Программные системы и комплексы

УДК 004.77
DOI:10.38028/ESI.2022.27.3.011

Создание инструментов сбора данных для анализа аспектов безопасности Интернета вещей

Исаева Ольга Сергеевна, Кулясов Никита Владимирович,
Исаев Сергей Владиславович

Институт вычислительного моделирования СО РАН,
Россия, Красноярск, *isaeva@icm.krasn.ru*

Аннотация: Целью работы является создание инструментов для сбора данных и исследования аспектов безопасности информационного взаимодействия распределённых устройств и приложений Интернета вещей (Internet of Things – IoT). Для достижения цели решены задачи: разработан специализированный исследовательский стенд, включающий все функциональные уровни архитектуры IoT, созданы инструменты сбора и агрегирования данных, построены показатели для выявления сетевых аномалий. Специализированный стенд включает сенсорный уровень, который состоит из измерительных устройств мониторинга окружающей среды, транспортный уровень, реализованный на основе инфраструктуры корпоративной сети, для сервисного уровня развёрнут кластер сбора и хранения данных, имеющий различные конфигурации настроек безопасности, на прикладном уровне размещено программное обеспечение для работы с данными. Инструменты выполняют сбор, агрегирование и анализ структурированных данных и неструктурированных журналов сетевого трафика, учитывая конфигурации настроек политик безопасности телекоммуникационных узлов. Построены показатели, отражающие активность и легитимность обращений с распределением по дням, странам и серверам. Инструменты предназначены для специалистов по кибербезопасности и позволяют анализировать влияние архитектуры IoT на обеспечение безопасности информационного взаимодействия элементов сети.

Ключевые слова: Интернет вещей, сетевые аномалии, кластер Kubernetes, протокол обмена сообщениями, Message Queuing Telemetry Transport (MQTT), Eclipse Mosquitto, Smart environments

Цитирование: Исаева О.С. Создание инструментов сбора данных для анализа аспектов безопасности Интернета вещей / О.С. Исаева, Н.В. Кулясов, С.В. Исаев // Информационные и математические технологии в науке и управлении. – 2022. – № 3(27). – С. 113-125. – DOI: 10.38028/ESI.2022.27.3.011.

Введение. Современные информационные технологии стали неотъемлемой частью любого производственного процесса. По данным аналитической компании Burning Glass Technologies, доля вакансий, требующих навыков работы в области искусственного интеллекта, больших данных, Интернета вещей, технологии блокчейн и облачных вычислений, существенно превышает объем квалифицированных IT-специалистов, ежегодно выпускаемых учебными заведениями в данном сегменте [1]. Особенностью программ обучения в области современных информационных технологий является ориентация не только на теоретическую проработку учебных материалов, но и на получение специализированных навыков проектирования и разработки информационных решений в различных прикладных областях. В этой связи требуется создание современных платформ, обеспечивающих повышение презентативности новых понятий и технологий для глубокого погружения в различные аспекты научно-исследовательской деятельности. Образовательные программы включают учебные стенды для исследования работы телекоммуникационного оборудования и систем управления, контроля или мониторинга данных [2]. В Красноярском математическом центре подготовка квалифицированных кадров сопровождается применением современных инструментов, обеспечивающих постановку и решение научных и прикладных задач. Одним из направлений исследований в этой области математический центр рассматривает создание инструмен-

тов практического изучения понятий, методов и технологий концепции промышленного Интернета вещей (Internet of Things – IoT).

Проводимое исследование направлено на создание масштабируемой инфраструктуры, включающей все функциональные уровни Интернета вещей, предназначенной для сбора данных и исследования аспектов безопасности информационного взаимодействия распределённых устройств и приложений. Методология проведения исследований строится на основе архитектурных решений специализированного стенда, объединяющего концептуальные понятия технологии Интернета вещей.

Интернет вещей представляет собой глобальную инфраструктуру информационного общества, обеспечивающую цифровые услуги за счёт организации связи между физическими или виртуальными объектами на основе совместимых информационных и коммуникационных технологий [3, 4]. Интернет вещей является интегратором «вещь-ориентированных», «интернет-ориентированных» и «семантико-ориентированных» технологий [5], определяя концепцию построения вычислительных сетей из распределённых устройств и приложений, где цифровые и физические объекты взаимодействуют между собой и с окружающей средой [6]. Массовое развёртывание огромного количества устройств IoT и неоднородность сценариев их использования требуют создания специализированных методов моделирования и анализа. Для такого направления инфокоммуникаций в настоящее время определены только общие подходы, которые постоянно корректируется и дополняется новыми спецификациями, на текущем этапе не способными в полной мере обеспечить надёжность и целостность огромных объёмов данных, получаемых от устройств.

Для проведения исследований выполнен обзор существующих технологических и архитектурных подходов к построению инфраструктуры IoT и рассмотрены проблемы безопасности на различных уровнях инфокоммуникационного взаимодействия распределённых устройств и приложений.

1. Подходы к построению функциональных уровней IoT. Архитектура IoT является естественной платформой для обеспечения автономности сбора данных, их совместимости и развёртывания независимых сервисов и приложений [7]. В общем случает она содержит четыре функциональных уровня: сенсорный, транспортный, сервисный и прикладной. Из-за неоднородности объектов IoT эти уровни могут включать промежуточное программное обеспечение между устройствами и приложениями, которое определяет абстрактное представление объектов для их эффективного обслуживания, управления и использования [8, 9]. Перечень функциональных уровней IoT приведён в табл. 1.

Таблица 1. Функциональные уровни IoT

Тип	Уровень	Функции	Примеры
Приложения	Прикладной	Предоставление информации, рекомендаций пользователю и взаимодействие с ним.	Интерактивные информационные панели в Grafana, Kibana. Приложения управления предприятием (ERP, ERM).
	Сервисный	Хранение данных, их аналитическая обработка	Облачные решения на основе инструментов Hadoop, ClickHouse, ES и др.
Инфраструктура	Транспортный	Передача данных между уровнями	Различные технологи и протоколы передачи данных (RS485, Ethernet, COM, CAN, RTU, Modbus, ZigBee, Z-Wave, Lora, BLE).
	Сенсорный	Взаимодействие с окружением (сбор, управление).	Системы мониторинга производственных помещений и устройств (протоколs MQTT, CoAP).

В таблице приведены примеры протоколов связи и приложений, используемые в современных архитектурах IoT. Выбор конкретных решений влияет на безопасность дальнейшего информационного взаимодействия уровней и приложений. Сенсорный уровень составляют устройства, интегрированные с датчиками, обеспечивающими сбор информации о состоянии наблюдаемых объектов в реальном масштабе времени. Транспортный уровень включает шлюзы и сети передачи данных. На этот уровень поступает большой объем данных, создаваемых устройствами IoT. Их объединение реализует инфраструктуру устройств. Данные, собираемые через сеть IoT, могут быть переданы службам, размещённым на облачных сервисах, которые позволяют интегрировать их с данными мобильных терминалов и других устройств [10].

