УДК 37.014 DOI: 10.34708/GSTOU. 2022.89.36.003

АНАЛИЗ ПРОБЛЕМ БЕЗОПАСНОСТИ БОЛЬШИХ ДАННЫХ В IOT И МЕТОДОВ ИХ ПРЕОДОЛЕНИЯ

© М.Я. Пашаев, С.И. Абдулаева, М.М. Намаева, Х.С. Халиева
ГГНТУ им. акад. М.Д. Миллионщикова, Грозный, Россия

Объем обмена данными в сети увеличивается за счет расширения пространства Интернета вещей и его все большей популярности в последние годы. В связи с ростом запросов на подключение к сети и использование сетевых технологий многократно возросла необходимость сохранения конфиденциальности и безопасности личных данных. В рассматриваемой проблемной плоскости пользователи сети, которые являются датчиками, также масштабно увеличивают объем обрабатываемых данных. В результате перед комплексом технического обеспечения – маршрутизаторы сети и сетевые серверы – ставятся серьезные задачи по контролю и обеспечению безопасности такого объема данных. Предлагаемая статья приводит анализ этих проблем, направленный на обеспечение безопасности пользователей и представление процедур, позволяющих решать проблемы и выражать стратегии их преодоления. Предлагаются некоторые методы использования функциональности корректного шифрования, в том числе и формирования безопасного протокола. Другой способ преодоления проблем с конфиденциальностью больших данных – это продуманный дизайн протокола связи.

Ключевые слова: большие данные, облачные серверы, интеллектуальные устройства, агрегация и интеграция данных, интеллектуальный анализ данных IoT, функции шифрования, конфиденциальность, криптографические функции, безопасность больших данных.

С каждым днем увеличивается не только количество людей, подключенных к Интернету, но и то же самое происходит с Интернет-вещами, подключенными к Интернету для обмена данными и взаимодействия друг с другом. В современную эпоху интеллектуальные устройства и Интернет-вещи могут передавать и получать информацию, а также взаимодействовать друг с другом. Это взаимодействие достигло такого уровня, что основной объем обмена данными через устройства связан с Интернетом вещей (IoT). Прогнозируется, что количество умных устройств, подключенных к Интернету, в 2022 году увеличится в 44 раза по сравнению с 2009 годом [1]. Датчики можно назвать доминирующими вещами, связанными с обменом данными через Интернет, которые считаются одним из самых больших ресурсов данных [2]. IoT следует рассматривать как интеллектуальную сеть, состоящую из вещей, взаимодействующих друг с другом.

IoT в сети производят и обмениваются данными, имеющими большой объем, скорость и разнообразие. Термин под названием «Big» был придуман для данных, воплощающих вышеупомянутые характеристики, при этом большие данные не имеют четкой категорийной констатации и обычно классифицируются на основе характеристических свойств.

IoT относится ко всем вещам, которые могут передавать или получать определенную информацию и могут взаимодействовать друг с другом. Интернет вещей можно назвать образным понятием, который практически распространяется и проникает во все области.

Обработка данных в области IoT осуществляется посредством сети, которая способна аккумулировать, обрабатывать, анализировать передаваемые данные всех элементов сети. Интернет вещей предсказывает, что в будущем мир будет контролироваться физическими устройствами, соединенными друг с другом через единую инфраструктуру. Другими словами, IoT – это гигантская сеть, в которой огромная масса устройств, датчиков и умных устройств взаимодействуют друг с другом [1].

Большая часть этой сети состоит из датчиков, наравне с такими, как смарт-карта, смартфон и компьютеры, которые также конфигурируют в сети для обмена необходимой информацией.

Большие данные (Big Data) также не имеют четкого и прозрачного определения [3]. В частности, можно привести такое определение – это набор данных из различных источников, от четко определенных до слабо определенных, полученных из человеческих или машинных источников. Было представлено множество определений, которые зависят от цели и области исследования. На данный момент 2 миллиарда человек подключены к Интернету, а 5 миллиардов пользуются мобильными телефонами. Ожидается, что к 2023 году количество устройств, подключенных к Интернету, превысит 50 миллиардов устройств. Прогнозируется, что количество сообщений, переданных к этому году (2023 г.), в 44 раза превысит количество сообщений, переданных к 2009 г. [1].

