REDS:

Телекоммуникационные устройства и системы

№3

2022

СОДЕРЖАНИЕ

ИССЛЕДОВАНИЕ ПРИНЦИПОВ ПОСТРОЕНИЯ И РАЗРАБОТКИ ОНТОЛОГИИ ИНЦИДЕНТОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Бойтунова Майя Дюлустановна,
МИРЭА, Москва, Россия
boytunova2@gmail.com

Шукенбаева Наиля Шаукатовна,
РГГУ, доцент, к.с.-х.н., Москва, Россия
nelshuk@mail.ru

Шукенбаев Айрат Бисенгалеевич,
МТУСИ, МИРЭА, доцент, к.т.н., Москва, Россия
shukenbaev@mail.ru

Аннотация

В работе проведено исследование принципов построения и разработки онтологии инцидентов информационной безопасности. В настоящее время онтология используется не повсеместно в информационных структурах, но рассмотренная модель наглядно показывает востребованность, выявляя скрытые возможности и слабые стороны со стороны информационной безопасности. Применение онтологии позволяет сократить временные затраты на обработку инцидентов информационной безопасности предприятия.

Ключевые слова: *Онтология, инцидент, информационная безопасность, система управления знаниями, управление инцидентами.*

Введение

Вопросы информационной безопасности давно играют особую роль в организациях и обществе, ведь информация – главная единица XXI века. Успех предпринимательской и производственной деятельности состоит не только в правильном распоряжении главным товаром – информацией, но и правильном определении стратегии безопасности. При проведении расследования инцидентов информационной безопасности специалисту для снижения показателей ущерба необходимо своевременно реагировать на атаки информационных ресурсов компании, обрабатывая большое количество информации и принимая незамедлительные решения.

Для управления большими данными об инцидентах необходимы более современные технологии, которые помогут обрабатывать большой объем данных. В настоящее время онтология используется во многих областях, особенно в лингвистике, но нельзя сказать, что в информационных структурах это широко распространенное явление. Обработка инцедентов информационной безопасности занимает достаточно много времени в ИТ-отделах, следовательно, уменьшение затрат времени на их обработку является одной из актуальных задач.

Основная часть

Управление инцидентами является важной составляющей частью управления информационной безопасностью (ИБ) и должно соответствовать требованиям системы менеджмента ИБ (СМИБ) (рис. 1) [1]. Основным двигателем прогресса в усовершенствовании имеющихся средств защиты информации в организации и выработки механизмов грамотного управления инцидентами и исключения рисков информационной безопасности являются накопленные данные при анализе инцидента [2]. Невыполнение безопасности вследствие нарушения целостности, конфиденциальности и доступности системы, сети или услуги будет предполагать собой инцидент, который можно идентифицировать, как сбой, нарушение защиты.

Любое событие, связанное с информационной безопасностью, представляет собой зафиксированный и определенный факт состояния системы, услуги или сети, указывающий на

возможное нарушение политики в области информационной безопасности или отказ средств управления, или ранее неизвестную ситуацию, которая может иметь отношение к безопасности [3].

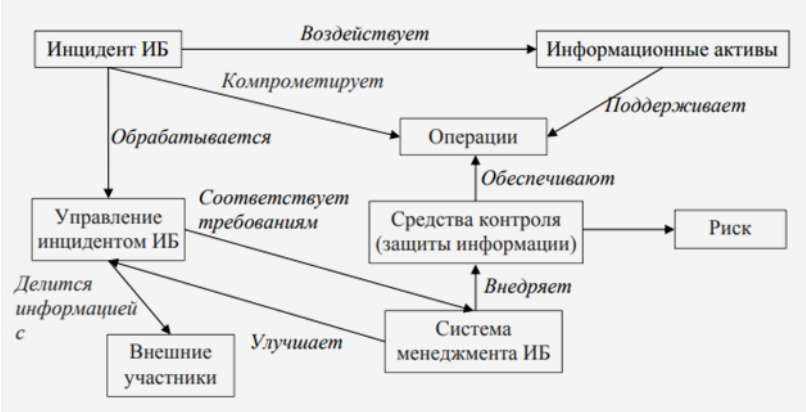

Рис. 1. Управления инцидентами ИБ

Событие, связанное с информационной защищенностью, не всегда означает, что выполненная попытка нарушения информационной безопасности была успешной, или существуют какие-либо изменившиеся результаты относительно состояния конфиденциальности, целостности и/или доступности. Не каждое событие информационной безопасности можно отнести к инцидентам. Событие является инцидентом, если наступило или наступит неблагоприятное последствие этого события.

В стандарте ISO/IEC 27035-1:2016 [6] описана связь объектов в цепи инцидентов ИБ (рис. 2). Когда угроза оказывает неблагоприятное воздействие и использует уязвимости ИС, услуг или сетей, это провоцирует образование событий ИБ и вероятно производит нежелательные инциденты информационных активов, которые также подвержены уязвимости. Затененные объекты на рисунке 2 – это уже существующие объекты; затрагиваемые незатененными объектами, которые вызывают инциденты ИБ.

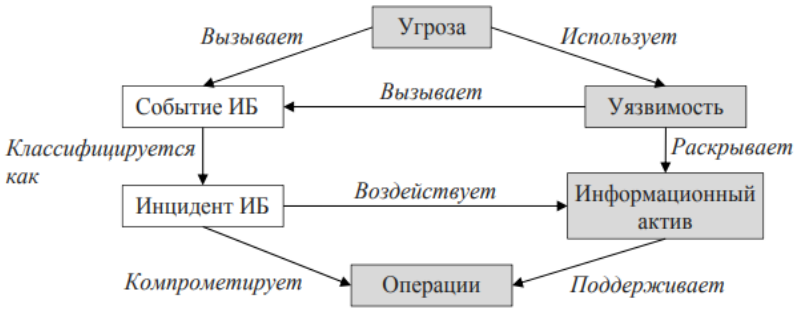

Рис. 2. Связь объектов в цепи инцидентов ИБ

Политики безопасности позволяют учесть различия между отдельными организациями и их профилями рисков. Для разных рассматриваемых предприятий одно и то же действие может быть позволительным для одного, а для другого являться инцидентом безопасности. Использование личных цифровых мобильных устройств при работе с корпоративным приложением, работающих с секретной информацией, является серьезной угрозой возникновения инцидента.

Инциденты ИБ бывают непреднамеренными и преднамеренными в первом случае это может быть следствием ошибки, связанной с человеческим фактором или природно-стихийном случаем, во втором случае осознанной атакой, нарушающей и наносящий вред ПО. Средства, вызывающие инциденты, классифицируются как нетехнические и технические: например, к первым относятся кража или потеря компьютерного оборудования, а ко вторым - заражение, выводящее из строя ПО [5].

Этапы реагирования на инцидент представляют собой циклический процесс (рис. 3). Периодический возврат на предыдущие этапы обеспечивает более точное идентифицированные и описание инцидента, а также вероятность их предотвращения.

Рис. 3. Этапы реагирования на инцидент (NIST SP 800-61)

Каждый инцидент имеет свою категорию, от определения которой зависит способ реагирования на него: будь эта оценка важности ИС и информации, зависимость от нарушения политик ИБ, размера финансового ущерба, степени тревоги и масштабности нарушения. Все мероприятия по удалению инцидента состоят из следующих основных шагов: подготовка, нахождение, удерживание, ликвидация, восстановление, подытоживания опыта и заключения. Как правило, организацией выполняется работа по снижению вероятности возникновения и реагирования инцидентов. В первую очередь все должно начинаться с разработки политики и построения деятельности по управлению инцидентами. Должна быть обязательно группа ответственных лиц, реагирующих на инциденты (IRT, incident response team). Также необходимо своевременное обучение кадров, обязательны настройка и внедрение систем журналирования и защитного резервного копирования, необходим постоянный сбор эмпирических данных о возможных уязвимостях и угрозах безопасности, как внутри системы, так и снаружи. Для этого необходим просмотр и изучение новостных сводок, сетей и систем, внедрение и применения различных средств обнаружения и оповещения об угрозах. Первоначально выявленный инцидент анализируется на предмет - считать ли его таковым. Есть сделаны такие выводы, то инцидент необходимо зарегистрировать.

Следующий шаг заключается в сканировании своей информационной инфраструктуры и выявление затронутых инцидентом участков, которые необходимо по возможности локализовать. При этом возможно отключение пораженных систем или их отдельных функций, передача атаки в «песочницу», смена паролей и т.п. Далее нужно устранить составляющие инцидента – уничтожение наносящих вред программ, обнаружение и нейтрализация затронутых инцидентом учетных записей, обнаружение уязвимостей и их элиминация. При необходимости следует восстановить систему до работоспособного состояния и восстановить все данные из резервных копий. Самым важным моментом после завершения операции является подведение итогов полученного опыта для профилактики реагирования в будущем в похожих случаях и оптимизации эффективности реагирования. Этот опыт может быть сохранен и использован в дальнейшем. А для этого наилучшим решением будет создание онтологии инцидентов информационной безопасности.

Появление онтологий и их быстрое развитие вызывает большое число факторов, среди которых:

1) увеличение в геометрической прогрессии объемов информации, которую нужно проанализировать, обработать кадрами различных отраслей деятельности;
2) чрезмерные насыщенность, повторяемость, противоречивость потоков информации;
3) потребность в сиюминутном использовании одного потока информации для разных целей разными кадрами;
4) глобальная интернетизация современной жизни, огромный хаос информации, требующий её структуризации и возможностей предоставления более результативного поиска;
5) обеспечения качества информационных интернет-услуг и необходимость мгновенного поиска нужной информации.

Все вышеперечисленные факторы служат драйверами к появлению такого понятия, как онтология во многих науках.

Цель создания онтологий определяется в первую очередь её спецификацией, из которой вырисовывается вид потенциальных пользователей и предполагаемого использования. Вид эксплицитной модели основан на концептуализации, опирающейся на структурирование предметных знаний. В основе трансформации (формальной или “вычислительной”) концептуальной модели лежит формализация накопленных знаний.

Выстраивание онтологий невозможно без поддержки процесса разработки её на основе действий, параллельных с разработкой.

Терминология конкретной предметной области описывается состоянием объектов и их связей между с собой – это есть онтология предметной области.

Для описания процесса рационального вывода важно ориентироваться на решение определенных проблем (начиная от отстраненных понятий, принадлежащих к схеме вывода до конкретизированных специальных методов) – это и есть онтология-задача.

Понятия, связанные с определенной проблемной областью и ее задачами, представляются прикладными онтологиями. Объекты предметной области при осуществлении конкретной деятельности играют определенную роль, присущих для обеих соединяемых онтологий

В данной работе описывается онтологию предметной области. Как сказано выше, это не отрицает другие виды онтологий, так как каждая из них является вытекающим следствием другой.

Анализ возможностей онтологических систем позволяет выявить следующие аспекты:

1) Явления, остающиеся неявными, предопределяют рост интеллектуальности системы управления знаниями (СУЗ);

2) Словарь, согласованный с потребителями, служит гарантом стандартизации на основе анализа целевой области;

3) Верификация информации между компьютерными системами и (или) между людьми влечет их повторное или совместное использование для вновь возникающих ситуаций;

4) Классификация информации, полученной и сведенной из разнородных источников, сохраняемой в общем словаре;

5) Воплощение метамодельной функциональности для построения, которая обеспечивается требуемыми понятиями, ограничениями, отношениями, выступающими в роли строительного материала для построения определённых моделей решения задач. [4]

Разработка онтологии включает в себя следующие этапы:

1) разработка требований для определения предметной области, по которой будет строится онтология;

2) определение задач, которые будут решаться при помощи онтологии;

3) определение объектов предметной области и связей между ними – концептуализация модели;

4) представление знаний в форме онтологии.

Три первые операции составляют всю алгебру онтологий.

Два контекста нового в онтологии при слиянии представляют собой пересечение, которое содержит в себе термины, эксплуатируемые в обоих контекстах. Операция содержит задачу выяснения минимального пересечения. Часть более широкого содержания онтологии определяется как контекст, который внутренне достоверен и согласован. На основании отображений для извлечения знаний организовываются контексты в границах каждой предметной области. Различные семантические исправления и уточнения при необходимости вносятся при реструктуризации для показа данных во внутреннем формате. Фрагменты онтологий образуют сформированные контексты, которые существенны для заданной прикладной задачи, которые также являются онтологиями и терминологически определены как соединяемые онтологии.

Организация новой непротиворечивой онтологии – это есть объединение терминов, которое содержит определения, используемые многими контекстами (как итог операции пересечения) и оставшиеся термины контекстов (которые остались за пределами пересечений).

Организация несимметричных подмножеств наборов определений, включенных в один контекст, образуется в результате выполнения операции вычитания одного контекста из другого. Дублирующие определения, используемые в нескольких контекстах, удаляются из результирующей онтологии. В конечном итоге имеются в онтологии только уникальные термины.

Далее представим ряд свойств алгебры онтологий: действия могут сочетаться, изменять последовательность, возможные переупорядочивания могут котироваться, допускается улучшение последовательности выполнения операций, а итоги предыдущих действий могут сохраняться в онтологии и вторично использоваться.

Для организации внутренне согласованного фрагмента начальной онтологии или ее среза предназначается выборка, которая будет вмонтирована в новую онтологию. Контекст — это срез в алгебре онтологий. В качестве исходных данных осознается, что фрагмент онтологии, который будет использоваться в новой онтологии, известен. Максимальный и оптимальный срезы организовываются касательно этой части.

После проведенных сравнительных оценок внешних и внутренних характеристик различных сред для создания онтологий, была выбрана система Protégé. Protégé разработана в Стэндфордском университете и распространяется по лицензии Mozilla Public License (MPL). Свободно распространяется и имеет открытый исходный код. Система позволяет существенно облегчить процесс разработки онтологии. [7]

Онтология прецедентов ИБ была разработана на языке OWL (Web Ontology Language). На рисунке 4 представлен фрагмент полученных результатов. С помощью установленных плагинов FaCT++ и HermiT после ее информационного наполнения для логического вывода проверяется целостность построения. Проверка выполняется через предположения (inferred), которое автоматически вычисляются за счет выстроенной иерархией классов и позволяют определить наличие либо отсутствие противоречий, а также наличие классов, не вошедших в иерархию, что является признаком нарушения целостности онтологии. Представленный на рисунке 4 результат вычисления иерархии классов (вкладка Class hierarchy) совпадает с иерархией, созданной на первом этапе построения OWL-онтологии, и содержит супер класс Thing, включающий в себя все пять классов онтологии. Класс Nothing, предназначенный для классов, не вошедших в иерархию, является пустым.

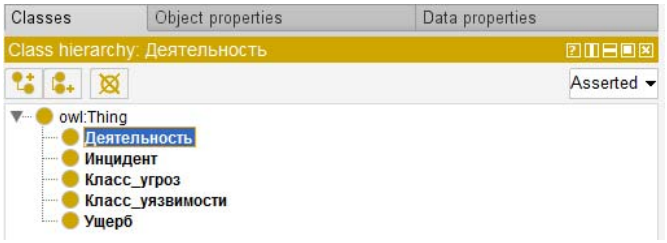

Рис. 4. Иерархия классов, вычисленная машиной вывода HermiT

Таким образом, в онтологии отсутствуют противоречия и классы, не вошедшие в общую иерархию, что подтверждает ее целостность и непротиворечивость. Другой функциональной возможностью среды Protégé является механизм выполнения специальных запросов, построенных в соответствии с синтаксисом OWL, к разработанной онтологии. Она позволяет не только проверять правильность структуры таксономии объектов и иерархии классов в онтологии, но и использовать ее для получения знаний о выбранной предметной области непосредственно при решении прикладных задач. В качестве примеров запросов к разработанной онтологии можно привести следующие:

– Какое решение и когда было вынесено после возникновения инцидента «Перед увольнением менеджер скопировал базу клиентов на свой почтовый ящик»? (рис. 5).
– Какие инциденты возникли 2 декабря 2019 г.? (рис. 6)
– Какие атомарные объекты могли стать причиной возникновения инцидента, на функционирование каких объектов это, в свою очередь, повлияло?
– Какие инциденты приносили высокий ущерб?
– И т.п.

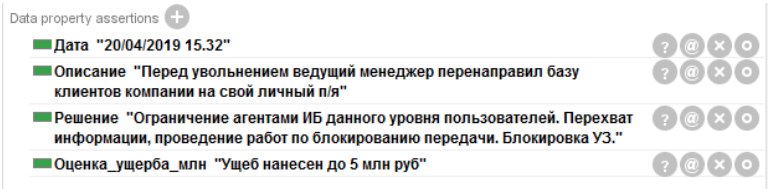

Рис. 5. Результат выполнения поискового запроса

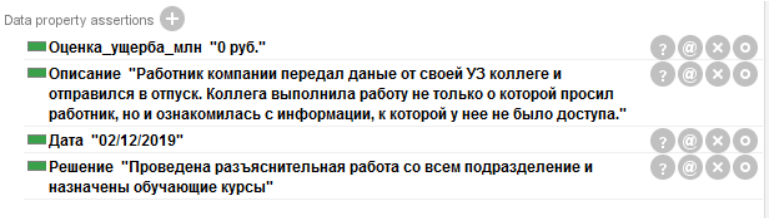

Рис. 6. Результат выполнения запроса

Заключение

Таким образом в работе были проанализированы принципы построения и разработка онтологии инцедентов информационной безопасности, реализован пример построения модели, ориентированной на идентификацию инциденты информационной безопасноти предприятия, а также выработки алгоритма действий в случае возникновения инцедента. Модель разработана для комплексного описания инцинидентов ИБ для их последующей автоматизированной обработки.

Существуют дальнейшие возможности для улучшения создания и качества онтологии предприятия, включая исследование более строгого и систематического подхода к моделированию текущего состояния инцидентов ИБ предприятия (as is) и возможных сценариев будущего состояния (to be) с использованием бизнес-возможностей организации на основе расчетов и выводов по эффективности предложений эффективности. Семантически обусловленные концептуальные модели предприятия также могут быть выражены в ключевых технологиях и системах безопасности, которые поддерживают организацию, формируя набор онтологически осведомленных технологий, которые коллективно реагируют на атаки в отказоустойчивой конфигурации.

Литература

1. ГОСТ Р ИСО/МЭК 27037-2014 Информационная технология. Методы и средства обеспечения безопасности. Руководства по идентификации, сбору, получению и хранению свидетельств, представленных в цифровой форме. – М.: Стандартинформ, 2014. – 48 с.

2. ГОСТ Р ИСО/МЭК 30121-2017 Информационные технологии (ИТ). Концепция управления рисками, связанными с проведением судебной экспертизы свидетельств, представленных в цифровой форме. М.: Стандартинформ, 2017. 12 с.

3. *Васильева И. Н*. Расследование инцидентов информационной безопасности: учебное пособие. СПб.: Изд-во СПбГЭУ, 2019. 113 с.

4. *Лапшин В.А.* Онтология в компьютерных системах. М.: Научный мир, 2011. 224 с.

5. *Шукенбаев А.Б., Шукенбаева Н.Ш., Деркач А.А.* Системы обеспечения безопасности охраняе-мых объектов. Сборник трудов XIV Международной отраслевой научно-технической конференции «Технологи информационного общества». (18-19 марта 2020 г. Москва, МТУСИ). М.: ИД Медиа Паблишер, 2020. 580 с. С. 463-466.

6. Стандарт ISO/IEC 27035-1:2016Информационная технология. Методы и средства обеспечения безопасности. Менеджмент инцидентов информационной безопасности. Часть 1. Принципы менеджмента инцидентов. [Электронный ресурс]. 2017. 28 с. URL: https://nd.gostinfo.ru/document/6260592.aspx (10.12.2021).

7. Официальная документация no Protege [Электронный ресурс]. URL: http://www.cselt.stet.it/fipa/fipa_rationale.htm. (10.10.2021).

ЦИФРОВАЯ ГИГИЕНА ИНФОРМАЦИОННОГО ОБЩЕСТВА: ВЛИЯНИЕ ПАНДЕМИИ COVID-19

Михалевич Игорь Феодосьевич,
Российский университет транспорта (МИИТ), доцент, к.т.н., с.н.с., Москва, Россия
mif-orel@mail.ru

Аннотация
Статья продолжает цикл публикаций, посвященных проблемам ускоренной цифровой трансформации информационного общества, вызванной, в том числе, пандемией COVID-19. В качестве важнейшего фактора безопасного существования и развития информационного общества рассмотрена проблема цифровой гигиены. Приведены примеры нарушения правил цифровой гигиены, повлекшие нарушения информационной (компьютерной, кибер-) безопасности, несущие существенные риски для развития информационного общества

Ключевые слова: *Вредоносное программное обеспечение, информационная безопасность, информационное общество, кибербезопасность, компьютерная безопасность, уязвимость, цифровая гигиена, цифровая трансформация*

Введение

Общемировой тенденцией развития информационного общества является цифровая экономика, переход к которой вызывает многообразные цифровые трансформации. Интегрируя различные взгляды [1-4] дадим следующее определение. Цифровая трансформация (в целом) – это процесс масштабных изменений во всех сферах жизнедеятельности людей посредством приоритетного применения цифровых устройств и цифровых технологий, направленный на повышение качества жизни, оперативности и качества принятия и реализации решений, эффективности производственных (бизнес) процессов, снижение негативного влияния решений (производства, бизнеса) на окружающую среду. Говоря об аппаратной и технологической базе будущего информационного общества, видится их высочайший уровень унификации, обеспечивающий не только функциональность, но и существенное снижение трудозатрат на решение однотипных задач, объективно существующих в различных сферах жизни людей, производства (бизнеса) и экономики в целом.

В российской программе создания цифровой экономики отдельно выделен проект «Информационная безопасность» [5], что указывает на важность данного направления. Целями реализации проекта являются обеспечение устойчивости и безопасности информационной инфраструктуры, создание эффективной системы защиты прав и законных интересов личности, бизнеса и государства от угроз информационной безопасности.

Однако решение задач информационной безопасности неправильно рассматривать узконаправленно, как задач, касающихся исключительно специалистов данного профиля. Это, например, как если бы обеспечение санитарного благополучия страны рассматривалось исключительно как задача санитарных врачей. Учитывая риски масштабных негативных последствий пренебрежительного отношения к проблемам информационной безопасности даже отдельных лиц, необходимо непрерывно формулировать и приводить в соответствие с текущими угрозами «гигиенические» правила поведения в цифровом информационном обществе.

Сферы распространения цифровой гигиены

Под «цифровой гигиеной» будем понимать набор общих правил поведения людей, направленных на снижение рисков нарушения информационной (компьютерной, кибер-) безопасности при массовом распространении и использовании цифровых устройств и цифровых технологий. Данные правила являются общеприменимыми для всех сфер информационного общества, занимающегося построением цифровой экономики, и сохранят свою актуальность в дальнейшем.

Говоря о цифровой гигиене, определим сферы ее распространения, для чего уточним тезаурус.

Согласно [6]:

– информационное общество – общество, в котором информация и уровень ее применения и дос-

тупности кардинальным образом влияют на экономические и социокультурные условия жизни граждан;

– информационное пространство – совокупность информационных ресурсов, созданных субъектами информационной сферы, средств взаимодействия таких субъектов, их информационных систем и необходимой информационной инфраструктуры;

– цифровая экономика – хозяйственная деятельность, в которой ключевым фактором производства являются данные в цифровом виде, обработка больших объемов и использование результатов анализа которых по сравнению с традиционными формами хозяйствования позволяют существенно повысить эффективность различных видов производства, технологий, оборудования, хранения, продажи, доставки товаров и услуг.

Киберпространство – комплексная среда, являющаяся результатом взаимодействия людей, программного обеспечения и услуг в Интернете посредством технологических устройств и сетей, к ним присоединенных, которая не существует в физической форме [7].

Киберпространство – среда информационного взаимодействия и обмена данными, реализуемая в компьютерных сетях и сетях связи. Элементами киберпространства являются сервера, компьютеры, телекоммуникационное оборудование, каналы связи, информационные и телекоммуникационные сети [8].

В [9], с соответствующими ссылками, приведены различающиеся детализированные определения терминов, ориентированные на специалистов соответствующих профилей, а потому не всегда однозначно воспринимаемые с общих позиций. Применительно к цифровой гигиене такая детализация избыточна и даже может быть вредна, учитывая возможности разночтения. Поэтому для общего применения можно предложить следующее объединяющее определение. Информационная безопасность (компьютерная безопасность, кибербезопасность, безопасность информационного пространства, безопасность киберпространства, интернет-безопасность) – это сохранение конфиденциальности, целостности и доступности информации в цифровом информационном пространстве, киберпространстве, компьютерной сети, сети Интернет.

Связи между различными областями безопасности, существующими и возникающими в цифровой среде, представлены на рисунке 1.

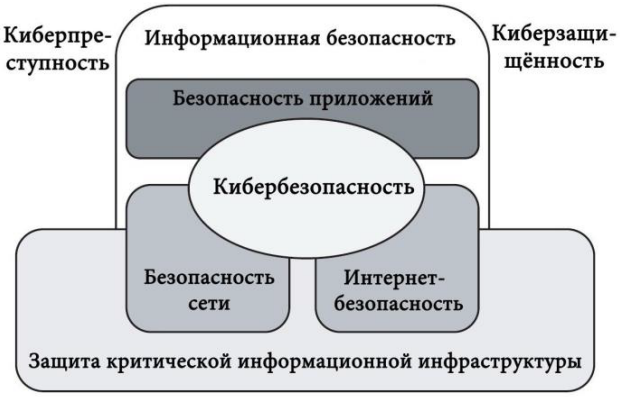

Рис. 1. Связи областей безопасности в цифровой среде (с учетом [7])

Для соблюдения «кибергигиенических» правил необходимо достичь понимании отношений, существующих и возникающих вновь в цифровой среде. Исходя из этого, требования цифровой гигиены необходимо формировать и уточнять (корректировать) в системе критериев, отражающих следующую цепочку состояний, важных для поддержания «здоровья» цифрового мира: ожидания – доверие – риски – ответственность – меры (реагирование).

