REDS:

Телекоммуникационные устройства и системы

№1

2022

СОДЕРЖАНИЕ

ПРЕДОБРАБОТКА ТЕКСТОВОЙ ИНФОРМАЦИИ НА ОСНОВЕ ТЕРМОВ ЕСТЕСТВЕННОГО ЯЗЫКА

Гадасин Денис Вадимович,
МТУСИ, доцент кафедры СИТиС, к.т.н., Москва, Россия
dengadiplom@mail.ru

Пак Екатерина Владимировна,
МТУСИ, магистрант гр. М112001(78), Москва, Россия
cthrnblck13@mail.ru

Коровушкина Вероника Максимовна,
МТУСИ, бакалавр гр. БСТ1801, Москва, Россия
vkorovushkina10@gmail.com

Мелькова Елизавета Константиновна,
МТУСИ, бакалавр гр. БСТ1801, Москва, Россия
emelkova1005@gmail.com

Аннотация

С бурным ростом количества передаваемой информации в сети и непрерывным ее увеличением неизбежно возрастает потребность в оптимизации не только оборудования и технологий передачи данных, но и разработке алгоритма предобработки информации на предмет ее ценности для конечного пользователя. Для разработки данного алгоритма будут рассмотрены сущности обработки естественных языков, исследование и проведение этапов подготовки текстовых данных для распознавания их машиной и вычисления, проводимые над текстом, позволяющие провести обширный анализ текстовых данных на предмет ее ценности и выявления закономерностей.

Ключевые слова: *NLP, обработка естественных языков, кодировка Шеннона-Фано, информационная энтропия, машинное обучение, предобработка текста, токенизация, слогоделение, ценность информации, разработка.*

Введение

В последние десятилетия в результате мощного развития науки и техники резко возросло значение информации в жизни и деятельности человека. Благодаря технологиям в мире образуется огромное количество информации, которое постоянно и быстро увеличивается, производя так называемый «информационный взрыв». Не только люди обмениваются информацией между собой, но происходит обмен между человеком и техникой, техникой и техникой. Данное обстоятельство привлекает к себе и внимание ученых. Традиционно с помощью информации передавались какие-либо сведения, т.е. они представлялись в определенной форме, исторически возникшей и развивающейся, передаваемой из поколения в поколение. В современном мире при развитии технологий понятие информации пришлось расширить. Понятие передачи информации перешло и на другие сферы бытия, в которых традиционно для обозначения взаимодействий использовались другие слова. Появилась информационная теория. Информатика исследует разные стороны ключевой сущности человечества – информации, исходя из практических задач и теории информации [9-16]. А уже через несколько сотен лет на человека обрушился настолько большой поток информации, что возникла проблема ориентации в этом потоке, поиска наиболее ценной информации в данный момент времени [5].

Известно, что возможности сетей стремительно развиваются в течение большого промежутка времени, это означает, что возрастает пропускная способность и емкость сетей передачи данных. Однако, даже стремительное развитие данных технологий не обеспечивает необходимые потребности рынка, это особенно остро стало заметно к 2019 году [7].

По данным Cisco Visiual Network Index, сетевой трафик растёт экспоненциально со среднегодовой скоростью роста 26% до 2023 года, и темпы роста по прогнозам не будут снижаться.

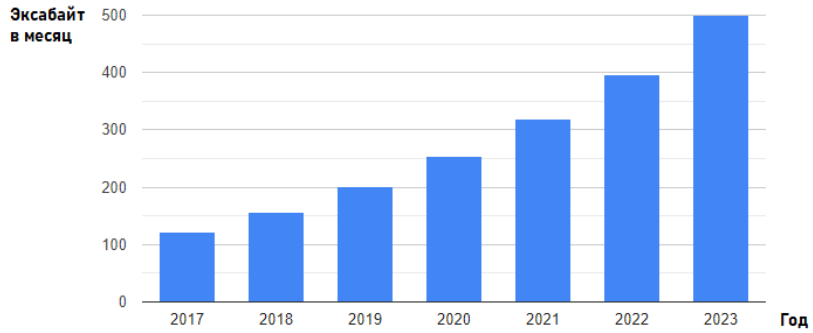

Рис. 1. Рост сетевого трафика 2017-2023 (источник: Cisco)

В связи с этими факторами, существует необходимость не только в том, чтобы увеличивать мощность оборудования и пропускную способность сетей передачи данных, но и разработать алгоритмы предобработки информации на предмет ее ценности для конечного пользователя. В работе будут анализироваться текстовые данные, для обработки которых требуется специальный алгоритм.

Возможности машинной обработки речи.

Существует определенное направление в математической лингвистике, занимающееся вопросом компьютерной обработки человеческого языка. Обработка языка Natural Language Processing, находится на пересечении нескольких наук и областей знаний. Одна область знаний принадлежит компьютерной сфере, а другая наукам о языке. Задача NLP в создании искусственного языка, понятного человеку и машине. С одной стороны, нужно понимать устройство компьютера и способы обработки им информации. С другой - представлять себе строение речи и то, как она может быть обработана машинными средствами и впоследствии представлена для восприятия человека [6].

С практической точки зрения наиболее большой трафик, который передается по сетям, в том или ином виде в своей основе представляет собой текстовую информацию, которую необходимо обработать с помощью математического аппарата. На начальном этапе вводятся данные для выполнения определенного алгоритма. Данные считываются компьютером и обрабатываются различными математическими алгоритмами. В зависимости от данных и от решаемой задачи выбирается алгоритм обработки. После обработки компьютер выводит эти же данные в виде информации, которые могут восприниматься человеком и быть ему полезны для определенных целей. Из обработанных данных также можно извлечь полезную информацию, получить знания. Для изучения информации приоритетны следующие характеристики:

- Скорость. Скорость двух процессов: увеличения количества информации и ее обработки для выполнения требований и задач. Большие данные создаются постоянно.
- Достоверность. Определение качества, достоверности данных. Качество данных может быть весьма различным, что отражается на конечном результат обработки или анализа.
- Объем – количество генерируемых и хранимых данных. Размер данных определяет ценность и потенциальное знание, а также может ли он считаться большими данными или нет.

Рассмотрим способы приближения к тому, чтобы компьютер понимал смысл текстовых документов. Прежде всего, необходимо адаптировать текст для его дальнейшей обработки. Это производится методами обработки естественного языка, называемого NLP. Текстовое представление необходимо для кластеризации или классификации документов, переводит документ в компактную форму. Предварительная обработка включают в себя такие этапы как: извлечение характеристик и их выбор [8].

Первым шагом предварительной обработки является извлечение признаков, в рамках этого шага текст разделяется на слова и представляется в виде отдельных слов. К методам обработки естественного языка относятся токенизация, удаление стоп-слов и стеммитизация [1].

Если попытаться разделить речевую информацию на единицы представления, то первым делом можно выделить предложения, слова, слоги, буквы.

Предложения в письменности разделяются точкой, восклицательными и вопросительными знаками. Данные знаки можно использовать в качестве условного обозначения конца предложения. Но в некоторых языках точки используются и в аббревиатурах, что также необходимо учитывать [2]. Если точка следует за аббревиатурой, она является неотъемлемой частью этой аббревиатуры. В данное время общепринятых стандартов для сокращений и аббревиатур не разработано. При обработке аббревиатур и сокращений может использоваться список наиболее известных аббревиатур. Если анализируется слово, которое заканчивается точкой, то его нужно будет сравнивать с таким списком для определения аббревиатура ли это или окончание предложения. Естественно, точность такого подхода зависит от того, насколько правильно список аббревиатур приспособлен к обрабатываемому тексту.

Далее речь делится на слова, которые при печати как правило разделяются пробелами. Это структурные единицы языка, которые также можно обозначить с помощью пробелов.

Разделение текста на слоги.

Слог – это наименьшая фонетическая структура, звуки при произношении и восприятии. Слово в речи делится на слоги. Слоги осознаются и произносятся в речи. Слоги состоят из гласных и согласных звуков. Когда мы слышим речь, то звуки слога воспринимаются как разные по громкости.

Слог – это сочетание звуков, в нем не заложено какого-либо определенного понятия, смысла или содержания. Слова, наделенные смыслом, образуются из слогов. Звуки, образующие слог имеют разную степень звучности. Наиболее сильные звуки называются слогообразующие, менее сильные – неслоговыми.

При произношении звуки внутри слога не отделяются друг от друга, поэтому слог считается наименьшей произносительной единицей.

О том, как правильно разделить слово на слоги существуют разные мнения. Выделяются четыре теории разделения слова на слоги. В слоге можно выделить более или менее сильные громкие звуки. Согласно первой рассматриваемой теории, они должны располагаться в порядке ослабления звучности. Датский лингвист Отто Эсперсен разработал шкалу сонорности из нескольких ступеней, в которой для определенных звуков устанавливается большая или меньшая степень звучности, на основе чего должно производиться слогоделение. Теория получила название сонорной.

В динамической теории выделение слога влияют много факторов: фонологических, просодических, акустических. Слог представляет собой волну с разной степенью интенсивности звуков. Теорию разрабатывали французский ученый лингвист Морис Граммон, русский и советский лингвист Л.В. Щерба).

Экспираторная теория: первый план выходит физическое усилие, необходимое для произнесения слога, слог образуется одним толчком выдыхаемого воздуха. На пик этого усилия приходятся самые громкие звуки. Представители: В.А. Богородицкий, Томпсон.

По теории открытого слога все слоги должны быть оканчиваться на гласные, т.е. быть открытыми. При сочетании согласного и последующего гласного скорее образуется слог, чем при обратной последовательности этих звуков. Только последний слог может заканчиваться на согласный звук. (Л.В. Бондарко, ПФШ).

В русском языке чаще всего применяется сонорная теория, в основе которой лежат критерии звучности, восприятия силы звуков. Разрабатывал ее и адаптировал для русского языка Р.И. Аванесов, советский лингвист, один из создателей московской фонологической школы. Он предложил собственную шкалу, по которой звуки делятся на три группы по степени звучности. Самыми слышимыми являются гласные звуки, затем по убыванию звучности следуют сонорные. Наименее звучными являются шумные согласные. Слог строится по принципу волны восходящей звучности. В русском языке самыми громкими являются гласные, за счет этого они и становятся главными в слоге, слогообразующими. Рядом со слогообразующим звуком стоят другие звуки, неслоговые или неслогообразующие. Таким образом слог – это сочетание слогового звука с неслоговыми. Слоговый звук один, и он является наиболее сильным. Вокруг него группируются другие звуки. Слог определяется как сочетание звуков с разной степенью звучности. Сонорные также могут быть слоговыми. Для выделения слога звучность важна слогообразующего, по отношению к примыкающим звукам.