Основное использование больших данных – выявление и оптимизация процессов в бизнесе, финансовой торговле, улучшении и оптимизации умных городов и стран, улучшении управления взаимоотношениями с клиентами, улучшении здравоохранения, спортивно-массовой деятельности, транспортных услуг и т. д. В настоящее время IoT гарантирует, что среда, в которой мы живем, может стать «умной», т. е. автоматизирует информационные потоки сообщений о состоянии какого-либо объекта, оснащенного датчиками на основе интеллектуальных систем.

Устройства IoT используют различные методы сетевых подключений к другим устройствам для обмена данными. Такие технологии, как стандартный Wi-Fi, Bluetooth с низким энергопотреблением, локальное оконечное оборудование (LTE), спутниковые соединения используются для подключения нескольких устройств на разных уровнях. Безопасность IoT связана не только с безопасностью устройства, но и с веб-интерфейсами, облачными сервисами и другими устройствами, которые с ним взаимодействуют. Существует множество методов, используемых для решения таких проблем, как конфиденциальность данных, вывод и агрегирование, что позволяет повторно идентифицировать некоторые персоналии, даже если они удалены из набора данных. В том числе следует отметить целый ряд уязвимостей безопасности, которые могут привести к небезопасному веб-интерфейсу.

Большие данные могут поступать из множества различных источников, таких как системы бизнес-транзакций, клиентские базы данных, медицинские записи, журналы посещений в Интернете, мобильные приложения, социальные сети, собранные результаты научных экспериментов, машинно-генерируемые данные и используемые датчики данных в реальном времени в среде Интернета вещей. Данные можно оставить в необработанном виде или предварительно обработать с помощью инструментов интеллектуального анализа данных или программного обеспечения для подготовки данных перед их анализом.

Большие данные также охватывают широкий спектр типов данных, включая структурированные данные в базах данных и хранилищах данных SQL, неструктурированные данные, такие как текстовые файлы и файлы документов, хранящиеся в кластерах Hadoop или системах NoSQL, и частично структурированные данные, такие как журналы веб-сервера или потоковой передачи данных с датчиков.

Информационные системы, основанные на IoT, хранятся и обрабатываются на серверах, обладающих сверхвысокой надежностью и мощностью. Облачные серверы собирают информацию с датчиков, радиочастотных идентификаторов и других интеллектуальных устройств и сохраняют ее в своей памяти [2]. Облачные вычисления подпитывались промышленностью, и мощные серверы таких компаний, как Google и Amazon, поддерживали эти серверы. Потребность в облачных сервисах в сети IoT остро ощущается, так как в этой сети острая необходимость с высокой скоростью обрабатывать и надежно хранить большие данные.

Технологии Big Data и IoT тесно взаимосвязаны. Несмотря на то, что социальные сети считались гигантскими производителями Big Data, но прогнозирование увеличения объема данных в мире показывает, что в ближайшем

будущем IoT получит приоритетную долю обмениваемой информации. Как упоминалось выше, будущее отношение к IoT сосредоточено на крупномасштабных умных городах [4]. Четвертая промышленная революция в мире в 2011 году стала пиком использования Интернета вещей. Датчики, которые имеют наибольшую долю в IoT, получили высокий рост выпуска в промышленности, вместе с этим увеличивая и объем передаваемой контролируемой информации.

Филип Чен описал еще один тип зависимости BIG DATA от IoT. Навигация, социальные сети, финансовая информация, информация о здоровье, астрономическая информация и интеллектуальный транспорт входят в число отраслей, которые производят много информации. Еще одна связь между IoT и Big Data, которая характеризуется сбором информации об окружающей среде, геоинформационными данными и астрономией через беспроводные датчики IoT. Со значительным ростом применения технологий IoT объем данных по этим областям также многократно увеличивается [5].