Сервисный уровень обеспечивает хранение данных, их аналитическую обработку и управление бизнес-процессами. Прикладной уровень составляют проблемно-ориентированные приложения, решающие задачи конкретной предметной области [11].

Исследования безопасности систем IoT имеют свои особенности (ввиду специфики архитектурных и функциональных подходов) и требуют адаптации или создания новых технологических решений. Инфраструктура, реализующая функциональные уровни IoT, гетерогенна, мобильна и динамична [12], что является причиной развития сетевых атак, направленных на маршрутизацию между устройствами. Защита, в этом случае, специфична для каждого функционального уровня.

На сенсорном уровне строятся системы обнаружения аномального поведения, основанные на профилировании IoT-устройств. В [13] профили устройств содержат статистические характеристики, отражающие интенсивность и продолжительность передачи пакетов. В [14, 15] предложен подход к выявлению аномального трафика методами машинного обучения. Применены методы: дерево решений, случайный лес, нейронная сеть прямого распространения, *k*-ближайших соседей и показано, что все они дают высокий результат обнаружения аномалий трафика. Недостатком подхода является необходимость предварительного поиска обучающих примеров как типичного, так и аномального поведения устройств, что сужает круг обнаруживаемых аномалий. Кроме того, для анализа применяются только структурированные данные о сеансах связи, без учёта дополнительных сведений из неструктурированных журналов сетевого трафика. Ввиду динамичности источников атак и постоянно меняющихся характеристик аномального поведения устройств, автоматического контроля на основе методов машинного обучения недостаточно.

Проблемы безопасности транспортного уровня вызваны используемыми протоколами межуровневого взаимодействия (облегчёнными по сравнению со стандартными протоколами). Протокол MQTT (Message Queuing Telemetry Transport) [16] работает по схеме «Издатель-Брокер-Подписчик», где издателями выступают датчики и другие измерительные устройства IoT, брокерами – сервера для сбора и хранения данных, а подписчики – это объекты, которые потребляют данные, такие, как приложения на смартфон, iPad и пр. Такая схема вызывает проблемы аутентификации различных устройств, проблемы целостности и конфиденциальности пакетов данных, собранных между узлами, или компрометации всей сети устройств, в случае влияния на доступность в инфраструктуре IoT [17]. Анализ проблем безопасности протокола MQTT выполнен в [18, 19]. Показано, что MQTT уязвим к лавинным атакам (разновидность DoS-атаки), заключающимся в захвате соединений с брокером; атакам SYN - flooding, когда устанавливается большое число незакрытых сеансов TCP, нагружающих брокер; атаки типа «отказ в обслуживании», когда устанавливаются несколько соединений с брокером и для каждого соединения отправляются как можно большее количество сообщений; проблемам компрометации или подмены издателей или подписчиков при запросе недопустимых сообщений [20]. Все атаки приводят к загрузке серверов, на которых

размещаются брокеры, до состояния невозможности обслуживать легитимных издателей и подписчиков. В [21] предложено включать флаги безопасности в пакеты передачи по MQTT и на их основе строить системы разграничения прав доступа. В [22] для этих целей предложены сервисы для авторизации пользователей, выполнена оценка производительности шлюза при высокой периодичности поступления данных. Показано, что по мере возрастания количества сообщений за единицу времени, их доставка за счёт анализа дополнительных заголовков, содержащих сведения о правах доступа, возрастает до 0.5-1 с, что существенно при соизмеримой частоте обновления данных. Необходимо соблюдать баланс между допустимым уровнем требований к безопасности и обеспечением скорости поступления данных. Для обеспечения упреждающего поведения систем мониторинга требуются инструменты, позволяющие специалисту по кибербезопасности проводить анализ статистики обращений, формировать показатели на основе неструктурированных журналов сетевого трафика и исследовать аномалии в данных в зависимости от настроек политики безопасности устройств в структуре IoT.

Созданы инструменты сбора данных и исследования аспектов безопасности информационного взаимодействия распределённых устройств и приложений IoT. Методы данного исследования включают проектирование архитектуры исследовательского стенда: состава и размещения измерительных устройств, организации коммутационной среды и кластера для хранения и анализа данных; его реализация позволит обеспечить совместную работу различных технологий и протоколов доступа в распределённой гетерогенной конфигурации. В архитектуру IoT предложено включить несколько брокеров данных, имеющих различные конфигурации политик безопасности и выполненных в виде физических или виртуальных серверов. Сбор и анализ данных с одновременно функционирующих брокеров, обладающих разной конфигурацией, позволит выявлять зависимости аномальных факторов от устройства сети IoT. Работа продолжает исследования в области кибербезопасности на основе сбора и анализа сетевого трафика [23, 24].

2. Построение исследовательского стенда IoT. Архитектура построенного для проведения исследований стенда IoT приведена на рис. 1. Показаны основные устройства и используемые технологические и инструментальные решения. Их выбор для каждого уровня архитектуры производился на основе литературы по IoT, рекомендаций со специализированных форумов разработчиков и выявления современных тенденций в данной предметной области.

Сенсорный уровень выполнен на основе измерительных устройств IoT (CL-210-E производства ICP DAS), выполняющих мониторинг показателей температуры, влажности, точки росы и концентрации мелкодисперсной пыли для устройств, которые размещены в специализированных технологических помещениях с телекоммуникационным оборудованием. Транспортный уровень реализуется на основе существующей инфраструктуры корпоративной сети Красноярского научного центра, в которую встраиваются устройства IoT. Для сбора, хранения и анализа данных развёрнут кластер Kubernetes [25] (K8s) на 10 узлов.

Вычислительная структура размещена на 3-х физических серверах и отказоустойчивом кластере виртуальных машин: на двух физических серверах запущены 8 виртуальных машин (система виртуализации Hyper-V), третий сервер используется как полноценный узел кластера, дополнительный узел развёрнут на виртуальной машине отдельного кластера Hyper-V. На всех узлах кластера K8s использована операционная система Ubuntu server 20.04. Кластер K8s развернут поверх системы контейнеризации Docker. Централизованное управление жизненными циклами контейнеров выполняется на платформе orchestration Rancher [26]. Узлы кластера выполняют функциональные роли, обеспечивающие распределение рабочих нагру-

зок: Worker – «Рабочее приложение», Etcd – «Хранилище ключей и конфигураций кластера» и ControlPlane – «Поддержка управления».