В русском языке выделяются различные типы слогов. Они бывают открытыми, закрытыми, прикрытыми, неприкрытыми. Открытые слоги заканчиваются гласным слогообразующим звуком. Закрытые заканчиваются на согласный. Прикрытые звуки начинаются с согласных. Неприкрытые – с гласных. Известно, что бывают слоги, состоящие из одного гласного звука. По вышеприведенной класси-

фикации они будут определятся как открытые и неприкрытые. Слоги, состоящие из фонем, могут не совпадать с морфемными частями слова, что создает дополнительную трудность при изучении вопросов, касающихся слогового строения.

Разделить слово на слоги можно, воспользовавшись определенными правилами. Существует закон восходящей звучности. В определенном месте один слог должен заканчиваться и начинаться другой. Для нахождения этого места можно воспользоваться указанным законом. Правила разрабатывались учеными лингвистами Р.И. Аванесовым и Л.В. Щербой. Различные теории согласны в одном: что для русского языка более характерны открытые слоги, чем закрытые. При чередовании согласных и гласных, слог заканчивается гласным и начинается со следующего согласного. Если вместе стоят две или более согласных, то делить на слоги нужно применяя закон восходящей звучности. Делить нужно так, чтобы сила звуков внутри слога возрастала при движении слева направо. Звуки Аванесов разделял на три группы по сонантности: гласные, сонорные и шумные. Недопустима последовательность звуков внутри слога сонорный-шумный. В такой ситуации сонорный должен быть отнесен к предыдущему слогу [3]. Эти правила не сложны, но не всегда применимы, возможно по причине их искусственного изобретения и многообразия и богатства нашего языка.

Л. В. Щерба наделяет ударение определяющим в разделении слов на слоги. Каждый слог образуется единичным мышечным движением, наибольшая сила которого приходится на гласный. А в месте наименьшего усилия полагается делить слово. Поэтому, когда между гласными стоит два согласных, первый из них необходимо отнести к предыдущему слогу, когда ударение падает на первый слог. Таким образом слог заканчивается на согласный звук и становится закрытым. Когда ударение падает на второй слог, согласный надо отнести ко второму слогу [4]. Данная теория также несовершенна, поскольку понятие физического усилия, мускульного не определено четко и не понятно, почему оно именно таким образом должно распределять звуки.

В современном русском языке принимается теория большей или меньшей сонорности звуков. Признается, что действительно делить на слоги следует по правилу восходящей звучности. Также учитывается закономерность тяготения к открытым слогам, которые заканчиваются на гласный звук. Чаще всего слог заканчивается гласным перед согласным.

- В последующий слог входит согласный, стоящий между гласными
- К следующему слогу примыкают сочетания шумных согласных между гласными
- Шумный согласный в сочетании с сонорным отходит также к следующему слогу
- При сочетании сонорных согласных между гласными они могут относится к следующему слогу. Или распределяться между предыдущим и последующим слогами.
- При сочетании сонорных согласных с шумным между гласными сонорный отходит к предшествующему слогу.
- К следующему слогу относятся два однородных согласных между гласными.
- При сочетании «й» с последующими сонорными и шумными согласными «й» отходит к предшествующему слогу.

Таким образом, можно сделать вывод, что конечный слог в русском языке оказывается в большинстве случаев открытым; закрытым он является тогда, когда оканчивается на сонорный.

Закон восходящей звучности заключается в том, что внутри слога при движении от звука к звуку в их последовательности внутри слова сонорность должна увеличиваться.

Одной из особенностей языка является то, что некоторые звуки не произносятся в определенных сочетаниях. Например, две одинаковые согласные при соединении морфем и перед третьим согласным внутри одного слога. Это чаще наблюдается на стыке корня и суффикса и реже – на стыке приставки и корня или предлога и слова. Слог обычно имеет главный, слогообразующий звук и второстепенные, неслоговые звуки.

В завершение темы слогоделения приведем также общепринятые правила деления на слоги в русском языке.

- К последующему слогу следует относить сочетание шумных или сонорных, сочетание шумного и сонорного согласных.
- При сочетании шумного и сонорного, шумный относится к предыдущему слогу, а сонорный – к последующему.

Использование известных правил для разделения слогов и теории Аванесова позволит однозначно разделить текстовый документ на слоги.

Буква – минимальная графическая единица, используемая в тексте. В отличии от разделения на слоги, разделение на буквы при обработке текста не будет составлять труда.

В машинном обучении выделение предложений, слов, знаков препинания, границ абзацев и пр. называется токенизацией [1].

Человеческая речь, переводимая в текстовые документы, содержит большое количество таких лингвистических элементов, которые могут быть проигнорированы при их компьютерной обработке. Соответственно этим словам выделяется тип данных. Требуется разработать метод, алгоритм для определения и удаления таких данных до начала обработки текста. Такие данные известны как стоп-слова. Понятие «стоп-слова» было впервые введено Г.П. Луном в 1958 году. Стоп-слова – это слова, которые чаще всего встречаются в документе и содержат небольшую информацию, которая как правило не требуется для машинного обучения. Удаление стоп-слов уменьшает векторное пространство, но и повышает производительность за счет увеличения скорости выполнения, вычислений, и в результате приводит к более высокой степени точности.

Стоп-слова имеют менее определяющее значение для данной структуры текста по сравнению с другими словами. Главные члены предложения, подлежащие, сказуемые и другие существительные и глаголы обычно не являются стоп-словами. Скорее ими могут быть предлоги, союзы и артикли. Стоп-слова являются служебными и не имеют какого-то конкретного содержания, предсказательной способности. Они не указывают на предмет. Стоп-слова часто употребляются, поэтому они могут повлиять на эффективность процесса обработки. Они влияют на процесс нахождения значимых ключевых слов, поскольку уменьшают влияние частотных различий между менее распространенными словами. Объем данных может быть уменьшен путем удаления стоп-слов и влияет на время обработки. Для удаления стоп-слов на этапе предобработки текстовой информации, воспользуемся классическим методом, при котором имеется подготовленный список стоп-слов. Текстовый документ проверяется на совпадения по этому списку. Стоп-слова удаляются.

При реализации и введении перечисленных этапов обработки текста, мы получаем четыре массива данных: текст, разделенных на предложения, слова, слоги и буквы, с расчетом количества символов, в том числе уникальных. Далее будем рассматривать вычисления, проводимые над каждым получившимся документом.

1. Информационная энтропия, при отсутствии информационных потерь, рассчитывается по формуле Хартли: $i = \log_2 N$, где N – мощность алфавита, i – количество информации в каждом символе сообщения. Для случайной величины x, принимающей n независимых случайных значений x_i с вероятностями $p_i = (i = 1, 2..., n)$, формула Хартли переходит в формулу Шеннона: $H(x) = -\sum_{i=1}^{n} p_i \cdot \log_2 p_i$

Эта величина также называется средней энтропией сообщения. Величина $H_i = -\log_2 p_i$ называется частной энтропией, характеризующей только i-е состояние.

2. Проведение кодировки методом Шеннона-Фано – алгоритм префиксного неоднородного кодирования. Относится к вероятностным методам сжатия (точнее, методам контекстного моделирования нулевого порядка). Данный алгоритм использует избыточность сообщения, заключённую в неоднородном распределении частот символов его (первичного) алфавита, то есть заменяет коды более частых символов короткими двоичными последовательностями, а коды более редких символов — более длинными двоичными последовательностями. Кодировка при помощи данного алгоритма осуществляется для приведения всех массивов данных к единой структуре, для дальнейшего проведения анализа и сравнения полученных значений.

3. Среднее число элементарных символов на символ сообщения. Данную величину можно представить как произведение количества символов кода, которым закодирован символ, на вероятность появления данного символа.

$$k = \sum_{i=1}^{n} l_{\text{кода } i} \cdot p_i$$

4. Количество информации, приходящееся на один символ, рассчитывается с помощью формулы:

$$I = \frac{H(x)}{k} \text{ бит/символ.}$$

Далее необходимо провести анализ нескольких текстовых данных в соответствии со всеми вычислениями и манипуляциями над текстом, чтобы отследить закономерность и понять, как изменяется энтропия в зависимости от точности описания предметной области в тексте.

В результате, чтобы провести данный эксперимент, была разработана программа, позволяющая проводить обработку текстовых данных, разделение, кодирование и вычисления над ними. Реализация программы представлена на рисунке 2.

Проведение эксперимента

В исследовании, при проведении эксперимента для анализа текстовых данных, было использовано три произведения:

- Л. Н. Толстой «Война и мир»
- Ф. М. Достоевский «Идиот»
- М. А. Шолохов «Тихий Дон»

Каждое из вышеперечисленных произведений было переведено в формат txt для загрузки в программу. При открытии программы требуется указать путь до текстового файла и указать путь для выгрузки полученных результатов. Также в программе была предусмотрена функция для разделения текста на равные отрезки: начиная от 5 и заканчивая 15. Для данных отрезков будет вычислены количество символов в отрезке и количество информации соответственно. Данная информация необходима для построения графиков. После запуска программы происходит анализ текстового документа, а именно:

- Токенизация, стоп-листинг
- Разбиение текста на предложения, слова, буквы и слоги
- Расчет количества предложений, слов, слогов и букв в текстовом документе
- Расчет количества уникальных предложений, слов, слогов и букв в текстовом документе
- Для полученных четырех типов данных с разделением по предложениям, словам, слогам и буквам рассчитывается информационная энтропия
- Кодировка всех полученных текстовых данных методом Шеннона-Фано
- Расчет количества элементарных символов для закодированных массивов текста
- Расчет количества информации, приходящееся на каждый элементарный символ.