По всей видимости, мы изначально ожидаем, что цифровой мир безопасен, так как он возник не на голом месте, им занимаются высокие профессионалы и о безопасности они точно подумали.

Это вызывает наше доверие к цифровому миру, который, как прогрессивный, наверняка безопаснее прежнего.

Эти два фактора приводят к тому, что из поля зрения выпадает фактор рисков. Раз что-то в цифровом мире возникло, то оно, наверное, подчинено определенным правилам и их нарушение влечет ответственность.

Ну, а раз, ответственность существует, то и меры примут те, кому положено.

В этой цепочке все может оказаться неверным, что легко объяснимо. Ведь психология человека, нормативное регулирование, реализация защитных мер существенно отстают по времени (не говоря о других факторах), чем скорость наращивания объемов цифрового мира и возникающих в нем угроз.

С учетом изложенного, для общего понимания в таблице 1 приведены основные функции информационной (компьютерной, кибер-) безопасности, сформулированные согласно [10] и многочисленных рекомендаций инженерного совета Интернета (IETF – Internet Engineering Task Force).

Что такое цифровая гигиена и для чего она нужна

Цифровая гигиена направлена на устранение или, по крайней мере, минимизацию негативных последствий угроз информационной безопасности, возникновение которых вызывает наличие различного рода уязвимостей.

Под уязвимостью понимается недостаток или слабость программного (программно-технического) средства или информационной системы в целом, которые могут быть использована для реализации угроз безопасности информации [11]. Уязвимости различаются по областям происхождения, типам и местам их возникновения (выявления).

К областям происхождения относят уязвимости программного кода, конфигурации, архитектуры или организационной структуры объекта. Происхождение уязвимостей может быть связано с комплексными недостатками объекта (многофакторные уязвимости).

Таблица 1

Основные функции информационной (компьютерной, кибер-) безопасности в цифровом мире

Функции безопасности	Назначение
Аутентификация источника данных	Проверка принадлежности данных их автору (источнику)
Аутентификация сторон	Проверка, что стороны информационного взаимодействия действительно является теми, за кого себя выдают
Защита от повтора	Гарантирование одним участником того, что аутентифицированные данные не являются старыми
Авторизация сторон	Представление сторон информационного взаимодействия друг другу и их полномочий доверенной третьей стороной
Конфиденциальность данных	Обеспечение невозможности несанкционированного получения доступа к данным или раскрытия данных
Доступность данных	Обеспечение невозможности прекратить доступ к данным
Целостность данных	Обеспечение возможности проверки, что информация не подверглась несанкционированной модификации или разрушению
Неотказуемость	Обеспечение невозможности отказа одной из сторон от факта участия в информационном обмене (полностью или в какой-либо его части: отправления и/или получения данных))
Подотчетность	Предоставление гарантий, что действия системных субъектов могут быть однозначно прослежены теми субъектами, кто отвечает за эти действия
Разграничение доступа	Обеспечение невозможности несанкционированного использования ресурсов цифрового устройства пользователя и/или домашней (корпоративной) сети

По типам уязвимости классифицируются исходя из недостатков, связанных, с неправильной настройкой параметров программного обеспечения, неполнотой проверки вводимых данных, подменой межсайтовых запросов, межсайтовым скриптингом (выполнением сценариев), переполнением буфера памяти, неконтролируемой форматной строкой, вычислениями, неправильными криптографическими преобразованиями, внедрением интерпретируемых операторов языков программирования, разметки или произвольного кода, команд операционной системы, а также с возможностью прослеживания пути доступа к каталогам и перехода по ссылкам, управлением учетными данными, аутентификацией, разрешениями, привилегиями, ресурсами и другими недостатками, приводящими к утечке (раскрытию) информации ограниченного доступа, модификации информации или ее недоступности (удалению, отказам в обслуживании) и т.п.

Местами возникновения (выявления) уязвимостей могут быть программное обеспечение (неважно, базовое, системное, прикладное или специальное), средства обработки информации, сетевое (коммуникационное, телекоммуникационное) оборудование, а также средства защиты информации.

Картина киберугроз постоянно меняется. Некоторые угрозы исчезают. Другие, наоборот, совершенствуются, приобретая дополнительные свойства, не утрачивая прежние (так называемые расширенные постоянные угрозы, Advanced Persis-tent Threats – APTs).

Активность и направленность действий киберпреступности во многом определяется состоянием защищенности активов (ресурсов), на которые направлены их интересы. Зачастую успешными становятся кибератаки на объекты, бреши в киберзащите которых вызваны «нецифрогигиеничным» поведением легальных пользователей. Такие пользователи плохо осведомлены о видах угроз, хотя бы наиболее распространенных, и способах их распространения. Такая информация регулярно доводится до сведения широкой публики в СМИ, но, к сожалению, чаще всего в связи с кибератаками, уже достигшими громкого успеха.

Такие «слуги» киберпреступности, как сетевые боты, давно у всех на слуху и с ними многие научились успешно бороться, занося сведения о них в спам или «черные» списки (списки для блокировки). Но беда в том, что источники, с которых может действовать один и тот же бот, постоянно меняются и, в конце концов, в атакуемую систему может попасть вредоносное программное обеспечение.

Киберпреступностью все чаще и все более изощренно используется сторонняя инфраструктура. Теперь это уже не только облачные хранилища, но и коммуникационные платформы, предоставляющие удобные для автоматизации сервисы, например, Google, Discord, Telegram [12].

Инфраструктура как сервис для вредоносных программ и ботнетов находится на подъеме благодаря использованию обычных маршрутизаторов и устройств IoT. Злоумышленники поняли, что гораздо проще использовать эти устройства в качестве прокси для сокрытия вредоносной активности, чем создавать специальное вредоносное программное обеспечение для такого разнообразия архитектур и устройств. Поэтому наблюдается продажа ботнетов «порабощенных» устройств в качестве услуги различным злоумышленникам. Примером является прошедшая в 2021 г. кампания под названием Meris, в рамках которой были использованы уязвимости роутеров Mikrotik для DDoS-атак на серверы Яндекса. Как оказалось, эта атака была лишь одной из кампаний, проводимых через ботнет MikroTik как сервис, предоставляющий анонимные прокси. Выяснилось, что ботнет состоит примерно из 200 тыс. «порабощенных» устройств с открытым SOCKS-прокси, готовых к найму на форумах «теневого» интернета (даркнета) [12]. (Примечание: SOCKS - это интернет-протокол, который позволяет незаметно для клиента и конечного сервера пересылать данные с помощью промежуточного сервера (прокси-) таким образом, что для конечного сервера данные, поступившие от SOCKS-прокси, воспринимаются в качестве данных, отправленных клиентом).

Примечательно, что этот ботнет существует с 2018 г. Несмотря на то, что его большая часть отключена, уязвимости в маршрутизаторах Mikrotik, похоже, еще не устранены. Вектор атаки хорошо известен и является общим для большинства устройств Интернета вещей и маршрутизаторов – это **неисправленная прошивка и учетные данные по умолчанию.**

Казалось бы, что может быть проще, чтобы данные каналы для кибератак устранить. Но нет. Недостатки в организации защиты информации (читай – обеспечения кибербезопасности) были и остаются крайне распространенными. Организационные меры защиты если не игнорируются полностью, то зачастую носят формальный характер. К сожалению, это проявляется не только в общей сфере, но иногда наблюдается и в подготовке специалистов по кибербезопасности и при выполнении работ, связанных с обязанностью принятия мер по киберзащите.

Что может быть более чувствительным для любого члена информационного общества, чем состояние здравоохранения. Поэтому организации здравоохранения все чаще становятся объектами кибератак, посредством которых может, например, блокироваться доступ персонала к медицинским записям, устройствам и другим цифровым инструментам с целью требования выкупа. Или может обманываться медицинский искусственный интеллект, которому будут передаваться неправильные снимки, что грозит постановкой неправильного диагноза. Понимание этого влечет увеличение вложений в кибербезопасность медицинских учреждений, однако угрозы безопасности, связанные с давно эксплуатируемым оборудованием, а также современными устройствами интернета-вещей и интернета медицинских вещей (IoT и IoMT), позволяют злоумышленникам успешно использовать программы-вымогатели и осуществлять другие кибератаки.

Исследование более 10 миллионов устройств в 300 больницах и медицинских учреждениях в США и по всему миру показало [13]:

– более 50% подключенных к Интернету устройств в типичной больнице подвержены критическим рискам и могут быть взломаны (рис. 2);

a) устройства IoT (Internet of Things), используемые при лечении пациентов

b) устройства IoMT (Internet of Medical Things)

c) устройства OT (Operational Technology, операционные технологии), используемые в системах отопления, вентиляции и кондиционирования воздуха, лифтов, электрических сетей и т.п.

d) подключенные устройства, не используемые непосредственно в лечении пациентов (IP-телефоны, выключатели света, кофеварки и т.п.)

Рис. 2. Подключенные устройства типичной больницы

– наиболее распространенными подключенными к интернету устройствами в больницах являются инфузионные насосы. На них приходится 12% от всех подключенных устройств и 38% от всех устройств IoMT, обнаруженных при исследовании. Их распространение обусловлено потребностью большинства пациентов во введении жидкостей во время пребывания в больнице. Уменьшить погрешность при совершении этих процедур медицинскому персоналу помогает применение инфузионных насосов, позволяющих регулировать количество и скорость вводимых жидкостей. Однако, 73% инфузионных насосов оказались уязвимы для взлома, что несет серьезные риски. Данные уязвимости позволяют удаленно получить несанкционированный доступ к электронным медицинским записям, внести в них изменения, например, изменить дозировку препарата, и, тем самым, причинить прямой вред здоровью пациента. При этом несанкционированный доступ к медицинским книжкам сам по себе является нарушением закона о защите персональных данных (их конфиденциальности);

– другими популярными подключенными к интернету медицинскими устройствами являются кардиомониторы и аппараты для УЗИ-диагностики, которые также входят в первую десятку по количеству уязвимостей. При этом 53% устройств в отделениях онкологии, 65% и 50% устройств отделениях фармакологии и лабораториях работают на старых версиях Windows, **которые уже не обновляются** (рис 3);

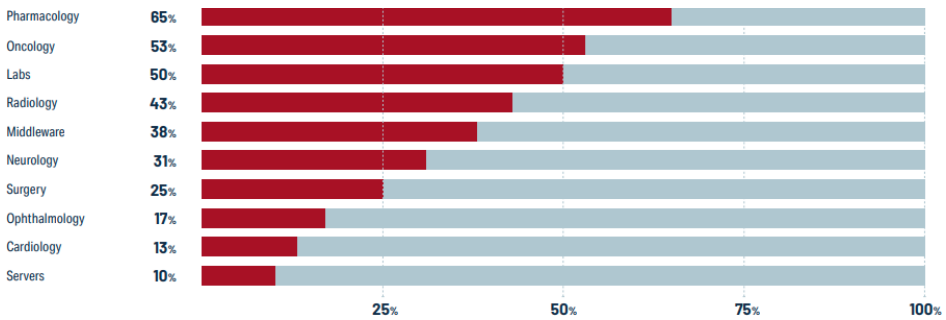

Рис. 3. Больничные отделения, использующие больше всего устройств с устаревшими версиями Windows (от общего числа устройств в отделении)

– наиболее распространенным риском для подключенных устройств остаются **ненадежные пароли**. Устройства поставляются с паролями и настройками по умолчанию, которые приведены в инструкциях и, как правило, общедоступны. **Пренебрежение рекомендацией изменить пароли и настройки, установленные производителями по умолчанию,** делает IoMT, IoT и другие подключенные устройства действующих объектов уязвимыми для атак.

Сведения о прямых атаках на подключенные к интернету медицинские устройства пока отсутствуют, но это не означает, что они невозможны или их не было. Понять истинную причину нарушения работы подключенного к интернету медицинского устройства зачастую является задачей нетривиальной, что обусловлено, в том числе, проблемой цифрового неравенства автоматизированных систем корпоративного и технологического управления, рассмотренной в [1,9]. Поэтому более известны атаки, связанные с проникновением через уязвимые устройства в информационные системы больниц и их блокировкой от сети. При этом медицинский персонал остается без доступа к медицинским записям, устройствам и другим цифровым инструментам, а злоумышленники требуют выкуп за их разблокировку.

Так, например, в сентябре 2020 г. кибератаке подверглась сеть больниц Universal Health Services (UHS), имеющая более 400 отделений, в основном в США [14]. Компьютеры пользователей перестали загружаться и на них появлялось требование о выкупе. Universal Health Services разместила на своем веб-сайте заявление о том, что ее общекорпоративная сеть «в настоящее время отключена из-за проблемы с ИТ-безопасностью. В марте 2020 г. компания сообщила, что эта атака обошлась ей в 67 миллионов долларов США, включая расходы на восстановление данных, упущенную прибыль из-за простоя и снижения потока пациентов [15].

В это же время была совершена кибератака в Германии [16]. Она привела к сбою в работе ИТ-системы крупной больницы Дюссельдорфа: медперсонал не мог получить доступ к данным, экстренных пациентов перевезли в другое место, операции отложили, а женщина, которая нуждалась в срочной госпитализации, умерла после того, как ее пришлось доставить в другой город для лечения. Расследование показало, что 30 серверов в больнице были зашифрованы, на одном из серверов была записка о вымогательстве, адресованная Университету имени Генриха Гейне, к которому относится дюссельдорфская больница, а не самой больнице. Полиция сообщила преступникам, что пострадала больница, а не университет, и это поставило под угрозу пациентов. После этого злоумышленники отозвали попытку вымогательства и предоставили цифровой ключ для расшифровки данных.

В связи с этим считается, что это может быть первая известная смерть от кибератаки вымогателей на медицинское учреждение.

Влияние пандемии COVID-19 на цифровой мир

Приведенные выше примеры свидетельствуют, что несоблюдение простых правил цифровой гигиены создает существенные риски не только для экономики, но и для жизни людей.

Пандемия COVID-19 вызвала массовый переход к удаленной работе и бурный рост электронной коммерции. При этом возможности специалистов по информационной безопасности влиять на полноценное обеспечение защитных мер были и остаются ограниченными, как минимум, временным фактором.

Влияние пандемии COVID-19 на кибербезопасность цифрового мира можно проследить с помощью данных рейтинга десяти наиболее распространенных типов уязвимостей веб-приложений, представленного на рис. 4 [17].

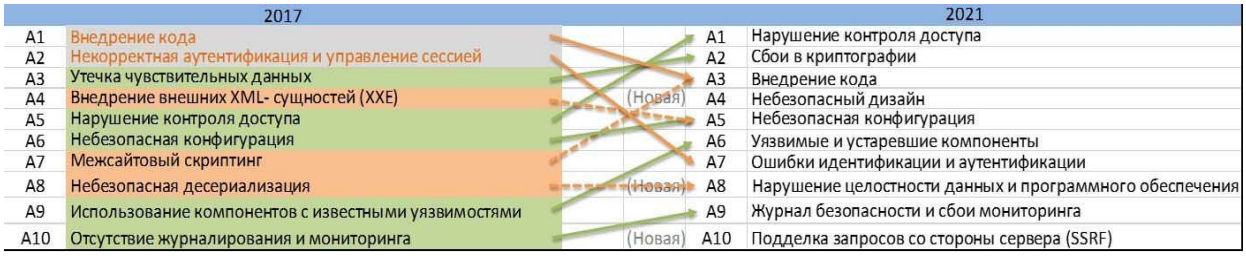

Рис. 4. Рейтинг Top 10 уязвимостей Web-приложений в 2021 г. по версии OWASP

Если в 2017 г. уязвимости, связанные с нарушением контроля доступа, находились на пятом месте рейтинга, то к 2021 г. они уверено заняли первое место. Повышение внимания к безопасной разработке программ привело к уменьшению числа уязвимостей кода, однако они по-прежнему весьма распространены и занимают третье место. В 2021 г. в рейтинг введена новая категория под названием «Небезопасный дизайн». В ней основное внимание уделяется рискам, связанным с недостатками проектирования, в частности, в области моделирования угроз, моделей и принципов безопасного проектирования и эталонных архитектур. Здесь же мы видим перемещение с девятого места в 2017 г. на шестое место к 2021 г. программного обеспечения, использующегося с известными уязвимостями или устаревшими компонентами.

Приведенные данные рейтинга легко объяснимы с позиций цифровой гигиены. В сфере информационной безопасности есть правила, не требующие специальных знаний, но обязательные к применению, как общие правила санитарной гигиены. Например, такое общее правило санитарной гигиены, как «Мойте руки перед каждым приемом пищи, с моющими средствами, чистой водой», можно интерпретировать в область надежности пароля.

Под надежностью пароля будем понимать его способность противостоять несанкционированному доступу к защищаемому цифровому ресурсу (устройству, программе, сети и т.д.). Для обеспечения надежности пароля должны соблюдаться три правила, по смыслу схожие с вышеприведенными гигиеническими правилами. Пароль должен удовлетворять критериям сложности, сменяемости и неповторяемости. Сложность определяется размерностью (длиной) и составом пароля. Например, требование сложности может быть выражено следующим образом: длина пароля – не менее 8 символов; состав – строчные и прописные буквы, цифры, знаки препинания и специальные символов, представленные в каждом виде в количестве не менее одного символа. Правило регулярной смены пароля необходимо с позиций времени его безопасной жизни, то есть времени, в течение которого взлом пароля считается практически невозможным при использовании злоумышленником обычных вычислительных средств и соблюдении легальными пользователями правил цифровой гигиены. Например, пароль нужно менять не реже одного раза в три месяца. И, наконец, нельзя повторно применять пароль, ранее хоть раз использовавшийся.

Заключение

Требования цифровой гигиены просты и легко применимы, но их несоблюдение несет риски не только для конкретного человека, но и для больших групп и, в отдельных случаях, общества в целом.

Есть правила, применение которых требует определенной подготовки, которые выходят за рамки данной статьи. Здесь же обращено внимание на важность неукоснительного соблюдения простейших правил, обеспечивающих, в частности, надежность парольной защиты.

Другим простейшим правилом цифровой гигиены является использование устройств с актуальными версиями прошивок, операционных систем и иных обновляемых программ. Если это правило соблюсти не представляется возможным, то на данном устройстве недопустимо использование программ, хранение и обработка чувствительной информации (например, банковских приложений, хранение учетных данных, данных бакновских карт и т.п.).

В список простейших правил цифровой гигиены входит также обязательность смены на всех без исключения устройствах заводских паролей и настроек на уникальные для конкретного объекта (пусть то организация или дом).

Соблюдение даже этих простейших правил защитит каждого члена информационного общества от риска, что его личная жизнь не станет достоянием общественности [17], персональный компьютер не станет распространителем ботов или зловредного программного обеспечения, а робот-пылесос, холодильник, утюг или иное «умное» устройство - участником DDOS-атаки.

Литература

1. *Михалевич И.Ф.* Цифровая трансформация систем управления в условиях пандемии COVID-19 // REDS: Телекоммуникационные устройства и системы. 2021. № 4. С. 25-31.

2. *Прохоров А., Коник Л.* Цифровая трансформация. Анализ, тренды, мировой опыт. М.: ООО «Альянс-Принт», 2019. 368 с.

3. *Буряк В. В.* Цифровая экономика, хактивизм и кибербезопасность: Монография. Симферополь: ИП Зуева Т.В., 2019. 140 с.

4. Цифровая экономика: Учебник / под ред. Л. А. Каргиной. М.: Прометей, 2020. 220 с.

5. Национальный проект «Национальная программа «Цифровая экономика Российской Федерации»» (утв. президиумом Совета при Президенте РФ по стратегическому развитию и национальным проектам, протокол от 04.06.2019 № 7). URL: https://legalacts.ru/doc/pasport-natsionalnoi-programmy-tsifrovaja-ekonomika-rossiiskoi-federatsii-utv-prezidiumom/ (дата обращения 19.01.2022).

6. Стратегия развития информационного общества РФ на 2017-2030 годы (утверждена Указом Президента РФ от 09.05.2017 г. № 203).

7. ISO/IEC 27032:2012. Information technology — Security techniques — Guidelines for cybersecurity.

8. СТО РЖД 08.021-2015. Автоматизированные системы управления технологическими процессами и техническими средствами железнодорожного транспорта. Требования к функциональной и информационной безопасности программного обеспечения. Порядок оценки соответствия.

9. *Михалевич И.Ф.* Проблема цифрового неравенства автоматизированных систем корпоративного и технологического управления // REDS: Телекоммуникационные устройства и системы. 2020. № 3. С. 43-47.

10. ГОСТ Р ИСО 7498-2-99. Информационная технология. Взаимосвязь открытых систем. Базовая эталонная модель. Часть 2. Архитектура защиты информации.

11. ГОСТ Р 56546-2015. Защита информации. Уязвимости информационных систем. Классификация уязвимостей информационных систем.

12. Avast Q3’21 Threat Report. URL: https://decoded.avast.io/threatresearch/avast-q321-threat-report/ (дата обращения 08.01.2022).

13. Research Report: The State of Healthcare IoT Device Security 2022. URL: https://www.cynerio.com/landing-pages/the-state-of-healthcare-iot-device-security-2022 (дата обращения 24.01.2022).

14. Major hospital system hit with cyberattack, potentially largest in U.S. history. URL: https://www.nbcnews.com/tech/security/cyberattack-hits-major-u-s-hospital-system-n1241254 (дата обращения 26.01.2022).

15. Шифровальщики: атаки на здравоохранение. URL: https://www.kaspersky.ru/blog/ransomware-vs-healthcare/30604/ (дата обращения 26.01.2022).

16. German hospital hacked, patient taken to another city dies. https://www.nbcnews.com/tech/security/german-hospital-hacked-patient-taken-another-city-dies-rcna125 (дата обращения 26.01.2022).

17. OWASP Top 10-2021. URL: https://owasp.org/Top10/ (дата обращения 19.01.2022).

18. Тысячи личных видео россиян с тайных камер наблюдения слили в Сеть. REN TV 30.11.2021. URL: https://ren.tv/news/kriminal/910018-tysiachi-lichnykh-video-rossiian-s-tainykh-kamer-nabliudenii-slili-v-set (дата обращения 09.01.2022).

АНАЛИЗ СПОСОБОВ КЛАССИФИКАЦИИ ТРАФИКА В ПРОГРАММНО-КОНФИГУРИРУЕМЫХ СЕТЯХ (SDN) С ПРИМЕНЕНИЕМ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Орешин Андрей Николаевич,
Академия ФСО России, сотрудник, к.т.н., доцент, Орёл, Россия

Андреев Сергей Юрьевич,
Академия ФСО России, сотрудник, к.т.н., Орёл, Россия
us12a@mail.ru

Аннотация
В данной работе рассматриваются вопросы классификации мультимедийного трафика данных в транспортных сетях с коммутацией пакетов оператора связи, которые базируются на технологии программно-конфигурируемой сети (SDN, Transport-SDN). Данная технология позволила применить методы машинного обучения с целью интеллектуального изучения трафика данных и на основе этого анализа решать вопросы управления инфраструктурой сети.

Ключевые слова: *машинное обучение, программно-конфигурируемая сеть (SDN), классификация, трафик данных, качество обслуживания (QoS), контроллер программно-конфигурируемой сети, приложение.*

Введение

В последние годы, в связи с быстрым развитием современных технологий интернета и мобильной связи, инфраструктура, устройства и сетевые ресурсы становятся все более сложными и разнородными. Развитие технологии Интернета вещей привело к высокому росту мультимедийного трафика по запросу (аудио, видео и изображения), который является чувствительным к задержкам и требует большей пропускной способности к мультимедийные приложения. Все это предъявляет новые требования с точки зрения обеспечения качества обслуживания и эффективной транспортировке данных к транспортным сетям с коммутацией пакетов, с учетом современных тенденций развития технологии. Для эффективной организации управления, обслуживания и оптимизации транспортных сетей необходимо использовать больше интеллектуальных данных. Однако в сетях связи с традиционной архитектурой управления методы машинного обучения (Machine Learning, ML) трудно применить с целью контроля и управления этими сетями. Построения транспортных сетей с коммутацией пакетов провайдера связи на базе технологии программно-конфигурируемой сети (Software Defined Networking, SDN) предоставила новые возможности для интеллектуального контроля и управления инфраструктурой. Возможности SDN такие, как логически централизованное управление, глобальное представление сети, анализ трафика на уровне программного обеспечения, динамическое обновление правил пересылки и другие облегчают применение методов машинного обучения.

Классификация трафика

Классификация трафика является важной сетевой функцией, которая обеспечивает способ выполнения детального управления сетью путем определения различных типов потоков трафика. С помощью классификации трафика сетевые операторы могут обрабатывать различные услуги и более эффективно распределять сетевые ресурсы.

Широко используемые методы классификации трафика включают подход на основе портов, глубокую проверку пакетов (Deep Packet Inspection, DPI) и машинное обучение [1-3]. Подход на основе портов использует номера портов TCP (Transmission Control Protocol) и UDP (User Datagram Protocol) для определения приложений. В прошлом многие приложения использовали хорошо известные порты, такие как TCP-порт 80 для протокола HTTP. В настоящее время большинство приложений работают на динамических портах, что делает подход на основе портов более неэффективным.

Метод глубокой проверки пакетов (DPI) сопоставляет полезную нагрузку потоков трафика с предопределенными шаблонами для определения приложений, которым принадлежат потоки трафика. Шаблоны определяются регулярными выражениями. Подход, основанный на DPI, как правило, обладает высокой точностью классификации. Однако у него есть некоторые недостатки. Во-первых, DPI может распознавать только приложения, шаблоны которых доступны. Экспоненциальный рост приложений делает обновление шаблона сложным и непрактичным [4]. Во-вторых, DPI сопряжен с высокими вычислительными затратами, поскольку необходимо проверять все потоки трафика. В-третьих, DPI не может классифицировать зашифрованный трафик в Интернете.