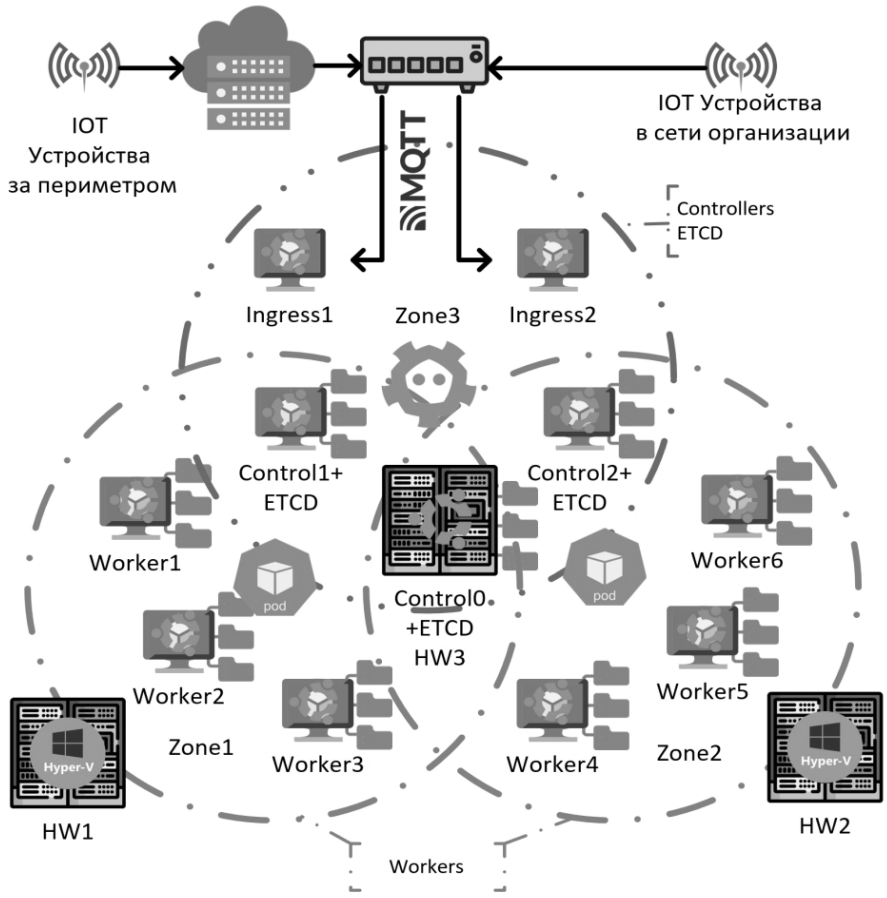


Рис. 1. Схема исследовательского стенда IoT

Выделены следующие области распределения ролей узлов кластера: 6 узлов Worker (ттх: 6 threads, 2gb dynamic RAM), на рисунке обозначены пунктиром с точкой, узел Etcd и ControlPlane (ттх: 2 threads, 2gb RAM), пунктир с двойной точкой, на пересечении областей находятся 3 узла, совмещающие роли Worker, Etcd и ControlPlane (ттх: 8 threads, 4gb RAM). Предполагается расширение узлов с индивидуальными настройками политик безопасности.

Накопление данных производится в брокерах, развёрнутых на Eclipse Mosquitto, поддерживающим протокол MQTT, предназначенным для обмена сообщениями между устройствами. Все сервера различаются конфигурацией настроек по способу доступа, использованию протокола шифрования и размещению, для исследования сетевых угроз и средства обеспечения безопасности доступа к данным. Настройки серверов: с авторизацией или без авторизации; с шифрованием (протокол Transport Layer Security, TLS) или без шифрования; с видимостью только из внутренней корпоративной сети или из сети Internet.

Построенная схема IoT реализована в виде исследовательского стенда, который обладает масштабируемостью и позволяет выполнять расширение каждого из уровней IoT.

3. Формализация схемы информационного взаимодействия IoT. Введём формализацию основных элементов схемы информационного взаимодействия и базовых операций для построения показателей анализа безопасности IoT.

3.1. Распределённые устройства IoT.

$P=\{p_1, p_2, \ldots, p_k\}$ – издатели (источники данных), k – количество устройств IoT.

$B=\{b_1, b_2, \ldots, b_m\}$ – брокеры (сервера, различающиеся настройками политик безопасности), m – количество виртуальных и физических серверов.

$S=\{s_1, s_2, \ldots, s_n\}$ – подписчики (потребители данных, системы анализа и мониторинга), k – количество приложений и источников обращений к данным (значение ограничено в каждый момент времени).

Брокер $b \in B$ характеризуется настройками политик безопасности, которые могут быть описаны кортежем признаков $Ser(b) = <Vie, Rol, Ac, Cod, Pl>$, где *Ser* – настройки сервера (брокера), *Vid* – виртуальный или физический сервер, значения: [*Vir, Phy*], *Rol* – роль из [*Work, Etcd, Contr*], *Ac* – способ доступа, принимает значение из [*Auth, Anon*], *Cod* – признак использования шифрования, принимает одно из значений [*Open, TLS*], *Pl* – размещение публичное или во внутренней сети, выбирается одно из значений: [*Priv, Publ*].

Максимальное число различных серверов брокеров определяется мощностью декартова произведения характеристик, то есть $|B|=|Ac \times Cod \times Pl|$, где $|\ |$ – мощность множества.

3.2. Пакеты с данными от устройств IoT и сетевые журналы. Распределённые устройства формируют пакеты данных, передают их брокерам, которые, в свою очередь, собирают данные и ведут сетевые журналы. От издателей поступают фактографические данные с результатами измерений (периодичность их поступления зависит от настроек издателей), от подписчиков – запросы на соединение и получение данных.

Множество J – журналов брокеров содержит сообщения от элементов из множества $\{P, S\}$ к брокеру $b \in B$. Обозначим $J_b(t) \in J$ – журнал брокера b в момент времени t, тогда $J = \cup_{t=0}^{T} J_b(t)$. Записи журналов формируются по шаблонам $Pt=\{Pt_l\ (c, b)\}$, где элемент $c \in \{P, S\}$ выступает как клиент, брокер $b \in B$ – как сервер. Шаблоны описываются кортежами из характеристик соединений. Для каждого задаётся название – *Name* (Pt_l) и служебное слово в сообщении *Theme* (Pt_l).

Pt_1: *Name* (Pt_1)=«Открытие соединения», *Theme* (Pt_1)=*"*New connection from". $Pt_1(c, b)=<t, client\_ip, client\_port, broker\_port>$, где t – отметка времени в Unix Timestamp, *client_ip* – адрес подключившегося клиента – элемента c, *client_port* – порт, с которого пришёл запрос на соединение, *broker_port* – порт, на котором брокер b получил запрос на соединение. Пример записи шаблона: *%t: Theme %client_ip:%client_port* on port *%broker_port*. Запись в журнале: *1655681432:* New connection from *172.16.0.209:42240* on port *1883*.