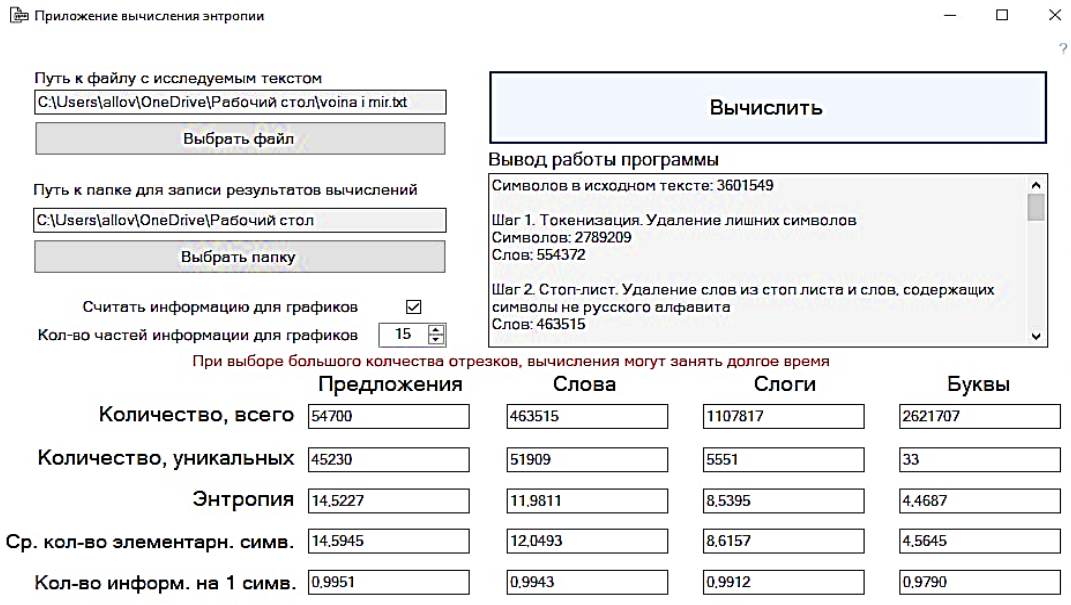

Рис. 2. Реализация программы

По окончании работы программы в указанный ранее путь записывается 5 отчетов: отчет по расчетам, информация для построения графиков, полученные результаты и кодировка предложений, слов, слогов и букв. Необходимо было произвести данные расчеты, чтобы провести наиболее обширный анализ текстовых документов и выявить закономерности, исследуя разные текстовые документы.

После прогона всех трех произведений через программу, были получены результаты расчетов, которые представлены в таблице 1.

Таблица 1

	Предложения	Слова	Слоги	Буквы
Л. Н. Толстой «Война и мир»				
Количество, всего	54700	463515	1107817	2621707
Количество, уникальных	45230	51909	5551	33
Энтропия	14,5227	11,9811	8,5395	4,4687
Ср. кол-во элем. символов	14,5945	12,0493	8,6157	4,5645
Кол-во инф-ии на 1 элем. символ	0,9951	0,9943	0,9912	0,979
Ф. М. Достоевский «Идиот»				
Количество, всего	14250	185352	419394	995003
Количество, уникальных	14019	26099	3998	33
Энтропия	13,7587	11,2071	8,4421	4,4688
Ср. кол-во элем. символов	13,8111	11,2912	8,5152	4,5615
Кол-во инф-ии на 1 элем. символ	0,9962	0,9925	0,9914	0,9796
М. А. Шолохов «Тихий Дон»				
Количество, всего	21945	172143	429633	1028933
Количество, уникальных	21491	44224	5616	32
Энтропия	14,3653	13,0687	8,8823	4,4997
Ср. кол-во элем. символов	14,4504	13,1481	8,9634	4,5733
Кол-во инф-ии на 1 элем. символ	0,9941	0,9939	0,9909	0,9839

Проанализировав полученные значения по трем произведениям, можно сделать следующие выводы. Текстовый документ – в данном случае идет рассмотрение крупных произведений, является описанием предметной области.

С точки зрения кодирования можно сделать вывод, что уникальная единица текста кодируется одним и тем же знаком, в то время как количество информации, которое будет приходится на уникальную единицу текста, будет разным. Это означает, что разное сочетание уникальных знаков и их положение напрямую влияет на количество информации.

Также можно сделать вывод, что чем крупнее является единица, описывающая предметную область, тем ближе к единице будет стремиться энтропия. Данный вывод можно сделать исходя из графика (рис. 3).

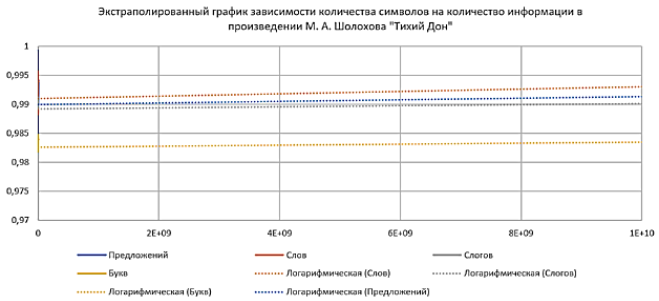

Рис. 3. График экстраполяции зависимости количества символов на количество информации в произведении М. А. Шолохова «Тихий Дон»

Было отмечено, исходя из таблицы, что с ростом количества единиц кодирования, уменьшается количество информации на один символ.

Заключение

При осуществлении этапа предобработки текста и проведении анализа трех крупных произведений, можно предположить, что при наиболее подробном описании предметной области в текстовом документе, порог по количеству информации на один элементарный символ не будет превышать 0,978 бит. Следовательно, на один элементарный символ будет приходиться 0,978 бит информации и 0,022 бита избыточного шума. При экстраполяции графиков зависимости количества символов исходного текста от количества информации видно, что с увеличением количества символов количество информации на один символ будет стремиться к единице. Одна буква содержит минимум 4.56 элементарных символов, тогда на одну букву приходится около 4,46 бит информации. При помощи машинной предобработки текста можно сделать вывод о содержании смысла в текстовом документе, исходя из количества информации, приходящейся на один символ — оно должно составлять не менее 4,46 бит.

Литература

1. *Боярский К.К.* Введение в компьютерную лингвистику. Учебное пособие. СПб: НИУ ИТМО, 2013. 474 с.
2. *Стельмах М.А., Миснянкин В.Г., Кунац А.Ю., Костина А.В.* Использование промежуточных языков представления для упрощения процесса перевода естественного языка в запросы к базе данных // Наука настоящего и будущего. 2017. Т. 1. С. 114-116.
3. *Аванесов, Рубен Иванович.* Фонетика современного русского литературного языка. // Р. И. Аванесов. - Leipzig: Zentralantiquariat der DDR, 1975. 239 с.
4. *Щерба Л.В.*. Избранные работы по языкознанию и фонетике; отв. ред. доц. Маргарита Ивановна Матусевич; Ленинградский ордена Ленина гос. ун-т им. А. А. Жданова. Ленинград: Изд-во Ленинградского ун-та, 1958. 182 с.
5. *Маклачкова В. В., Гадасин Д. В., Волкова М. Д., Вакурин И. С.* Лексический и семантический поиск статей в научной библиотеке // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 21-30.
6. *Полякова А. Н., Кольцова А. В., Гадасин Д. В*. Место искусственного интеллекта в сетях хранения данных // Шаг в будущее: искусственный интеллект и цифровая экономика: Smart Nations: экономика цифрового равенства: Материалы III Международного научного форума, Москва, 09-10 декабря 2019 года. М.: Государственный университет управления, 2020. С. 4-6.
7. *Гадасин Д. В., Пак Е. В*. Применение модели Бэкмэна для распределения потоков в сетях с сегментной маршрутизацией // REDS: Телекоммуникационные устройства и системы. 2020. Т. 10. № 4. С. 18-23.
8. *Гадасин Д. В., Юдина А. А*. Сложные сети как симбиоз современных сетевых технологий и искусственного интеллекта // Технологии информационного общества: Сборник трудов XIV Международной отраслевой научно-технической конференции, Москва, 18-19 марта 2020 года. М.: Издательский дом Медиа Паблишер, 2020. С. 270-272.
9. *Вакурин И. С., Гадасин Д. В.* Аспекты легальности принятия решений системами искусственного интеллекта // Искусственный интеллект и цифровая экономика: взгляд студенчества: материалы I Всероссийской студенческой научно-практической конференции, Москва, 13 ноября 2019 года / Министерство науки и высшего образования Российской Федерации, Государственный университет управления. – Москва: Государственный университет управления, 2020. С. 235-237.
10. *Литвин Я.С., Гадасин Д.В*. Семантическая сеть как инструмент обработки визуальной информации // Телекоммуникации и информационные технологии. 2018. Т. 5. № 2. С. 111-118.
11. *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.
12. *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.
13. *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.
14. *Павлов С.В., Докучаев В.А.* О разработке методологических основ построения модели технических средств радиомониторинга // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 7. С. 48-51.
15. *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.
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.

ПЕРЕСТРОЙКА РЕЗОНАНСНЫХ ТУННЕЛЬНЫХ УРОВНЕЙ В ПРОЦЕССЕ ОБРАЗОВАНИЯ СВЕРХРЕШЕТКИ

Дегтярев Владимир Фомич,
Московский технический университет связи и информатики (МТУСИ), доцент, к.ф.-м.н., Москва, Россия
vfsteel2008@gmail.com

Жилинский Алексей Петрович,
Московский технический университет связи и информатики (МТУСИ), профессор, д.ф.-м.н., Москва, Россия
zhilinsk@yandex.ru

Аннотация

Установлено, что при образовании сверхрешетки, состоящей из последовательности потенциальных ям и барьеров, возникают резонансные уровни, энергия которых определяется числом волн де Бройля, укладывающихся на ширине ямы. Для частиц с энергией, равной энергии уровней, прозрачность структуры равна единице. С увеличением числа звеньев происходит расщепление этих уровней на близкие подуровни. Рассмотрены схема и механизм перестройки уровней в цепочке. В основе этого механизма лежат представления о точках смены фаз колебаний осцилляторов при образовании цепи. Установлено, что параметры этих подуровней (энергия, полуширина и волновая функция) зависят от параметров барьеров и числа ячеек в цепочке. Предложена модель, позволяющая определить характеристики этих подуровней, в частности их энергию и волновые функции.

