEE Press, Da Nang, Vietnam; 2017. p. 712-717. (In Eng.) doi: https://doi.org/10.1109/ICOIN.2017.7899588

[4] Alom Z., Bontupalli V.R., Taha T.M. Intrusion Detection Using Deep Belief Network and Extreme Learning Machine. In: *Artificial Intelligence: Concepts, Methodologies, Tools, and Applications*; ed. by Management Association, Information Resources. Hershey, PA: IGI Global; 2017. p. 357-378. (In Eng.) doi: https://doi.org/10.4018/978-1-5225-1759-7.ch014

[5] Kim Ji., Kim Ja., Thi Thu H.L., Kim H. Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection. *2016 International Conference on Platform Technology and Service (PlatCon).* IEEE Press, Jeju, Korea (South); 2016. p. 1-5. (In Eng.) doi: https://doi.org/10.1109/PlatCon.2016.7456805

[6] Anderson B., McGrew D. Identifying Encrypted Malware Traffic with Contextual Flow Data. *Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security (AISec'16).* Association for Computing Machinery, New York, NY, USA; 2016. p. 35-46. (In Eng.) doi: https://doi.org/10.1145/2996758.2996768

[7] Anderson B., Paul S., McGrew D. Deciphering malware's use of TLS (without decryption). *Journal of Computer Virology and Hacking Techniques.* 2018; 14(3):195-211. (In Eng.) doi: https://doi.org/10.1007/s11416-017-0306-6

[8] Boneh D., Sahai A., Waters B. Functional encryption: a new vision for public-key cryptography. *Communications of the ACM*. 2012; 55(11):56-64. (In Eng.) doi: https://doi.org/10.1145/2366316.2366333

[9] Rivest R.L., Adleman L., Dertouzos M.L. On Data Banks and Privacy Homomorphisms. In: DeMillo R.A. (Ed.) *Foundations of Secure Computation*. Academic Press, New York; 1978. p. 169-179. (In Eng.)

[10] Gentry C., Halevi S. Implementing Gentry's Fully-Homomorphic Encryption Scheme. In: Paterson K.G. (ed.) *Advances in Cryptology – EUROCRYPT 2011. EUROCRYPT 2011. Lecture Notes in Computer Science.* Vol. 6632. Springer, Berlin, Heidelberg; 2011. p. 129-148. (In Eng.) doi: https://doi.org/10.1007/978-3-642-20465-4_9

[11] Brakerski Z., Gentry C., Vaikuntanathan V. (Leveled) Fully Homomorphic Encryption without Bootstrapping. *ACM Transactions on Computation Theory*. 2014; 6(3):13. (In Eng.) doi: https://doi.org/10.1145/2633600

[12] Park H.A., Lee D.H., Lim J., Cho S.H. PPIDS: Privacy Preserving Intrusion Detection System. In: Yang C.C., et al. (eds.) Intelligence and Security Informatics. PAISI 2007. *Lecture Notes in Computer Science*. Vol. 4430. Springer, Berlin, Heidelberg; 2007. p. 269-274. (In Eng.) doi: https://doi.org/10.1007/978-3-540-71549-8_27

[13] Domingo-Ferrer J. A new privacy homomorphism and applications. *Information Processing Letters*. 1996; 60(5):277-282. (In Eng.) doi: https://doi.org/10.1016/S0020-0190(96)00170-6

[14] Domingo-Ferrer J. A Provably Secure Additive and Multiplicative Privacy Homomorphism. In: Chan A.H., Gligor V. (eds.). *Information Security. ISC 2002. Lecture Notes in Computer Science*. Vol. 2433. Springer, Berlin, Heidelberg; 2002. p. 471-483. (In Eng.) doi: https://doi.org/10.1007/3-540-45811-5_37

[15] Yao A.C. Protocols for secure computations. *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982).* IEEE Press, Chicago, IL, USA; 1982. p. 160-164. (In Eng.) doi: https://doi.org/10.1109/SFCS.1982.38

[16] Goldreich O., Micali S., Wigderson A. How to play ANY mental game. *Proceedings of the nineteenth annual ACM symposium on Theory of computing (STOC'87).* Association for Computing Machinery, New York, NY, USA; 1987. p. 218-229. (In Eng.) doi: https://doi.org/10.1145/28395.28420