Pt_2: *Name*(Pt_2)=«Подтверждение соединения», *Theme*(Pt_2)=*"*New client connected from". $Pt_2(c, b)=<t, client\_ip, client\_port, client\_alias, protocol\_version, session\_status, keepalive, username>$, где *client_alias* – псевдоним клиента, *protocol_version* – версия протокола, *session_status* – статус сессии, *keepalive* – продолжительность соединения, *username* – имя пользователя при авторизации, остальные параметры описаны в шаблоне Pt_1.

Пример записи шаблона: %*t*: *Theme* %*client_ip*: %*client_port* as %*client_alias* (p%*protocol_version*, c%*session_status*, k%*keepalive*, u%*username*). Запись в журнале: *1655681432*: New client connected from *172.16.0.209:42240* as *mqtt2mysql_k8s* (*p2, c1, k0, u'razor'*).

Pt_3: *Name* (Pt_3)=«Подписка», *Theme*(Pt_3)=*"*Received SUBSCRIBE from". $Pt_3(c, b)=<t, client\_alias, topick>$, где *topick* – тема подписки, остальные параметры описаны в шаблоне Pt_2. Пример записи шаблона: *%t: Theme %client_alias. %t %client_alias %topic*. Запись в журнале: *1655681517*: Received SUBSCRIBE from *mqtt2mysql*. *1655681517*: *mqtt2mysql 'modules/topic/412'*.

Структуры шаблонов Pt_4 и Pt_5 совпадают. *Name* (Pt_4)=«Приём данных», *Them* (Pt_4)=*"*Received PUBLISH from". *Name* (Pt_5)=«Отправка данных», *Theme* (Pt_5)=*"*Sending PUBLISH to". $Pt_4(c, b) = Pt_5(c, b)=<t, client\_alias, duplicate, qos, retain, source\_mid, topic, size>$, где *duplicate* – признак дублирования, *qos* – параметр QOS, *retain* – удержание, *source_mid* – идентификатор сообщения, *size* – размер пакета. Пример записи шаблонов: %*t*: *Theme* %*cli-*

ent_alias (d%*duplicate*, q%*qos*, r%*retain*, m%*source_mid*, '%*topic* ' (%*size* bytes)). Запись в журнале (по шаблону Pt_4): *1655681549*: Received PUBLISH from *CL-210-E_163B59* (*d0, q0, r0, m0, 'modules/mqtt114-434/All'* (232 bytes)). Запись в журнале (по шаблону Pt_5): *1655681665*: Sending PUBLISH to *mqtt2mysql* (*d0, q0, r0, m0, 'modules/mqtt111-206/All'*, (231 bytes)).

Описаны шаблоны всех действий, включая запрос или подтверждение состояния, разрыв соединения, отмену подписки, пропуск данных, сообщения об ошибках и уведомлениях.

Каждый журнал J_b содержит записи о прохождении пакетов, соответствующие шаблонам $Pt=\bigcup_{i=1}^{|Pt|} Pt_i$, где $|Pt|$ определяет количество шаблонов в заданной реализации. Для анализа журналов выполняется их разбор. Брокеры разбирают пакеты данных и наполняют базу данных с результатами измерений и журналы.

Обозначим D – множество результатов измерений, выполненных в заданные моменты времени. Для каждого издателя $p \in P$ подмножество $D_p(t) \in D$ содержит данные, полученные от устройства p в момент времени t. Тогда $D = \bigcup_{t=1}^{T} D_p(t)$ описывает все данные от источника p за период наблюдений T, формируемый из отсчётов с интервалом в $\Delta t+\tau$, *где* Δt – период измерений, τ – вариация периода.

3.3. Операции над данными журналов. Сетевые аномалии могут наблюдаться на всех уровнях информационного взаимодействия IoT. Для их выявления выполняется анализ множества *{D, J}*. Автоматизировать в полной мере анализ возможно только для определения крупных событий, с выраженными признаками, определяемыми по прецедентам, но для превентивного реагирования и мониторинга состояния устройств и коммутационной среды требуются инструменты извлечения данных из неструктурированных источников, группировки показателей по признакам и их визуализации в графических представлениях.

Критерии анализа взаимодействия между брокером и издателями (устройствами IoT) основаны на статистике, отражающей частоту получения данных от каждого из устройств, и их объем. Для $p \in P$ и $D_p(t_1)$, $D_p(t_2)$,…, $D_p(t_i)$, где $t_i=t_{i-1}+\Delta t+\tau_i$ выполняется: $0 < D_p(t_i) \le D_p(t_{i-1})$, то есть поступающие данные ограничены.

Для каждого $\tau_i = (t_i - t_{i-1} - \Delta t)$ выполняется $0 \le \tau_i \le \Delta t$, то есть задержки сети при получении данных не превышают заданный интервал их обновления. Кроме того, сравнение значений τ_i с настройками $Ser(b_j)$ для $b_j \in B$ позволит выявить влияние политик безопасности на качество доступа к данным IoT.

Критерии анализа взаимодействия между брокером и подписчиками (приложения IoT) основаны на разборе неструктурированных сетевых журналов. Используются стандартные операции объединения, пересечения множеств и вводятся операции выбора подмножества по условию, замены и группировки, которые реализуются в структурах шаблонов множества *Pt*:

$\bigcup_{b \in B} J_b$ – объединение журналов для подмножества брокеров. $\bigcap_{b \in B,A} J_b$ – пересечение журналов для устройств из *B* по полю *A*. Например, для выбора запросов от одного источника.

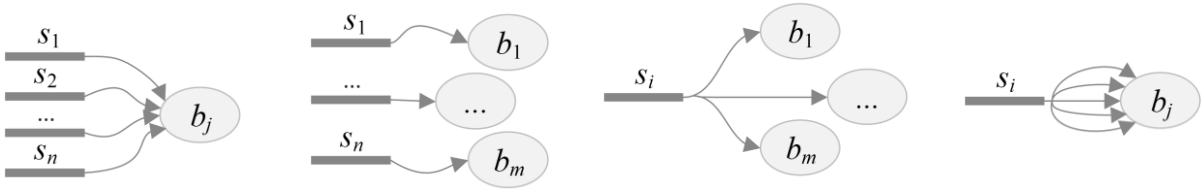
Sel (J_b, θ) – множество элементов из журнала J_b, удовлетворяющих условию θ. Если $\theta=\emptyset$, то функция *Sel* выбирает все данные журналов, если $\theta = Pt_i$, то выбираются данные журналов по условию по шаблону Pt_i.