Ключевые слова: *квантовая механика, квантовый барьер, волновая функция, прозрачность, сверхрешетка, туннелирование.*

Введение

Известно, что уменьшение активных размеров полупроводниковых структур приводит к проявлению новых кантовомеханических явлений. Характерные размеры структур, при которых проявляются эти эффекты, составляют 1...100 нм [1]. В этом диапазоне в полной мере начинают проявляться квантовые эффекты, и физика проводимости определяется квантово-механической интерференцией электронных волн. Вследствие сильного пространственного ограничения носителей заряда в этих структурах *(a* ~ 1 нм), энергия размерного квантования имеет порядок $\Delta E \sim \frac{\hbar^2}{2m^* a^2} = 340 meV$. Эта величина сравнима с шириной запрещенной зоны типичных полупроводников и на порядок превосходит тепловую энергию носителей заряда при комнатной температуре. Таким образом, в полупроводниковых наноструктурах эффекты размерного квантования будут играть существенную роль, определяя их основные электрофизические свойства. При этом возникает возможность использования квантовых эффектов для качественно новых технологий. Физическое и математическое описание таких структур и их свойств достаточно широко представлено в литературе, как в периодической, так и в различных монографиях [2-4]. Особый интерес представляет резонансное прохождение заряженных частиц сквозь периодическую структуру. Исследованию этой проблемы посвящен ряд публикаций, например [2,3,5]. В этих работах рассмотрены важные вопросы, связанные с поведением электронов в идеальных бесконечных структурах. Основное внимание при этом было уделено зонной теории, энергетическому спектру электронов в идеальных и неидеальных системах. В настоящее время эффект резонансного туннелирования в тонкопленочных гетероструктурах является основой создания целого ряда новых приборов

Задачи, связанные с распространением волн (электромагнитных, электронных) в слоистых средах возникают и во многих других разделах науки и техники. В частности, такие среды, как плазма, ионосфера, атмосфера, океан, содержат слоистые структуры. Решение задач о прохождении волн в этих средах, расчет коэффициентов отражения и прохождения при распространении электромагнитных

волн имеют большое значение как для расчета радиотрасс с отражением от ионосферы, так и для многих задач дистанционной диагностики ионосферной плазмы [6, 14-17]. Поэтому вопросы, рассмотренные в данной работе, могут представлять интерес и для специалистов этих направлений.

Методика моделирования

Решение задач наноэлектроники невозможно без использования современных математических методов. В настоящей работе нахождение волновых функций и коэффициентов прозрачности системы барьеров проводилось путем непосредственного решения уравнения Шредингера для заданного потенциала с соответствующими граничными условиями в системе компьютерной алгебры MAPLE. Эта система обладает широкими возможностями для численного и аналитического решения многих задач, а также для графического представления полученных результатов. В работе в качестве примера рассмотрена сверхрешетка, состоящая из ям и барьеров с параметрами: ширина барьера (a) – 1 нм, его высота (U) – 2 эВ, ширина ямы (b) – 1 нм, период d = 2 нм. Амплитуда падающей волны принималась равной единице ($\Psi_{in}(x) = e^{ikx}$). В работе исследовалась зависимость прозрачности структуры от энергии частицы. По зависимости $T(E)$ (T – коэффициент прозрачности цепочки) определялось положение максимума пика (E_0), его энергетическая ширина ΔE_{05} на половине его высоты и добротность. Отметим, что результаты, полученные в данной работе, а также развитая методика, относятся и к цепочкам с другими параметрами.

Резонансно-туннельные уровни в элементарном звене цепочки

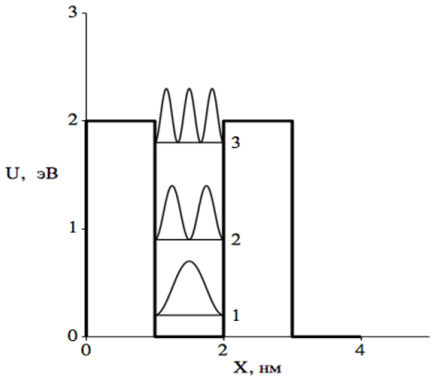

Рис. 1. Схема элементарного звена цепочки и волновые функции в случае, когда на ширине ямы укладывается 1 – одна полуволна, 2 – две полуволны, 3 – три полуволны

В рассматритваемой структуре элементарным звеном является двухбарьерная квантовая структура (ДБКС), показанная на рисунке 1. Рассмотрен случай, когда энергия частицы меньше высоты барьера. Резонансные уровни и волновые функции для этого случая показаны на рисунке. Основные результаты компьютерного моделирования перечислены ниже.

1. В элементарном звене, также как и в потенциальной яме конечной глубины (толщина барьера при этом считается бесконечно большой) при $E < U$ образуется система резонансно-туннельных уровней (**РТУ**) [7], положение которых определяется условием $b = n\frac{\lambda}{2}$, где b– ширина ямы, λ-длина волны для частицы с энергией E, n= 1, 2,… – число полуволн, укладывающихся на ширине ямы (рис.1 кривые 1, 2, 3). Прозрачность структуры для частиц с такой энергией равна единице.

2. В элементарном звене рассматриваемой структуры число таких уровней равно 3 (рис. 1). Как показывают расчеты, их энергетическое положение следующее: E_{01} = 0.228949 эВ, E_{02} = 0.887202 эВ и E_{03} = 1.818140 эВ. Энергия этих пиков соответствует положению энергетических уровней в яме такой же ширины и глубины, что неудивительно при малой прозрачности барьеров, разделяющих ячейки. Это легко понять, если учесть, что оба типа пиков соответствуют условию, что на ширине ямы умещается одинаковое число полуволн.

Расщепление резонансных уровней на подуровни при увеличении числа звеньев

Рассмотрим движение электрона в сверхрешетке. Движение частицы в цепочке описывается уравнением Шредингера [7,8]

$$\frac{d^2\Psi(x)}{dx^2}+\frac{2m}{\hbar^2}(E-U(x))\Psi(x)=0) \quad (1)$$

U – потенциальная энергия взаимодействия электрона с барьером. Так как цепочка состоит из N одинаковых звеньев, то U и Ψ- периодические функции. В соответствии с принципом суперпозиции [7,8] $\Psi=\sum_n C_n\Psi_n$ Здесь C_n – амплитуда вероятности нахождения электрона у n-го атома.

Поскольку размеры цепочки ограничены, то необходимо учесть граничные условия

$$\Psi(0)=0 \text{ и } \Psi(x_{N+1})=0 \quad (2)$$

Будем учитывать взаимодействие электрона только с ближайшими соседями. Тогда уравнение, описывающее поведение электрона в n-ой ячейке принимает вид [7-9]:

$$(E_m-E_{0m})C_n=-A_mC_{n-1}-A_mC_{n+1} \quad (3)$$

Здесь для уровня с номером m (m=1, 2, 3 на рис. 1) $-A_m=-H_{n,m}=\int\Psi_n^*H\Psi_m dx$, энергия исходного уровня составляет E_{0m}. Поскольку размеры цепочки ограничены, то происходит отражение волн от передней и задней границы и образуется стоячая волна. Поэтому искать решение уравнения (3) следует в виде

$$C_n=(\alpha\exp(ikx)+\beta\exp(-ikx)) \quad (4)$$

Первое слагаемое в (4) представляет собой прямую волну, а второе – отраженную. Значения коэффициентов α и β определяются из граничных условий. Подставляя выражение (4) в (3), после простых преобразований получаем

$$E_m=E_{0m}-2A_m\cos(kd) \quad (5)$$

Это соотношение дает связь между энергией электрона и его волновым числом (дисперсионное уравнение).

Используя граничные условия (2), получим

$$\alpha+\beta=0$$
$$\alpha\exp(ik(N+1)d)+\beta\exp(-ik(N+1)d)=0$$

Отсюда

$$\exp(ik(N+1)d)-\exp(-ik(N+1)d)=0.$$

или

$$\sin(k(N+1)d)=0 \quad (6)$$

В решетке, состоящей из N звеньев, существует N возможных решений, соответствующих определенным модам колебаний. Этим модам (j – номер моды колебаний) соответствуют значения волнового числа

$$k_j=\frac{\pi j}{(N+1)d}, \quad \text{где } j=1, 2,\ldots N. \quad (7)$$

Подставляя (7) в (5), получаем для энергии подуровней следующее соотношение

$$E_{mj}=E_{0m}-2A_m\cos(\frac{\pi j}{N+1}) \text{ , где } j=1, 2,\ldots, N \quad (8)$$

Проведенное рассмотрение приводит к зонному характеру распределения состояний по энергии и позволяет найти энергетический спектр частиц в зависимости от количества звеньев в цепи [8,9].

Эта модель, однако, не вскрывает физической сути процесса. Сделать это можно, рассмотрев конфигурацию волновой функции (рис. 2) [4,9,10]. Заметим, что здесь речь идет именно о самой волновой функции (рис. 2Б), а не о квадрате ее модуля, как показано на рис. 2А.

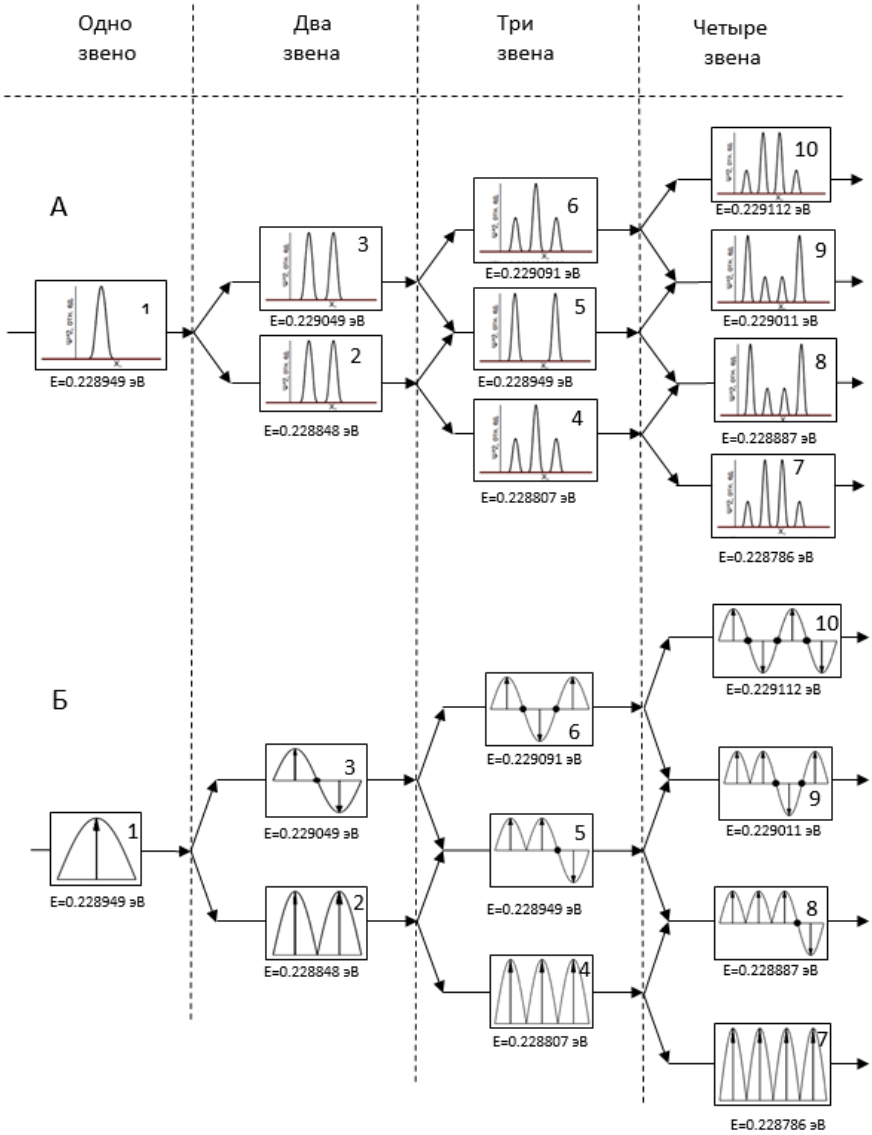

Рис. 2. А. Схема расщепления уровня с энергией 0.228949 эВ в простой цепочке (в квадратиках показаны квадраты модуля соответствующих волновых функций).
Б. Диаграммы, иллюстрирующие возможные конфигурации волновых функций. При изменении фазы колебаний на противоположную энергия колебаний возрастает и соответственно увеличивается энергия подуровней. Места смены фаз колебаний отмечены точкой. На рисунке конфигурация волновой функции в элементарном звене условно изображена значком

Схема расщепления нижнего уровня (уровень *1* на рис. 1) в процессе образования цепочки рассмотрена ниже. В этом случае на ширине ямы при резонансе укладывается одна полуволна. Пусть теперь к данному звену присоединили второе. В результате взаимодействия осцилляторов в звеньях образуется новая волновая функция и происходит расщепление этого уровня на два подуровня. В первом случае колебания в обеих полуволнах происходят в одной фазе, а во втором – в противоположных. Это проиллюстрировано на рис. 2Б на диаграммах 2 и 3 (место смены фаз колебаний обозначено на диаграммах точкой).