Подходы, основанные на машинном обучении, могут правильно распознавать зашифрованный трафик и требуют гораздо меньших вычислительных затрат, чем подход, основанный на DPI. Таким образом, подходы, основанные на ML, были широко изучены. Для классификации трафика сначала собирается большое количество транспортных потоков, а затем применяются методы ML для извлечения знаний из собранных транспортных потоков. В SDN контроллер имеет представление глобальной сети, что облегчает сбор и анализ трафика. Таким образом, подходы на основе ML, как правило, реализуются в контроллере. Было проведено много исследований для классификации трафика с разных точек зрения, таких как классификация трафика с учетом потока, приложений и QoS. В этом подразделе мы обобщим соответствующие исследования.

Классификация трафика с учетом потока слонов

Классификация трафика с учетом потока «слонов» (Elephant Flow-aware Traffic Classification) направлена на идентификацию потоков «слонов» (elephant) и потоков «мышей» (mice). Потоки «слонов» – это долговременные потоки трафика данных, требующие большой пропускной способности, в то время как потоки «мышей» – это кратковременные потоки трафика данных, требовательные к задержкам. В центре обработки данных 80% потоков трафика составляют потоки «мышей». Однако большая часть байтов переносится в потоках «слонов» [5]. Для эффективного управления потоками трафика в центрах обработки данных необходимо идентифицировать потоки «слонов».

Исследование проблемы планирования потока трафика в сети гибридного центра обработки данных представлены в [6]. Во-первых, методы машинного обучения используются для классификации трафика с учетом потока на границе сети. Затем централизованный контроллер SDN может использовать результат классификации для реализации эффективных алгоритмов оптимизации потоков трафика данных.

В SDN предлагается метод обучения, учитывающий затраты, для обнаружения потоков «слонов» [7]. Предлагаемая стратегия обнаружения потока «слонов» состоит из двух этапов. На первом этапе используется измерение «головных» пакетов, чтобы отличить подозрительные потоки «слонов» от потоков «мышей». На втором этапе используется дерева принятия решений (Decision Tree, DT) [8] в качестве метода обнаружения для анализа того, являются ли эти подозрительные потоки слонов потоками слонов или нет.

Классификация трафика с учетом приложений

Классификация трафика с учетом приложений (Application-aware Traffic Classification) направлена на определение приложений транспортных потоков. В [1] представлены материалы исследования классификации трафика с учетом приложений в корпоративной сети. Сеть на базе технологии SDN с использованием протокола OpenFlow развертывается в корпоративной сети для сбора данных о трафике. Затем применяются алгоритмы классификации с целью отнесения потоков трафика к различным приложениям. В исследованиях [9] предлагается мультиклассификатор для идентификации приложений путем объединения классификатора на основе машинного обучения (ML) и классификатора на основе глубокой проверки пакетов (DPI). При поступлении нового потока сначала выбирается классификатор на основе машинного обучения (ML) для выполнения классификации. Если точность результата классификатора на основе машинного обучения (ML) превышает пороговое значение, то это будет результат мультиклассификатора напрямую. В противном случае будет выполнена классификация на основе глубокой проверки пакетов (DPI). Если классификатор на основе глубокой проверки пакетов (DPI) не возвращает в качестве результата «неизвестно», то его результат будет выбран в качестве результата мультиклассификатора.

Для классификации приложений, работающих по протоколу UDP, предлагается механизм поведенческой классификации, обеспечивающий точную классификацию трафика с учетом приложений [10]. В частности, алгоритм опорных векторов (Support Vector Machine, SVM) используется для классификации UDP–трафика в соответствии с записями Netflow (например, количеством принятых пакетов и байтов). Результаты моделирования показывают, что точность классификации предлагаемого механизма классификации на основе алгоритма опорных векторов SVM составляет более 90%.

Для идентификации мобильных приложений предлагается структура, называемая Atlas [11]. Для сбора достоверных данных от конечных устройств используется подход краудсорсинга. Собранные данные используются для обучения дерева решений. Обученная модель способна идентифицировать мобильные приложения транспортных потоков. Результаты моделирования показывают, что средняя точность классификации Atlas составляет более 94% для 40 лучших приложений в Google Play [2].

Глубокая нейронная сеть (Deep Neural Network, Deep NN) используется для идентификации мобильных приложений. Трафик мобильной сети собирается из экспериментальной сети. Пять критериев потока такие, как адрес назначения, порт назначения, тип протокола, поле TTL и размер пакета выбираются для обучения восьмиуровневой модели глубокой нейронной (Deep NN). Результаты моделирования показывают, что обученная модель может достичь точности 93,5% для идентификации 200 мобильных приложений.

Мобильное приложение часто генерирует различные типы потоков. Например, приложение Facebook может генерировать потоки видео, голоса, мгновенных сообщений (Instant Messaging, IM) и обмена файлами и так далее. Чтобы обеспечить детальную классификацию трафика с учетом мобильных приложений, необходимо определить, как сами мобильные приложения, так и типы потоков. В беспроводной пограничной сети с поддержкой SDN предлагается детализированная система классификации трафика, поддерживаемая мобильными приложениями, называемая Traffic Vision [12]. Процессор Traffic Vision (телевизионный движок) является основным компонентом Traffic Vision. Высокоуровневая архитектура Traffic Vision с поддержкой SDN и рабочий процесс телевизионного движка показаны на рисунке 1. Телевизионный движок выполняет три основные задачи: сбор, хранение и извлечение статистики потоков и данных наземного обучения с конечных устройств и устройств доступа; классификатор на основе алгоритма дерева решений используется для идентификации имени приложения, такого как YouTube, Facebook, Amazon и т.д.; классификатор на основе алгоритма k-ближайших соседей (k-Nearest Neighbor) применяется для идентификации типов потоков, таких как видеоконтент, аудиофайл, видеочат и т.д.

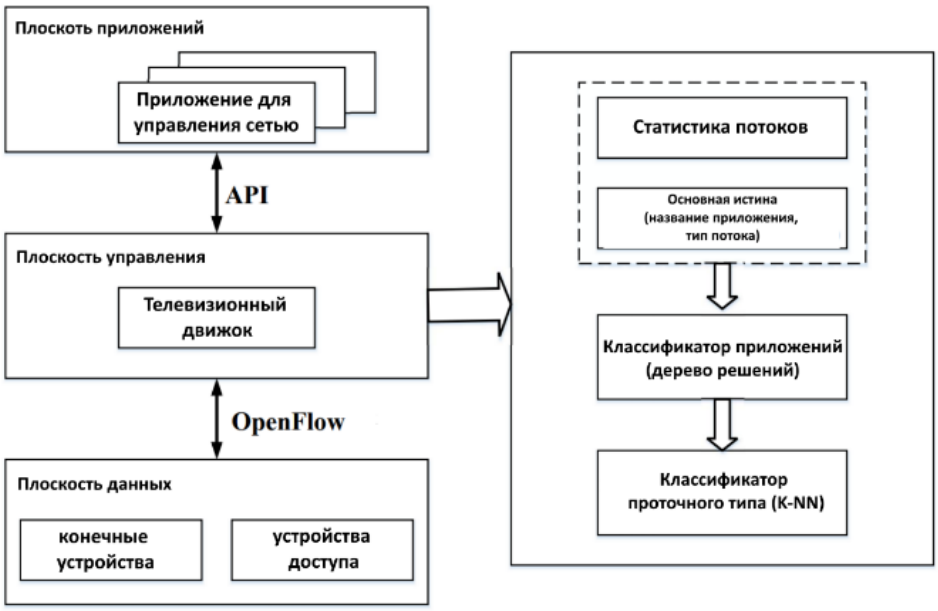

Рис. 1. Высокоуровневая архитектура TrafficVision с поддержкой SDN и рабочий процесс телевизионного «движка»

Классификация трафика с учетом QoS

Классификация трафика с учетом качества обслуживания QoS (QoS-aware traffic classification) направлена на определение классов качества обслуживания (QoS) транспортных потоков. В условиях экспоненциального роста числа приложений в Интернете идентифицировать все приложения сложно и непрактично. Однако приложения могут быть разделены на различные классы обслуживания (QoS) в соответствии с их требованиями к качеству обслуживания (QoS), например, задержка пакетов, дрожание (джиттер) пакетов и потери пакетов. К классу обслуживания (QoS) может принадлежать множество различных приложений. Таким образом, данная классификация является более эффективным способом классификации транспортных потоков в соответствии с их требованиями к качеству обслуживания (QoS) [3]. Также предложена система классификации трафика, учитывающая качество обслуживания (QoS), с использованием алгоритмов контролируемого обучения и глубокой проверки пакетов (DPI). Алгоритм глубокой проверки пакетов (DPI) применяется для обозначения части потоков трафика известных приложений. Затем помеченный обучающий набор данных используется алгоритмами контролируемого обучения с, такими как алгоритм опорных векторов (SVM) для классификации потоков трафика неизвестных приложений. Таким образом, потоки трафика как известных, так и неизвестных приложений подразделяются на различные классы QoS. Результаты моделирования показывают, что данная подход обладает высокой точностью классификации (около 90%).

Заключение

Классификация трафика с учетом потоков «слонов» часто применяется в центрах обработки данных. Методы детальной классификации трафика, т. е. классификации трафика с учетом приложений и качества обслуживания (QoS) могут увеличить задержку обработки трафика данных, поэтому они не подходят для центров обработки данных.

Классификация трафика с учетом приложений часто применяется для мелкозернистого трафика и управления сетью с использованием этой информации. Однако в условиях экспоненциального роста числа приложений в Интернете идентифицировать все приложения нецелесообразно. Существующие работы определяют только наиболее популярные приложения.

Классификация трафика с учетом качества обслуживания (QoS) может использоваться сетевыми операторами с целью оптимизации распределения сетевых ресурсов для потоков трафика в соответствии с их классами обслуживания (QoS).

В целом, для классификации трафика могут использоваться обучающие алгоритмы под наблюдением и полу-под наблюдением. Для алгоритмов супервизионного обучения требуется помеченный обучающий набор данных, в котором потоки трафика помечены известными классами, такими как поток elephant, приложения или классы QoS. DPI-это распространенный метод маркировки потоков трафика, но он требует больших вычислительных затрат при маркировке большого количества потоков трафика. Более того, постоянно растущее число новых приложений также делает алгоритмы контролируемого обучения менее эффективными. Напротив, алгоритмам полууправляемого обучения требуется лишь небольшая часть помеченных данных, поэтому они более эффективны для детальной классификации трафика.

Поскольку эксперименты проводятся на основе разных обучающих наборов данных, то невозможно напрямую сравнивать производительность различных алгоритмов обучения. Эксперименты показывают показали, что точность классификации напрямую от размера и объема обучающих наборов данных. В целом, с увеличением размеров и объема обучающих наборов данных точность классификации может быть повышена. По сравнению с обычными алгоритмами машинного обучения, алгоритмы глубокого обучения более подходят для обработки больших многомерных обучающих наборов данных. Однако, чем глубже нейронная сеть, тем больше времени требуется для обучения модели нейронной сети.

Литература

1. *Amaral P., Dinis J., Pinto P., Bernardo L., Tavares J., Mamede H. S.* Machine learning in software defined networks: Data collection and traffic classification // Proc. IEEE ICNP'16, Singapore, Singapore, Nov. 2016, pp. 1-5.

2. *Qazi Z. A., Lee J., Jin T., Bellala G., Arndt M., Noubir G.* Application-awareness in SDN // Proc. ACM SIGCOMM'13, Hong Kong, China, 2013, pp. 487-488.

3. *Wang P., Lin S. C., Luo M.* A framework for QoS-aware traffic classification using semi-supervised machine learning in SDNs // Proc. IEEE SCC'16, San Francisco, CA, USA, June. 2016, pp. 760-765.

4. *Гласснер Э.* Глубокое обучение без математики. Т. 2: Практика / пер. с анг. В. А. Яроцкого. М.: ДМК Пресс, 2020. 610 с.

5. *Benson T., Akella A., Maltz D. A.* Network traffic characteristics of data centers in the wild // Proc. ACM IMC'10, Melbourne, Australia, 2010, pp. 267-280.

6. *Glick M., Rastegarfar H*. Scheduling and control in hybrid data centers // Proc. IEEE PHOSST'17, San Juan, Puerto Rico, July. 2017, pp. 115-116.

7. *Xiao P., Qu W., Qi H., Xu Y., Li Z.* An efficient elephant flow detection with cost-sensitive in SDN // Proc. IEEE INISCom'15, Tokyo, Japan, March. 2015, pp. 24-28.

8. Li Y., Li J. MultiClassifier: A combination of DPI and ML for application-layer classification in SDN // Proc. IEEE ICSAI'14, Shanghai, China, Nov. 2014, pp. 682-686.

9. *Гласснер Э*. Глубокое обучение без математики. Т. 1: Основы / пер. с анг. В. А. Яроцкого. М.: ДМК Пресс, 2019. 578 с.

10. *Rossi D., Valenti S.* Fine-grained traffic classification with Netflow data // Proc. ACM IWCMC'10, Caen, France, 2010, pp. 479-483.

11. *Nakao A., Du P.* Toward in-network deep machine learning for identifying mobile applications and enabling application specific network slicing // IEICE Trans. Communications, p. 2017CQI0002, 2014.

12. *Uddin M., Nadeem T.* TrafficVision: A case for pushing software defined networks to wireless edges // Proc. IEEE MASS'16, Brasilia, Brazil, Oct. 2016, pp. 37-46.

АНАЛИЗ ЭФФЕКТИВНОСТИ СУЩЕСТВУЮЩЕЙ СИСТЕМЫ ОЦЕНКИ КАЧЕСТВА ОКАЗАНИЯ УСЛУГ СОТОВОЙ СВЯЗИ В СОВРЕМЕННЫХ УСЛОВИЯХ

Пальцин Денис Анатольевич,
ФГУП «Ордена Трудового Красного Знамени Российский научно-исследовательский институт радио имени М.И. Кривошеева», Директор центра исследований сетей доступа, Москва, Россия
Palcin@niir.ru

Фень Алексей Сергеевич,
ФГУП «Ордена Трудового Красного Знамени Российский научно-исследовательский институт радио имени М.И. Кривошеева», начальник отдела качества центра исследований сетей доступа, Москва, Россия

Ступницкий Михаил Михайлович,
ФГУП «Ордена Трудового Красного Знамени Российский научно-исследовательский институт радио имени М.И. Кривошеева», Ученый секретарь, к.т.н., доцент, Москва, Россия

Аннотация

Рассмотрены вопросы формирования показателей качества услуг сотовой связи в зависимости от поколений сотовой связи. Кратко описаны международные тенденции в регулировании качества полного спектра услуг сотовой связи, отражающие достигнутый технологический, социальный и экономический уровни развития. Дано описание показателей качества для различных технологий сотовой связи. Определены тенденции формирования базовых пакетов услуг сотовой связи. Сформулирован набор показателей качества сотовой связи для базовых пакетов услуг и определен переход к интегральным показателям качества.

Ключевые слова: *Сотовая связь, базовые пакеты услуг сотовой связи, качество услуг сотовой связи, группы показателей качества сотовой связи, интегральные показатели качества сотовой связи.*

1. Введение

Стремительное и масштабное внедрение цифровых технологий во все сферы жизни общества и государства существенно изменило требования к номенклатуре телекоммуникационных услуг и оперативности доступа к ним. Будущее отрасли напрямую связано с такими понятиями, как интернет вещей, облачные сервисы, переход на 5G и 6G стандарты, технологии Big Data, мобильные финансы, развитие конвергентных услуг, способы монетизации контента и др.

Значение цифровых технологий в нашей жизни достигло в последние годы новых высот, все больше людей используют интернет не только для развлечений и общения, а имеют возможность получить хорошую работу, образование, качественные медицинские услуги.

В России соответствуя мировой тенденции наблюдается интенсивный рост потребления интернет трафика. Динамика роста потребления интернет-трафика в стране за последнее время и прогноз на 2022-2023 годы показаны на рисунке 1 [1].

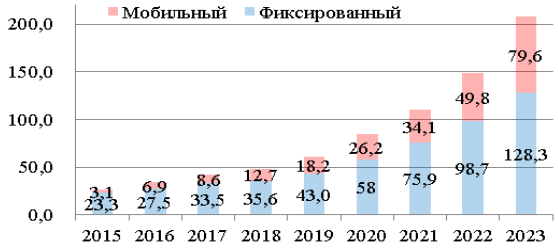

Рис. 1. Динамика роста потребления интернет – трафика в России в эксабайтах

Одними из ключевых требований к современной и перспективной связи являются обеспечение мобильного и он-лайн информационного обмена. Учитывая определяющее влияние уровня выполнения этих требований на качество реализации новых технологий и принятие решений в различных системах управления, в том числе в системах искусственного интеллекта и в системах интернета вещей, особое значение приобрела проблематика повышения качества оказания услуг связи и контроля этого качества на различных этапах жизненного цикла сетей связи. Наибольшую актуальность вопросы качества услуг приобретают в сфере сотовой связи, которая находится на острие проблемы обеспечения названной выше мобильности и он-лайн оперативности связи.

Современные сети сотовой связи стали неотъемлемым элементом государственной инфраструктуры, без которого невозможно существование современного общества, повышение уровня жизни граждан и развитие экономики страны.

Бурное развитие российской цифровой экономики и «жажда» потребления трафика ставит перед операторами связи серьезные вызовы, чтобы обеспечить абонентов современными услугами сотовой связи с заданным уровнем качества.

Тема качества услуг связи знакома каждому пользователю услуг мобильной связи, когда возникает потребность позвонить по мобильному телефону или выйти в сеть Интернет, а возможности нет – отсутствует радиосигнал сотовой связи или не доступна конкретная услуга. Также хорошо знакома ситуация, когда вызов прерывается не по желанию абонента, а из-за технических сбоев сети сотовой связи. Или скорость доступа в сеть Интернет на столько мала, что не позволяет загрузить требуемый контент.

В последнее время практически все операторы сотовой связи предпринимают значительные усилия для улучшения доступности и качества сотовой связи. При этом даже в крупных городах нашей страны остаются еще белые пятна, где не доступен сигнал сотовой связи. За пределами городов, если удалиться от крупного населенного пункта несколько километров возникают проблемы с доступом в сеть Интернет. Это происходит по причине финансовой непривлекательности для операторов сотовой связи развития современных сетей сотовой связи в полном объеме на всей лицензионной территории. При этом в настоящее время 25% населения Российской Федерации проживает в сельской местности.

В современных условиях требуется оперативная оценка качества и принятие адекватных мер либо в масштабах пакета услуг, а следовательно сегмента сети, либо в масштабах всей сетевой инфраструктуры. Остро стоит вопрос о разработке и оценке применимости интегральных показателей качества. Такие показатели стали бы основой не только оценки эффективности сетей связи как таковых на этапе их эксплуатации, но и прогнозирования их развития с учетом взаимосвязи большого количества зачастую конфликтно влияющих на результат факторов

Таким образом, тема управления качеством оказания услуг сотовой связи вышла на первое место по значимости как для граждан – потребителей услуг, так и для операторов связи.

2. Эволюция сетей сотовой связи.

С момента появления подвижных технологий связи человечество приобрело свободу общения в любой точке пространства, где есть радиосигнал. Запуск коммерческих аналоговых сетей сотовой связи произошел в начале восьмидесятых годов 20 века.

Подвижная сотовая связь прошла глобальную трансформацию от момента появления до настоящего времени за достаточно короткий промежуток времени. Технологии сотовой связи - набор функциональных и технических возможностей сетей от простой голосовой связи стремительно расширились до предоставления различных цифровых услуг и сервисов.

Развитие систем сотовой связи рассматривается как эволюционный переход от одного поколения к последующим, от 1G до 5G и перспективному 6G.

Сети сотовой связи первого поколения (1G) были построены с использованием аналоговых технологий. Технологии первого поколения сотовой связи позволяли передавать только голос. На закате технологии первого поколения позволяла организовать канал передачи данных с ничтожно малой скоростью.

Отсутствие единых международных стандартов технологий сотовой связи приводило к невозможности пользоваться услугами сотовой связи при поездках заграницу страны, отсутствовал роуминг в современном понимании. На данном этапе развития технологий сотовой связи оценка качества услуг сотовой связи была не возможна.

Сети сотовой связи второго поколения стали глобальным рывком в развитии технологий, осуществился переход от аналогового к цифровому виду передачи информации, стали использоваться элементы шифрования, добавилась новая услуга связи по передаче и получению коротких текстовых сообщений (SMS), скорость передачи данных увеличилась до 9,6 кбит/с. Стремительный набор популярности сети Интернет привел к созданию новых абонентских устройств и нового протокола беспроводной передачи данных (WAP) для обеспечения возможности выхода в сеть Интернет.

Сети сотовой связи второго поколения основываются на технологии коммутации каналов связи, которая предусматривает при осуществлении сеанса связи закрепления за абонентом постоянного дуплексного канала. Для увеличения скорости доступа в сеть Интернет, повышения емкости сети сотовой связи и эффективности использования радиочастотного спектра к стандартным возможностям сетей второго поколения добавилась поддержка технологий пакетной передачи данных (GPRS) с пропускной способностью до 270 кбит/с, а позднее ее более скоростная модификация (EDGE) с пропускной способностью до 810 кбит/с.

Снижение стоимости услуг сотовой связи привело к существенному увеличению количества абонентов. В силу массовости услуги появилась потребность в оценке качества услуг сотовой связи изначально для самих сотовых операторов с целью модернизации и развития сотовых сетей. Технические специалисты сотовых операторов стали использовать показатели качества, характеризующие доступность услуг связи: степень покрытия территории, долю успешно установленных голосовых соединений и долю успешно доставленных SMS сообщений.

Основной отличительная стороной сетей третьего поколения является высокая скорость передачи данных. Для сетей сотовой связи третьего поколения смещен акцент с голосовой связи в сторону потребления интернет трафика и возможностей сети передачи данных. Сети третьего поколения соединяют в себе передачу как голосового трафика, так и пакетов данных.

Сети третьего поколения стали использовать широкий пропускной канал, когда каждому пользователю назначается индивидуальный цифровой код, который распространяется по всей полосе выделенных радиочастот.

Эволюционное развитие сетей третьего поколения является технология HSPA, которая использует многокодовую передачу данных, что позволило разогнать скорость передачи данных до 84 Мбит/с. Такие сети получили название 3,5G и являются переходными к сетям сотовой связи четвертого поколения.

Параллельно с развитием технологий и ростом услуг сотовой связи увеличивалось количество показателей качества услуг сотовой связи. Техническими специалистами дополнительно стали использоваться следующие показатели качества: доля успешно установленных соединений для передачи данных, среднее значение скорости передачи данных в направлении к абоненту, среднее значение скорости передачи данных в направлении от абонента.

Взрывной рост количества цифровых сервисов и он-лайн услуг, переход к предоставлению мультисервисных услуг связи, все возрастающее количество абонентских терминалов и устройств стали катализатором дальнейшей эволюции сетей сотовой связи и появлению сетей четвертого поколения.

Главное отличие сетей сотовой связи четвертого поколения от предыдущих заключается в том, что они полностью основаны на протоколах пакетной передачи данных. Сети четвертого поколения - это широкополосные высокоскоростные подвижные сети передачи данных.

Технологии сетей сотовой связи четвертого поколения позволяют интегрировать все прибыльные секторы телекоммуникационного рынка: голосовую связь, мобильный доступ в Интернет, доставка тяжелого видео контента, подключение устройств Интернета вещей (IoT), создание беспилотного транспорта и роботизация производств. При этом отсутствие возможности удовлетворить все возрастающую потребность в пропускной способности сотовых сетей и обеспечить стабильность высоких сетевых показателей, рост числа устройств, подключённых к сети Интернет, появление новых виртуальных «пользователей» привели к эволюционному появлению сетей сотовой связи пятого поколения.

Основное отличие сетей пятого поколения в том, что они являются сверхширокополосными мультисервисными сетями с элементами виртуализации сетевых элементов. Услуги голосовой связи и доступа в сеть Интернет, предоставляемые людям, стали незначительной частью сервисов, которые могут быть оказаны в сетях пятого поколения. Сети пятого поколения стали основой для взаимодействия между устройствами.

С учетом такой трансформации услуг и глобальное расширение пользователей – устройств к сетям пятого поколения предъявляются более жесткие требования по параметрам: высокие скорости передачи данных (скорость к пользователю до 20 Гбит/с, скорость от пользователя до 10 Гбит/с), значительное повышение энергоэффективности абонентских устройств, минимальные временные задержки на радиоинтерфейсе менее 1 мс, возможность пользования услугами на сверхбыстрой скорости до 500 км/ч, большее число подключенных устройств до 1 млн. устройств на км^2 [2].

Некоторые из этих показателей, например, пиковая скорость передачи данных и автономность технологически несовместимы. Подходы в сетях пятого поколения заключаются в разделении сценарных моделей оказания услуг сотовой связи в зависимости от важности того или иного показателя. В концепции построения сетей пятого поколения физическая архитектура сети разделена на отдельные виртуальные слои (Network Slicing) для оказания определенного перечня услуг с определенными гарантированными показателями.

Поэтому сети пятого поколения можно разделить на три составляющие в зависимости от требований по набору и качеству предоставляемых услуг.

Усовершенствованная сотовая широкополосная связь (eMBB).

Массовая межмашинная связь (mMTC).

Сверхнадежная межмашинная связь с низкими задержками (URLLC).

Это привело к существенному увеличению количества параметров качества услуг связи и цифровых сервисов и потребность их объединения в отдельные группы для оценки эффективности сетей связи на этапе их эксплуатации, прогнозирования их развития и повышения качества предоставляемых услуг, учитывая взаимосвязь большого количества зачастую конфликтно влияющих на результат оценки факторов [3]. Например, можно выделить основные наборы параметров качества, объединенные в соответствующие группы.

Группа показателей доступности услуг связи: степень покрытия сети связи, доля успешно установленных голосовых соединений, доля успешно доставленных SMS сообщений, доля успешно установленных соединений для передачи данных, доступность внешних сетевых ресурсов.

Группа показателей непрерывность оказания услуг связи: доля обрывов голосовых соединений, доля неуспешных сессий передачи данных, средняя доля потери пакетов, средняя разборчивость речи на соединение.