Sub (J_b, *A*, *V*) – множество элементов из журнала J_b, в котором значения параметра *A* заменены на значения *V* из дополнительных справочников. Если в качестве *A* указывается параметр журнала *client_ip*, а *V* – столбец из справочника с географической привязкой ip-адресов, то функция позволяет выбирать и в дальнейшем группировать источники запросов.

Group (J_b, *A*) – группировка данных из журнала J_b по признаку *A*.

Введённые функции позволяют автоматизировать обработку данных, строить выборки и их графические представления, объединяя данные по источникам, странам, датам, типам

обращений, а также настройкам политик безопасности и размещения серверов с брокерами данных. Выбор последовательности записей для каждого источника и рассмотрение по ним результатов соединений по заданным шаблонам позволяют определять легитимность событий. В качестве условия θ может быть задана одна из схем взаимодействия данных. Основные схемы, вызывающие интерес, показаны на рис. 2.



а) $n \to 1$ б) $n \to m$ в) $1 \to m$ г) $1 \to 1$

Рис. 2. Схемы взаимодействия с брокером

Выявление показанного на схемах взаимодействия напрямую не свидетельствует о наличии на сервере брокера аномальных соединений, но большое число подобных запросов за ограниченное время требует включения этого показателя в мониторинг для выявления источников и результатов соединений, а в объединении с функциями выбора и объединения данных по шаблонам с результатами источникам позволяют ограничивать трафик.

4. Результаты применения стенда IoT. Созданный исследовательский стенд позволяет наглядно демонстрировать основные понятия технологии IoT и подходы к обеспечению кибербезопасности. Инструменты автоматически разделяют данные от устройств и статистику обращений к серверам. Собираемые данные предназначены для анализа сетевых аномалий, отражающих такие угрозы, как перехват данных и управляющих сообщений, подмена информации, перегрузка оборудования ложными пакетами, определяющих свойства надёжности транспортной среды и безопасности межуровневого взаимодействия устройств и приложений.

Рассмотрены показатели сетевого трафика для мониторинга активности и безопасности обращений. На рис. 3 показана динамика роста запросов к устройствам IoT и серверам сбора данных.

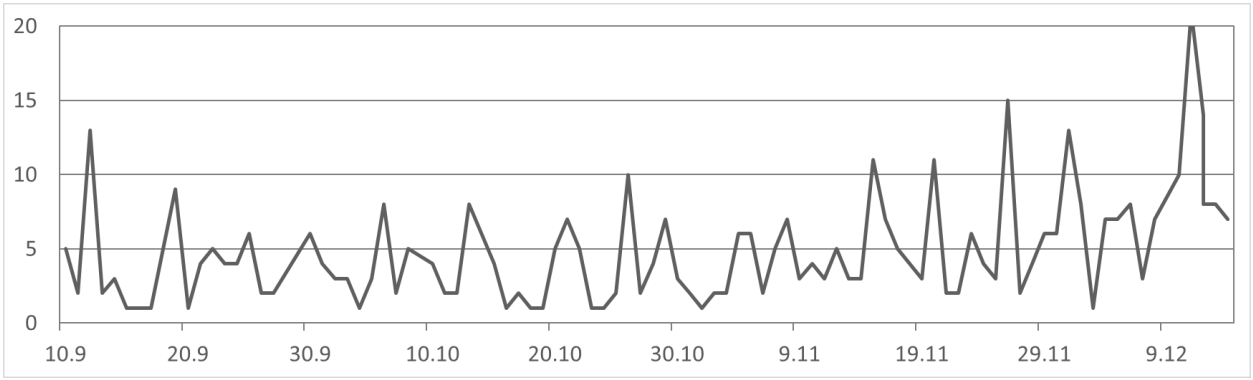


Рис. 3. Активность обращений (оси графика: количество обращений, дата обращения)

Исследования показывают, что после подключения стенда IoT к сети Internet наблюдается постепенный рост количества соединений. Причём факт подключения стенда не был анонсирован в публичном пространстве. Это показывает постоянное сканирование различных устройств и сервисов, поддерживающих технологию IoT.

На рис. 4 показано обобщённое распределение обращений к серверам по странам-источникам. Выполнено объединение данных для всех серверов, независимо от параметров их конфигурации.

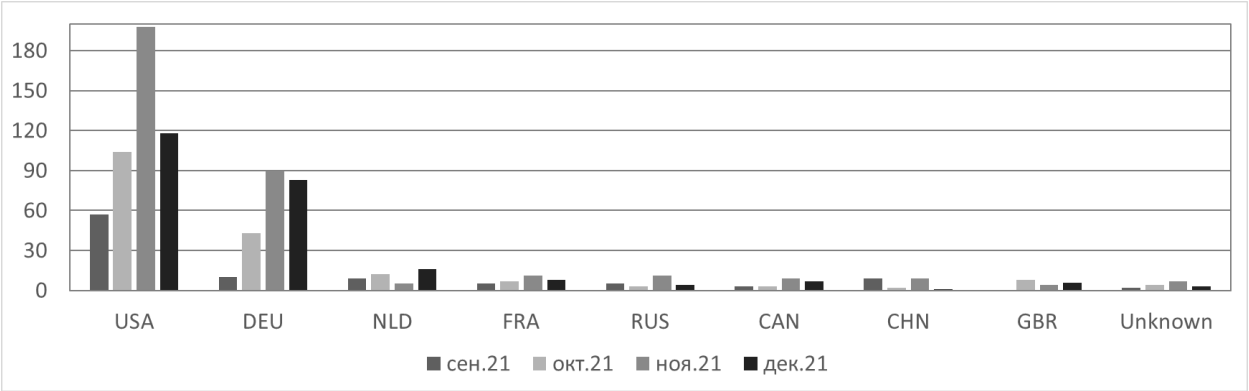


Рис. 4. Распределение обращений по странам

Гистограмма обращений показывает преобладающую активность хостов из США и Германии, что может свидетельствовать не только о популярности технологий IoT в этих странах, но и распространении в них агентов сканирования устройств IoT.

Для исследования безопасности различных конфигураций серверов выполнены выбор нелегитимных обращений и их группировка по серверам, имеющим различные настройки. На рис. 5 показана гистограмма выделенных обращений за несколько месяцев по всем серверам.

Из построенной гистограммы видно, что журналы содержат зафиксированные нелегитимные обращения как из внешней сети, так и из внутренней. Дальнейшее исследование направлено на выявление источников и причин соединений для хостов из внутренней сети.