Обмен данными в социальных сетях и масштабный рост объемов информации в этих сетях также повлияли на многократное увеличение количества выявляемых фактов использования данных в злонамеренных и мошеннических целях. Необходимость защиты информации пользователей, в том числе и конфиденциальности, диктует применение адекватных мер формирования стратегии информационной безопасности в условиях развития IoT и Big Data [6]. Итак, два важных фактора решения названной проблемы заключаются в следующем:

1. *Конфиденциальность*. Личная безопасность пользователей указывает на уровень доступности для других пользователей. Некоторые исследователи разделили концепцию конфиденциальности на три раздела: безопасность инфраструктуры, информационная безопасность и управление информацией [5]. Каждый пользователь (частное лицо или устройство) может предоставить доступ к своей конфиденциальной информации другим, однако при этом гарантировано ограничение доступа и обеспечение его конфиденциальности.

2. *Безопасность*. Большой объем данных и скорость их обработки затрудняют обеспечение безопасности. Принципы формирования защиты информации в условиях BIG DATA по технологиям реализации довольно сложны и до сих пор недостаточно изучены [3]. Гигантские IT-компании, такие как Google, Microsoft, YouTube, Skype, использовали большой спектр программных и аппаратных разработок для обеспечения информационной безопасности, что демонстрирует сложность защиты на гигантских серверах [7].

Угрозы безопасности состоят из нескольких разделов. Атака Dennial of Service (DoA) формируется в инфраструктуре, а атака на функцию шифрования и контроля доступа формируется в разделе конфиденциальности [4]. Защита пользователей от этих атак является проблемой безопасности.

С помощью ежедневного анализа поведения пользователей киберзащитники и эксперты могут отмечать и сообщать об аномальных онлайн-действиях, которые могут быть подозрительными и аномальными. Однако точный анализ профилей пользователей по-прежнему затруднен из-за сложной, разнообразной и динамичной природы поведения людей в Интернете. Большинство предыдущих исследований сосредоточены только на технологической перспективе защиты от кибератак. Ограниченная доступность соответствующих исследований создает большие трудности как для исследователей, так и для промышленных участников.

Опишем некоторые особенности BIG DATA:

- Объем. Объем данных (террабайт и более) представляет собой серьезную проблему для BIG DATA. Разнообразие типа информации, форматов данных и их предметных областей (Facebook, например, производит 500 террабайт данных каждый день). Сложно регламентировать определенные границы для объемов BIG DATA, но это не область мелкомасштабных данных.
- Разнообразие. Например, соответствующие данные, «считываемые» с датчиков, имеют множество вариаций представляемых форматов.