В случае противофазных колебаний энергия увеличивается, а случае синфазных уменьшается. Заметим, что аналогичным образом изменяется частота колебаний и в связанных электрических контурах [11,12].

При увеличении числа звеньев взаимодействие осцилляторов в соседних ячейках происходит подобным образом. При этом число точек смены фаз колебаний может изменяться от 0 до максимального значения *N*-1, где *N* – число звеньев в цепи. Число подуровней, таким образом, будет равно числу звеньев в цепочке. Подуровень, который соответствует колебаниям, происходящим синфазно, обладает наименьшей энергией. С появлением точек смены фаз энергия подуровня возрастает пропорционально числу этих точек. На рисунке 2Б приведены соответствующие диаграммы. Следует отметить, что положение точек смены фаз в цепочке не влияет на энергию подуровней. С увеличением длины цепочки энергетическое расстояние между подуровнями уменьшается, что связано с ослаблением взаимодействия между звеньями.

Перестройка волновых функций при образовании цепочки

При образовании слоистой квантово-размерной структуры существенные изменения испытывают не только энергетические уровни, но и волновые функции. Рассмотрим, например, волновые функции, соответствующие первому уровню в элементарной ячейке, в четырехзвенной цепочке. Квадраты модуля этих функций показаны на рисунке 3. Рисунок 3А соответствует первой моде колебаний, а рисунке 3Б – второй моде. Длина стрелок на рисунке пропорциональна вероятности обнаружения частицы в данном звене.

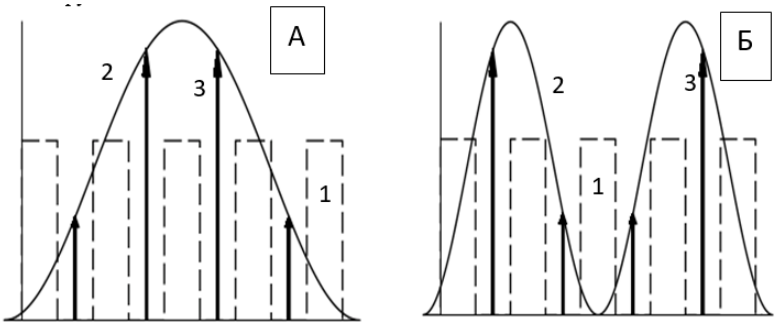

Рис. 3. Схема образования волновой функции в цепочке из четырех звеньев: А-первая мода (Е=0.228786 эВ, Б-вторая мода (Е=0.228887 эВ). 1-схема барьеров; 2-огибающая $\Psi^2(x)$ для первой (А) и второй (Б) моды ; 3-стрелки, длина которых пропорциональна $\Psi^2(x)$

В элементарном звене квадрат модуля волновой функции имеет характерный колоколообразный вид (рис. 2А, кривая 1). При образовании цепочки волновая функции приобретает сложный вид. Конфигурация волновой функции при этом в значительной степени определяется структурой слоёв цепочки. Важную роль при этом играют развитые в предыдущем разделе представления о точках смены фаз колебаний. Число мод колебаний определяется числом ячеек. Наименьшую энергию имеют колебания, соответствующие первой моде, когда на длине цепочки укладывается одна полуволна. Колебания осцилляторов здесь во всех звеньях происходят в одной фазе (точки смены фаз отсутствуют (рис. 2)). Для второй моды колебаний на длине цепочки укладывается две полуволны. Такой конфигурации соответствует одна точка смены фаз колебаний и энергия уровня *E*=0.228887 эВ. Квадрат модуля волновой функции в этом случае показан на рисунке 2Б.

Рассмотренная ситуация сохраняется и для последующих мод колебаний. Здесь также росту номера моды соответствует увеличение числа точек смены фаз и увеличение энергии подуровней. Обращает на себя внимание симметричный вид волновых функций относительно середины цепочки. Именно такой вид волновой функции обеспечивает, по нашему мнению, равенство потоков частиц слева направо и справа налево, что и приводит к высокой прозрачности цепочки. Следует отметить также, что волновые функции *m*-ой и (*N*+1-*m*)-ой мод имеют совершенно одинаковый вид.

Действительно из (4) и (6) вытекает, что вероятность обнаружить частицу в ячейке с номером n можно представить в виде

$$\Psi^2 = a \cdot \sin^2(kx_n) = a \sin^2\left(\frac{\pi jn}{N+1}\right) \qquad (9)$$

Здесь учтено, что $x_n = nd$, n- номер ячейки (n=1…N), d -период структуры, j-номер моды (j=1..N).

Из этого соотношения видно, что при $j_1 + j_2 = N + 1 \quad \Psi_1^2 = \Psi_2^2$, и при $n_1 + n_2 = N + 1 \quad \Psi_1^2 = \Psi_2^2$

Таким образом, рассмотренная модель позволяет объяснить основные особенности перестройки волновых функций при образовании слоистой структуры.

Заключение

В работе изучен процесс перестройки резонансно-туннельных уровней (РТУ) при образовании слоистой квантово-размерной структуры. Установлено, что существенным изменениям при этом подвергается и волновая функция.

1. Установлено, что при образовании цепочки РТУ расщепляются на систему подуровней, число которых равно числу звеньев в цепочке.
2. Определены значения энергии этих подуровней и соответствующие им волновые функции в зависимости от числа звеньев. Прозрачность цепочки для этих значений энергии равна единице. Предложена методика, позволяющая рассчитать эти энергии и построить соответствующие волновые функции.
3. Для объяснения механизма перестройки уровней развиты представления о точках смены фаз колебаний. Это такие точки, в которых фазы колебаний в соседних звеньях изменяются на противоположные. Число таких точек изменяется в пределах от нуля до N-1. Соответственно число подуровней, на которые расщепляется исходное состояние, равно числу звеньев в цепи. Чем больше таких точек, тем выше энергия колебаний.

Литература

1. *Усанов Д.А., Скрипаль Ал.В., Скрипаль Ан.В., Абрамов А.В.* Компьютерное моделирование наноструктур: Учеб. пособие. Саратов, 2013. 100 с.
2. *Драгунов В.П., Неизвестный И.Г., Гридчин В.А*. Основы наноэлектроники. М.: Логос, 2006. 496 с.
3. *Демиховский В.Я, Вугальтер Г.А*. Физика квантовых низкоразмерных структур. М.: Логос, 2000. 248 с.
4. *Херман М.* Полупроводниковые сверхрешетки. М.: Мир, 1989. 240 с.
5. *Аладышкин А.Ю.* Туннельные явления в нанофизике. Н. Новгород: Нижегород. гос. ун-т., 2011, 32 с.
6. *Голант В.Е., Жилинский А.П., Сахаров И.Е*. Основы физики плазмы. СПб.: Лань, 2011. 448 с.
7. *Левич В.Г., Вдовин Ю.А., Мямлин В.А*. Курс теоретической физики. Т. 2. М.: Наука, 1971, 936 с.
8. *Фейнман Р., Лейтон Р., Сэндс М.* Фейнмановские лекции по физике. Т. 9, Квантовая механика (II). М.: Мир, 1967. 259 с.
17. *Дегтярев В.Ф., Жилинский А.П.* Наноструктуры. Математическая физика и моделирование. 2020. Т.21, № 2. С. 33.
18. *Киттель Ч*. Введение в физику твердого тела. М.: Наука, 1978. 791 с.
19. *Стрелков С.П.* Введение в теорию колебаний. СПб: Лань, 2005. 440 с.
20. *Крауфорд Ф.* Берклеевский курс физики. Т. 3. Волны. М.: Наука, 1984. 521 с.
21. *Жилинский А.П., Дегтярев В.Ф*. Некоторые особенности изучения взаимодействия микрочастиц с потенциальным барьером в курсе квантовой физики в техническом университете // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2018. Т. 7. № 1. С. 12-14.
22. *Жилинский А.П., Дегтярев В.Ф.* Некоторые свойства резонансных пиков прозрачности при образовании цепочки барьеров // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 5. С. 46-51.
23. *Дегтярев В.Ф., Жилинский А.П*. Туннельное уширение резонансных уровней в слоистых квантово-размерных структурах // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11. № 3. С. 19-26.
24. *Жилинский А.П., Дегтярев В.Ф*. Особенности взаимодействия микрочастиц с прямоугольным и трапециевидным потенциальным барьером // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 8. С. 10-16.

О РАЗРАБОТКЕ ДАТАСЕТА ДЛЯ ОБНАРУЖЕНИЯ СЕТЕВЫХ АТАК

Ерохин Сергей Дмитриевич,
Московский технический университет связи и информатики, ректор, к.т.н., доцент, Москва, Россия
esd@mtuci.ru

Борисенко Борис Борисович,
Московский технический университет связи и информатики, ведущий научный сотрудник, к.т.н., доцент, Москва, Россия
fepem@yandex.ru

Фадеев Александр Сергеевич,
Московский технический университет связи и информатики, научный сотрудник, Москва, Россия
aleksandr-sml@mail.ru

Мартишин Иван Дмитриевич,
Московский технический университет связи и информатики, научный сотрудник, Москва, Россия
martishinid@gmail.com

Аннотация

В ходе работы был проведен обзор существующих классов сетевых атак. Проанализированы наиболее популярные и актуальные датасеты. На основе датасета CSE-CICIDS-2018 за счет выделения основных информативных признаков произведено сокращение размерности входного вектора датасета. Проведена классификация компьютерных атак (КА), используемых в датасетах.