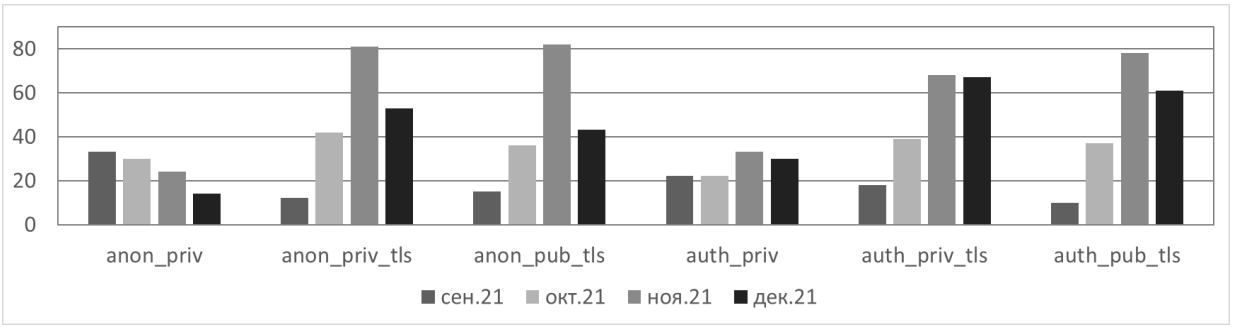


Рис. 5. Количество нелегитимных обращений к серверам

Заключение. В работе получены следующие результаты:

– разработан специализированный исследовательский стенд, включающий все функциональные уровни архитектуры IoT;

– выполнена формализация схемы информационного взаимодействия основных элементов архитектуры IoT в рамках исследовательского стенда;

– предложена формализация базовых операций для построения показателей анализа безопасности IoT;

– созданы инструменты сбора и агрегирования сетевых журналов и контролируемых устройствами IoT показателей;

– построены показатели для выявления сетевых аномалий.

Разработанный исследовательский стенд представляет собой сконфигурированные в единую платформу по технологии Интернета вещей устройства с датчиками мониторинга окружающей среды, телекоммуникационные устройства и сервера для сбора и анализа данных. В состав стенда входит кластер на 10 узлов, вычислительная структура которого объединяет физические сервера и виртуальные машины. Структура стенда обеспечивает совместную работу различных технологий и протоколов доступа в распределённой гетероген-

ной конфигурации. Результаты его работы содержат данные измерений с датчиков и журналы сетевого трафика, которые размещаются в хранилище данных.

Научные и практические результаты работы заключаются в создании формализованных основ для исследования структурированных данных, контролируемых устройствами IoT и неструктурированных журналов сетевого взаимодействия между брокером данных, издателями и подписчиками. Это позволит выполнять мониторинг состояния устройств и коммутационной среды для анализа различных аспектов безопасности IoT.

Стенд IoT позволяет проводить научные исследования и обучение студентов по дисциплинам «Информационные и вычислительные сети» и «Интеллектуальные системы», обеспечивая погружение в тематическую среду, что способствует успешной реализации процесса подготовки квалифицированных IT-специалистов. Данные, генерируемые устройствами IoT, позволят проводить исследования методов интеллектуального анализа больших данных. Собираемые сетевые журналы позволят исследовать аспекты безопасности при конфигурировании межуровневого взаимодействия IoT.

Планируется развитие данных работ для исследования архитектурных и функциональных аспектов технологии IoT и основ обеспечения безопасного взаимодействия устройств и приложений. Все элементы исследовательского стенда имеют потенциал для масштабирования, что позволяет подключать новые наборы устройств IoT и использовать новые сетевые протоколы.

Благодарности. Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ в рамках мероприятий по созданию и развитию региональных НОМЦ (Соглашение 075-02-2022-873).

Список источников

1. Lyu W., Liu J. Artificial Intelligence and emerging digital technologies in the energy sector. Applied Energy, 2021, no. 303(9), pp. 117615, DOI: 10.1016/j.apenergy.2021.117615.
2. Training stand for learning the basics of embedded systems programming. Available at: https://www.dlink.ru/stand.html (дата обращения: 28.03.2022).
3. Росляков А.В. Интернет вещей: учебное пособие / А.В. Росляков, С.В. Ваняшин, А.Ю. Гребешков. – Самара: ПГУТИ, 2015. – 200 с.
4. Atzori L., Iera A., Morabito G. The internet of things: A survey. Computer Networks, 2010, no. 54(15), pp. 2787–2805, DOI: 10.1016/j.comnet.2010.05.010.
5. Angelo G. Simulation of the Internet of Things. Proceedings of the IEEE 2016 International Conference on High Performance Computing and Simulation, 2016, DOI: 10.1109/HPCSim.2016.7568309.
6. Мальцева Н.С. Моделирование гибридной топологической инфраструктуры сети «Интернета вещей» / Н.С. Мальцева, А.А. Сорокин, П.С. Резников, В.М. Дорохов // Наука, образование, инновации: пути развития, 2019. – № 10. – С. 89–93.
7. Javed A., Heljanko K., Buda A., Främling K. CEFIoT: A fault-tolerant IoT architecture for edge and cloud IEEE World forum on Internet of things, 2018, pp. 813-818, DOI: 10.1109/WF-IoT.2018.8355149.
8. Botta A., Donato W., Persico V., Pescapé A. Integration of cloud computing and internet of things: A survey. Future Generation Computer Systems, 2016, no. 56, pp. 684–700, DOI:10.1016/ j.future.2015.09.021.
9. Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 2013, no. 29, pp. 1645–1660, DOI: 10.1016/j.future.2013.01.010.
10. Miorandi D., Sicari S., Pellegrini F., Chlamtac I. Internet of things: Vision, applications and research challenges. Ad Hoc Networks, 2012, no. 10(7), pp. 1497-1516.
11. Sanabria-Russo L., Pubill D., Serra J., Verikoukis C. IoT data analytics as a network edge service. IEEE Conference on computer communications workshops, 2019, pp. 969-970, DOI: 10.1109/INFCOMW.2019.8845207.
12. Кучерявый А.Е. Интернет вещей / А.Е. Кучерявый // Электросвязь, 2013. – № 1. – С. 21-24.
13. Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Breitenbacher D., Shabtai A., Elovici Y. N-BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Computing, Special Issue. Securing the IoT, 2018, no. 17(3), pp. 12–22.