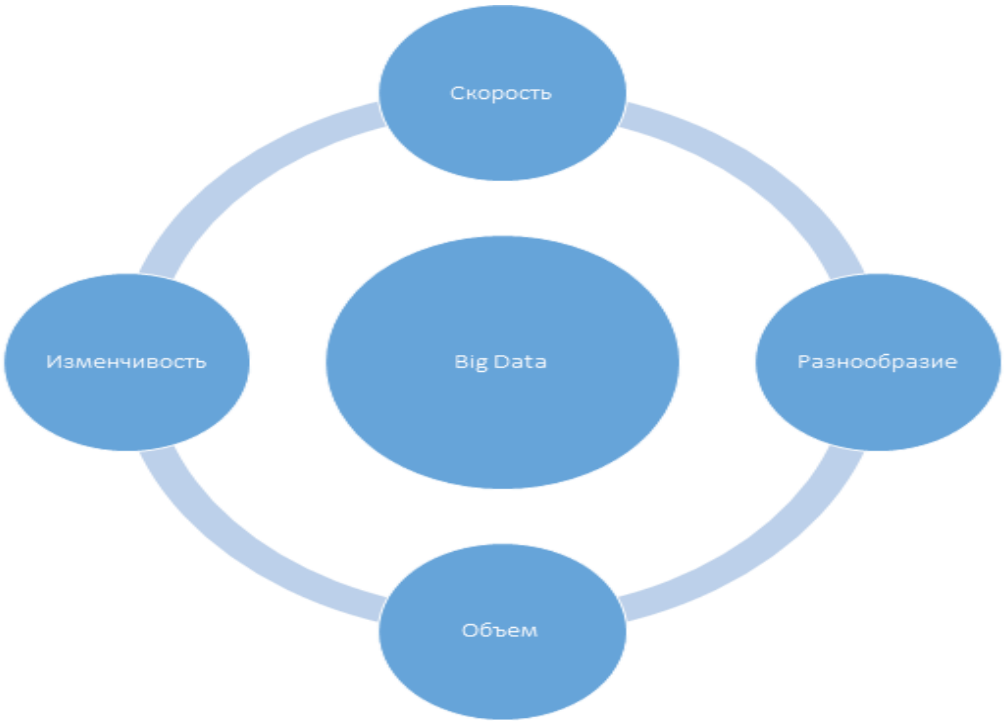

Рис. 1. Основные проблемы внедрения больших данных

• Скорость. Из описания BIG DATA выше акцентируем внимание на разнообразии и сложности структуры данных, если при этом добавляются высокоскоростные и объемные данные, обработка которых представляет определенные трудности. Скорость – это еще одна особенность BIG DATA, которая порождает одновременно и проблему обработки высокоскоростных данных.

• Изменчивость. Передача разнообразных данных множеством пользователей осуществляется из большого числа источников. Аналитические исследования данной темы определяют изменчивость данных одной из четырех первоначальных проблем, представленных на рисунке 1.

– Правдивость. Структура данных достаточно сложна, и достоверность данных в BIG DATA не может быть строго объективно установлена.

– Интеллектуальный анализ данных и оптимизация. Одна из самых больших задач в области BIG DATA – очистить гигантский пул информации от дополнительных «избыточных» данных и выбрать оптимизированный вариант комплекса данных.

– Сложность: как видно из других объяснений, по мере увеличения объема данных возрастает и сложность. Это можно разделить на три части:

1) сложность данных;

2) вычислительная сложность;

3) сложность системы.

Анализ данных и моделирование: эта задача указывает на некоторые типы подмножеств: разделение информации и сбор потерянных данных.

Интерпретация данных: этот шаг выглядит как этап построения изображения, важный для принятия решения о полученных данных.

Конфиденциальность: как описывалась эта проблема выше, конфиденциальность – самая сложная задача в эпоху цифровых технологий.

Безопасность: сохранение конфиденциальности пользователей и предотвращение распространения вредоносных программ. В соответствии с важностью обработки BIG DATA следует также учитывать их безопасность.

• Управление данными: по мере роста BIG DATA компании и организации потенци-

ально должны быть готовы управлять большими данными, улучшение качества данных и сохранение ценности данных являются одними из ключевых элементов информации.

• Право собственности на данные. Помимо безопасности и конфиденциальности, важно также обеспечивать право собственности на данные, что особенно актуально проявляется при их совместном использовании данных.

В этой статье мы проанализировали проблемное поле для безопасности и конфиденциальности данных в BIG DATA, а также перечень адекватных способов преодоления этих проблем в IoT. Самая основная и примитивная цель в области безопасности и конфиденциальности – обеспечить все три функции конфиденциальности, честности и аутентификации [8].

Многие устройства имеют ограниченную вычислительную мощность и не могут быть зашифрованы с помощью ключей. Но сегодня разработано множество методов шифрования, которые одновременно являются быстрыми, безопасными и недорогими [7]. Есть также много программ IoT, которые все еще находятся в зачаточном состоянии в области удостоверения личности, цифровой подписи, и каждая из них имеет сильные и слабые стороны. Иногда путем простых изменений программа могла быть практически полезной [9].

Чтобы обеспечить конфиденциальность данных пользователей и предотвратить нарушения этики информационной безопасности в сети, следует использовать безопасные протоколы. Существует множество протоколов для передачи информации, аутентификации пользователей, цифровых подписей и так далее. Каждое из них представлено с целью свойств, каждое из которых имеет сильные и слабые стороны, которые можно зафиксировать, проанализировав их детально. Облегченные протоколы и использование криптографических функций сетевой безопасности будут использоваться не всегда, а иногда необходимо использовать асимметричное шифрование, цифровую подпись и другие функции шифрования, значительно увеличивающие вычислительную нагрузку сети [10]. Следовательно, необходимо спроектировать такой протокол, чтобы основная нагрузка по его вычислению на сервере серверов и пользователей (объектов) была достаточно мощной.