Ключевые слова: *Сетевые атаки, набор данных (датасет), CSE-CICIDS-2018, выделение признаков, системы обнаружения вторжений (СОВ, IDS).*

Введение

В современном мире сетевая безопасность имеет решающее значение. За последние годы произошло множество кибератак. Крупнейшими, наиболее сложными и серьезными кибератаками стали: SolarWinds, Microsoft Exchange, Kaseya, Log4j [1], WannaCry [2], Petya [3] взлом бюро кредитных историй Equifax [4], утечка данных Facebook [5]. В результате пострадали миллионы пользователей и сотни предприятий.

Современная стратегия обеспечения сетевой безопасности должна учитывать ряд таких факторов, как увеличение надежности сети, эффективное управление безопасностью и защита от постоянно эволюционирующих угроз и новых видов атак. Одним из подходов к превентивному решению проблем безопасности корпоративных систем является моделирование угроз, которое включает определение основных ресурсов в системе и их угроз. Данный подход может быть объединен с имитацией атак для получения вероятностных оценок безопасности.

В данной работе на основе анализа существующих датасетов проведена оценка признаков.

Обзор существующих классов сетевых атак

Сетевая атака – это попытка причинить вред, раскрыть, изменить, уничтожить, украсть или получить незаконный доступ к ресурсу сетевой системы [6].

Применительно к датасетам наиболее часто используемыми классами атак являются [7]:

Атаки типа "отказ в обслуживании" (DoS) основаны на временном блокировании нормального использования сетевых ресурсов путем генерации большого объема трафика. Примерами DoS-атак являются атаки botnet, Slowloris, smurf и SYN flood.

Распределенные DoS-атаки (DDoS) основаны на ограничениях пропускной способности серверов, перегружая их запросами с целью вызвать отказ в обслуживании. Примерами DDoS-атак являются атаки типа "отказ локальной сети" (LAND), ping-of-death, RUDY и teardrop.

User-to-Root (U2R) атаки с целью обнаружения уязвимостей системы, при которых обычный пользователь несанкционированно получает root-права. Примерами атак U2R являются переполнение буфера, rootkit, Perl и loadmodule.

Атаки Remote-to-Local (R2L) направлены на использование удаленной системы для получения не-

санкционированного доступа к целевой системе и нанесения ей ущерба. Примеры атак R2L включают перебор Secure Shell (SSH), warezmaster, multihop, imap и шпионские атаки (spy attacks).

Probe атаки основаны на поиске уязвимостей во всей сети путем отправки сканирующих пакетов и получения информации о системе. Примерами Probe атак являются атаки Satan, IP sweep и port sweep.

Перебор по словарю (password, словарные атаки) – атака, заключающаяся в получение несанкционированного доступа к системе, используя методы слов из словаря для кражи паролей. Примерами атак являются перебор FTP-Patator и перебор SSH-Patator.

Атаки-инъекции (Injection) используют скрипты, которые вводят команды/запросы с целью получения несанкционированного доступа и кражи информации. Примерами инъекционных атак являются SQL-инъекции и межсайтовый скриптинг (XSS).

При разработке СОВ необходимо учитывать дополнительных факторы:

1. Атака одного типа может быть началом другой атаки другого типа.
2. Некоторые характеристики атак могут меняться с течением времени.
3. Некоторые типы атак могут иметь схожий характер.

Признаки, используемые для обнаружения атак типа probe, U2R и R2L, имеют высокую степень сходства, что объясняет, почему эти три типа атак часто неправильно классифицируются между собой [8]. Помимо указанных выше атак, в датасетах классифицируют: фаззеры, бэкдоры, эксплойты, шеллкоды, черви, ботнет, веб-атаки.

Анализ современных датасетов, используемых при обнаружении вторжений

Датасет для обнаружения вторжений может быть разработан путем сбора информации из различных источников, таких как потоки сетевого трафика, содержащие информацию о хосте, поведении пользователя и конфигурации системы [9].

Для обнаружения вторжений и оценки производительности СОВ исследователями используются датасеты, созданные на основе данных сетевого трафика. Среди них: DARPA, KDD CUP 99, NSL-KDD, DEFCON, CAIDA, LBNL, CDX, Kyoto, Twente, ISCX2012, ADFA, CTU-13, UNSW-NB15, CIDDS-001, UGR-16, CIC-IDS2017, CSE-CICIDS-2018. Основные характеристики датасетов указаны в таблице 1 [10-13].

Таблица 1

Описание датасетов

№	Датасет	Количество признаков	Классы сетевых атак
1	DARPA (1998-2000)	41	DoS, R2L, U2R, Surveillance/probing attacks
2	KDD CUP 1999	41	DoS, R2L, U2R, Surveillance/probing attacks
3	DEFCON (2000)	Flag traces	Telnet Protocol Attacks
4	Kyoto (2006)	24	Различные атаки на honeypots (backscatter, DoS, exploits, malware, port scans, shellcode)
5	NSL-KDD (2009)	43	DoS, R2L, U2R, Surveillance/probing attacks
6	CAIDA (2002-2016)	20	DDoS
7	LBNL	Internet traces	Следы вредоносного воздействия (malicious traces)
8	CDX (2009)	5	Buffer Overflow
9	Twente (2009)	Потоки IP	Malicious traffic, Side-effect traffic, Unknown traffic, and Uncorrelated alerts
10	ISCX 2012	19	Brute Force SSH, HTTP, DoS, DDoS using an IRC Botnet, Infiltrating the network from inside
11	ADFA (2013, 2017)	System call traces	Zero-day attacks, Stealth attack, C100, Webshell attack
12	CTU-13 (2011)	33	botnets (Menti, Murlo, Neris, NSIS, Rbot, Sogou, Virut)
13	UNSW-NB15 (2015)	49	Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconaissance, Shellcode, Worms
14	CIDDS-001	14	Port scanning, DoS, BruteForce, Ping Scan
15	UGR-16	132 (Feature as Counter), 13 (в CSV файле)	low- and high-rate DoS, Port scanning, UDP port scanning, SSH scanning, Botnet, Spam
16	CIC-IDS2017	85	DoS Hulk, PortScan, DDoS, DoS GoldenEye, FTP-Patator, SSHPatator, DoS slowloris, DoS Slowhttptest, Bot, Infiltration, Heartbleed, Web Attack – Brute Force, Web Attack – XSS, Web Attack – SQL Injection
17	CSE-CICIDS-2018	80	Brute Force, Heartbleed, Botnet, DoS, DDoS, Web attacks, Infiltration of the network from inside

Выделение основных признаков, используемых в датасетах

В данной работе использован датасет CSE-CIC-IDS2018 [14], как наиболее полный, актуальный и практически применимый. Датасет включает 7 различных сценариев атак [15-18].

Он состоит из нескольких csv файлов, содержащих записи признаков сетевых сессий, которые были смоделированы в определенный день. Так каждый день был сгенерирован определенный тип атаки, который включал несколько ее видов (табл. 2). Совместно со вредоносным был сгенерирован трафик легального пользователя сети.

Таблица 2

Имя атаки	Дата проведения	Время начала	Время окончания
FTP-BruteForce	Wednesday-14-02-2018	10:32	12:09
SSH-Bruteforce	Wednesday-14-02-2018	14:01	15:31
DoS-GoldenEye	Thursday-15-02-2018	9:26	10:09
DoS-Slowloris	Thursday-15-02-2018	10:59	11:40
DoS-SlowHTTPTest	Friday-16-02-2018	10:12	11:08
DoS-Hulk	Friday-16-02-2018	13:45	14:19
DDoS attacks-LOIC-HTTP	Tuesday-20-02-2018	10:12	11:17
DDoS-LOIC-UDP	Tuesday-20-02-2018	13:13	13:32
DDOS-LOIC-UDP	Wednesday-21-02-2018	10:09	10:43
DDOS-HOIC	Wednesday-21-02-2018	14:05	15:05
Brute Force -Web	Thursday-22-02-2018	10:17	11:24
Brute Force -XSS	Thursday-22-02-2018	13:50	14:29
SQL Injection	Thursday-22-02-2018	16:15	16:29
Brute Force -Web	Friday-23-02-2018	10:03	11:03
Brute Force -XSS	Friday-23-02-2018	13:00	14:10
SQL Injection	Friday-23-02-2018	15:05	15:18
Infiltration	Wednesday-28-02-2018	10:50	12:05
Infiltration	Wednesday-28-02-2018	13:42	14:40
Infiltration	Thursday-01-03-2018	9:57	10:55
Infiltration	Thursday-01-03-2018	14:00	15:37
Infiltration	Thursday-01-03-2018	14:00	15:37
Bot	Friday-02-03-2018	10:11	11:34
Bot	Friday-02-03-2018	14:24	15:55

Каждая запись представляет из себя совокупность признаков. Их общее количество равно 80. Это число признаков велико и при обработке датасета с таким признаковым пространством потребуются огромные вычислительные и временные ресурсы. Поэтому существует необходимость сокращения признакового пространства с минимальным снижением качества распознавания атак [19].

В первую очередь необходимо очистить набор от записей с пустыми значениями признаков. Поскольку количество записей довольно большое, были удалены записи с бесконечными и некорректными значениями. Эти операции были проделаны для каждого csv файла. После было выполнено слияние наборов из всех файлов в один общий набор.

Изучив содержание набора, можно сделать вывод, что данные не сбалансированы по классам трафика (задача распознавания атак сводит к разделению всего трафика на класс обычных сессий и классы всевозможных атак, т.е. классификация) (рис. 1).

```
Benign                    4934548
Bot                        285933
SSH-Bruteforce             187589
FTP-BruteForce             187589
DDOS attack-HOIC           180386
Infilteration              158841
DoS attacks-Hulk           139890
DoS attacks-SlowHTTPTest   139890
```

Рис. 1. Соотношения классов анализируемого трафика

Произведена балансировка путем удаления некоторых записей из переполненных классов. В качестве соотношения выбрано:

$$\begin{cases} 2k, \text{если это класс } Benign \\ k, \text{если это класс атаки} \end{cases}, \text{где } k - \text{количество записей в наименьшем классе.}$$

При этом, в новом наборе будет $(n+1)k$ записей(n – количество классов) (рис. 2).

```
Benign                      279780
FTP-BruteForce              139890
DDOS attack-HOIC            139890
Bot                         139890
DoS attacks-Hulk            139890
Infilteration               139890
SSH-Bruteforce              139890
DoS attacks-SlowHTTPTest    139890
```

Рис. 2. Соотношение классов в новом наборе

Далее необходимо отбросить те признаки, по которым не будет происходить классификация. Это признаки Timestamp, Dst Port и Protocol. Использование этих признаков при обучении приведет к потере качества распознавания модели на новых данных и также сильно ухудшит качество классификации. Кроме того, данные признаки описывают адресацию в сети и могут быть легко подделаны. Таким образом, отбор будет проводиться из оставшихся 76 признаков. Признак Label определяет класс трафика и участвует в оценке качества моделей.