14. Татарникова Т. М. Обнаружение атак в сетях интернета вещей методами машинного обучения / Т. М. Татарникова, П. Ю. Богданов // Информационно-управляющие системы, 2021. – № 6. – С. 42–52. – DOI: 10.31799/1684-8853-2021-6-42-52.
15. Vaccari I., Chiola G., Aiello M., Mongelli M., Cambiaso E. MQTTset, a new dataset for machine learning techniques on MQTT. Sensors, 2020, no. 20, pp. 1-17, DOI:10.3390/s20226578.
16. Saxena S., Jain S., Arora D., Sharma P. Implications of MQTT connectivity protocol for IoT based device automation using home assistant and OpenHAB. 6th International conference on computing for sustainable global development, 2019, pp. 475-480.
17. Roldán-Gómez J., Carrillo-Mondéjar J., Castelo Gómez J.M., Martínez J.L. Security assessment of the MQTT-SN protocol for the Internet of Things. Journal of Physics: Conference Series, 2021, no. 2224, DOI:10.1088/1742-6596/2224/1/012079.
18. Haripriya A., Kulothungan K. Secure-MQTT: An efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for Internet of Things. EURASIP Journal on Wireless Communications and Networking, 2019, DOI: 10.1186/s13638-019-1402-8.
19. Firdous S.N., Baig Z., Valli C. A. Ibrahim Modelling and evaluation of malicious attacks against the IoT MQTT protocol. Proceedings of the IEEE International Conference on Internet of Things,2017, pp. 748–755, DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData.2017.115.
20. Wood A.D., Stankovic J.A. Denial of service in sensor networks. Computer, 2002, no. 35, pp. 54–62, DOI: 10.1109/MC.2002.1039518.
21. Munshi A. Improved MQTT secure transmission flags in smart homes. Sensors, 2022, no. 22(6), pp. 2–15, DOI: 10.3390/s22062174.
22. Дикий Д.И. Протокол передачи данных MQTT в модели удалённого управления правами доступа для сетей Интернета / Д.И. Дикий, В.Д. Артемьева // Научно-технический вестник информационных технологий, механики и оптики, 2019. – № 19(1). – С. 109–117. – DOI: 10.17586/2226-1494-2019-19-1-109-117.
23. Isaev S.V., Kononov D.D. Analysis of the dynamics of Internet threats for corporate network web services. CEUR Workshop Proceedings, 2021, vol. 3047, pp. 71–78, DOI 10.47813/sibdata-2-2021-10.
24. Kononov D.D., Isaev S.V. Development of secure automated management systems based on web technologies. IOP Conference Series: Materials Science and Engineering, 2019, vol. 537(5), DOI:10.1088/1757–899X/537/5/052024.
25. Лукша М. Kubernetes в действии / М. Лукша. – М: ДМК Пресс, 2018. – 672 с.
26. Rancher 1.6 Docs. Available at: https://rancher.com/docs/rancher/v1.2/en/ (accessed: 28.03.2022).

Исаева Ольга Сергеевна. *К.т.н, с.н.с. отдел Прикладной информатики Институт вычислительного моделирования СО РАН. Область научных интересов: методы искусственного интеллекта, анализ данных, цифровые двойники. SPIN-код: 8412-5807. ORCID: 0000-0002-5061-6765. Researcher ID: A-8905-2018. isaeva@icm.krasn.ru. Россия, г. Красноярск, ул. Академгородок, д. 50, стр. 44.*

Кулясов Никита Владимирович. *Программист первой категории в ИВМ СО РАН. Область научных интересов: аналитика данных, информационная безопасность, распределённые вычислительные системы. AuthorID: 884843, SPIN-код: 9017-5937, ORCID: 0000-0001-5582-9498, WOS Research ID: ADA-6587-2022, razor@icm.krasn.ru. Россия, г. Красноярск, ул. Академгородок, д. 50, стр. 44.*

Исаев Сергей Владиславович. *К.т.н, доцент базовой кафедры вычислительных и информационных технологий Сибирского федерального университета, заведующий отделом Информационно-телекоммуникационных технологий и заместитель директора Института вычислительного моделирования СО РАН по научной работе. Область научных интересов: кибербезопасность и защита информации, интернет-технологии, распределённые информационные системы, интеллектуальные системы. SPIN-код: 1552-8542, AuthorID: 1623, ORCID: 0000-0002-6678-0084, si@icm.krasn.ru, Россия, г. Красноярск, ул. Академгородок, д. 50, стр. 44.*

UDC 004.77
DOI:10.38028/ESI.2022.27.3.011

Creating data collection tools to analyze security aspects Internet of Things

Olga S. Isaeva, Nikita V. Kulyasov, Sergey V. Isaev
Institute of Computational Modeling of SB RAS, Russia, Krasnoyarsk, *isaeva@icm.krasn.ru*

Abstract. The goal of investigation is to create data collection tools for studying the security aspects of information interaction between distributed devices and applications of the Internet of Things (IoT). To achieve the goal, the tasks were solved: a specialized research stand including all functional levels of the IoT architecture was developed, tools for collecting and aggregating data were created and indicators to detect network anomalies were built. The specialized stand includes a sensor level, which consists of measuring devices for monitoring the environment, a transport level implemented on the basis of a corporate network infrastructure, a data collection and storage cluster with various configurations of security settings is deployed for the service level, software for working with data is placed at the application level. The tools collect, aggregate and analyze structured data and logs unstructured on network traffic, taking into account the configuration settings of the security policies of telecommunications nodes. The indicators reflecting the activity and legitimacy of requests with distribution by days, countries and servers have been constructed. The tools are designed for cybersecurity specialists and allow you to analyze the impact of the IoT architecture on the security of the information interaction of network elements.

Keywords: Internet of Things, network anomalies, Kubernetes cluster, messaging protocol, Message Queuing Telemetry Transport (MQTT), Eclipse Mosquitto, Smart environments

Acknowledgements: This work is supported by the Krasnoyarsk Mathematical Center and financed by the Ministry of Science and Higher Education of the Russian Federation in the framework of the establishment and development of regional Centers for Mathematics Research and Education (Agreement No. 075-02-2022-873).