Группа показателей времени ожидания услуги: среднее время установления голосового соединения, среднее время передачи сообщения SMS, среднее время регистрации абонентского терминала в сети пакетной передачи данных.

Группа показателей целостность: показатель сетевой задержки, среднее значение скорости передачи данных в направлении к абоненту, среднее значение скорости передачи данных в направлении от абонента.

3. Управление качеством услуг сотовой связи.

Международные тенденции в регулировании качества полного спектра услуг связи отражают достигнутый технологический, социальный и экономический уровни развития телекоммуникаций.

Если на заре своего развития сети сотовой связи строились с применением уникальных технологий, характерными для определенной страны, и являлись премиальными сегментами на рынке услуг связи с минимальным количеством пользователей и отсутствием конкуренции, то с момента появления сетей сотовой связи второго поколения стали применяться унифицированные принципы, что привело к возникновению конкуренции на рынке сотовых услуг, начала снижаться цена на услуги и увеличиваться абонентская база. Указанные предпосылки, а также ускоренные темпы научно-технического прогресса, все увеличивающееся разнообразие услуг, технологий и средств связи, увеличение объемов, передаваемых данных потребовали стандартизации подходов к объективной оценке и управлению качеством услуг сотовой связи.

На международном уровне проблемой регламентации стандартов качества телекоммуникаций занимается ряд крупных международных организаций, главной из которых является Международный союз электросвязи (МСЭ, ITU). К процессу стандартизации качества услуг связи привлечены регуляторы, представители отраслевой промышленности и операторы связи

Концепция качества предоставления услуг связи определена в серии рекомендаций МСЭ в 2008 году. В рекомендациях определена терминология в области качества услуг связи, требования к пока-

зателям качества обслуживания, сетевым показателям и нормам на эти показатели (Рекомендации МСЭ-Т E.800, E.802, G. 1000, Y. 1540, Y. 1541 и др.).

В соответствии с требованиями МСЭ параметры оценки качества предоставляемых услуг должны количественно и качественно оцениваться, быть удобными для аудита, иметь стандарты для сравнения.

В международных стандартах качество услуг связи включает в себя:

– качество функционирования сети (NP);

– качество предоставления услуги (QoS);

– качество восприятия услуги (клиентский опыт) (QoE).

Качество функционирования сети связи (NP) является необходимой основой для обеспечения качества услуг связи и неразрывно связано с целостностью, устойчивостью и безопасностью функционирования сети связи общего пользования.

Качество функционирования сети связи непосредственно обеспечивается операторами связи при планировании, строительстве и эксплуатации сетей связи.

Качество функционирования сетей связи оценивается по данным сетевой статистики, собираемым системами технологического мониторинга с применением пробников и других инструментов сбора данных. Такие технологические системы, ориентированные на операторское оборудование конкретного производителя, отслеживают ряд важнейших параметров функционирования сети, позволяя операторам получать детальную информацию о функционировании сетей и реагировать наиболее эффективно. При этом сравнение между собой результатов технологического мониторинга различных сетей связи при использовании оборудования различных производителей не представляется возможным.

Качество предоставления услуги связи (QoS) отражает результат взаимодействия используемого абонентом (пользователем) пользовательского (оконечного) оборудования с сетью связи в процессе оказания услуги связи.

Оценка качества предоставления услуги связи производится в процессе сквозного тестирования «из конца в конец» (E2E, End-to-End) с применением:

– пользовательского (оконечного) оборудования, оснащенного специальным программным обеспечением (программный агент);

– испытательных (тестовых) комплексов, применяемых в ходе специально организованных драйв-тестов.

Тестирование с применением пользовательского (абонентского) оборудования является наиболее массовым и проводится силами самих абонентов при повседневном использовании ими своего пользовательского (оконечного) оборудования, оснащенного самостоятельно установленным ими специальным программным обеспечением (программным агентом). Такое тестирование может проводиться как автоматически, через регулярные промежутки времени, так и в ручном режиме по инициативе абонента, в случаях неудовлетворенности абонента качеством услуг.

Полевые испытания (Drive test) являются наиболее объективным способом оценки и выполняются по единым утвержденным методикам с применением технических испытательных средств – тестовых комплексов. Тестовые комплексы по имеющимся функциям являются специализированным абонентским оборудованием, имеющим наиболее полную функциональность и поддерживающим все реализованные на сети связи технологии, режимы работ, диапазоны частот. Драйв-тесты проводятся квалифицированным персоналом, а методики испытаний предписывают вполне определенный порядок их проведения при большом количестве тестовых проб для достижения требуемой достоверности результатов. Этим обеспечивается возможность получения оценок, наиболее полно отражающих потенциально возможное качество услуг связи, предоставляемое оператором. Драйв-тесты являются весьма затратным по стоимости и времени способом оценки качества, что ограничивает возможности их регулярного или массового использования.

Клиентское восприятие качества услуги связи (QoE) формируется на основе сравнения качества фактически оказываемых услуг связи с заявленным оператором связи уровнем качества или с ожиданиями абонентов (пользователей).

Основным методом оценки восприятия качества услуги связи является опрос, при котором абоненты (пользователи) дают совокупную оценку качества услуг связи, полученных ими в течение некоторого времени на определенной территории обслуживания сети связи.

Результаты опросов и оценок восприятия качества услуги связи и жалобы абонентов (пользователей) на качество используются операторами связи для своевременного и эффективного реагирования на возникающие проблемы с качеством услуг связи [4].

Показатели и параметры, характеризующие клиентский опыт, качество услуг связи и качество сети связи, отражают разные аспекты качества связи и служат разным целям, однако между ними существует внутренняя взаимосвязь и взаимозависимость.

4. Подходы к расчету показателей качества услуг сотовой связи.

Системы передачи цифровой информации передают цифровые данные, обрабатываемые на уровне различного программного обеспечения и преобразовывающие их в соответствующий вид данных. Виды данных могут быть абсолютно любыми: от голоса до видеофайлов. Обработка и воспроизведение информации может выполняться как на уровне программного обеспечения сотового телефона, универсальных обозревателей цифрового контента - «браузеров» (browser), так и на уровне специализированных программ, таких как мультимедиа проигрыватели, программы общения (Skype, WhatsApp, Viber и др.), программы просмотра или редактирования текстовых, графических или видеоданных.

Сети связи отдельного оператора, как правило, являются многоуровневой системой, включающей систему коммутации с первичной или вышестоящей вторичной сетью связи, систему управления и коммутации внутри сети связи оператора, включающей узлы связи, линии связи, устройства коммутации и распределения сигналов, а также оконечные устройства связи оператора, формирующие абонентские линии связи со стороны оператора.

При этом с ростом нагрузки на телекоммуникационную инфраструктуру вследствие увеличения количества пользователей и роста количества передаваемой информации, возникает целый спектр задач поиска эффективных способов управления и контроля качества, предотвращения отказов, прогнозирования трафика с целью дальнейшей оптимизации структуры сети.

Как показал проведенный анализ рассматриваемая нами проблема оценки качества услуг сотовой связи, индикации состояния и доступности телекоммуникационной инфраструктуры на сетях сотовой связи связана с рядом проблем, в числе которых:

– неравномерное территориальное распределение элементов инфраструктуры связи, их территориальной удаленности и ряда других факторов;

– отсутствие общего подхода к оценке и мониторингу качества услуг, а также различных нормативов качества.

Сложившаяся система многочисленных и многофакторных по своей сути показателей качества оказания услуг связи достаточно хорошо работает на уровне узких технических специалистов операторов сотовой связи и разработчиков систем связи. При этом, учитывая разнообразие технологий построения сетей сотовой связи, разнообразия предоставляемых услуг связи, параметров качества для различных технологий необходимо разработать инструментарий для более гибкого анализа качества услуг телекоммуникационной инфраструктуры на основе систем ключевых показателей качества, а также параметрических моделей, связывающих показатели качества доступности, непрерывности, времени ожидания услуги, целостности [5].

Группа показателей качества, характеризующих доступность услуг связи.

Группа показателей качества, отражающая доступность услуг связи для абонента, является основной группой качества, вносящей наибольший вклад в интегральный показатель качества. Поскольку если услуги связи не доступны для абонента совсем, то говорить о дальнейших показателях качества услуг связи не имеет смысла.

Основным показателем качества, не связанным с конкретной услугой, оказывающим максимальное влияние на доступность услуг является степень покрытия сети связи ($C_{покр}$).

В общем виде степень покрытия территориального района рассчитывается по формуле:

$$C_{покр} = \frac{P_{покр}}{P_O} * 100\%$$

$C_{покр}$ – показатель степени покрытия связи территориального района;

$P_{покр}$ – показатель площади покрытия связи в км², суммарная площадь участков, на которых уровень принимаемого сигнала сотовой связи достаточен для обеспечения возможности пользоваться услугами связи;

P_O – общая площадь территориального района в км².

Показатель доли успешно установленных голосовых соединений ($C_{УСГ}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент попыток установления голосовых соединений, завершившихся успешно.

$$C_{УСГ} = \frac{N_{УСГ}}{N_Г} * 100\%$$

$C_{УСГ}$ – показатель доли успешно установленных голосовых соединений;

$N_{УСГ}$ – количество успешно установленных голосовых соединений по данным сетевой статистики, либо полевых испытаний (Drive test), за определенный период времени;

$N_Г$ – общее количество попыток установления голосовых соединений за определенный период времени.

Данный показатель агрегирует все причины, по которым соединения не состоялись, (блокировка по причине перегрузки коммутационной системы или системы радиодоступа, внутренние технические ошибки и т.д.).

Показатель доли успешно доставленных SMS сообщений ($C_{УСС}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент попыток отправки SMS сообщений, завершившихся успешной доставкой адресату.

$$C_{УСС} = \frac{N_{УСС}}{N_С} * 100\%$$

$C_{УСС}$ – показатель доли успешно доставленных SMS сообщений;

$N_{УСС}$ – количество успешно доставленных SMS сообщений по данным сетевой статистики, либо полевых испытаний (Drive test), за определенный период времени;

$N_С$ – общее количество отправленных SMS сообщений за определенный период времени.

Данный показатель агрегирует все причины, по которым SMS сообщения были не доставлены адресату (блокировка по причине перегрузки центра отправки SMS или системы радиодоступа, внутренние технические ошибки и т.д.).

Показатель доли успешно установленных соединений для передачи данных ($C_{УСД}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент успешных попыток установления соединений для обеспечения возможности передачи данных (голосовых, видео, файлов и др.).

$$C_{УСС} = \frac{N_{УСД}}{N_Д} * 100\%$$

$C_{УСД}$ – показатель доли успешно установленных соединений для передачи данных;

$N_{УСД}$ – количество успешно установленных соединений для обеспечения возможности передачи данных по всем типам протоколов по данным сетевой статистики, либо полевых испытаний (Drive test), за определенный период времени;

$N_Д$ – общее количество попыток установления соединений для передачи всех видов данных за определенный период времени.

Данный показатель агрегирует все причины, по которым соединения не состоялись (блокировка по причине перегрузки коммутационной системы или системы радиодоступа, внутренние технические ошибки и т.д.).

Доступность внешних сетевых ресурсов (C_{PSR}) характеризует вероятность доступности внешних сетевых ресурсов. Рассчитывается как усредненный за определенный период времени (например, ме-

сяц) процент успешного получения ответа на запросы Ping, направленные с абонентского оборудования.

$$C_{PSR} = \frac{N_{PSR}}{N_P} * 100\%$$

C_{PSR} – показатель доступности внешних сетевых ресурсов (PSR - Ping Success Rate);

N_{PSR} – количество полученных ответов на запросы Ping по данным полевых испытаний (Drive test), за определенный период времени;

N_P – общее количество отправленных запросов Ping за определенный период времени.

Группа показателей качества, характеризующих непрерывность оказания услуг связи.

Доля обрывов голосовых соединений ($C_{ОГ}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент голосовых соединений, закончившихся не по инициативе абонента.

$$C_{ОГ} = \frac{N_{ОГ}}{N_Г} * 100\%$$

$C_{ОГ}$ – показатель доли обрывов голосовых соединений;

$N_{ОГ}$ – количество голосовых соединений, закончившихся не по инициативе абонента по данным сетевой статистики, либо полевых испытаний (Drive test), за определенный период времени;

$N_Г$ – общее количество голосовых вызовов за определенный период времени.

Данный показатель агрегирует все причины, по которым соединения были разорваны не по инициативе абонента (внутренние технические ошибки, не успешные попытки переключения абонента внутри сети между сотами (Handover) и т.д.).

Доля неуспешных сессий передачи данных ($C_{ОПД}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент сессий передачи данных, закончившихся не по инициативе абонента.

$$C_{ОПД} = \frac{N_{ОПД}}{N_{ПД}} * 100\%$$

$C_{ОПД}$ – показатель доли неуспешных сессий передачи данных;

$N_{ОПД}$ – количество сессий передачи данных, закончившихся не по инициативе абонента по данным сетевой статистики, либо полевых испытаний (Drive test), за определенный период времени;

$N_{ПД}$ – общее количество сессий передачи данных за определенный период времени.

Данный показатель агрегирует все причины, по которым сессии были разорваны не по инициативе абонента (внутренние технические ошибки, не успешные попытки переключения абонента внутри сети между сотами (Handover) и т.д.).

Средняя доля потери пакетов ($C_{ПП}$) рассчитывается как усредненный за определенный период времени (например, месяц) процент отношения полученных и переданных пакетов по сети пакетной передачи данных.

$$C_{ПП} = \frac{N_{ПолП}}{N_{ПерП}} * 100\%$$

$C_{ПП}$ – показатель средней доли потери пакетов за определенный промежуток времени, рассчитывается по данным полевых испытаний (Drive test) или End-to-End тестирования;

$N_{ПолП}$ – количество полученных пакетов по сети пакетной передачи данных;

$N_{ПерП}$ – количество переданных пакетов по сети пакетной передачи данных.

Средняя разборчивость речи на соединение (MOC) рассчитывается как средняя оценка качества передачи речи по шкале MOS путем прямого измерения значения с использованием алгоритма PESQ/POLQA.

$$MOC = \frac{\sum_1^{N_i} MOC_i}{N_i}$$

MOC – среднее значение качества передачи речи по шкале MOS, рассчитывается по данным полевых испытаний (Drive test) между мобильным абонентом и автоответчиком или End-to-End тестирования мобильный абонент - мобильный абонент;

MOC_i – оценка качества передачи речи MOS при i-й речевой последовательности по шкале (1…5) в соответствии с Рек. ITU-T P.862.1 (P.862.2 и P.863);

N_i – общее количество речевых последовательностей по успешно установленным соединениям при контрольных вызовах.

Группа показателей качества, характеризующих времени ожидания услуги.

Среднее время установления голосового соединения ($T_{ГС}$) рассчитывается как усредненная за определенный период времени (например, месяц) величина времени задержки ответа сигнала сети.

$$T_{ГС} = \frac{\sum_1^{N_{ГСi}} (t_{ОГСi} - t_{ЗГСi})}{N_{ГСi}}$$

$T_{ГС}$ – среднее значение времени с момента отправки запроса на голосовое соединение и моментом получения ответа сети, интервал времени между моментом начала получения сигнала КПВ (контроля посылки вызовов) и ответом вызываемого абонента не учитывается, рассчитывается по данным полевых испытаний (Drive test);

$t_{ЗГСi}$ – время отправки i-го запроса на голосовое соединение;

$t_{ОГСi}$ – время получения сигнала ответа сети на i-й запрос на голосовое соединение;

$N_{ГСi}$ – общее количество отправленных запроса на голосовое соединение за определенный период времени.

Среднее время передачи SMS сообщения ($T_{ДС}$) рассчитывается как усредненная за определенный период времени (например, месяц) величина времени доставки SMS сообщения от передающего терминала к приемному.

$$T_{ДС} = \frac{\sum_1^{N_{ОСi}} (t_{ПСi} - t_{ОСi})}{N_{ОСi}}$$

$T_{ДС}$ – среднее значение времени доставки SMS сообщения от передающего терминала к приемному, рассчитывается по данным полевых испытаний (Drive test) или End-to-End тестирования;

$t_{ОСi}$ – время отправки i-го SMS сообщения;

$t_{ПСi}$ – время получения i-го SMS сообщения;

$N_{ОСi}$ – общее количество отправленных SMS сообщений за определенный период времени.

Среднее время регистрации абонентского терминала в сети пакетной передачи данных ($T_{РПД}$) рассчитывается как усредненная за определенный период времени (например, месяц) величина времени, необходимого для регистрации абонентского терминала в сети пакетной передачи данных.

$$T_{РПД} = \frac{\sum_1^{N_{РПДi}} (t_{ОПДi} - t_{ЗПДi})}{N_{РПДi}}$$

$T_{РПД}$ – среднее значение времени регистрации абонентского терминала установления в сети пакетной передачи данных, рассчитывается по данным полевых испытаний (Drive test) и End-to-End тестирования;

$t_{ЗПДi}$ – время отправки i-го запроса для регистрации абонентского терминала в сети пакетной передачи данных;

$t_{ОПДi}$ – время получения сигнала ответа сети на i-й запрос для регистрации абонентского терминала в сети пакетной передачи данных;

$N_{РПДi}$ – общее количество отправленных запроса на голосовое соединение за определенный период времени.

Группа показателей качества, характеризующих целостность услуг связи.

Показатель сетевой задержки (C_{Ping}) рассчитывается по данным полевых испытаний (Drive test) или End-to-End тестирования как усредненное за определенный период времени (например, месяц) время сетевой задержки, вносимой инфраструктурой сотовой сети пакетной передачи данных.

$$C_{Ping} = \frac{\sum_1^{N_i} (t_{ППi} - t_{ОПi})}{N_i}$$

C_{Ping} – показатель сетевой задержки;

$t_{ОПi}$ – время отправки короткого пакета длиной 32 байта (процедура Ping);

$t_{ППi}$ – время приема ответного пакета при выполнении процедуры Ping;

N_i – общее количество отправленных пакетов за определенный период времени.

Данный показатель агрегирует все причины, по которым могут быть внесены задержки инфраструктурой сотовой сети пакетной передачи данных (перегрузка на сетевом оборудовании, неоптимальный маршрут, технические сбои и др.).

Среднее значение скорости передачи данных в направлении к абоненту (V_{DL}) рассчитывается как усредненная за определенный период времени (например, месяц) величина скорости передачи данных в направлении от тестового сервера передачи данных к абонентскому терминалу.

$$V_{DL} = \frac{\sum_1^{N_{DLi}} V_{DLi}}{N_{DLi}}$$

V_{DL} – среднее значение скорости передачи данных в направлении к абоненту;

V_{DLi} – величина скорости передачи данных в направлении от тестового сервера передачи данных к абонентскому терминалу при i-ом по данным полевых испытаний (Drive test) или End-to-End тестирования:

$$V_{DLi} = \frac{P_i}{(t_{ЗПi} - t_{НПi})}$$

P_i – размер тестового файла с данными;

$t_{НПi}$ – время начала передачи тестового файла с данными;

$t_{ЗПi}$ – время завершения передачи тестового файла с данными;

N_{DLi} – общее количество тестовых измерений скорости передачи данных в направлении к абоненту за определенный период времени.

Среднее значение скорости передачи данных в направлении от абонента (V_{UL}) рассчитывается как усредненная за определенный период времени (например, месяц) величина скоростей передачи данных в направлении от абонентского терминала к тестовому серверу передачи данных.

$$V_{UL} = \frac{\sum_1^{N_{ULi}} V_{ULi}}{N_{ULi}}$$

V_{UL} – величина скорости передачи данных в направлении от тестового сервера передачи данных от абонентского терминала при i-ом по данным полевых испытаний (Drive test) или End-to-End тестирования:

$$V_{ULi} = \frac{P_i}{(t_{ЗПi} - t_{НПi})}$$

P_i – размер тестового файла с данными;

$t_{НПi}$ – время начала передачи тестового файла с данными;

$t_{ЗПi}$ – время завершения передачи тестового файла с данными;

N_{ULi} – общее количество тестовых измерений скорости передачи данных в направлении от абонента за определенный период времени [6].

5. Основные пакеты базовых услуг сотовой связи.

Рост конкуренции, востребованность конвергентных услуг, принципиальная комплексность информационного обеспечения граждан и экономики, необходимость совершенствования качества услуг на телекоммуникационном рынке привели к формированию основных пакетов базовых услуг. Основные пакеты базовых услуг сотовой связи были сформированы в целях обеспечения потребностей фокус групп абонентов в тех или иных услугах с учетом экономической и технической целесообразности операторами связи. Набор услуг в разных пакетах может пересекаться, но с точки зрения критических показателей качества являются уникальными. С учетом развития цифрового общества и технологий сотовой связи операторами связи были сформированы разные пакеты базовых услуг в соответствии с потребностями абонентов [7]. Из них можно выделить четыре основных пакета базовых услуг сотовой связи, показатели качества для них представлены в таблице 1.

Таблица 1

Матрица пакетов услуг и параметров качества

Показатели качества услуг сотовой связи	Пакет базовых услуг			
	№ 1	№ 2	№ 3	№ 4
Группа показателей доступности услуг связи				
степень покрытия сети связи	•	•	•	•
доля успешно установленных голосовых соединений	•	•		
доля успешно доставленных SMS сообщений	•			
доля успешно установленных соединений для передачи данных		•	•	•
доступность внешних сетевых ресурсов		•	•	
Группа показателей непрерывность оказания услуг связи				
доля обрывов голосовых соединений	•	•		
доля неуспешных сессий передачи данных		•	•	•
средняя доля потери пакетов			•	•
средняя разборчивость речи	•	•		
Группа показателей времени ожидания услуги				
среднее время установления голосового соединения	•	•		
среднее время передачи SMS	•			
среднее время регистрации в сети пакетной передачи данных		•	•	•
Группа показателей целостности услуги связи				
показатель сетевой задержки			•	
среднее значение скорости ПД в направлении к абоненту		•	•	
среднее значение скорости ПД в направлении от абонента		•	•	

Пакет базовых услуг сотовой связи № 1 для голосовой телефонной связи.

Включенные в пакет услуги сотовой связи отвечают самым простым потребностям абонента в голосовой связи и отправки SMS или ограничены техническими возможностями абонентского оборудования. К основным показателям качества для этого пакета услуг можно отнести:

– степень покрытия сети связи,
– доля успешно установленных голосовых соединений;
– доля обрывов голосовых соединений;
– среднее время установления голосового соединения;
– средняя разборчивость речи на соединение;
– доля успешно доставленных SMS сообщений;
– среднее время передачи SMS сообщения.

Пакет базовых услуг сотовой связи № 2 для голосовой телефонной связи и базового доступа к сети Интернет.

Включенные в пакет услуги сотовой связи удовлетворяют потребности абонента не только в голосовой телефонной связи, но и в базовом уровне доступа к сети Интернет с низким уровнем потребления трафика. К основным показателям качества для подобного пакета услуг можно отнести:

– степень покрытия сети связи,
– доля успешно установленных голосовых соединений;
– доля обрывов голосовых соединений;
– среднее время установления голосового соединения;
– средняя разборчивость речи на соединение;
– доля успешно установленных соединений для передачи данных;
– доступность внешних сетевых ресурсов;
– доля неуспешных сессий передачи данных;
– среднее время регистрации абонентского терминала в сети пакетной передачи данных;
– среднее значение скорости передачи данных в направлении к абоненту;
– среднее значение скорости передачи данных в направлении от абонента.

Пакет базовых услуг сотовой связи № 3 для высокоскоростной передачи данных.

Включенные в пакет услуги сотовой связи обеспечивают потребности абонентов в получении услуг по организации высокоскоростного доступа к сети передачи данных, в том числе к сети Интернет. К основным показателям качества для подобного пакета услуг можно отнести:

– степень покрытия сети связи,
– доля успешно установленных соединений для передачи данных;
– доступность внешних сетевых ресурсов;
– доля неуспешных сессий передачи данных;
– средняя доля потери пакетов;
– среднее время регистрации абонентского терминала в сети пакетной передачи данных;
– показатель сетевой задержки;
– среднее значение скорости передачи данных в направлении к абоненту;
– среднее значение скорости передачи данных в направлении от абонента.

Пакет базовых услуг сотовой связи № 4 для интернета вещей и M2M взаимодействия.

Включенные в пакет услуги сотовой связи отвечают потребностями абонента в организации каналов связи с (между) различными устройствами. К основным показателям качества для подобного пакета услуг можно отнести:

– показатель степени покрытия сотовой связи
– доля успешно установленных соединений для передачи данных
– доля успешных сессий для передачи данных
– среднее время регистрации абонентского терминала в сети пакетной передачи данных;
– средняя доля потери пакетов.

Анализ содержания пакетов базовых услуг связи и приведенных выше применяемых в настоящее время показателей качества оказания услуг связи позволяет заключить, что для обеспечения высоких качественных характеристик современных конвергентных сетей связи наряду с новыми техническими решениями необходим переход к показателям качества нового информационного, системного уровня – интегральным показателям. В настоящее время работоспособная система интегральных показателей качества оказания услуг связи отсутствует.

6. Заключение

Каждый из описанных линейных показателей качества характеризует лишь отдельную составляющую базовых услуг сотовой связи в пакете. Вопросы значимости каждого отдельного показателя качества и однозначного определения качества услуг для каждого пакета услуг сотовой связи являются наиболее востребованным для абонента. В связи с большим количеством показателей качества по каждому сформированному пакету базовых услуг сотовой связи, необходим переход к формированию интегральных показателей качества пакета базовых услуг. Базируясь на линейных показателях качества интегральные показатели должны в упрощенной и удобной для оценки форме отражать характеристики качественности всего пакета базовых услуг. Для каждого из пакетов услуг необходимо разработать отдельный интегральный показатель качества и набора значений для каждого типа абонентов – пользователей различных услуг и сервисов. Такие интегральные показатели качества базовых пакетов услуг сотовой связи должны обеспечивать абонентам возможность самостоятельно оценивать качество предоставленных услуг сотовой связи различных операторов сотовой связи с точки зрения используемых абонентом услуг и сервисов.