Исходя из особенностей реализации проектов IoT с подходом майнинга BIG DATA, в данной работе определены приоритетные задачи повышения эффективности реализации безопасности и конфиденциальности данных пользователей. Предложены методы, которые могли бы обеспечить более безопасное пространство для маломощных пользователей. Наиболее важной из предложенных нами методов является грамотная разработка коммуникационных протоколов, чтобы вычислительная нагрузка ложилась на поставщиков услуг (облачных серверов).

ЛИТЕРАТУРА

1. *Хеглунд Д., Линдемер С., Фурухед М., Раза Ш.* IoT: на пути к инфраструктуре открытых ключей для Интернета вещей, компьютеров и безопасности. 2020.
2. *Миркин Б.Г.* Введение в анализ данных: учебник и практикум / Б.Г. Миркин. Москва: Издательство Юрайт, 2022. 174 с.
3. *Вьющенко О.О., Маслова М.А.* Об обеспечении безопасности в сфере интернета вещей / О.О. Вьющенко, М.А. Маслова // Информационные технологии. Т. 6. №3. 2021. С. 33-39.
4. *Роуз М.* Интернет вещей (IoT). Повестка дня ИОТ. М. Роуз [Электронный ресурс]. Режим доступа: https://internetofthingsagenda. techtarget. com/definition/Internet-of-Things-IoT (дата обращения: 14.10.2022).
5. *Алисултанова Э.Д., Моисеенко Н.А., Юсупова Р.В., Тасуев У.Р.* Система оптимизации работы клиентских служб на основе нейронных сетей // Современные наукоемкие технологии. 2021. №12 (часть 1). С. 15-20.

6. *Довгаль В. А., Довгаль Д. В.* Управление ресурсами в Интернете Вещей / В. А. Довгаль, Д. В. Довгаль // Дистанционные образовательные технологии: материалы II Всерос. науч.-практ. конф., г. Ялта, 2017 г. Симферополь: АРИАЛ, 2017. С. 168-173.
7. *Хаджиева Л. К., Хаджиев М. Р., Исрахимова А. Т.* Концепция систем IOT для сельскохозяйственных услуг с высокой степенью автономии // Вестник ГГНТУ. Технические науки. Том XVII. №3 (25). Грозный, 2021.
8. *Хаджиева Л. К., Хаджиев М. Р., Хашумов И. У.* Перспективы внедрения технологии 5G и взаимодействие с системой «УМНЫЙ ДОМ» // Вестник ГГНТУ. Технические науки. Том XVII. №4 (26). Грозный, 2021.
9. *Хашумов И. У., Хаджиев М. Р., Мальцагов Х. Х.* Информационная безопасность и уязвимости ИОТ // Вестник ГГНТУ. Технические науки. Том XVI. №2 (20). Грозный, 2020.
10. *Моисеенко Н. А., Джабраилов И. С.* Проектирование информационной системы управления выборами: инициатива современности // Вестник ГГНТУ. Технические науки. Том XV. №2 (16). Грозный, 2019.

BIG DATA SECURITY ANALYSIS IN IOT AND METHODS FOR OVERCOMING THEM

© M. Ya. Pashaev, S. I. Abdullaeva, M. M. Namaeva, Kh. S. Khaliyeva
GSTOU named after acad. M. D. Millionshchikov, Grozny, Russia

The volume of data exchange in this network is increasing due to the expansion of the Internet of Things and its increasing prominence in recent years. With increasing requests to connect to the network and use its services, the need to maintain privacy and security is felt like never before, and it has all become a challenge. Keeping users safe on smaller networks seems easier and threats more predictable. Because network users who are sensors also increase the amount of data. As a result, network routers and network servers have become a major challenge in controlling and securing this amount of data. We analyze these issues in this article. This document aims to ensure the safety of users and to provide procedures for solving problems and expressing coping strategies. Some methods are suggested, such as the use of an encryption function and a secure protocol. Another way to overcome problems with big data is through thoughtful design of the communication protocol. Finally, we point out some of the existing problems associated with big data in the Internet of Things.