Для оценки признаков будет использоваться алгоритм машинного обучения Random forest. Но для начала необходимо перевести строковые метки классов в численные в столбце Label. Затем происходит разделение набора на два непересекающиеся класса: обучающая выборка (70% от всего набора) и тестовая выборка (оставшиеся 30%).

Сначала проверены показатели значимости признаков на одном дереве решения, используя алгоритм DecisionTreeClassifier. Дерево ограничено 15 листовыми узлами. На рисунке 2 представлено, как оно будет выглядеть после тренировки на обучающей выборке.

```
|--- Fwd Seg Size Min <= 30.00
|   |--- Init Fwd Win Byts <= 32719.50
|   |   |--- Bwd Pkt Len Mean <= 35.05
|   |   |   |--- Init Fwd Win Byts <= 2051.50
|   |   |   |   |--- Fwd Seg Size Min <= 22.00
|   |   |   |   |   |--- class: 0
|   |   |   |   |--- Fwd Seg Size Min >  22.00
|   |   |   |   |   |--- class: 6
|   |   |   |--- Init Fwd Win Byts >  2051.50
|   |   |   |   |--- Fwd IAT Mean <= 2520.75
|   |   |   |   |   |--- class: 1
|   |   |   |   |--- Fwd IAT Mean >  2520.75
|   |   |   |   |   |--- class: 0
|   |   |--- Bwd Pkt Len Mean >  35.05
|   |   |   |--- class: 0
|   |--- Init Fwd Win Byts >  32719.50
|   |   |--- Fwd IAT Min <= 209.50
|   |   |   |--- class: 0
|   |   |--- Fwd IAT Min >  209.50
|   |   |   |--- Flow Duration <= 73696.00
|   |   |   |   |--- class: 2
|   |   |   |--- Flow Duration >  73696.00
|   |   |   |   |--- class: 0
|--- Fwd Seg Size Min >  30.00
|   |--- Init Bwd Win Byts <= 228.50
|   |   |--- Fwd IAT Min <= 69.50
|   |   |   |--- Fwd Pkts/s <= 833333.34
|   |   |   |   |--- Bwd Header Len <= 26.00
|   |   |   |   |   |--- Fwd Pkts/s <= 366666.67
|   |   |   |   |   |   |--- class: 4
|   |   |   |   |   |--- Fwd Pkts/s >  366666.67
|   |   |   |   |   |   |--- class: 5
|   |   |   |   |--- Bwd Header Len >  26.00
|   |   |   |   |   |--- class: 0
|   |   |   |--- Fwd Pkts/s >  833333.34
|   |   |   |   |--- class: 5
|   |   |--- Fwd IAT Min >  69.50
|   |   |   |--- Fwd IAT Min <= 101400.00
|   |   |   |   |--- class: 3
|   |   |   |--- Fwd IAT Min >  101400.00
|   |   |   |   |--- class: 6
|   |--- Init Bwd Win Byts >  228.50
|   |   |--- class: 7
```

Рис. 3. Дерево решений после обучения

При апробации модели на тестовой выборке можно сделать вывод о качестве классификации. Для этого используется матрица ошибок (рис.4). Видно, что в целом классы распознаются однозначно, за исключение класса Benign, который иногда моделью воспринимается как класс Infilteration и двух классов DoS attacks-SlowHTTPTest и FTP-BruteForce, которые иногда путаются между собой.

```
[[40846,   308,    49,   298,     5,     0,   319,    67],
 [ 1492, 19374,     0,     0,     0,     0,     0,     0],
 [   40,     0, 21011,     0,     0,     0,     0,     0],
 [   57,     0,     0, 20955,    48,     1,     3,   209],
 [    0,     0,     0,     0, 11442,  9437,     0,     0],
 [    0,     0,     0,     0,  4549, 16507,     0,     0],
 [16228,   336,    43,     3,    21,     0,  4245,    38],
 [    0,     0,     0,     0,     1,     3,     0, 20917]],
```

Рис. 4. Матрица ошибок для дерева решений

С помощью метода отбора признаков по весам SelectFromModel получаем распределение весов признаков как характеристики их значимости. В итоге из 80 признаков значимыми алгоритм выявил 9 (рис. 5 и 6).

```
[0.026, 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.069, 0.   , 0.   , 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.02 , 0.   , 0.   , 0.216, 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.026, 0.04 ,
 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.22 , 0.171, 0.   , 0.212, 0.   , 0.   , 0.   , 0.   ,
 0.   , 0.   , 0.   , 0.   , 0.   ])
```

Рис. 5. Веса признаков в модели дерева решений

```
0.      Init Fwd Win Byts - 0.21969184441300585
1.      Fwd IAT Min - 0.21582095982922678
2.      Fwd Seg Size Min - 0.21228238742608013
3.      Init Bwd Win Byts - 0.17142515501231795
4.      Bwd Pkt Len Mean - 0.06932672575185797
5.      Fwd Pkts/s - 0.040224147558710815
6.      Flow Duration - 0.026025982215246952
7.      Bwd Header Len - 0.025700513056640995
8.      Fwd IAT Mean - 0.019502284736912574
```

Рис. 6. Отобранные признаки

Но при таком решении есть недостатки. В работе модели присутствует элемент рандомизации, поэтому при разных случайных числах рейтинг отбираемых признаков может немного изменяться. Поэтому возникает задача усреднить значения весов и вывести новый общий список отобранных признаков. Для решения этой задачи используется алгоритм RandomForestClassifier, представляющий из себя ансамбль деревьев решений. После обучения получаем значения весов признаков. Отобраны первые 20 наиболее значимых признаков (рис. 7). Для визуального сравнения соотношения весов признаков отображено на рисунке 8.

```
1.      #64     0.099   Init Fwd Win Byts
2.      #67     0.092   Fwd Seg Size Min
3.      #33     0.051   Fwd Header Len
4.      #0      0.047   Flow Duration
5.      #36     0.042   Bwd Pkts/s
6.      #17     0.042   Flow IAT Max
7.      #35     0.041   Fwd Pkts/s
8.      #14     0.040   Flow Pkts/s
9.      #15     0.040   Flow IAT Mean
10.     #18     0.036   Flow IAT Min
11.     #19     0.030   Fwd IAT Tot
12.     #23     0.029   Fwd IAT Min
13.     #22     0.029   Fwd IAT Max
14.     #20     0.026   Fwd IAT Mean
15.     #65     0.024   Init Bwd Win Byts
16.     #34     0.023   Bwd Header Len
17.     #53     0.018   Bwd Seg Size Avg
18.     #11     0.016   Bwd Pkt Len Mean
19.     #1      0.016   Tot Fwd Pkts
20.     #60     0.015   Subflow Fwd Pkts
```

Рис. 7. Признаки, отобранные случайным лесом и значения их весов

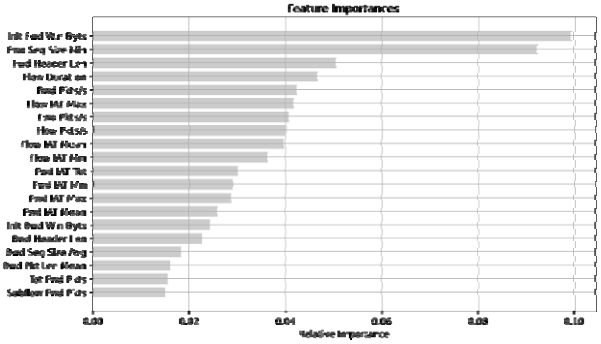

Рис. 8. Представление отобранных признаков

После апробации модели на тестовых данных при анализе матрицы ошибок, можно заметить, что проявляется тенденция предыдущей модели (рис. 9).

```
[[36162,    51,    24,     1,     0,     0,  5654,     0],
 [    8, 20853,     1,     0,     0,     0,     4,     0],
 [    4,     0, 21044,     0,     0,     0,     3,     0],
 [    0,     0,     0, 21273,     0,     0,     0,     0],
 [    0,     0,     0,     0, 10777, 10102,     0,     0],
 [    0,     0,     0,     0,  2417, 18639,     0,     0],
 [ 6386,    17,    26,     0,     2,     0, 14483,     0],
 [    0,     0,     0,     0,     1,     3,     0, 20917]],
```

Рис. 9. Матрица ошибок для случайного леса

Для наглядности можно представить распределение значений каждого из признаков в виде гистограммы(рис.10). Признаки, которые стоят выше в списке значимости имеют более «разбросанное» распределение, в отличие от тех, что находятся ниже, и их значения в основном сосредоточены в некой части области значений. Так же в признаках, которые ниже, содержится больше нулевых значений.