Литература

1. Статистика отрасли // интернет ресурс Минцифры России. URL: https://digital.gov.ru/ru/pages/statistika-otrasli/#section-398. (дата обращения: 26.01.2022).
2. *Степутин А.Н., Николаев А.Д.* Мобильная связь на пути к 6G. Том 2. Москва, Вологда. 2021. С. 14-329.
3. *Тихвинский В.О.* Перспективы сетей 5 G и требования к качеству их обслуживания. Электросвязь. Москва. 2014. С. 40-43.
4. *Бабков В.Ю., Полынцев П.В., Устюжанин В.И.* Качество услуг мобильной связи, оценка, контроль и управление. Москва. 2005. С. 100-160.
5. *Пальцин Д.А.* Качество услуг будет на госконтроле. ИКС № 04. Москва. 2013.
6. Методика оценки качества услуг подвижной радиотелефонной связи, утвержденная Минкомсвязью России 4 декабря 2014 года № НН-П19-21799.
7. *Сушков А.А.* Тарификация современных услуг передачи данных в мобильных сетях // CONNECT! Мир связи № 1. 2012. С. 2-6.

РЕКОМЕНДАЦИИ ПО СНИЖЕНИЮ ВЛИЯНИЯ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ СЕТЕЙ 5G НА ЧЕЛОВЕКА И ОКРУЖАЮЩУЮ СРЕДУ

Панкратов Денис Юрьевич,
МТУСИ доц. каф. СиСРТ к.т.н., Москва, Россия
dpankr@mail.ru

Сизов Дмитрий Викторович,
ФГУП РТРС, инженер средств радио и телевидения, Сахалинская обл. г. Углегорск, Россия
deeemon8@mail.ru

Аннотация
В данной статье рассматривается, какое возможное негативное влияние могут оказывать на здоровье человека и окружающую среду сети пятого поколения (5G) использующие сверхвысокочастотное излучение (десятки ГГц), а также предлагаемые в научном сообществе пути к снижению влияния излучений сетей 5G на человеческий организм и окружающую среду.

Ключевые слова: *5G, New Radio, влияние на человека, электромагнитное поле, радиочастотное излучение.*

Введение

В ближайшее время подходит к завершению стандартизация сетей 5G, представляющих дальнейшее развитие и усовершенствование уже действующих сетей 4G [1]. 5G – Это телекоммуникационный стандарт мобильной связи нового поколения.

Международный союз электросвязи в середине 2015 года разработал план по развитию стандарта мобильной связи 5G.

В течение нескольких лет появились и прошли испытания первые разработки. Во второй половине 2018 года американская компания Verizon запустила первую в мире коммерческую сеть пятого поколения. Однако эта сеть обеспечивала население только домашним интернетом, а не полноценной мобильной связью [2].

В начале 2018 года тестовую сеть запустили в Китайском городе Чунцин. В Южной Корее новую технологию опробовали в феврале того же 2018 года, во время Зимней Олимпиады в Пхенчхане, а в апреле 2019 года коммерческая сеть 5G уже работала в Сеуле [2]. В Европейских странах технологии 5G уже доступны в Германии, Венгрии, Австрии, Италии, Великобритании [2].

С 2019 года в России появляются первые сети стандарта 5G, работающие в тестовом режиме. В начале октября 2020 года МТС и ГТРК «Владивосток» совместно провели первый на территории России прямой эфир с помощью сети 5G с использованием оборудования компании Huawei.

Для передачи сигнала использовалась уникальная техническая разработка от Huawei – рюкзак с помещёнными в него модемами 5G, позволяющими транслировать в телевизионный эфир изображение и звук в реальном времени. В апреле 2021 года российский оператор МТС и Huawei ввели в эксплуатацию первые пользовательские зоны 5G, доступные для коммерческих абонентов [3].

Согласно программе "Цифровая экономика", устойчивое покрытие сетями 5G десяти городов-миллионников должно быть обеспечено к 2022 году, а в городах России с населением свыше миллиона человек – к 2024 году [4].

Мобильные сети 5G существенно расширяют функциональные возможности мобильных сетей связи предыдущих поколений. Основными особенностями сетей 5G являются следующие:

Усовершенствованный мобильный широкополосный доступ (Enhanced Mobile BroadBand, eMBB);

Сверхнадёжные коммуникации с низкой задержкой (Ultra Low Latency Reliable Communication, ULLRC);

Массовые межмашинные коммуникации (massive Machine Type Communication, mMTC) [5].

На основе приведённых выше трёх видов сценариев строится многообразие услуг и возможностей сетей IMT2020 (сетей 5G), наиболее характерные из которых:

Скорости передачи данных больше 10 Гбит/с;

«Умный дом» (Smart Home) и «Умное здание» (Smart Building);

Умный город;

Новые видеоуслуги качества 4К/8К;

Работа в облаке;

Дополненная и виртуальная реальность (AR/VR)

Промышленная автоматизация;

Бизнес-критичные приложения (Mission Critical Applications);

Беспилотные транспортные средства (Driverless Vehicles) [6].

В сети подвижной связи стандарта 5G/IMT-2020 используется новый радиоинтерефейс (New Radio, NR) согласно спецификациям 3GPP серии 38, а также поддерживается обратная совместимость со стандартом LTE-Advanced[5].

Интерфейс NR разрабатывался для обеспечения высоких скоростей передачи данных при меньших задержках, а также более эффективного использования радиочастотных ресурсов за счет:

применения сигналов с большей шириной полосы частот (до 100 МГц в диапазоне до 6 ГГц и до 400 МГц в диапазоне свыше 6 ГГц);

обеспечения минимальных задержек на радиоинтерфейсе за счет возможности увеличения частоты следования временных слотов кадровой структуры и за счет модификации протокола управления радиоресурсами;

применения адаптивного к нагрузке временного дуплекса;

применения более эффективных по помехоустойчивости полярных кодов;

использования активных антенных систем в миллиметровых диапазонах с большим количеством элементов, узкой диаграммой направленности излучения и высокой избирательностью;

реализации индивидуальных сценариев использования ресурсов полосы частот канала NR для абонентских терминалов различных типов и производительности (широкополосных / узкополосных абонентских терминалов, абонентских терминалов с агрегацией несущих) [7].

Радиоинтерфейс 5G New Radio (NR) подразумевает два частотных диапазона: FR1 (410 – 7125 МГц) и FR2 (24250 МГц-52600 МГц), каждый с различными возможностями [8,9]. Используется как сантиметровый, так и миллиметровый частотные диапазоны. Использование новых более высокочастотных диапазонов вызывает обеспокоенность научного сообщества, а также общественности о том каким образом будет влиять излучения сетей 5G на человека [7].

В этой статье мы рассмотрим, как воздействуют на здоровье человека и окружающую среду сети 5G использующие высокочастотные диапазоны FR1 и FR2, а также пути к минимизации негативного воздействия излучений от сетей 5G на человека и окружающую среду по материалам [10].

Сети 5G и действующие нормы на излучения

В наше время беспроводные технологии глубоко проникли во все сферы жизнедеятельности человека. Ежегодно большими темпами растёт количество устройств, использующих беспроводные технологии. Это мобильные телефоны, теле- и радио- приёмо-передающие устройства, охранные сигнализации, системы навигации, системы слежения, роутеры, различные модемы, системы «Умный дом». Быстро развивающаяся система – интернет вещей (internet of things, IoT) использующая различные протоколы и среды для передачи данных и взаимодействия с различными устройствами [1].

Технология Интернета вещей (IoT), предлагает особую среду, в которой большое количество датчиков и электронных устройств взаимодействуют посредством радиоинтерфейсов 5G и 4G. В сетях 5G заложены большие возможности для организации взаимодействия со всем многообразием беспроводных устройств различного назначения, и объединения их в одну большую сеть. Всё это однозначно указывает на то, что рассматривать влияние сетей 5G необходимо в совокупности с технологиями Wi-Fi, Wi-MAX, а также 3G и 4G.

Для каждого типа радиоэлектронного устройства выделен определённый частотный диапазон. В целях снижения негативного влияния электромагнитных излучений на человека и окружающую среду, введены санитарные нормы на излучение [11].

В современном мире электромагнитные поля техногенного происхождения превращаются в опасный экологический фактор, на который стоит обращать пристальное внимание [12,13]. Существенный вклад в суммарное ЭМП создается антеннами базовых станций, расположенных на крышах домов или антенно-мачтовых сооружениях, расположенных практически на каждом шагу.

Законодательство Российской Федерации подразумевает нормирование ЭМП отдельно для обслуживающего персонала радиотехнических объектов и населения. Санитарные нормы учитывают многие факторы, например, то, что население будет облучаться круглосуточно, а персонал обслуживающий радиотехнические установки оказывается под воздействием ЭМП только в течение рабочего дня – 8 ч. Так же важно обратить внимание на то, что диапазон состояний организма населения довольно широк (от ребенка до пожилого человека, от здорового человека до тяжело больного), чем у персонала обслуживающего радиотехнические объекты, так как работники находятся под постоянным медицинским контролем. Поэтому предельно допустимые уровни ЭМИ для производственного персонала несколько выше, чем для населения.

Значения уровней излучения ЭМП в пределах жилой застройки, а также внутри помещений (жилых, общественных, производственных), составляют не более, чем: 10 В/м для диапазонов частот от 27 до 30 МГц; 3 В/м для диапазонов частот от 30 МГц до 0,3 ГГц и 10 мкВт/см2 для диапазонов частот от 0,3 до 2,4 ГГц [14].

При работе различных беспроводных устройств даже на разных частотах может возникнуть эффект аддитивности. Определяется этот эффект как сумма плотности потоков энергии, создаваемых каждым передатчиком в определяемой точке. Аддитивность необходимо рассматривать раздельно для эффектов теплового и электрического возбуждения [15].

Приведённые выше факты указывают на то, что следует изучать влияние электромагнитного излучения не в отдельности для различных систем и технологий, а в совокупности, так как при взаимодействии различных устройств негативное влияние на человеческий организм может усиливаться, особенно вследствие эффекта аддитивности.

Биологическое воздействие сетей 5G

Биологические ткани – это материал с очень сложной структурой [15]. На микроскопическом уровне живые ткани состоят из заряженных частиц. Незначительные изменения, связанные с изменением электрических параметров заряженных частиц живых тканей, могут привести к существенным физиологическим изменениям в органах и системах человека. Неблагоприятное воздействие на живые организмы оказывает весь спектр электромагнитного излучения, но с различной интенсивностью.

Внешние электромагнитные поля меняют физико-химические характеристики живых организмов. Заряженные частицы, из которых состоят элементы, под воздействием электрического поля будет поляризоваться. При высоком постоянном напряжении возможен электрофорез, т.е. перемещение таких крупных заряженных частиц, как клетки и молекулы.

С повышением частоты электромагнитного поля происходит изменение свойств тканей. Например, можно заметить, что с повышением частоты диэлектрические потери в тканях существенно возрастают. На частотах около 1 ГГц диэлектрические потери составляют около 50 % от общих потерь, а при 10 и 30 ГГц составляют 90 % и 98 %, соответственно [12,15].

Электромагнитные поля, воздействующие на ткани, вызывают их нагрев. Контролируемое воздействие радиоволн дециметрового и сантиметрового диапазона средней и высокой интенсивности широко используются в физиотерапии для лечения заболеваний, в частности онкологических и сердечно-сосудистых. При локальном воздействии на живые ткани создаётся требуемый температурный режим, при котором меняются условия функционирования пораженного органа [15].

Однако неконтролируемое воздействие излучения на человеческий организм может спровоцировать возникновению различных заболеваний. С ростом частоты электромагнитного излучения эффект от воздействия увеличивается. Так, в диапазоне частот 300 МГц – 300 ГГц поглощается от 40% до 50 % падающей волны (остальное отражается), глубина проникновения в биологические ткани равна примерно 1/10 длины волны [13]. При этом электромагнитная энергия трансформируется в кинетическую и вызывает общий нагрев тканей по всей глубине проникновения внутрь организма.

Рекомендации по снижению негативного влияния сетей 5G.

Введение в эксплуатацию новых сетей связи, предоставление различных услуг и сервисов по беспроводному интерфейсу даёт толчок к повышению уровня жизни населения, общественной безопасности, развитию экономики и здравоохранения. Несмотря на все риски, присущие воздействию электромагнитных полей высокой частоты на человеческий организм, как население, так и бизнес требуют устойчивой и бесперебойной связи 24 часа в сутки.

Применяя современные достижения науки, техники и знания в области защиты от высокочастотных электромагнитных полей, можно добиться баланса для достижения преимуществ технологий без ущерба для здоровья граждан.

Рекомендации по повышению осведомленности о воздействии излучения на здоровье сотовых телефонов, беспроводной связи и сетей 5G, а также рекомендации, которые могут снизить облучение населения подробно рассмотрены в [16]. В целях снижения негативного влияния полей высокой частоты от базовых станций 5G на человека и окружающую среду приведём некоторые рекомендации.

Рекомендация 1. Размещать на специально созданных веб-сайтах ссылки, содержащие информацию и предупреждения о радиочастотном излучении от всех источников, особенно от небольших сот 5G, развернутых в местах массового скопления людей. Также на веб-сайтах должны демонстрироваться материалы освещающие правильное использования сотовых телефонов для минимизации воздействия РЧ-излучения, вестись пропаганда по радио, телевидению, в печатных СМИ и в интернете, предупреждающая о рисках для здоровья, связанных с воздействием радиации от устройств 5G. Важное значение имеют предупреждения, касающиеся детей и подростков, а также беременных женщин. Население должно быть предупреждено о потенциальных опасностях радиочастотного излучения и ему должны быть даны простые рекомендации для уменьшения рисков облучения. Пример предупреждения о радиочастотном излучении приведен на рисунке 1.

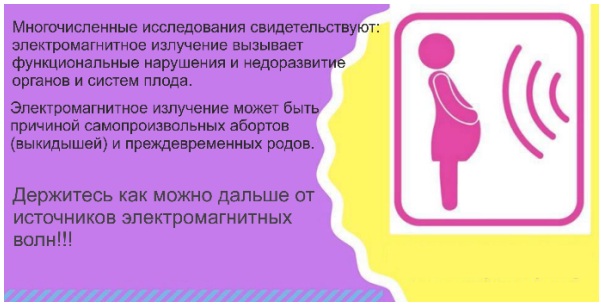

Рис. 1. Пример плакат с предупреждением о радиочастотном излучении [17]

Вебсайт должен предоставлять посетителям возможность регистрировать свое мнение о текущих нормах воздействия. Основное использование данных, собранных в этом реестре, будет заключаться в оценке уровня интереса к радиационному облучению со стороны граждан.

Рекомендация 2. Каждое сооружение, на котором размещается антенна 5G, должно иметь маркировку, указывающую, что присутствует повышенное РЧ-излучение. Плакат должна быть на уровне глаз и хорошо виден для прочтения (см. рис. 2).

Рис. 2. Пример предупреждающего плаката [17]

Рекомендация 3. Школы, библиотеки, детские сады должны переходить с беспроводных радиочастотных соединений для компьютеров, ноутбуков, планшетов и других устройств на проводные или оптические соединения. Существуют доказательства того, что чем младше ребенок, тем он более восприимчив к негативному воздействию радиочастотного излучения. Проводные соединения или оптическая беспроводная связь не подвергают детей воздействию радиочастотного излучения [17].

Новые оптические сетевые решения для учебных и офисных помещений (например, типа Li-Fi) предлагают более быстрое, здоровое и безопасное соединение, чем 5G и Wi-Fi. Технологии типа Li-Fi

используют видимый свет, которому организм может противостоять без какого-либо вреда при гораздо более высоких уровнях интенсивности (например, прямой солнечный свет), чем требуется для передачи данных. Оптическая беспроводная система может быть включена в модернизацию экономичного светодиодного освещения помещений, что может сэкономить значительные затраты на электроэнергию [17].

Рекомендация 4. Измерения мощности сигнала должны проводиться на всех беспроводных объектах в рамках процесса ввода в эксплуатацию и в соответствии с действующими нормативными документами. Измерения также необходимо проводить, когда в систему вносятся изменения, которые могут повлиять на ее излучение, например, изменения в управляющем ею программном обеспечении. Уровень сигнала следует оценивать в худшем случае, в местах, расположенных вблизи башни, а также в местах скопления людей. Результаты по сбору данных должны быть доступны для общественности через веб-сайты. В случае если измеренная мощность беспроводного объекта превышает пороговые значения излучения, администрация имеет право немедленно отключить объект. Измерения должны выполняться независимым подрядчиком, а стоимость измерений будет нести установщик на месте.

Признано, что теоретические расчеты позволяют оценить уровни электромагнитных полей от вводимого в эксплуатацию объекта. Однако бывают случаи, когда излучение от вышек может быть сфокусировано зданиями, местностью и антеннами в режиме формирования лучей, в результате чего уровни сигнала будут значительно выше, чем можно было бы ожидать в теоретических расчетах. Проведение полевых измерений – единственный верный подход к определению реальных уровней напряжённости и плотности потока мощности электромагнитного поля в данном месте [17].

Администрация города должна иметь право вводить в действие правила зонирования, регулирующие размещение средств индивидуальной беспроводной связи в пределах географических границ городов.

Администрации должны установить приоритет размещения сотовых вышек так, чтобы в наиболее безопасном месте было проще получить разрешение на установку базовой станции и, наоборот, в тех местах, где расположение базовой станции будет наиболее сильно облучать население, получение разрешения должно быть наиболее трудным.

Рекомендация 5. Протоколы для выполнения измерений мощности сигнала в зонах вокруг беспроводных устройств, должны оценивать характеристики сигнала, которые, как известно, влияют на здоровье человека. Эти протоколы должны учитывать импульсивный характер излучения с высокой скоростью передачи данных, который, как показывает растущее количество доказательств, оказывает значительно большее негативное влияние на здоровье человека, чем непрерывное излучение. Протоколы также должны учитывать суммарное воздействие нескольких источников излучения [17].

Рекомендация 6. Желательно чтобы базовые станции и опоры с антеннами 5G были удалены от жилых домов, предприятий и школ. Это должно осуществляться администрацией города во время процесса выдачи разрешения, если только владельцы жилых домов, предприятий или школьных округов не откажутся от этого ограничения. Также владельцы собственности должны быть в праве принимать решения о размещении приёмопередающих объектов перед их собственностью.

Важно уделять приоритетное внимание безопасности граждан, особенно потому, что 5G – это модернизация, а не предоставление беспроводных услуг в необслуживаемых районах.

Рекомендация 7. Начать работу по измерению интенсивности радиочастотного излучения в новых миллиметровых частотных диапазонах с целью разработки и уточнения постоянно обновляемой карты уровней радиочастотного облучения. Данные должны быть собраны таким образом, чтобы идентифицировать географические районы с особенно высоким уровнем радиочастотного излучения, места, где радиочастотный сигнал для беспроводной связи отсутствует (мертвые зоны), и места, где радиочастотный сигнал находится в допустимых пределах. Пользователями этих данных будут покупатели, арендаторы недвижимости или население в целом, использующее контрольные значения для сравнения и принятия собственных решений об аренде или проживании в данном районе.

Через некоторое время, при накоплении определённого количества данных возможно создание обширной базы данных по излучению сетями 5G, которая предоставит полезные карты и информацию для будущих исследований в области общественного здравоохранения.

Рекомендация 8. Предлагается, чтобы все новые сотовые телефоны и другие продаваемые беспроводные устройства были оснащены обновленным программным обеспечением, которое может предотвратить излучение телефона, когда он расположен рядом с телом. Сотовые телефоны содержат

датчики приближения, которые позволяют сотовому телефону излучать сигналы только тогда, когда он находится на определенном расстоянии от тела, например, когда его держат пальцами или кладут на стол. Это не влияет на функциональность устройства, а только на способ его использования, в частности, не прижимая его к голове или телу.

Реализовать данную функцию можно путём обновления программного обеспечения сотового телефона, так как в этих телефонах уже есть датчик приближения для отключения экрана и программных клавиш при приближении телефона к телу [17]. При этом изменении экран и ВЧ-цепь автоматически отключаются. Это снижает вероятность возникновения, например, рака мозга. Сотовые телефоны должны поставляться с этой функцией, и с инструкциями в руководстве по его отключению. На устройстве должна быть программная кнопка, чтобы легко снова включить подавление излучения, например, если устройство передано ребенку. Во всех случаях включить ограничение должно быть проще, чем отключить его. Сотовые телефоны, предназначенные специально для детей, при любых обстоятельствах не должны излучать радиоволны, если они расположены вплотную к телу. Также приветствуется установка таких датчиков приближения в ноутбуки и планшеты.

Рекомендация 9. Поощрять внедрение волоконно-оптических линий связи, для внутренних проводных соединений и оптической беспроводной связи для обслуживания всех коммерческих и общественных объектов на территории города. По сравнению с беспроводной связью оптоволокно обладает более лучшими характеристиками для передачи данных: скорость, безопасность и надежность сигнала, при этом избегая биологического воздействия на людей и окружающую среду.

Рекомендация 10. Необходимы дальнейшие фундаментальные научные исследования в сотрудничестве с медицинским сообществом, чтобы выявить характеристики выраженных клинических симптомов, связанных с воздействием радиочастотного излучения. Дальнейшие исследования только начинают изучать квантово-механические механизмы, которые являются фундаментальной основой для понимания биологических изменений, происходящих во время взаимодействия радиочастотного излучения и молекул.

Медицинское сообщество также может помочь определить соответствующие меры защиты. Все эти усилия (фундаментальная наука, клиническая оценка, эпидемиологические исследования) должны быть полностью независимыми и проводиться вне коммерческого влияния.

Рекомендация 11. Использовать предупреждающие знаки о воздействии, которые должны быть вывешены в коммерческих и общественных зданиях. Кроме того, создавать коммерческие и некоммерческие зоны, особенно в медицинских учреждениях, свободные от радиочастотного излучения, где сотрудники и посетители изолированы от воздействия беспроводных радиочастотных излучений. Многие граждане сообщают о чувствительности к электромагнитному излучению, исходящему от устройств, используемых для предоставления услуг сотовой и фиксированной беспроводной связи внутри зданий. Предлагается, чтобы владельцы коммерческих и общественных зданий, особенно в медицинских учреждениях, добровольно размещали у входов указатели, касающиеся уровней радиочастот и зон, свободных от радиочастотного излучения, в пределах этих структур, для оповещения входящих в здание.

Рекомендация 12. Необходимо привлечь организации, обладающие соответствующими научными знаниями, включая экологические знания, для разработки пределов безопасности радиочастотного излучения, которые защитят деревья, растения, птиц, насекомых. Текущие ограничения безопасности были установлены с намерением защитить людей только от краткосрочных последствий, но не защищать флору или фауну от вреда. Необходимо обеспечить защиту окружающей среды и дикой природы с помощью эффективных стандартов безопасности. Ветви деревьев, птицы и опылители будут ближе, чем люди к антеннам сотовой связи 5G и связанной с ней уплотненной инфраструктуре 4G. Фактически, беспроводное излучение от сотовых антенн очень велико в пространстве, окружающем антенну. Оно может превышать пределы на некотором расстоянии от антенны, именно здесь находятся листья деревьев, птицы и насекомые. Таким образом, они более подвержены воздействию ЭМП, находясь в зоне прямой видимости беспроводных радиочастотных лучей.

Заключение

Потребность в беспроводных мультимедийных услугах связи быстро растёт. Новейшие стандарты и технологии сетей пятого поколения IMT2020/5G призваны решить эту проблему [19-32]. Благодаря высоким скоростям передачи данных и гибкости настройки таких сетей связи возможны такие услуги

как телемедицина, дистанционное образование, создание интеллектуальных систем жизнеобеспечения, Интернет вещей. Благодаря этому человек намного повысить уровень и комфорт своей жизни.

Однако на ровне с очевидным рывком технологий и повышения уровня и безопасности жизни существуют и минусы. Например, в сетях 5G будут использоваться новые частотные диапазоны – сантиметровые и миллиметровые. Использование новых более высокочастотных диапазонов вызывает обеспокоенность научного сообщества, а также общественности о том каким образом будет влиять излучения сетей 5G на человека и окружающую среду. Даже небольшое локальное повышение излучения может привести к головной боли и бессоннице, а сильное воздействие может вызвать более серьезные последствия для здоровья. Организовать жилое пространство и выбрать оборудование таким образом, чтобы источники приносили только пользу – задача непростая.

При использовании радио технологий 5G необходимо особенно тщательно контролировать уровень электромагнитного излучения, особенно в режиме направленной передачи (Beamforming), так как возрастает направленность лучей в сторону абонента [18]. Решить эти задачи поможет повышение уровня осведомлённости людей об опасности электромагнитных полей и мер по защите от их воздействия с учетом приведенных выше рекомендаций. Рекомендации, приведённые в данной статье нацелены на минимизацию вредного воздействия излучения сетей 5G на здоровье населения и окружающую среду, а также повышению экологической безопасности связи 5G и качества жизни населения.