References

1. Lyu W., Liu J. Artificial Intelligence and emerging digital technologies in the energy sector. Applied Energy, 2021, no. 303(9), pp. 117615, DOI: 10.1016/j.apenergy.2021.117615.
2. Training stand for learning the basics of embedded systems programming, Available at: https://www.dlink.ru/stand.html (accessed: 28.03.2022).
3. Roslyakov A.V., Vanyashin S.V., Grebeshkov A.Yu. Internet veshchej: uchebnoe posobie [Internet of things: a tutorial], Samara: PGUTI, 2015, 200 p. (In Russian)
4. Atzori L., Iera A., Morabito G. The internet of things: A survey. Computer Networks, 2010, no. 54(15), pp. 2787–2805, DOI: 10.1016/j.comnet.2010.05.010.
5. Angelo G. Simulation of the Internet of Things. Proceedings of the IEEE 2016 International Conference on High Performance Computing and Simulation, 2016, DOI: 10.1109/HPCSim.2016.7568309.
6. Malceva N.S., Sorokin A.A., Reznikov P.S. [et al.] Modelirovanie gibridnoj topologicheskoj infrastruktury seti «Interneta veshchej» [Modeling the hybrid topological infrastructure of the Internet of Things network]. Nauka, obrazovaniye, innovatsii: puti razvitiya [Science, education, innovation: ways of development], 2019, no. 10, pp. 89–93. (In Russian)
7. Javed A., Heljanko K., Buda A., Främling K. CEFIoT: A fault-tolerant IoT architecture for edge and cloud. IEEE World forum on Internet of things, 2018, pp. 813-818, DOI: 10.1109/WF-IoT.2018.8355149.
8. Botta A., Donato W., Persico V., Pescapé A. Integration of cloud computing and internet of things: A survey. Future Generation Computer Systems, 2016, no. 56, pp. 684–700, DOI:10.1016/ j.future.2015.09.021.
9. Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 2013, no. 29, pp. 1645–1660, DOI: 10.1016/j.future.2013.01.010.
10. Miorandi D., Sicari S., Pellegrini F., Chlamtac I. Internet of things: Vision, applications and research challenges. Ad Hoc Networks, 2012, no. 10(7), pp. 1497-1516.
11. Sanabria-Russo L., Pubill D., Serra J., Verikoukis C. IoT data analytics as a network edge service. IEEE Conference on computer communications workshops, 2019, pp. 969-970, DOI: 10.1109/INFCOMW.2019.8845207.
12. Kucheryavy A.E. Internet veshchej [Internet of Things]. Electrosvyaz [Telecommunications], 2013, no. 1, pp. 21-24. (In Russian)
13. Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Breitenbacher D., Shabtai A., Elovici Y. N-BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Computing, Special Issue. Securing the IoT, 2018, no. 17(3), pp. 12–22.

14. Tatarnikova T. M., Bogdanov P. Yu. Obnaruzhenie atak v setyah interneta veshchej metodami mashinnogo obucheniya [Intrusion detection in internet of things networks based on machine learning methods]. Informatsionno-upravliaiushchie sistemy [Information and Control Systems], 2021, no. 6, pp. 42–52, doi:10.31799/1684-8853-2021-6-42-52. (In Russian)
15. Vaccari I., Chiola G., Aiello M., Mongelli M., Cambiaso E. MQTTset, a new dataset for machine learning techniques on MQTT. Sensors, 2020, no. 20, pp. 1-17, DOI:10.3390/s20226578.
16. Saxena S., Jain S., Arora D., Sharma P. Implications of MQTT connectivity protocol for IoT based device automation using home assistant and OpenHAB. 6th International conference on computing for sustainable global development, 2019, pp. 475-480.
17. Roldán-Gómez J., Carrillo-Mondéjar J., Castelo Gómez J.M., Martínez J.L. Security assessment of the MQTT-SN protocol for the Internet of Things. Journal of Physics: Conference Series, 2021, no. 2224, DOI: 10.1088/1742-6596/2224/1/012079.
18. Haripriya A., Kulothungan K. Secure-MQTT: An efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for Internet of Things. EURASIP Journal on Wireless Communications and Networking, 2019, DOI: 10.1186/s13638-019-1402-8.
19. Firdous S.N., Baig Z., Valli C. A. Ibrahim Modelling and evaluation of malicious attacks against the IoT MQTT protocol. Proceedings of the IEEE International Conference on Internet of Things, 2017, pp. 748–755, DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData.2017.115.
20. Wood A.D., Stankovic J.A. Denial of service in sensor networks. Computer, 2002, no. 35, pp. 54–62, DOI: 10.1109/MC.2002.1039518.
21. Munshi A. Improved MQTT secure transmission flags in smart homes. Sensors, 2022, no. 22(6), pp. 2–15, DOI: 10.3390/s22062174.
22. Dikii D.I., Artemeva V.D. Protokol peredachi dannyh MQTT v modeli udalyonnogo upravleniya pravami dostupa dlya setej Interneta [MQTT data protocol in remote access control management model for Internet networks]. Nauchno-tekhnicheskij vestnik informacionnyh tekhnologij, mekhaniki i optiki [Scientific and Technical Journal of Information Technologies, Mechanics and Optics], 2019, no. 19(1), pp. 109–117, DOI: 10.17586/2226-1494-2019-19-1-109-117. (In Russian)
23. Isaev S., Kononov D. Analysis of the dynamics of Internet threats for corporate network web services. CEUR Workshop Proceedings, 2021, no. 3047, pp. 71–78, DOI: 10.47813/sibdata-2-2021-10.
24. Kononov D.D., Isaev S.V. Development of secure automated management systems based on web technologies. IOP Conference Series: Materials Science and Engineering, 2019, no. 537(5), DOI: 10.1088/1757–899X/537/5/052024.
25. Luksha M. Kubernetes v dejstvii [Kubernetes in action], Moscow, DMK Press, 2018, 672 p. (In Russian)
26. Rancher 1.6 Docs. Available at: https://rancher.com/docs/rancher/v1.2/en/ (accessed 28 March 2022).

***Isaeva Olga Sergeevna**. Ph.D., senior researcher Department of Applied Informatics Institute of Computational Modeling SB RAS. Research interests: artificial intelligence methods, data analysis, digital twins. SPIN: 8412-5807, ORCID: 0000-0002-5061-6765. Researcher ID: A-8905-2018, isaeva@icm.krasn.ru. Russia, Krasnoyarsk, st. Akademgorodok, d. 50, building 44.*

***Kulyasov Nikita Vladimirovich.** Programmer of the first category at the ICM SB RAS. Research interests: data analytics, information security, distributed computing systems. AuthorID: 884843, SPIN code: 9017-5937, ORCID: 0000-0001-5582-9498, WOS Research ID: ADA-6587-2022, razor@icm.krasn.ru. Russia, Krasnoyarsk, st. Akademgorodok, d. 50, building 44.*

***Isaev Sergey Vladislavovich.** Candidate of Technical Sciences, Associate Professor of the Basic Department of Computing and Information Technologies of the Siberian Federal University, Head of the Department of Information and Telecommunication Technologies and Deputy Director of the Institute of Computational Modeling of the Siberian Branch of the Russian Academy of Sciences for scientific work. Research interests: cybersecurity and information protection, Internet technologies, distributed information systems, intelligent systems. SPIN: 1552-8542, AuthorID: 1623, ORCID: 0000-0002-6678-0084, si@icm.krasn.ru. Russia, Krasnoyarsk, st. Akademgorodok, d. 50, building 44.*

Статья поступила в редакцию 11.04.2022; одобрена после рецензирования 16.09.2022; принята к публикации 19.09.2022.

The article was submitted 04/11/2022; approved after reviewing 09/16/2022; accepted for publication 09/19/2022.