[17] Mohassel P., Zhang Y. SecureML: A System for Scalable Privacy-Preserving Machine Learning. *2017 IEEE Symposium on Security and Privacy (SP).* IEEE Press, San Jose, CA, USA; 2017. p. 19-38. (In Eng.) doi: https://doi.org/10.1109/SP.2017.12

[18] Peikert C., Vaikuntanathan V., Waters B. A Framework for Efficient and Composable Oblivious Transfer. In: Wagner D. (ed.). *Advances in Cryptology – CRYPTO 2008. CRYPTO 2008. Lecture Notes in Computer Science.* Vol. 5157. Springer, Berlin, Heidelberg; 2008. p. 554-571. (In Eng.) doi: https://doi.org/10.1007/978-3-540-85174-5_31

[19] Niksefat S., Sadeghiyan B., Mohassel P., Sadeghian S. ZIDS: A Privacy-Preserving Intrusion Detection System Using Secure Two-Party Computation Protocols. *The Computer Journal.* 2014; 57(4):494-509. (In Eng.) doi: https://doi.org/10.1093/comjnl/bxt019

[20] Fang L., Xia J. Full Security: Fuzzy Identity Based Encryption. *IACR Cryptology ePrint Archive.* 2008. Article number: 307. 22 p. Available at: https://eprint.iacr.org/2008/307 (accessed 15.08.2021). (In Eng.)

[21] Canetti R. Universally composable security: a new paradigm for cryptographic protocols. *Proceedings 42nd IEEE Symposium on Foundations of Computer Science.* IEEE Press, Newport Beach, CA, USA; 2001. p. 136-145. (In Eng.) doi: https://doi.org/10.1109/SFCS.2001.959888

[22] Damgard I., Geisler M., Kroigard M. Homomorphic encryption and secure comparison. *International Journal of Applied Cryptography.* 2008; 1(1):22-31. Available at: https://www.inderscienceonline.com/doi/abs/10.1504/IJACT.2008.017048 (accessed 15.08.2021). (In Eng.)

[23] Gilad-Bachrach R., Dowlin N., Laine K., Lauter K., Naehrig M., Wernsing J. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. *Proceedings of the 33rd International Conference on Machine Learning (PMLR).* Vol. 48. New York, NY, USA; 2016. p. 201-210. Available at: https://proceedings.mlr.press/v48/gilad-bachrach16.html (accessed 15.08.2021). (In Eng.)

[24] Dhote Y., Agrawal S., Deen A.J. A Survey on Feature Selection Techniques for Internet Traffic Classification. *2015 International Conference on Computational Intelligence and Communication Networks (CICN).* IEEE Press, Jabalpur, India; 2015. p. 1375-1380. (In Eng.) doi: https://doi.org/10.1109/CICN.2015.267

[25] Gao N., Gao L., Gao Q., Wang H. An Intrusion Detection Model Based on Deep Belief Networks. *2014 Second International Conference on Advanced Cloud and Big Data.* IEEE Press, Huangshan, China; 2014. p. 247-252. (In Eng.) doi: https://doi.org/10.1109/CBD.2014.41

Поступила 15.08.2021; одобрена после рецензирования 06.10.2021; принята к публикации 24.11.2021.
Submitted 15.08.2021; approved after reviewing 06.10.2021; accepted for publication 24.11.2021.

Об авторе:

Полянская Марина Сергеевна, магистрант кафедры информационной безопасности факультета вычислительной математики и кибернетики, ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова» (119991, Российская Федерация, г. Москва, ГСП-1, Ленинские горы, д. 1), **ORCID: https://orcid.org/0000-0002-6848-0413,** m.s.polyanskaya@mail.ru

Автор прочитал и одобрил окончательный вариант рукописи.

About the author:

Marina S. Polyanskaya, Master degree student of the Chair of Information Security, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1 Leninskie gory, Moscow 119991, GSP-1, Russian Federation), **ORCID: https://orcid.org/0000-0002-6848-0413,** m.s.polyanskaya@mail.ru

The author has read and approved the final manuscript.