Литература

1. Сайт Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации. Режим доступа: URL:https://digital.gov.ru/uploaded/files/proekt-kontseptsii-sozdaniya-i-razvitiya-setej-5g-imt-2020-v-rossijskoj-federatsii.pdf (дата обращения: 10.08.2021).
2. Оператор мобильной связи Tele2: официальный сайт. Режим доступа: URL: https://msk.tele2.ru/journal/article/what-is-5G (дата обращения: 15.08.2021).
3. Сайт Huawei. Режим доступа: URL: https://huawei.ru/news/mts-i-huawei-zapustili-pervye- kommercheskie-zony-5g-v-moskve/ (дата обращения: 14.08.2021).
4. Официальный сайт РИА Новости. Режим доступа: URL: https://ria.ru/20210902/5g1748401537.html (дата обращения: 14.08.2021).
5. *Anwer Al Dulaimi,* 5G Networks: Fundamental Requirements, Enabling Technologies, and Operations Management. Wiley IEEE Press. 2018. P. 784.
6. Сайт ITU. Режим доступа: URL:www.itu.int/ru/mediacentre/backgrounders/Pages/5G-fifth-generation-of-mobile-technologies.aspx (дата обращения: 10.08.2021).
7. *Сизов Д.В., Панкратов Д.Ю.* Оценка влияния электромагнитных полей сетей 5G на человека // Телекоммуникации и информационные технологии. 2021. Т. 8. № 1. С. 13-20.
8. Веб-сайт ETSI. Режим доступа: https://www.etsi.org/deliver/etsi_ts/138100_138199/13810101/16.05.00_60/ts_13810101v160500p.pdf. (дата обращения: 18.09.2021). – Текст: электронный.
9. СанПиН 2.1.8/2.2.4.1190-03 "Гигиенические требования к размещению и эксплуатации средств сухопутной подвижной радиосвязи".
10. *Панкратов Д.Ю., Сизов Д.В.* // «Радиоэлектронные устройства и системы для инфокоммуникационных технологий» («РУЭС-2020»). Москва, 2020. 5 с.
11. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Технологии в системах радиосвязи на пути к 5G. М.: Горячая линия – Телеком, 2018. 280 с.
12. *Тляубердина А. Ш., Антипин Б. М., Виноградов Е. М*. Оценка электромагнитной безопасности путем измерения уровней радиоизлучений в УВЧ-ОВЧ диапазонах // Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2019.
13. *Будашева Н. В*. Оценка воздействия электромагнитного излучения объектов связи на человека: учебно-методическое пособие. Орёл: Академия ФСО России, 2018. 56 с.
14. Руководства ICNIRP по ограничению воздействия переменных электрических, магнитных и электромагнитных полей (до 300 ГГЦ).
15. Оценка риска для здоровья населения при воздействии переменных электромагнитных полей (до 300 ГГц) в условиях населенных мест. Методические рекомендации МР 2.1.10.0061-12 Москва: Федеральный центр гигиены и эпидемиологии Роспотребнадзора, 2013. 35 с.
16. Сайт всемирной организации здравоохранения. Режим доступа: https://www.who.int/ru. (дата обращения: 18.09.2021).
17. Final Report on Commission to Study the Environmental and Health Effects of Evolving 5G Technology (RSA 12-K: 12-14, HB 522, Ch. 260, Laws of 2019). MEMORANDUM. State of New Hampshire.

18. *MinKeun C., Liang L., Andreas J.* - Millimeter-Wave Massive MIMO Testbed with Hybrid Beamforming// Sony Research Center Lund, Sweden – 2012.

19. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Анализ пропускной способности канала MIMO в условиях замираний // Системы синхронизации, формирования и обработки сигналов. 2018. Т. 9. № 2. С. 13-20.

20. *Крейнделин В.Б., Старовойтов М.Ю.* Повышение помехоустойчивости системы связи MIMO с пространственным мультиплексированием методом додетекторного сложения // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 4. С. 4-13.

21. *Бакулин М.Г., Крейнделин В.Б.* Проблема повышения спектральной эффективности и емкости в перспективных системах связи 6G // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 2. С. 25-31.

22. *Крейнделин В.Б., Резнёв А.А*. Матрица пространственно-временного кода высокой размерности типа "Голден" // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 6. С. 34-40.

23. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Исследование вероятностных моделей радиоканала MIMO с учетом взаимной корреляции передающей и приемной сторон с помощью компьютерного моделирования // REDS: Телекоммуникационные устройства и системы. 2017. Т. 7. № 1. С. 64-68.

24. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Алгоритмы нелинейной фильтрации двоичной ЛРП со случайной задержкой и случайной начальной фазой // Системы синхронизации, формирования и обработки сигналов. 2019. Т. 10. № 2. С. 45-51.

25. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Методы приема псевдослучайных последовательностей в системах радиосвязи // REDS: Телекоммуникационные устройства и системы. 2018. Т. 8. № 1. С. 108-112.

26. *Крейнделин В.Б., Григорьева Е.Д.* Анализ быстрого алгоритма умножения матриц и векторов для банка цифровых фильтров // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 1. С. 4-10.

27. *Бакулин М.Г., Бен Режеб Т.Б.К., Крейнделин В.Б., Смирнов А.Э.* Способы минимизации объёма передаваемой информации в обратном канале многоантенных систем MIMO // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 3. С. 17-24.

28. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Применение технологии MIMO в современных системах беспроводной связи разных поколений // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 4. С. 4-12.

29. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Исследование вероятностных моделей радиоканала MIMO с учетом взаимной корреляции передающей и приемной сторон с помощью компьютерного моделирования // REDS: Телекоммуникационные устройства и системы. 2017. Т. 7. № 1. С. 64-68.

30. *Крейнделин В.Б., Григорьева Е.Д*. Реализация банка цифровых фильтров с пониженной вычислительной сложностью // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 7. С. 48-53.

31. *Панкратов Д.Ю., Степанова А.Г.* Компьютерное моделирование технологии MIMO для систем радиосвязи // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 12. С. 33-37.

32. *Панкратов Д.Ю., Сердюков А.А*. Моделирование системы MIMO в режиме Beamforming. DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 2. С. 12-21.

О МНОГОКАНАЛЬНОЙ ТЕОРИИ ПРИЕМНЫХ АНТЕНН

Смирнов Евгений Владимирович,
МТУСИ, ст. преподаватель, Москва, Россия
smirnovmtuci@rambler.ru

Аннотация

В наших предыдущих работах было показано, что поле рассеяния произвольной приемной антенны можно рассматривать как сумму информационной составляющей поля рассеяния (ИСПРА) и ортогональной составляющей (ОСПРА). Каждая из них состоит из основной и дополнительной компоненты. Эти компоненты образуют каналы взаимодействия, по которым происходит отбор мощности приемной антенной от плоской волны. Настоящая работа посвящена исследованию связи компонент ИСПРА и ОСПРА приемной антенны с падающей на нее волной. Показано, что дополнительные компоненты ИСПРА создают каналы взаимодействия по которым от плоской волны отбирается мощность, расходуемая только на образование взаимной мощности между основной и дополнительными компонентами ИСПР и не отбирается мощность, передаваемая от плоской волны к нагрузке антенны. Мощность, поступающая в нагрузку антенны, передается только по каналу взаимодействия основной ИСПРА. В свою очередь мощности управляемой и неуправляемой компонент ИСПРА создаются за счет мощности, отбираемой по каналам основной ИСПРА и ОСПРА.

Ключевые слова: *диаграмма рассеяния, оптическая теорема, основная информационная составляющая, приемная антенна, дополнительная информационная составляющая, ортогональная составляющая.*

Введение

При падении плоской волны на приемную антенну часть ее мощности рассеивается антенной, а часть поступает в нагрузку. Полная мощность, отбираемая при этом от плоской волны называется мощностью экстинкции (P_{ext}) и может быть определена с использованием оптической теоремы [3,4]. Причем для определения P_{ext} нет необходимости знать поле рассеяния антенны во всех направлениях. Достаточно знать только ее диаграмму рассеяния (ДР) в направлении распространения падающей на антенну волны $\vec{A}(\vec{n}_0,\vec{n}_0)$. В [1,2] было показано, что только часть полного поля рассеяния антенны (ППРА) связана с передачей мощности от плоской волны к нагрузке антенны. Эта составляющая была названа информационной составляющей поля рассеяния антенны (ИСПРА), форма которой центрально симметрична диаграмме направленности антенны в режиме передачи. Разбиение ППРА на компоненты, одна из которых связана с передачей информации позволило с этой компонентой связать свой канал взаимодействия, по которому передается мощность от плоской волны к нагрузке антенны. Поскольку в общем случае рассогласованной нагрузки в ППРА возникает составляющая, совпадающая по форме с диаграммой направленности антенны (ДН) в режиме передачи, то и с этой составляющей можно связать свой канал взаимодействия, который был назван диаграммным каналом взаимодействия. На этой основе в [5] была предложена трехканальная модель приемной антенны, в которой в качестве третьего канала взаимодействия рассматривается канал, образованный ортогональной по отношению к ИСПРА составляющей ППРА (ОСПРА). Во всех упомянутых выше работах полагалось, что ИСПРА и ОСПРА не имеют структуры и состоят из одной компоненты. Позже в [6] было показано, что в общем случае ИСПРА и ОСПРА обладают собственной структурой и сами могут состоять из основной компоненты и двух дополнительных. Причем одна из них не зависит от нагрузки (неуправляемая дополнительная компонента) (НДК), а другая зависит (управляемая дополнительная компонента) (УДК). Понимание тонкой структуры ППРА в дальнейшем [7] позволило исследовать вопрос о связи каждой компоненты рассеянного поля с нагрузкой антенны. Было установлено, что токи, создаваемые дополнительной ИСПРА и ОСПРА в нагрузке, компенсируют друг друга. И результирующий ток, протекающий по нагрузке, определяется только током основной ИСПРА. При-

чем этот ток полностью совпадает с током нагрузки, который дает применение теоремы взаимности к приемной антенне.

Таким образом, можно считать, что связи, существующие между токами нагрузки и всеми компонентами поля рассеяния антенны, довольно глубоко изучены. Чего нельзя сказать о связях компонент поля рассеяния с полем плоской волны, падающей на приемную антенну. Поскольку трехканальная модель приемной антенны не учитывала наличие дополнительных компонент ИСПРА и ОСПРА в ее поле рассеяния, а, следовательно, и дополнительных каналов взаимодействия рассеянного поля с полем плоской волны, то появляется необходимость в уточнении этой модели. Решению такой задачи и посвящена настоящая работа. В общем случае в виду того, что рассеянное поле представляется как сумма трех компонент ИСПРА и трех компонент ОСПРА возникает шесть каналов взаимодействия. Для понимания процесса взаимодействия каждой компоненты ППРА с плоской волной необходимо во-первых определить сколько мощности отбирается от волны по каждому каналу взаимодействия, а во-вторых ответить на вопрос на что расходуется эта мощность. Здесь следует отметить, что поскольку нас интересует, прежде всего, физическая картина процесса взаимодействия приемной антенны с полем, падающей на нее плоской волны, то рассматривать отдельно работу всех шести каналов не всегда целесообразно. Например, составляющая ППРА, совпадающая по форме с ДН антенны является суммой управляемых дополнительных компонент ИСПРА и ОСПРА. И поэтому логично рассматривать процесс взаимодействия диаграммной составляющей ППРА в виде единого канала взаимодействия диаграммной составляющей с полем плоской волны, а не в виде двух каналов для дополнительных компонент ИСПРА и ОСПРА. Также целесообразно отдельно рассматривать канал взаимодействия, по которому мощность от плоской волны передается к нагрузке антенны. Ниже приведены основные соотношения, характеризующую работу приемной антенны в многоканальном режиме.

Основная часть

При падении плоской волны на приемную антенну возникает рассеянное антенной поле $\dot{\vec{E}}_s(\vec{n}_0,\vec{r}_0)$, которое в общем виде можно записать в форме

$$\dot{\vec{E}}_s(\vec{n}_0,\vec{r}_0)=\dot{\vec{A}}_s(\vec{n}_0,\vec{r}_0)\frac{\exp(-ikr)}{r}, \tag{1}$$

где $\dot{\vec{A}}_s(\vec{n}_0,\vec{r}_0)$ - комплексная диаграмма рассеяния антенны, $\vec{r}_0$ - единичный вектор в сферической системе координат (r,θ,φ), а $\vec{n}_0$ - единичный вектор направления распространения, падающей волны.

Как указывалось выше, основным соотношением, характеризующим энергетику процесса взаимодействия приемной антенны с падающей плоской волной, является оптическая теорема

$$P_L+P_s=-\frac{2\pi}{kZ_0}Jm\left(\vec{e}_0^{\,*},\dot{\vec{A}}_s(\vec{n}_0,\vec{n}_0)\right), \tag{2}$$

В (2) P_L – мощность, выделяемая в нагрузке характеризуемой коэффициентом отражения Γ, а P_s – мощность, рассеиваемая антенной, Z_0 – волновое сопротивление свободного пространства, а $k=\frac{2\pi}{\lambda}$ - волновое число. Единичный вектор $\vec{e}_0$ – характеризует поляризацию плоской волны, $\dot{\vec{A}}_s(\vec{n}_0,\vec{n}_0)$ – диаграмма рассеяния антенны (ДРА) в направлении распространения плоской волны. В [5] ППРА представлялось в виде суммы только ИСПРА, ОСПРА и управляемой диаграммной составляющей поля рассеяния антенны (УДСПРА), поэтому подставляя в (2) соответствующие составляющие $\dot{\vec{A}}_s(\vec{n}_0,\vec{n}_0)$ для общего случая рассогласованной с антенной нагрузки была сформулирована трехканальная модель работы приемной антенны в виде трех оптических теорем, каждая из которых описывает свой канал взаимодействия

$$-\frac{2\pi}{kZ_0}Jm\left(\overset{*}{\vec{\mathrm{e}}}_0,\dot{\vec{A}}_{\inf}(\vec{n}_0,\vec{n}_0))\right)=P_L+P_d+P_{\inf}=2P_L^{\max} \tag{3}$$

$$-\frac{2\pi}{kZ_0}Jm\left(\overset{*}{\vec{\mathrm{e}}}_0,\dot{\vec{A}}_d(\vec{n}_0,\vec{n}_0))\right)=-2P_L^{\max}\,\mathrm{Re}\left(\frac{\tilde{A}-\overset{*}{\gamma}}{1-\gamma\tilde{A}}\overset{*}{\vec{A}}_{\perp}^{0}\right)=P_{mut} \tag{4}$$

$$-\frac{2\pi}{kZ_0}Jm\left(\overset{*}{\vec{\mathrm{e}}}_0,\dot{\vec{A}}_{\perp}(\vec{n}_0,\vec{n}_0))\right)=P_{\perp} \tag{5}$$

где $\dot{\vec{A}}_{\inf}(\vec{n}_0,\vec{n}_0))$, $\dot{\vec{A}}_d(\vec{n}_0,\vec{n}_0))$ и $\dot{\vec{A}}_{\perp}(\vec{n}_0,\vec{n}_0))$ - диаграммы рассеяния информационной, диаграммной и ортогональной составляющих в направлении падения плоской волны, $\tilde{A}$ и γ – коэффициенты отражения от нагрузки и входа антенны соответственно, P_d, $P_{\inf}$, $P_{\perp}$ - мощности УДСПРА, ИСПРА и ОСПРА, P_{mut} – взаимная мощность между ДСПРА и оставшейся частью ППРА, а $P_L^{\max}$ – мощность, выделяемая в нагрузке при $\Gamma=\overset{*}{\gamma}$, определяемая соотношением

$$P_L^{\max}=\frac{\pi D}{2k^2Z_0}\left|(\dot{\vec{F}}(-\vec{n}_0)\cdot\dot{\vec{\mathrm{e}}}_0)\right|^2, \tag{6}$$

где $\dot{\vec{F}}(\vec{r}_0)$ – ДН приемной антенны в режиме передачи, а D – максимальное значение коэффициента направленного действия (КНД) антенны.

Соотношения (3-5) имеют простое физическое содержание. Первое из них описывает работу информационного канала взаимодействия, по которому мощность от плоской волны передается к антенне и расходуется в ней на мощность, выделяемую в нагрузке и на мощности рассеяния информационной и диаграммной составляющих поля рассеяния. Мощность, отбираемая по диаграммному каналу взаимодействия, как видно из (4), расходуется только на взаимную мощность между УДСПРА и оставшейся частью ППРА. В случае согласования антенны и нагрузки ($\Gamma=\overset{*}{\gamma}$) ДСПРА, вызванная отражениями от нагрузки и антенны, пропадает, взаимная мощность исчезает и диаграммный канал выключается. Модель приемной антенны из трехканальной превращается в двухканальную. Как следует из (5) мощность, поступающая в антенну по ортогональному каналу взаимодействия, расходуется только на мощность ОСПРА. Поскольку в общем случае ИСПРА и ОСПРА обладают собственной структурой, которая была изучена нами ранее в [6], ППРА может быть представлено в виде

$$\dot{\vec{\mathrm{E}}}_s(\vec{n}_0,\vec{r}_0)=\dot{\vec{E}}_{\inf}(\vec{n}_0,\vec{r}_0)+\dot{\vec{E}}_{\perp}(\vec{n}_0,\vec{r}_0), \tag{7}$$

где ИСПРА $\dot{\vec{E}}_{\inf}(\vec{n}_0,\vec{r}_0)$ и ОСПРА $\dot{\vec{E}}_{\perp}(\vec{n}_0,\vec{r}_0)$ определяются следующими соотношениями

$$\dot{\vec{E}}_{\inf}(\vec{n}_0,\vec{r}_0)=\dot{\vec{E}}_{\inf}^{bas}(\vec{n}_0,\vec{r}_0)+\dot{\vec{E}}_{\inf}^{add1}(\vec{n}_0,\vec{r}_0)+\dot{\vec{E}}_{\inf}^{add2}(\vec{n}_0,\vec{r}_0), \tag{8}$$

$$\dot{\vec{E}}_{\perp}(\vec{n}_0,\vec{r}_0)=\dot{\vec{E}}_{\perp}^{bas}(\vec{n}_0,\vec{r}_0)+\dot{\vec{E}}_{\perp}^{add1}(\vec{n}_0,\vec{r}_0)+\dot{\vec{E}}_{\perp}^{add2}(\vec{n}_0,\vec{r}_0), \tag{9}$$

В (8) $\dot{\vec{E}}_{\inf}^{bas}(\vec{n}_0,\vec{r}_0)$, $\dot{\vec{E}}_{\inf}^{add1}(\vec{n}_0,\vec{r}_0)$ и $\dot{\vec{E}}_{\inf}^{add2}(\vec{n}_0,\vec{r}_0)$ - основная и дополнительные компоненты ИСПРА, определяемые соотношениями

$$\vec{\dot{E}}^{bas}_{\text{inf}}(\vec{n}_0,\vec{r}_0)=\frac{D}{2ik}(\vec{\dot{F}}(-\vec{n}_0)\cdot\dot{\vec{\text{e}}}_0)\overset{*}{\vec{\dot{F}}}(-\vec{r}_0)\frac{\exp(-ikr)}{r}, \tag{10}$$

$$\vec{\dot{E}}^{add1}_{\text{inf}}(\vec{n}_0,\vec{r}_0)=\alpha_2\vec{\dot{E}}^{bas}_{\text{inf}}(\vec{n}_0,\vec{r}_0), \tag{11}$$

$$\vec{\dot{E}}^{add2}_{\text{inf}}(\vec{n}_0,\vec{r}_0)=\left(\alpha_3\frac{\tilde{A}-\overset{*}{\gamma}}{1-\gamma\tilde{A}}\right)\vec{\dot{E}}^{bas}_{\text{inf}}(\vec{n}_0,\vec{r}_0). \tag{12}$$

В (11) и (12) коэффициенты α_2 и α_3 определяются соотношениями

$$\alpha_2=\pm\sqrt{1-\frac{D}{D_s\alpha_1^2}\left|(\vec{\dot{F}}(-\vec{n}_0)\cdot\dot{\vec{\text{e}}}_0)\right|^2}, \tag{13}$$

$$\alpha_3=\frac{D}{4\pi}\oint_{4\pi}\vec{\dot{F}}(\vec{r}_0)\vec{\dot{F}}(-\vec{r}_0)\partial\Omega. \tag{14}$$

В (13) D_s и D – это максимальные значения КНД диаграммы рассеянного антенной поля при $\Gamma=\overset{*}{\gamma}$ и диаграммы направленности антенны соответственно, а α_1 определяется выражением

$$\alpha_1=Jm\left(\overset{*}{\dot{\vec{\text{e}}}}_0,\ \text{e}^{i\varphi_s}\dot{\overline{F_s}}(\vec{n}_0,\vec{n}_0))\right), \tag{15}$$

где $\dot{\overline{F_s}}(\vec{n}_0,\vec{n}_0)$ - комплексная нормированная к максимуму диаграмма рассеяния антенны в направлении падения плоской волны, а φ_s – аргумент комплексной амплитуды диаграммы рассеяния $\dot{\vec{A}}_s(\vec{n}_0,\vec{r}_0)$.

С учетом того, что управляемая диаграммная составляющая ППРА $\vec{\dot{E}}_d(\vec{n}_0,\vec{r}_0)$ равна сумме вторых дополнительных компонент ИСПРА и ОСПРА

$$\vec{\dot{E}}_d(\vec{n}_0,\vec{r}_0)=\vec{\dot{E}}^{add2}_{\text{inf}}(\vec{n}_0,\vec{r}_0)+\vec{\dot{E}}^{add2}_{\perp}(\vec{n}_0,\vec{r}_0) \tag{16}$$

и соотношений (8) и (9) ППРА может быть представлено в виде

$$\dot{\vec{\text{E}}}_s(\vec{n}_0,\vec{r}_0)=\vec{\dot{E}}^{bas}_{\text{inf}}(\vec{n}_0,\vec{r}_0)+\vec{\dot{E}}^{add1}_{\text{inf}}(\vec{n}_0,\vec{r}_0)+\dot{\vec{\text{E}}}_d(\vec{n}_0,\vec{r}_0)+$$
$$+\vec{\dot{E}}^{bas}_{\perp}(\vec{n}_0,\vec{r}_0)+\vec{\dot{E}}^{add1}_{\perp}(\vec{n}_0,\vec{r}_0). \tag{17}$$

Мощность, рассеянная антенной P_s складывается из мощностей отдельных компонент поля, входящих в (17) и взаимной мощности между ними. Взаимная мощность между ортогональными и информационными компонентами отсутствует, так как они удовлетворяют условию ортогональности, физическим содержанием которого является отсутствие взаимной мощности между этими компонентами. Поэтому полная взаимная мощность будет состоять только из трех составляющих, одна из которых нам известна. Это взаимная мощность между УДСПРА $\dot{\vec{\text{E}}}_d(\vec{n}_0,\vec{r}_0)$ и оставшейся частью ППРА, которая возникает из-за того, что в состав УДСПР согласно (16) входят вторые дополнительные информационные и ортогональные компоненты.

Они взаимодействуют с основными и первыми дополнительными компонентами информационной и ортогональных составляющих, которые входят в оставшуюся часть поля рассеяния. Эта мощность P_{mut} описывается правой частью оптической теоремы для диаграммного канала взаимодействия (4) и равна мощности, передаваемой от плоской волны к антенне по диаграммному каналу взаимодействия. Второй составляющей полной взаимной мощности является взаимная мощность между основной и первой дополнительной составляющей ИСПРА P_{inf}^{mut}. Поскольку для этих составляющих известны их представления в явном виде (10) и (11), то найти взаимную мощность между ними не представляет труда

$$P_{inf}^{mut} = \alpha_2 \frac{\pi D}{k^2 Z_0} \left| (\dot{\vec{F}}(-\vec{n}_0) \cdot \dot{\vec{e}}_0) \right|^2 . \tag{18}$$

Подставляя в правую часть оптической теоремы выражение для первой дополнительной ИСПРА из (11) получим мощность, отбираемую от плоской волны по каналу взаимодействия, который образовался за счет учета первой дополнительной ИСПРА

$$-\frac{2\pi}{kZ_0} Jm\left(\overset{*}{\vec{e}}_0, \dot{\vec{A}}_{\text{inf}}^{add1}(\vec{n}_0, \vec{n}_0) \right) = \alpha_2 \frac{\pi D}{k^2 Z_0} \left| (\dot{\vec{F}}(-\vec{n}_0) \cdot \dot{\vec{e}}_0) \right|^2 . \tag{19}$$

Из сравнения выражений (18) и (19) можно сделать вывод о том, что источником взаимной мощности P_{inf}^{mut} является мощность, отбираемая от плоской волны по первому дополнительному информационному каналу. Таким образом, исходная трехканальная модель приемной антенны за счет учета дополнительного компонента ИСПРА превратилась в четырехканальную модель. Третьей составляющей ($P_{\perp}^{mut}$) полной взаимной мощности является взаимная мощность между основной и первой дополнительной ОСПРА. Так как в общем случае для этих составляющих отсутствуют аналитические представления, то провести исследования, аналогичные предыдущему исследованию для компонент ИСПРА не представляется возможным. Поэтому целесообразно объединить основную и первую дополнительную компоненты ОСПРА и рассматривать единый ортогональный канал взаимодействия $\dot{\vec{A}}_{\perp}^{\Sigma}(\vec{n}_0, \vec{n}_0)$. При этом несложно показать, что мощность, отбираемая от плоской волны по этому каналу, будет расходоваться на мощность основной $P_{\perp}^{bas}$ и первой дополнительной ОСПРА $P_{\perp}^{add1}$, взаимную мощность между ними $P_{\perp}^{mut}$ и мощность первой дополнительной ИСПРА. В итоге многоканальная модель произвольной приемной антенны может быть записана в виде

$$-\frac{2\pi}{kZ_0} Jm\left(\overset{*}{\vec{e}}_0, \dot{\vec{A}}_{\text{inf}}(\vec{n}_0, \vec{n}_0) \right) = P_L + P_d + P_{\text{inf}} = 2P_L^{\max}, \tag{20}$$

$$-\frac{2\pi}{kZ_0} Jm\left(\overset{*}{\vec{e}}_0, \dot{\vec{A}}_d(\vec{n}_0, \vec{n}_0) \right) = -2P_L^{\max} \operatorname{Re}\left(\frac{\tilde{A} - \gamma}{1 - \gamma \tilde{A}} \overset{*}{\tilde{A}}_{\perp}^0 \right) = P_{mut}, \tag{21}$$

$$-\frac{2\pi}{kZ_0} Jm\left(\overset{*}{\vec{e}}_0, \dot{\vec{A}}_{\text{inf}}^{add1}(\vec{n}_0, \vec{n}_0) \right) = \alpha_2 \frac{\pi D}{k^2 Z_0} \left| (\dot{\vec{F}}(-\vec{n}_0) \cdot \dot{\vec{e}}_0) \right|^2 , \tag{22}$$

$$-\frac{2\pi}{kZ_0} Jm\left(\overset{*}{\vec{e}}_0, \dot{\vec{A}}_{\perp}^{\Sigma}(\vec{n}_0, \vec{n}_0) \right) = P_{\perp}^{bas} + P_{\perp}^{add1} + P_{\perp}^{mut} + P_{\text{inf}}^{add1}. \tag{23}$$

Заключение

В настоящей работе исследовались рассеивающие свойства приемных антенн. В результате проведенного исследования удалось выяснить тонкую структуру процесса взаимодействия составляющих рассеянного антенной поля с полем, падающей на нее плоской волны. Известная ранее трехканальная модель приемной антенны была построена в предположении, что информационная и ортогональные составляющие имеют простую структуру. Дальнейшие исследования показали, что каждая из этих составляющих имеет дополнительные компоненты, с которыми можно связать новые каналы взаимодействия антенны и плоской волны. Поэтому исходная трехканальная модель была преобразована в многоканальную модель. По дополнительному каналу взаимодействия от плоской волны отбирается мощность, расходуемая только на образование взаимной мощности между основной и неуправляемой дополнительной компонентой ИСПРА. В свою очередь мощности управляемой и неуправляемой компонент ИСПРА создаются за счет мощности, отбираемой по каналам ИСПРА и ОСПРА. Мощность, поступающая в нагрузку антенны, передается только по каналу взаимодействия основной ИСПРА.