Keywords: big data, big data analysis, IoT, big data security.

REFERENCES

1. Heglund, D. Lindemer S., Furuhed M. and Raza S. (2020) IoT: na puti k infrastrukture otkrytykh klyuchei dlya Interneta veshchei, komp'yuterov i bezopasnosti/ [IoT: towards a public key infrastructure for the internet of things, computers and security].
2. Mirkin, B. G. (2022) *Introduction to data analysis: textbook and workshop.* [Vvedenie v analiz dannykh: uchebnik i praktikum], Yurayt Publishing House, Moscow, 174 p.
3. Vyushchenko, O. O. and Maslova, M. A. (2021) 'Ob obespechenii bezopasnosti v sfere interneta veshchei'. *Informatsionnye tekhnologii.* [About ensuring security in the sphere of the Internet of Things. Information technologies]. V. 6, №3, pp. 33-39.

4. Rose, M. Internet veshchei (IoT). Povestka dnya IOT [Internet of things (IoT). IOT agenda], available at: https://internetofthingsagenda. techtarget. com/definition/Internet-of-Things-IoT (accessed 10/14/2022).
5. Alisultanova, E. D., Moiseenko, N. A., Yusupova, R. V. and Tasuev, U. R. (2021) 'Sistema optimizatsii raboty klientskikh sluzhb na osnove neironnykh *setei' Sovremennye naukoemkie tekhnologii*. [System for optimizing the work of client services based on neural networks. Modern high technologies]. No. 12 (part 1), pp. 15-20.
6. Dovgal, V. A. and Dovgal, D. V. (2017) 'Upravlenie resursami v Internete Veshchei'. *Distantsionnye obrazovatel'nye tekhnologii: materialy II Vseros. nauch. – prakt. konf., g. Yalta, 2017 g. Simferopol'* [Resource management in the Internet of Things. Distance learning technologies: materials of the II All-Russia. scientific-pract. Conf., Yalta, 2017 Simferopol], ARIAL, pp. 168-173
7. Khadzhieva, L. K., Khadzhiev, M. R. and Israkhimova, A. T. (2021) 'Kontseptsiya sistem IOT dlya sel'skokhozyaistvennykh uslug s vysokoi stepen'yu avtonomii'. *Vestnik GGNTU. Tekhnicheskie nauki.* [The concept of IOT systems for agricultural services with a high degree of autonomy. Herald of GSTOU. Technical sciences]. Volume XVII, №3 (25), Grozny.
8. Khadzhieva, L. K., Khadzhiev, M. R. and Khashumov, I. U. (2021) 'Perspektivy vnedreniya tekhnologii 5G i vzaimodeistvie s sistemoi «UMNYI DOM»'. *Vestnik GGNTU. Tekhnicheskie nauki.* [Prospects for the introduction of 5G technology and interaction with the "SMART HOUSE" system. Herald of GSTOU. Technical sciences], Volume XVII, №4 (26), Grozny.
9. Khadzhieva, L. K., Maltsagov, Kh. Kh. (2019) 'Informatsionnaya bezopasnost' i uyazvimosti IOT'. *Vestnik GGNTU. Tekhnicheskie nauki.* [Analysis of the technology "Internet of Things" (IOT) and its role in the "Smart Home". Herald of GSTOU. Technical sciences]. Volume XV, №4 (18), Grozny.
10. Moiseenko, N. A. and Dzhabrailov, I. S. (2019) 'Proektirovanie informatsionnoi sistemy upravleniya vyborami: initsiativa sovremennosti'. *Vestnik GGNTU. Tekhnicheskie nauki*. [Designing an information system for managing elections: an initiative of the present. Herald of GSTOU. Technical sciences]. Volume XV, №2 (16), Grozny.