Литература

1. *Смирнов Е.В.* Об информационной составляющей поля рассеяния приемных антенн // INTERMATIC 2014 Материалы Международной научно-технической конференции “Фундаментальные проблемы радиоэлектронного приборостроения”, 1-5 декабря 2014 г., М.: МИРЭА, 2014. Часть 5. С. 19-23.

2. *Смирнов Е.В.* Исследование информационного канала взаимодействия произвольных приемных антенн // T-Comm: Телекоммуникации и транспорт. 2015. Т. 9. № 7. С. 41-46.

3. *Ерохин Г.А.* О достижимых характеристиках в трехмерных задачах синтеза пассивных рассеивателей // РЭ. 1986, Т.31. С. 1447-1450.

4. *Ерохин Г.А.* Оптическая теорема для приемных антенн и ее следствия // РЭ. 1990, Т. 35. С. 2065-2071.

5. *Смирнов Е.В.* Исследование трехканальной модели произвольных приемных антенн // T-Comm: Телекоммуникации и транспорт. 2016. Т. 10. №9. С. 9-13.

6. *Смирнов Е.В.* Исследование структуры информационной составляющей поля рассеяния приемных антенн // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 8. С. 19-26.

7. *Смирнов Е.В.* О связи компонент информационной и ортогональной составляющих поля рассеяния приемной антенны с ее нагрузкой // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11. №3. С. 67-73.

АНАЛИЗ ФУНКЦИОНАЛЬНЫХ ВОЗМОЖНОСТЕЙ СОВРЕМЕННЫХ МЕССЕНДЖЕРОВ

Фатхулин Тимур Джалилевич,
Московский технический университет связи и информатики, старший преподаватель, Москва, Россия
t.d.fatkhulin@mtuci.ru

Куликова Алена Андреевна,
Московский технический университет связи и информатики, студентка группы БВТ1802, Москва, Россия
alena0365@gmail.com

Аннотация

В работе рассматриваются функциональные возможности современных приложений для обмена сообщениями. Предметом исследования являются клиент-серверные приложения для обмена сообщениями. Цель работы – рассмотреть наиболее популярные современные мессенджеры, их функциональные возможности, выявить достоинства каждого из рассматриваемых мессенджеров и информацию об их уровнях безопасности. Методологическую основу работы составляют метод теоретического анализа и описательный метод, метод обобщения и сравнительный метод.

Ключевые слова: *Мессенджер, обмен сообщениями, корпоративный мессенджер, безопасность мессенджеров, клиент-серверные приложения, функциональные возможности мессенджеров.*

Введение

В современном мире пользователям смартфонов сложно представить жизнь без возможности мгновенного обмениваться сообщениями, совершать аудио- и видеозвонки. Решением для осуществления данных потребностей клиентов стали приложения для обмена сообщениями (мессенджеры, от англ. messenger — курьер, посланник). Они также предоставляют возможность осуществлять обмен фотографиями, видео, документами и аудиозаписями. В работе рассматриваются этапы развития технологий обмена сообщениями. Представлены наиболее значимые программные реализации мессенджеров, которые определили основные функции современных мессенджеров. Приведена информация по мессенджерам, которые используются для общения сотрудников различных корпораций. Указан актуальный рейтинг популярности приложений для обмена сообщениями. Определены достоинства рассматриваемых мессенджеров. В заключении работы сделаны выводы про проведенному анализу функциональных возможностей современных мессенджеров.

Развитие технологий обмена сообщениями

Первым приложением для обмена сообщениями, которое получило широкое распространение по всему миру, было ICQ. Оно появилось в 1996 году. Однако сама идея обмена сообщениями в реальном времени появилась в 60-х годах, когда появились первые сети. В Массачусетском Технологическом институте была создана система CTSS – Compatible Time-Sharing System. Эта система позволяла одновременно нескольким пользователям обмениваться текстовыми сообщениями. Ей стали пользоваться работники института. Однако возможности данной системы ограничивались лишь текстовыми сообщениями, вычислительные системы того времени большего позволить не могли [4].

В 70-х годах появилась возможность обмениваться и текстовыми сообщениями и файлами. В 80-х годах, когда персональные компьютеры все больше стали распространяться по миру, системы для общения стали активно развиваться. Пользователям предоставлялись различные бесплатные и платные сервисы. Например, BBS – bulletin board system – электронная доска сообщений. Проблема состояла в том, что подобные сервисы требовали от пользователей определенных знаний в области компьютерных и сетевых технологий, о современной простоте использования мессенджеров речи не шло.

Во второй половине 90-х годов Интернет стал более обычным и вполне доступным явлением, компьютеры стали использовать все большее количество людей. Израильская компания Mirabilis, основанная четырьмя студентами, запустила программу Интернет-пейджер – ICQ, тем самым совершив революцию в области общения в Интернете. Данная программа совмещала в себе возможности электронной почты и чата. Этот мессенджер и стал основой для всех современных мессенджеров.

Важным этапом в развитии мессенджеров стало создание Skype в 2003 году. С появлением данного мессенджера стало возможным осуществление голосовых звонков. В 2005 году после выхода версии Skype 2.0 появилась возможность осуществлять видеозвонки.

В 2009 году компания WhatsApp Inc. выпустила самый популярный на сегодняшний день мессенджер – WhatsApp.

Возможности современных мессенджеров

В настоящее время мессенджер предоставляет возможность осуществлять обмен не только текстовыми сообщениями и файлами. Современный мессенджер имеет ряд следующих основных функций:

1. мгновенная передача текстовых сообщений;
2. архивирование переписок;
3. отправка файлов различного расширения;
4. прикрепление фото, аудио, видео и других документов;
5. отправка геолокации;
6. голосовые и видео звонки;
7. запись голосовых и видео сообщений;
8. создание групповых чатов;
9. возможность использования смайликов/стикеров;
10. создание фотографий и видео непосредственно в приложении;
11. настройка параметров под нужды пользователя.

Актуальность мессенджеров определяется их простотой в использовании, доступностью и практичностью. Благодаря мессенджерам у людей появилась возможность работать удаленно, что в настоящее время очень востребовано из-за пандемии. Приложения для обмена сообщениями стали активно применяться не только для личного общения, но и стали полноценным рабочим инструментом. Например, интеграция с Telegram на сегодняшний день стала очень важной для множества пользователей различных CRM-систем. Она предоставляет возможность наладить взаимодействие с клиентам, а также решать вопросы внутри компании.

Мессенджер – это программы или приложения, которые можно установить на смартфон, персональный компьютер, ноутбук или планшет. Мессенджер позволяет мгновенно обмениваться с друзьями текстовыми сообщениями, совершать аудио- и видеозвонки, отправлять фотографии, видео, аудиосообщения и различные файлы.

В современном обществе подобные приложения стали чем-то обыденным в жизни. Мессенджеры стали самым популярным средством для общения между собеседниками. Количество таких приложений на рынке растет с каждым годов.

Пользователям мессенджеров важно, чтобы в приложении был прозрачный и понятный интерфейс, позволяющий за пару нажатий на кнопки оказаться в нужном пользователю разделе приложения. Также немаловажным условием является отсутствие надоедающих рекламных баннеров.

После того, как стабильность и скорость работы беспроводных сетей позволили людям обмениваться сообщениями не только с помощью компьютеров и ноутбуков, основные приложения для общения выпустили версии для мобильных браузеров и версии ПО для смартфонов. Данные нововведения изменили восприятие мессенджеров, как средств коммуникации [2].

Эти изменения произошли после появления смартфонов на ОС Android от Google и на IOS от Apple. WhatsApp, который был создан компанией WhatsApp Inc., был первым мессенджером, имеющим привязку к номеру мобильного телефона. Первая версия приложения была заменой стандартного контакт-листа, где пользователи могли установить свой статус (свободен или занят). Вскоре WhatsApp получил поддержку push-уведомлений для смартфонов на IOS от Apple. Это стало переломным моментом для WhatsApp’а. Пользователи начали использовать эти уведомления как метод общения между разными ОС.

Для работы пользователя с приложением для обмена сообщениями необходимо обеспечить минимальные требования:

1. доступ к сети Интернет (для использования мессенджеров на смартфонах нужен доступ к сети Wi-Fi или доступ к мобильному Интернету);
2. для использования мессенджера нужно зарегистрироваться в приложении, обычно требуется указать номер телефона;
3. на устройстве собеседника также должен быть установлен мессенджер, с которого отправляют сообщения;

Приложения для обмена сообщениями имеют множество преимуществ. Рассмотрим некоторые из них:

1. обмен текстовыми сообщениями;
2. использование приложений для обмена сообщениями – бесплатно;
3. имеется возможность отправлять фотографии, видео, аудиосообщения и различные файлы;
4. создание групповых чатов;
5. возможно совершать аудио- и видеозвонки;
6. можно заносить переписки в архив;
7. возможно отправить геолокацию;

Многие мессенджеры в настоящее время поддерживают функцию создания резервных копий в облачных хранилищах. Данная функция очень полезна по следующим причинам:

1. в случае серьезной неисправности смартфона не всегда есть возможность сохранить все данные, но с помощью резервного копирования можно восстановить переписку на другом смартфоне;
2. при покупке нового смартфона также есть возможность восстановить переписку;
3. при удалении программы или сбросе смартфона к заводским настройкам можно будет восстановить переписку.

Приложения для обмена сообщениями как корпоративный способ коммуникации

Корпоративный мессенджер – это приложение, с помощью которого сотрудники одной компании могут общаться друг с другом: обмениваться текстовыми сообщениями, документами, файлами, создавать групповые чаты, информационные каналы, а также решать множество различных рабочих задач. Приложение может иметь мобильную версию, версию для ПК и Web-версию.

Внедрение корпоративных приложений для обмена сообщениями помогает упростить коммуникацию между сотрудниками компании. Особенно актуально это решение в настоящее время, когда большинство компаний переходит на удаленную работу из-за пандемии. Весь персонал организации получает одновременный доступ к информации, за счет чего решение задач происходит намного быстрее, чем по электронной почте или с помощью телефонных звонков.

Ввиду актуальности корпоративных приложений для обмена сообщениями рассмотрим их некоторые преимущества:

– удобство общения и взаимодействия, а также мгновенный доступ к информации: сотрудники компании всегда находятся на связи, все данные переписок сохраняются, что позволяет просмотреть информацию в любое удобное для сотрудника время;

– возможность создания групповых чатов по темам: в мессенджерах есть возможность создавать групповые чаты, каждый из них может служить чатом для обсуждения определенной темы;

– возможность совершать групповые видео-звонки: сотрудники могут создать групповую видео-конференцию.

Корпоративные приложения для обмена сообщениями являются наиболее оптимальным вариантом для поддержания связи между сотрудниками компании.

Подобного рода приложения имеют высокий уровень защищенности. Высокий уровень защищенности обуславливается тем, что в большинстве случаев корпоративный мессенджер работает в локальной сети и размещается на собственных серверах компании, что позволяет защитить приложение от взломов и несанкционированного доступа к информации компании.

Существует множество корпоративных приложений для обмена сообщениями, рассмотрим по уровню безопасности наиболее популярные и распространенные из них.

Telegram. В настоящий момент является одним из самых безопасных мессенджеров. У Telegram имеется собственный протокол шифрования MTProto [5] Данный метод шифрования считается наи-

более безопасным: все сообщения шифруются на устройстве отправителя и расшифровываются у получателя. Компания усердно работает над усилением безопасности, раз в год она проводит конкурс с денежным вознаграждением тем, кто сможет найти уязвимость в мессенджере.

Skype. Безопасность данного приложения для обмена сообщениями не самая сильная его сторона. Большая часть мер безопасности относятся к стандартным [6]. Например, HTTPS – протокол для веб-версии, шифрование с помощью алгоритма AES – 256. Недавно Skype начал поддерживать создание секретных чатов с шифрованием end-to-end. Несмотря на все меры, у данного приложения для обмена сообщениями есть проблемы с безопасностью. В 2018 году была обнаружена уязвимость, позволяющая вредоносному программному обеспечению получать доступ к компьютерам пользователей.

Discord. Данное приложение для обмена сообщениями изначально создавалось для людей, которые играют в видеоигры, сейчас Discord активно используется в бизнес-сфере. В нем есть возможность создавать групповые чаты, видеоконференции. В мессенджере можно создать общий чат для всей компании и отдельный для каждой из команд. Discord по безопасности не уступает Telegram. Он разработан с продвинутыми алгоритмами шифрования, позволяющими сделать общение более безопасным. В приложении также есть защита от DDoS-атак (Distributed Denial of Service) и утечек IP. Пользователь может настроить двухфакторную аутентификацию, с помощью которой вход в аккаунт пользователя возможен только после подтверждения на телефоне.

Slack. Данное приложение для обмена сообщениями было первым корпоративным мессенджером на рынке. Он работает по HTTPS-протоколу, как и Discord поддерживает двухфакторную аутентификацию. На платформе Slack для всех пользователей по умолчанию настроено шифрование данных при передаче и хранении. В качестве дополнительной защиты используются такие инструменты, как журналы аудита, Slack EKM (Enterprise Key Management – корпоративное управление ключами) и интеграция с ведущими поставщиками DLP (Data Leak Prevention – средства предотвращения утечек конфидециальный информации) [7]. Также в Slack есть функция kill switch, с помощью которой при хакерской атаке администратор чата может одновременно сделать запрос на смену паролей у всех пользователей.

Hangouts. Приложение для обмена сообщениями, созданное компанией Google. Он используется для конфиденциального обмена информацией. Сообщения Hangouts шифруются по умолчанию в момент передачи [8]. Шифрование соответствует стандартам Инженерного совета Интернета (IETF) в отношении протоколов DTLS (Datagram Transport Layer Security – протокол датаграмм безопасности транспортного уровня) и SRTP (Secure Real-time Transport Protocol – безопасный протокол передачи данных в реальном времени). Приложение для обмена сообщениями поддерживает двухфакторную аутентификацию.

Результаты проведенного сравнительного анализа мер безопасности вышеприведенных корпоративных приложений для обмена сообщениями приведены в таблице 1.

Таблица 1

Информация о безопасности мессенджеров

Приложение для обмена сообщениями	Безопасность
Telegram	Собственный протокол шифрования MTProto, 256-битное шифрования, возможность создания секретных чатов
Skype	Большинство мер безопасности – стандартные, вследствие чего было много утечек данных. Возможность создавать чаты с шифрованием end-to-end
Discord	Продвинутые алгоритмы шифрования, защита от DDoS-атак, утечек IP, возможность настройки двухфакторной аутентификации
Slack	HTTPS-протокол, поддержка двухфакторной аутентификации
Hangouts	Шифрование сообщений в момент передачи, протоколы DTLS и SRTP, двухфакторная аутентификация

Рейтинг популярности приложений для обмена сообщениями

В наше время все большее количество людей стало использовать приложения для обмена сообщениями. Рынок приложений для обмена приложениями растет, рассмотрим наиболее популярные и востребованные из них в настоящий момент, а также выявим их достоинства.

Одним из самых популярных приложений для обмена сообщениями является WhatsApp. Количество его пользователей составляет более 2 миллиардов. Это приложение бесплатно и очень простое в использовании. Для связи используется мобильный Интернет или Wi-Fi. С помощью этого приложения можно отправлять различные файлы, обмениваться сообщениями, создавать групповые чаты, совершать видео и аудио-звонки.

Достоинства WhatsApp:

1. осуществление входа со смартфона;
2. наличие веб-версии мессенджера;
3. возможность создать резервную копию чатов;
4. простой и удобный интерфейс;
5. возможность перенести данные с одного устройства на другое.

Facebook Messenger уступает WhatsApp по количеству пользователей – в нем зарегистрирована около 1.3 миллиардов. Изначально он был частью социальной сети Facebook (с 28 октября 2021 компания, владеющая Facebook была переименована в Meta Platforms Inc.). В 2011 году Facebook Messenger стал отдельным сервисом. Если в случае с WhatsApp приложение для обмена сообщениями просто в использовании, то с Facebook Messenger использование приложения более сложное. Однако это не мешает его пользователям общаться между собой.

Достоинства Facebook Messenger:

1. наличие оптимизированной под слабые Интернет-соединения версии – Messeger Lite;
2. возможность создания групповых чатов;
3. возможность отправлять голосовые сообщения;
4. все данные привязаны к аккаунту пользователя (при смене устройства все переписки сохраняются).

WeChat. Приложение для обмена сообщениями, которое распространено в большей степени в странах Азии, в особенности в Китае. Количество его пользователей составляет 1.2 миллиарда. WeChat позволяет обмениваться сообщениями, видеосообщениями и стикерами. Данное приложение является не просто мессенджеров, а полноценной официальной платежной системой для коммунальных услуг, заказа еды, такси, товаров.

Достоинства WeChat:

1. возможность отправлять голосовые, музыкальные сообщения;
2. возможность создания группового чата;
3. возможность делиться контактными данными с помощью Bluetooth;
4. наличие системы поиска собеседника, находящегося неподалеку от абонента;
5. удобный и понятный интерфейс.

Telegram резко стал набирать популярность с 2020 года. Это продолжается и в настоящий момент. Количество пользователей, зарегистрированных в Telegram, составляет около 600 миллионов. Большой прирост аудитории произошел 4 октября 2021 года, когда такие приложения для обмена сообщениями, как WhatsApp и Facebook Messenger стали недоступны по всему миру. В тот день в Telegram зарегистрировалось более 70 миллионов новых пользователей.

Достоинства Telegram:

1. высокая степень защиты данных;
2. возможность создания чатов и каналов на сотни тысяч участников;
3. возможность обмениваться файлами большого размера;
4. корректная работа мессенджера при слабом Интернет-соединении;
5. возможность создания секретных чатов;
6. наличие таймера для автоуничтожения сообщений спустя заданное время;

Существует большое количество различных приложений для обмена сообщениями, поэтому у пользователей есть возможность выбрать, чем им будет удобнее и проще пользоваться. На рисунке 1 представлены данные о популярности различных приложений для обмена сообщениями, взятые с ресурса [9].

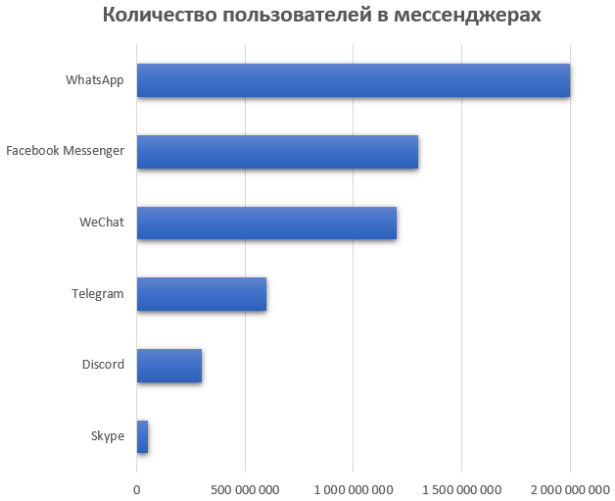

Рис. 1. Количество пользователей в приложениях для обмена сообщениями по данным на 2021 год

Заключение

Рассмотрены этапы развития приложений для обмена сообщениями, их основные функции, возможности и минимальные требования, которые необходимы для корректной работы мессенджеров. Показано, что благодаря мессенджерам у людей появилась возможность работать удаленно, что в настоящее время очень востребовано из-за пандемии. Приложения для обмена сообщениями стали активно применяться не только для личного общения, но и стали полноценным рабочим инструментом. Например, интеграция с Telegram на сегодняшний день стала очень важной для множества пользователей различных CRM-систем. Она предоставляет возможность наладить взаимодействие с клиентам, а также решать вопросы внутри компании. Также были приведены примеры самых распространенных и популярных приложений для обмена сообщениями, для каждого из них были рассмотрены их достоинства и недостатки, для корпоративных мессенджеров было проведено сравнение по уровню их безопасности.

Литература

1. *Городничев М.Г., Мосева М.С., Полянцева К.А.* Мобильное приложение многопараметрического сбора основных характеристик транспортных потоков под управлением ОС iOS // Свидетельство о государственной регистрации программы для ЭВМ № 2020612021 от 13.02.2020.

2. *Фатхулин Т.Д., Денисова М.А., Колосова Е.Р.* Анализ основных технологий виртуализации с целью выбора гипервизора с требуемыми характеристиками // Сборник трудов XIII Международной отраслевой научно-технической конференции «Технологии информационного общества». В 2-х томах. Том 2. М.: ИД Медиа Паблишер, 2019. С. 111-114.

3. *Жаббаров И.Ш., Фатхулин Т.Д.* Обоснование выбора системы виртуализации, предоставляющей необходимый функционал для предприятия заданного уровня // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2019. № 1. С. 233-239.

4. Эволюция мессенджеров – короткая история индустрии сообщений [Электронный ресурс]: // Сетевое издание «Нескучные технологии» – Режим доступа: https://itcrumbs.ru/istoriya-evolyutsiya-messendzherov_23557, свободный. Загл. с экрана. (дата обращения 17.01.2022).

5. MtProto Mobile Protocol [Электронный ресурс]: // Официальный сайт Telegram - Режим доступа: https://core.telegram.org/mtproto, свободный. Загл. с экрана. (дата обращения 17.01.2022).

6. Безопасность в Skype [Электронный ресурс]: // Свободная энциклопедия Википедия - Режим доступа: Безопасность в Skype — Википедия (wikipedia.org), свободный. – Загл. с экрана. (дата обращения 17.01.2022).

7. Security at Slack [Электронный ресурс]: // Официальный сайт Slack - Режим доступа: https://slack.com/trust/security?geocode=en-ru, свободный. Загл. с экрана. (дата обращения 17.01.2022).

8. Центр безопасности Google [Электронный ресурс]: // Официальный сайт Google - Режим доступа: https://safety.google/security/built-in-protection/, свободный. Загл. с экрана. (дата обращения 17.01.2022).

9. 10 самых популярных мессенджеров в мире [Электронный ресурс]: // Сетевое издание Kanobu - Режим доступа: https://kanobu.ru/articles/10-samyih-populyarnyih-messendzherov-v-mire-376082/, свободный Загл. с экрана. (дата обращения 17.01.2022).

СРЕДСТВА РЕАЛИЗАЦИИ ПОИСКОВЫХ И КОНТЕКСТНЫХ МЕХАНИЗМОВ ДЛЯ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ

Яковенко Наталья Викторовна,
МТУСИ, старший преподаватель каф. СИТиС, Москва, Россия
nv1906.iakovenko@yandex.ru

Шведов Андрей Вячеславович,
МТУСИ, старший преподаватель каф. СИТиС, Москва, Россия
a.v.shvedov@mtuci.ru

Пантелеева Ксения Александровна,
МТУСИ, каф. СИТиС, Москва, Россия
ksenya2013333@gmail.com

Гадасин Даниил Денисович,
МТУСИ, каф. СИТиС, Москва, Россия
gadasin115@gmail.com

Аннотация

С развитием цифровой экономики объем данных неудержимо растет. Для работы с большими данными необходимы определенные инструменты, которые рассмотрены в данной статье. Статья представляет исследования компонентов фреймворка Hadoop на различные метрики производительности и возможности.

Ключевые слова: *Большие данные, обработка данных, хранение данных, алгоритмы обработки данных, Big Data, Data Development, Data Processing, Data Storage, Data Processing algorithms, Apache Hadoop.*

Введение

Высокие темпы развития цифровой экономики достигаются, в частности, за счет сбора, использования и анализа огромного объема цифровых данных из всех сфер жизнедеятельности человека [1]. Все эти данные собираются и аккумулируются множеством платформ на основе цифровых профилей и следов пользователей в сети Интернет [12-24], и если раньше данная опция (доступ в сеть Интернет) была доступна лишь ограниченному кругу лиц, то сейчас большинство населения планеты имеют возможность доступа в глобальную сеть, о чем свидетельствует анализ объемов глобального трафика интернет-протокола IP – с 1992 года он увеличился с примерно 100 гигабайт в день до 150 700 гигабайт в секунду за счет роста количества пользователей сети и использования новых концепций, например, интернета вещей (Internet of Things, IoT) [2], технологии облачных (Cloud Ccomputing) и граничных (Edge computing) вычислений, виртуализации сетевых функций (NFV), а также программно-конфигурируемых сетей (SDN) [3].

Каждый день происходит генерация примерно 500 миллионов новых сообщений в социальной сети Twitter, 4 миллиона гигабайт данных в социальной сети Facebook, 294 миллиона электронных писем, 65 миллионов сообщений в сервисе WhatsApp и 720 тысяч часов нового видео-контента в системе видеохостинга YouTube. В 2007 году объем хранимой информации во всем мире составлял 295 эксабайт (295 миллиардов гигабайт) [4]. В 2017 году в «Архиве интернета» – сервисе, который занимается поиском и сохранением важных цифровых данных – хранилось 30 петабайт информации – это примерно 300 миллиардов веб-страниц, 12 миллионов книг, 4 миллиона аудиозаписей, 3,3 миллиона видеороликов, 1,5 миллиона фотографий и 170 тысяч различных дистрибутивов ПО. Это количество выросло до 64,2 зеттабайт в 2020 году по данным International Data Corporation (IDC) [5].

Влияние данных на сферы человеческой жизни зависит от их типа – персонализированные или обезличенные, общедоступные или закрытые, полученные с согласия пользователя, с помощью наблюдения за его действиями, или экстраполируемых аналитически, используемые в коммерческих или государственных целях, конфиденциальные или не конфиденциальные. В связи с необходимо-

стью сбора, обобщения, хранения, анализа и моделирования данных появились компании, которые создают стоимость данных на рынке в результате их преображения в цифровой интеллект и получения денежных средств за их коммерческое использование.

В своем большинстве исходные данные представляют из себя неструктурированный, сырой массив, а традиционные методы обработки данных столкнулись с большими трудностями и привели к ряду потенциальных проблем, таких как высокие затраты на хранение массивных данных, недостаточная производительность пакетной обработки данных, отсутствие потоковой обработки данных, ограниченная масштабируемость и дополнительные активы данных. Все это выявило следующие недостатки: неясные требования к большим данным для сервисов, серьезная разрозненность данных на предприятиях, низкая доступность и их качество, технологии и архитектура управления, связанные с данными, безопасность, трудность нахождения баланса между открытостью данных и конфиденциальностью. Каждое вычислительное устройство ограничено своей вычислительной мощностью, за счет которой время обработки данных может как уменьшаться, так и увеличиваться.

Жизнедеятельность человека определена в пространственно-временном измерении. При взаимодействии пространства и времени происходит событие, а фиксация этого события ведет к появлению информации [6]. Чтобы можно было воспользоваться информацией, она должна быть приведена в необходимую для понимания форму (структурирована) и если обработку данных, которые составляют базу данных, можно легко представить, то в случае больших данных такая обработка происходит намного сложнее.

Большими данными будем считать такие данные, время обработки которых всеми вычислительными устройствами стремится к бесконечности. В свою очередь бесконечность можно трактовать и как предел времени, по прошествии которого данные потеряют свою актуальность.

Об актуальности развития больших данных говорит динамика роста глобального рынка больших данных (рисунок 1).

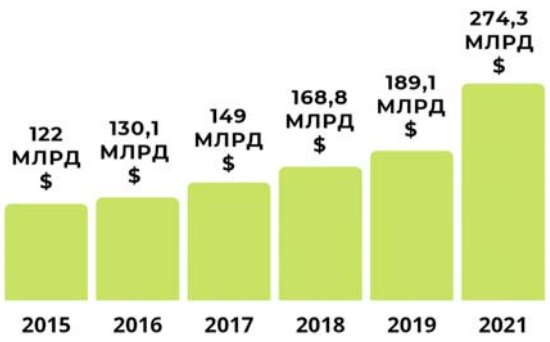

Рис. 1. Динамика роста глобального рынка Больших данных в миллиардах долларов

На начало 2021 года населения нашей планеты насчитывало порядка 7,83 миллиарда человек. По данным ООН, человечество растет каждый год на 1%. Это значит, что с 2020 года численность увеличилась на более 80 миллионов человек. Также на момент написания статьи по анализу прошедшего, 2021, года мобильным телефонным устройством пользовались порядка 5,22 миллиарда человек — что составляет 66,6% всего человечества. Количество уникальных пользователей мобильных устройств выросло на 1,8% (93 миллиона) с января 2020 года, тогда как общее количество подключений с мобильного устройства увеличилось на 72 миллиона (0,9%) и к началу 2021 года стало равным 8,02 миллиарда. Количество пользователей мирового интернета в январе 2021 года составляло порядка 4,66 миллиарда человек, что больше предшествующего года на 316 миллионов (7,3%). Уровень возможности доступа к интернету по всему миру на момент написания составляет 59,5%. Также на момент написания в мире количество пользователей социальных сетей насчитывает порядка 4,20 миллиарда, что составляет 53,6% от мирового населения. В 2020 году это количество было меньше на 490 миллионов, что показывает, что количество пользователей социальных сетей растет на более чем 13% в год.

За счет всего вышеперечисленного объем данных быстро растет, а для хранения данных необходимы Центры Обработки Данных (ЦОДы), и их количество также растет. Если данный рост не сдерживать, то это грозит огромными проблемами экологии из-за выброса парникового газа в атмосферу. Объем же ресурсов, сдерживающих рост производства средств хранения, ограничен, в том время как объем данных неудержимо растет. Таким образом, можно говорить об необходимости повысить качество как алгоритмов обработки данных, так и алгоритмов сжатия данных.

Результаты исследования

По данным компании IDC, более 60% корпоративных хранилищ данных (КХД) занимает информация, не приносящая организации никакой пользы (информация, к которой давно никто не обращался, которая потеряла свою актуальность, «корпоративный мусор» или неструктурированные данные) [7]. Поэтому эффективность использования данных и повышение их качества зависит от способа их обработки. Процесс повышения качества данных является последовательностью преобразований полученной информации с аномалиями (дубликаты, несоответствия, нестандартные представления, нарушения целостности, пропущенные значения, отсутствие стандартов значений) с целью получения целостного представления и необходимого качества для работы приложений и формирования точных результатов, что в среднем занимает 5,56 мс при одном запросе в секунду.

Интуитивно понятно, что сокращение времени запроса зависит от уровня вычислительной мощности. Повышение мощности можно достичь двумя способами:

– Экстенсивным – увеличение производительности вычислительных устройств;

– Интенсивным – улучшение архитектуры устройств.

В рамках повышения производительности устройства чаще всего повышают тактовую частоту процессоров, что ведет к повышению тепловыделения, в связи с чем приходится прикладывать немалые усилия для совершенствования систем охлаждения, а это ведет к сильному удорожанию самих устройств.

Еще одна возможность увеличения производительности предполагает распараллеливание процесса вычислений, что может приблизить мощность вычислительных систем к предельной мощности мозга человека (примерно $10^{18} - 10^{20}$ флопс) – на данный момент производительность самого мощного суперкомпьютера составляет $4{,}42 \cdot 10^{17}$ флопс, что ниже производительности мозга человека на 1-3 порядка. Однако, подобная архитектура не в полной мере будет удовлетворять решению реальных задач, так как в зависимости от решаемой задачи и в силу разнородности архитектурных подходов для каждой решаемой задачи необходимо выбирать свою архитектуру. В работе с изображением, звуком, естественным языком и смыслом человеческой речи наиболее эффективными являются нейроподобные вычислительные структуры. Но зачастую мощность оценивают по максимальной вычислительной способности аппаратных компонентов системы, тогда как вычислительная мощность нейронной сети этих компонентов вероятнее всего значительно ниже.

Создание новых алгоритмов и способов кодирования и обработки данных также повышает производительность вычислительных устройств.

Как отмечалось выше, вычислительные устройства могут обрабатывать только структурированные данные, работа с неструктурированными данными делится на два этапа: приведение неструктурированных данных к структурированным и обработка структурированных данных. Для обработки данных необходима жесткая структура, которая должна работать по общим и единым для всех правилам, поэтому были разработаны и внедрены стандарты в области обработки больших данных. Росстандартом были утверждены два новых стандарта, а именно: ГОСТ «Информационные технологии. Эталонная архитектура больших данных. Часть 2: Варианты использования и производные требования» и ГОСТ «Информационные технологии. Большие данные. Техническое задание. Требования к содержанию и оформлению» [8,9].

В стандарте ISO/IEC 20547-3:2020 «Информационные технологии – Эталонная архитектура больших данных – Часть 3: Эталонная архитектура», который на момент написания находится на стадии публикации международного стандарта, содержится информация о рамочной структуре (концепции) использования аналитики больших данных в большинстве служб и подразделений организации, определяется эталонная модель процесса аналитики больших данных (Big Data Analytics Process Reference Model, BDA PRM), а также определяется эталонная модель оценки процесса (Big Data Analytics Process Assessment Model, BDA PAM), которая содержит два измерения: размерность процесса, включающая процессы, определенные на основе набора PRM-моделей, в том числе модель BDA PRM, и размерность возможностей процесса, определяемых на основе системы измерения процесса.

Фреймворк для обработки и хранения больших данных Apache Hadoop и его компоненты

Самой распространенной экосистемой для работы с большими данными на сегодняшний день считается Apache Hadoop. Экосистема предназначена для обработки больших объемов данных, превышающих допустимый объем на одном узле.

Компонент для хранения HDFS

Для хранения данных используется компонент Hadoop Distributed File System (HDFS). Он предназначен для хранения большого объема данных на сотнях компьютеров, и для хранения нужна только

одна файловая система. Файл разделяется на один или несколько блоков данных и хранится в группе DataNodes. В реальных кейсах использования существует много кластеров хранимых данных, которые достигают уровня PB (петабайт), а размер узла более 10 KB (килобайт). Согласно официальному сайту Hadoop кластер Hadoop имеет около 100 000 процессоров и работает на 40 000 машинных узлов. Но так как управляющий узел имен namenode хранит объекты блоков и имен в памяти, размер пространства имен (и, следовательно, количество файлов) ограничен объемом памяти кучи. В настоящее время куча объемом 14 GB (гигабайт) может хранить 60 миллионов объектов блоков и имен. Следовательно, если у пользователя содержится 2 блока на файл, то он ограничен 20 миллионами файлов. Это существенное ограничение для больших кластеров. Для проверки состояний 2 тысячи узлов каждые 3 секунды отправляют отчет о блоке.

При проведении экспериментальных тестов на исследование скорости передачи данных, во время которых файл размером 292,2 Мбайт отправлялся из локальной файловой системы (Local File System) в распределенную файловую систему (HDFS), были получены результаты, которые отображены в таблице 1.

Таблица 1

Результаты исследования скорости передачи файла с данными из локальной файловой системы (Local File System) в распределенную файловую систему (HDFS)

Способ хранения данных	**Размер файла с учетом избыточности/ репликации (Мб)**	**Коэффициент избыточности/ репликации**	**Количество проведенных итераций**	**Среднее время передачи файла (мс)**	**Средняя скорость передачи данных (Мб/с)**
RS (3,2)	488,2	1,67	90	297,3	88,62
RS (6,3)	439,2	1,50	90	922,0	100,00
XOR (2,1)	438,4	1,50	90	683,3	108,89
Replication	876,6	3,00	90	506,3	64,84

Вычислительный двигатель MapReduce

Обработка и анализ данных может проводиться с помощью фреймворка MapReduce, Tez, Spark или Flink. MapReduce – это вычислительный двигатель первого поколения, в то время как Tez и Spark – вычислительные двигатели второго поколения. Они используются для параллельных и автономных вычислений крупномасштабных наборов данных (размером более 1 TB). Процесс MapReduce состоит из двух этапов: обобщение большого объема неупорядоченных данных на основе объекта (Map) и получение результата после обработки (Reduce). Программы на основе MapReduce не всегда работают быстро, но данные всегда оптимизировано распределяются между узлами, в связи с чем вычислительный движок удобен в использовании. Однако некоторые функции могут сильно повлиять на производительность, например, функция Combiner, которая помогает уменьшать количество данных при записи на диск и передачи через сеть. Для быстрорешаемых на нераспределенных системах задач, включающих также наличие входных данных с оперативной памяти одного персонального компьютера или кластера с небольшим объемом, MapReduce малоэффективен. Разработчикам сложно разрабатывать программный код для этих вычислительных движков. Чтобы упростить разработку и повысить эффективность, для описания алгоритмов и процессов обработки данных разработан уровень абстрактного языка более высокого уровня. Для этого доступны Pig и Hive. Pig описывает MapReduce аналогично скриптам, в то время как Hive использует SQL. И Pig, и Hive преобразуют скрипт и SQL в программы MapReduce, а затем вычислительные механизмы обрабатывают данные.

Компонент для анализа Hive

Компонент Hive используется для извлечения, преобразования и загрузки данных. Это механизм, который может хранить, запрашивать и анализировать большой объем данных, хранящихся в Hadoop. Инструмент хранилища данных Hive может отображать файлы структурированных данных в таблицу базы данных, предоставлять функцию запроса SQL и преобразовывать инструкции SQL в задачи MapReduce для выполнения. Hive подходит для статистического анализа хранилищ данных, например, анализ и запрос данных за определенный период времени.

Компонент для хранения и анализа HBase

HBase – сокращение от базы данных Hadoop, представляет собой распределенную базу данных, ориентированную на столбцы, построенную на файловой системе Hadoop (HDFS) [10]. Как часть файловой системы Hadoop, HBase обеспечивает случайный доступ к данным для чтения/записи в реальном времени. HBase реализует фильтры Блума и алгоритмы сжатия, а также операции с памятью

для столбцов, которые описаны в тезисе BigTable. Таблицы HBase можно использовать в качестве входных и выходных данных задач MapReduce. Каждая таблица в HBase разделена на несколько подтаблиц (HRegions) на основе определенного диапазона ключей. По умолчанию, если размер региона превышает 256 МБ, регион делится на два региона. Метаданные таблиц хранятся в ZooKeeper, поэтому загрузка данных невелика.

Возможности компонента HBase ограничены только возможностями аппаратных компонентов.

На рисунке 4 приведены результаты исследования загрузки на кластере, который содержит 16 NodeManager и 16 RegionServers, в то время как тактовая частота процессора CPU Xeon E5-2680 v4 вычислительного устройства, на котором проходили испытания, составляет 2.40 ГГц.

В таблице 2 приведены результаты исследования компонента HBase в задаче загрузки на кластер.

Таблица 2

Результаты исследования HBase в задаче загрузки на кластер

Number of records	Size (of 1 record)	Volume (MB)	Executors	Partitions	Seconds	Speed (MB/sec)	Speed (Record/sec)
50 000 000	100	4768	370	512	119	40	420 168
10 000 000	1000	9537	94	8	125	76	80 000
10 000 000	1000	9537	94	16	76	125	131 579
10 000 000	1000	9537	99	32	62	154	161 290
10 000 000	1000	9537	99	64	55	173	181 818
20 000 000	1000	19 073	245	128	84	227	238 095
50 000 000	1000	47 684	214	128	197	242	253 807
50 000 000	1000	47 684	325	128	149	320	335 570
40 000 000	1000	38 147	233	128	115	332	347 826
10 000 000	5000	47 684	99	32	140	341	71 429
50 000 000	1000	47 684	361	256	127	375	393 701
10 000 000	10 000	95 367	99	32	248	385	40 323
50 000 000	2000	95 367	326	256	169	564	295 858
50 000 000	2000	95 367	370	128	161	592	310 559
10 000 000	10 000	95 367	99	64	153	623	65 359
50 000 000	5000	238 419	370	256	261	913	191 571
50 000 000	10 000	476 837	321	128	463	1030	107 991
50 000 000	10 000	476 837	322	256	451	1057	110 865
50 000 000	10 000	476 837	367	256	438	1089	114 155
50 000 000	10 000	476 837	325	128	408	1169	112 549

Проанализировав результаты загрузки, можно увидеть, что, увеличивая количество партиций и экзекуторов, возможно получить увеличение скорости загрузки, которая в том числе зависит от объема загруженной записи. Более крупные блоки могут дать прирост в измерении Мбайт/сек, тем временем мелкие – в количестве записей при вставке в единицу времени, при условии, что прочие параметры равны.

Также исследование HBase показало, что обработка одиночных запросов, такие как Get, Put, Delete, оказываются довольно дорогими для производительности, поэтому по возможности их следует объединять в List, что позволит получить прирост производительности.

На рисунке 2 показан график результатов исследования затрат времени при чтении 50000 записей из MemStore.

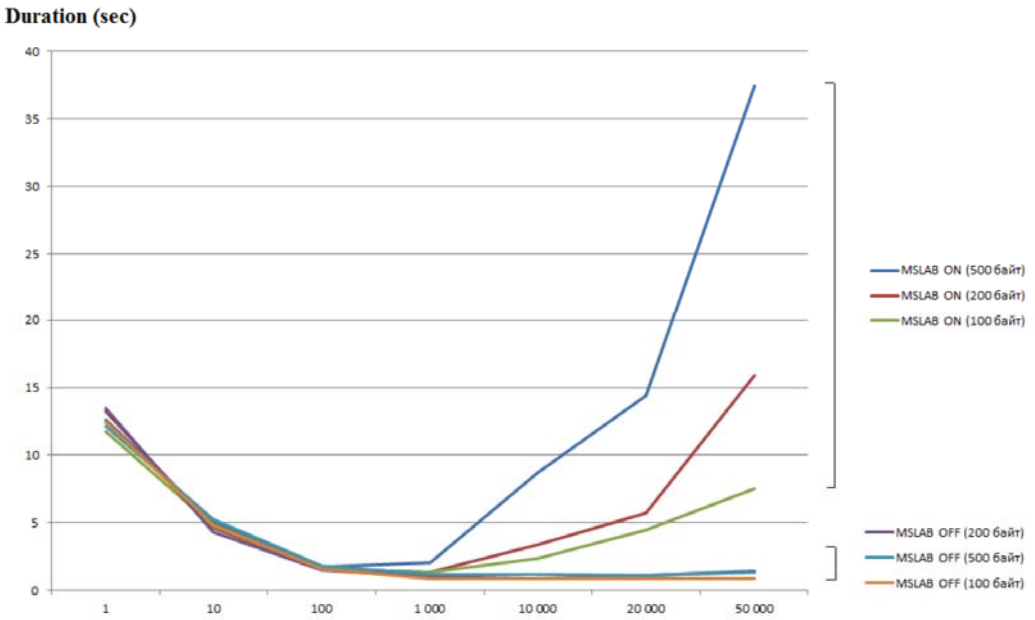

Рис. 2. График результатов исследования затрат времени при чтении 50000 записей из MemStore

В рамках данного исследования, чтение записей производилось в рамках одного потока. По оси абсцисс представлено количество ключей, которые содержатся в запросе. При увеличении количества ключей в рамках одного запроса до значения 1000 время выполнения запроса уменьшается, то есть скорость обработки запроса увеличивается [11]. Но когда включается режим MSLAB, то после прохождения порога в 1000 записей начинается резкое падение значений производительности, при этом видно, что чем больше объем данных в одной записи, тем больше время обработки. Исследование проводилось на виртуальной машине, содержащей 8 ядер.

Инструмент управления потоками данных Flume

Инструменты Sqoop и Flume разработаны используются для восполнения пробелов, когда традиционные инструменты сбора данных не могут получать массивные данные.

На рисунке 3 показаны результаты исследования затрат времени на решение такой задачи, как отправка блока, состоящей из 500 событий, на узлы Flume с помощью файловых каналов.

Рис. 3. График результатов затрат времени на решение задачи отправки блока на узлы Flume через файловые каналы

При этом исследовании для хранения данных канала были использованы два различных интерфейса обмена дынных с накопителями, а именно: один узел использовал накопитель SSD, другой – последовательный интерфейс SATA.

Инструмент для распределенной обработки потоков Flink

Инструмент Flink в основном используется для вычислений в реальном времени. Его конвейерная архитектура обеспечивает высокую пропускную способность, в том числе за счет собственной подсистемы управления памятью и ее эффективного использования. Он обрабатывает данные с молниеносной скоростью, его также называют 4G Big Data. Он может работать во всех распространенных кластерных средах и выполнять вычисления с любой скоростью и масштабом памяти.

Заключение

Технологии больших данных можно разделить на эксплуатационные, которые обеспечивают эксплуатационные возможности для диалоговых рабочих нагрузок в режиме реального времени, а также в случаях использования данные для сбора и хранения, и аналитические системы, которые обеспечивают аналитические возможности для исторического и сводного анализа, а также в случаях использования большей части или всех данных.

Актуальность больших данных подтверждает зависимость человека от мобильного телефона, интернета и социальных сетей. Если тенденция зависимости сохранится, то в 2022 году все пользователи в мире проведут в социальных сетях в общей сложности $3{,}7\ 10^{12}$ часов, что равносильно более 420 миллионам лет человеческого существования в сумме, что обеспечит рост объема собираемых данных.

Литература

1. *Гадасин Д.В., Шведов А.В., Клыгина О.Г., Гадасин Д.Д.* Реализация платформы туманных вычислений для предоставления сервисов IoT // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11. № 2. С. 65-75.
2. *Гадасин Д. В., Шведов А. В., Ермалович А. В.* Концепция "интернет вещей" как вектор развития информационно-коммуникационных технологий на пути к "индустрии 4.0" // Технологии информационного общества : XI Международная отраслевая научно-техническая конференция: сборник трудов, Москва, 15-16 марта 2017 года. – Москва: ООО "Издательский дом Медиа паблишер", 2017. С. 352-353.
3. *Докучаев В.А., Кальфа А.А., Мытенков С.С., Шведов А.В.* Анализ технических решений по организации современных центров обработки данных // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 6. С. 16-24.
4. *Hilbert, M., López, P.:* The world's technological capacity to store, communicate, and compute information. Science 332(6025), pp. 60-65 (2011).
5. Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts [Электронный ресурс]. Режим доступа: https://www.idc.com/getdoc.jsp?containerId=prUS47560321 (дата обращения: 14.12.2021).
6. *Гадасин Д. В*., *Ермалович А. В., Шведов А. В.* Цифровое неравенство и социальный аспект цифровой трансформации // Перспективные технологии в средствах передачи информации – ПТСПИ-2017 : Материалы 12-ой международной научно-технической конференции, в 2-х томах, Суздаль, 05-07 июля 2017 года. Суздаль: Владимирский государственный университет им. Александра Григорьевича и Николая Григорьевича Столетовых, 2017. С. 77-78.
7. *Гадасин Д. В., Шведов А. В.* Проблемы интеграции концепции "интернет вещей" и облачных вычислений // Технологии информационного общества: Материалы XIII Международной отраслевой научно-технической конференции, Москва, 20-21 марта 2019 года. М.: Издательский дом Медиа Паблишер, 2019. С. 22-23.
8. ГОСТ Р ИСО/МЭК 20547-2-2020. Информационные технологии. Эталонная архитектура больших данных. Часть 2: Варианты использования и производные требования. (ISO/IEC TR 20547-2:2018, IDT); введ. – 2022 – 03 – 01. М.: Стандарт-Информ, 2020. 509 с. (Национальный стандарт Российской Федерации).
9. ГОСТ Р ИСО/МЭК 24668. Информационные технологии. Эталонная архитектура больших данных. Часть 3: Эталонная архитектура. (ISO/IEC DIS 24668, IDT); Москва: Стандарт-Информ, 2021. 130 с. (Национальный стандарт Российской Федерации).
10. Apache HBase Reference Guide [Электронный ресурс]. Режим доступа: https://hbase.apache.org/book.html - (дата обращения: 15.12.2021).
11. *Гадасин Д. В., Шведов А. В., Каледина А. В., Юдина А. А.* Мониторинг и анализ log-файлов инфокоммуникационной сети с помощью комплексного инструментария elk stack // Актуальные проблемы и перспективы развития экономики: Труды XVIII Всероссийской с международным участием научно-практической конференции, Симферополь-Гурзуф, 24–26 октября 2019 года / Под редакцией Н.В. Апатовой. Симферополь-Гурзуф: ИП Зуева Т.В., 2019. С. 297-298.

12. *Шведов А.В., Гадасин Д.В., Цыгулёва А.В., Вакурин И.С.* Разгрузка очереди сети при помощи Гамильтонова цикла // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11. № 3. С. 45-53.

13. *Kalmykov N.S., Dokuchaev V.A.* Segment routing as a basis for software defined network // T-Comm. 2021. Т. 15. № 7. С. 50-54.

14. *Dokuchaev V.A., Maklachkova V.V., Statev V.Yu.* Classification of personal data security threats in information systems // T-Comm. 2020. Т. 14. № 1. С. 56-60.

15. *Докучаев В.А., Маклачкова В.В., Статьев В.Ю.* Цифровизация субъекта персональных данных // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 6. С. 27-32.

16. *Pavlov S.V., Dokuchaev V.A., Mytenkov S.S.* Model of a fuzzy dynamic decision support system // T-Comm. 2020. Т. 14. № 9. С. 43-47.

17. *Pavlov S.V., Dokuchaev V.A., Maklachkova V.V., Mytenkov S.S.* Features of supporting decision making in modern enterprise infocommunication systems // T-Comm. 2019. Т. 13. № 3. С. 71-74.

18. *Докучаев В.А., Ерёменко В.А., Маклачкова В.В., Мытенков С.С., Шевелёв С.В.* Профессиональные квалификации специалистов по контролю качества информационно-коммуникационных систем // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 11. С. 62-67.

19. *Гадасин Д.В., Кольцова А.В., Гадасин Д.Д., Полякова А.Н.* Оценка вероятности формирования виртуального кластера // Системы синхронизации, формирования и обработки сигналов. 2021. Т. 12. № 1. С. 4-12.

20. *Кузин И.А., Гадасин Д.В.* Модель контейнера данных для минимизации трафика при передаче субъективных характеристик объектов на изображении трехмерной сцены // Телекоммуникации и информационные технологии. 2021. Т. 8. № 2. С. 96-100.

21. *Гадасин Д.В., Кольцова А.В., Полякова А.Н.* Модель построения кластера для пограничных вычислений // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2020. № 1. С. 86-92.

22. *Усачева Д.И., Шишкин М.О., Гадасин Д.В., Гузеев А.В.* Применение OLAP-технологий для анализа многомерных данных в контакт-центре // Телекоммуникации и информационные технологии. 2019. Т. 6. № 1. С. 142-149.

23. *Гадасин Д.В., Кузин И.А.* Модель представления цветовых и глубинометрических характеристик объектов на изображении // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 31-38.

24. *Гадасин Д.В., Нестерова Е.А.* Особенности проведения практических занятий по дисциплине мультимедийные информационные системы для стадии "исследование и обоснование создания информационной системы" // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2021. Т. 10. № 1. С. 15-21.