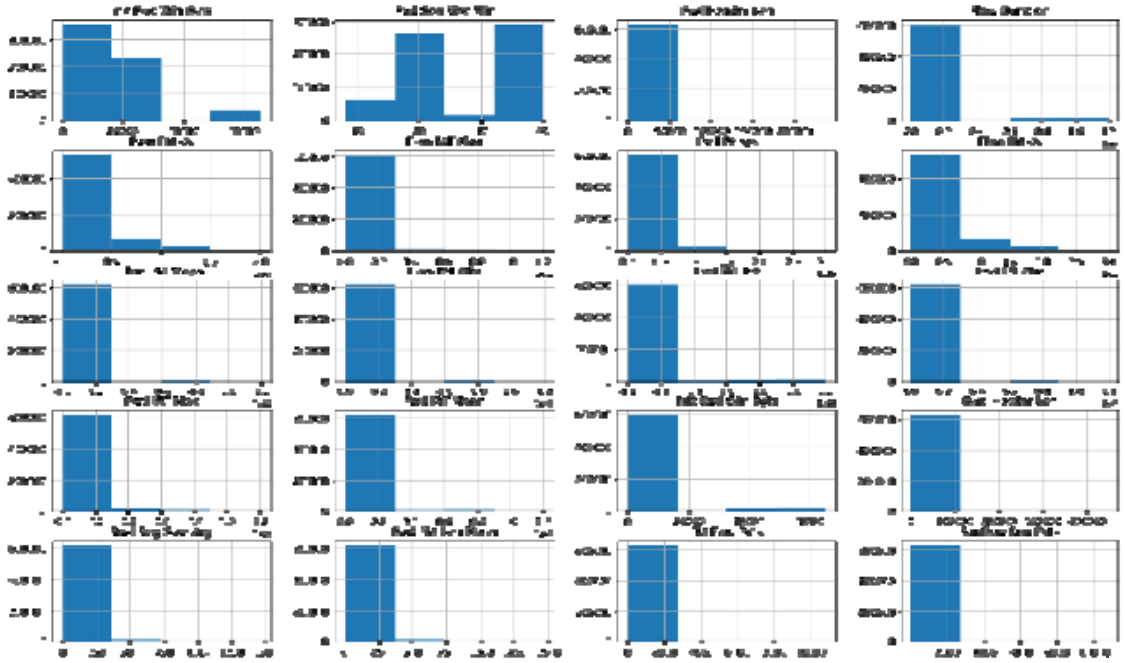

Рис. 10. Распределение значений признаков

Для проверки корреляции между признаками рассчитана корреляционная матрица (рис. 11). Из матрицы видно, что некоторые признаки сильно коррелируют с другими. Это может отрицательно отразиться на обучении моделей в будущем на данном датасете, так как возможно снижение производительности обобщения данных из-за высокой дисперсии и таким образом – меньшая интерпретируемость модели.

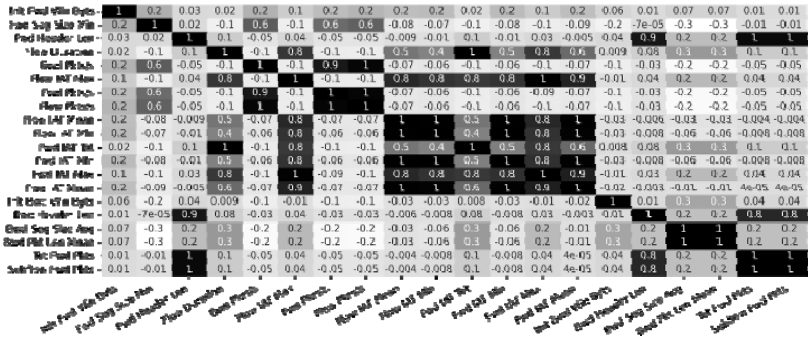

Рис. 11. Корреляционная матрица признаков

Для исправления ситуации необходимо отделить признаки, которые сильно коррелируют с другими и при этом находятся ниже в рейтинге. Интерпретируя это на матрицу, следует изучить верхний треугольник матрицы и убрать признаки, которые подписаны снизу. Результат очистки представлен на рисунке 12. Следует отметить, что удалялись не все сильно коррелированные признаки, а только те, что коллинеарны более чем с одним другим признаком, и коэффициент которых больше 0,9.

В итоге осталось 14 признаков.

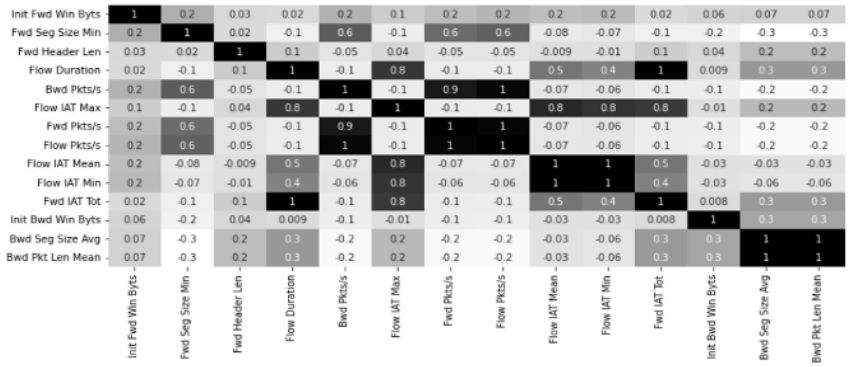

Рис. 12. Корреляционная матрица после удаления признаков

Данный способ снижения размерности реализуется через комплексный подход. Применяется встроенный отбор признаков, реализованный внутри модели случайного леса. После признаки были отсеяны с помощью изучения характеристик корреляции признаков между собой, что является методом фильтрации при отсеивании признаков. Было замечено, что даже после отсеивания признаки, находящиеся внизу списка значимости, имеют много нулевых значений. Это может быть результатом сбоев в работе сборщика признаков CICFlowMeter v3[20].

Кроме того, поскольку все виды атак были искусственно созданы в разные дни и записаны в разные файлы, их совмещение не вполне отражает реальную картину, которая может возникнуть в атакуемой среде, так как могут применяться разные сценарии атак одновременно. Причем при таких атаках, как отказ в обслуживании, могут значительно изменяться параметры среды передачи, в следствии чего и значения соответствующих им признаков тоже изменятся. По той же причине нельзя применять модель, обученную на данном наборе данных и использовать ее в сети, где архитектура и технические возможности сильно отличаются.

Заключение

В работе были приведены данные о характерных признаках атак, которые играют роль в обнаружении отдельных компьютерных атак. Проведен анализ современных датасетов, используемых при обнаружении сетевых атак. На основе датасета CSE-CIC-IDS2018 была проведена оценка признаков. В ходе проведенной работы было выявлено, что лишь 14 признаков имеют важное значение при классификации компьютерных атак.

Литература

1. *Hardcastle J. L.* Блог компании SDxCentral, URL: https://www.sdxcentral.com/articles/news/worst-cyberattacks-of-2021-so-far/2021/12/ (дата обращения:18.01.2022).

2. *Наместников Ю.* Чему должен научить нас WannaCry Журнал "Information Security/ Информационная безопасность" № 3, 2017, стр. 4-5, URL: https://lib.itsec.ru/articles2/focus/chemu-dolzhen-nauchit-nas-wannacry (дата обращения: 20.12.2021).

3. Новая эпидемия шифровальщика Petya / NotPetya / ExPetr. Блог Kaspersky Lab, URL: https://www.kaspersky.ru/blog/new-ransomware-epidemics/17855/ (дата обращения: 20.12.2021).

4. *Лукацкий А.* Взлом Equifax: разбор полетов, URL: https://www.securitylab.ru/blog/personal/Business_without_danger/344730.php (дата обращения: 20.12.2021).

5. *Журавлев К.* Крупнейшие технологические неудачи 2021 года: рейтинг CNN. Газета.ru, URL: https://www.gazeta.ru/tech/2021/12/28_a_14367523.shtml?updated (дата обращения: 20.12.2021).

6. *Aljabri M., Aljameel S.S., Mohammad R.M.A., Almotiri S.H., Mirza S., Anis F.M., Aboulnour M., Alomari D.M., Alhamed D.H., Altamimi H.S.* Intelligent Techniques for Detecting Network Attacks: Review and Research Di-

rections. Sensors 2021, 21, 7070, URL: https://doi.org/10.3390/s21217070 (дата обращения: 30.12.2021).

7. *Gümüşbaş D., Yıldırım T., Genovese A., Scotti F.* A Comprehensive Survey of Databases and Deep Learning Methods for Cybersecurity and Intrusion Detection Systems, in IEEE Systems Journal, vol. 15, no. 2, 2021, pp. 1717-1731, URL: https://doi.org/10.1109/JSYST.2020.2992966 (дата обращения: 27.01.2022).

8. *Mishra P., Varadharajan V., Tupakula U., Pilli E. S.* A detailed investigation and analysis of using machine learning techniques for intrusion detection, IEEE Commun. Surveys Tuts., vol. 21, no. 1, 2019, pp. 686–728.

9. *Koch R.* Towards Next-Generation Intrusion Detection. 2011 3rd International Conference on Cyber Conflict, pp. 151–168.

10. *Thakkar A., Lohiya R.* A Review of the Advancement in Intrusion Detection Datasets, Procedia Computer Science, Volume 167, 2020, Pages 636-645, ISSN 1877-0509, URL: https://doi.org/10.1016/j.procs.2020.03.330 (дата обращения: 17.01.2022).

11. *Гетьман А.И., Горюнов М.Н., Мацкевич А.Г., Рыболовлев Д.А.* Методика сбора обучающего набора данных для модели обнаружения компьютерных атак. Труды ИСП РАН, том 33, вып. 5, 2021 г., стр. 83-104, URL: https://doi.org/10.15514/ISPRAS-2021-33(5)-5 (дата обращения: 17.01.2022).

12. *Sharafaldin I., Gharib A., Lashkari A.H., Ghorbani A.A.* Towards a reliable intrusion detection benchmark dataset. Software Networking, 2018(1), pp. 177–200, URL: https://doi.org/10.13052/jsn2445-9739.2017.009 (дата обращения: 17.01.2022).

13. *Ерохин С. Д., Журавлев А. П.* Сравнительный анализ открытых наборов данных для использования технологий искусственного интеллекта при решении задач информационной безопасности // Системы синхронизации, формирования и обработки сигналов. 2020. Т. 11. № 3. С. 12-19.

14. CSE-CIC-IDS2018 on AWS. A collaborative project between the Communications Security Establishment (CSE) & the Canadian Institute for Cybersecurity (CIC), URL: https://www.unb.ca/cic/datasets/ids-2018.html (дата обращения: 17.01.2022).

15. *Ravikumar D.* Towards Enhancement of Machine Learning Techniques Using CSE-CIC-IDS2018 Cybersecurity Dataset, 2021, Thesis, Rochester Institute of Technology.

16. *Borisenko B.B., Erokhin S.D., Fadeev A.S., Martishin I.D.* Intrusion detection using multilayer perceptron and neural networks with long short-term memory, Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2021 - Conference Proceedings, Svetlogorsk, Kaliningrad Region, 2021, URL: https://doi.org/10.1109/SYNCHROINFO51390.2021.9488416 (дата обращения: 18.01.2022).

17. *Erokhina O. V., Borisenko B. B., Martishin I. D., Fadeev A. S.* Analysis of the multilayer perceptron parameters impact on the quality of network attacks identification, Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2021 - Conference Proceedings, Svetlogorsk, Kaliningrad Region, 2021, URL: https://doi.org/10.1109/SYNCHROINFO51390.2021.9488344 (дата обращения: 18.01.2022).

18. *Erokhin S., Borisenko B., Fadeev A.* Reducing the dimension of input data for ids by using match analysis, 28th Conference of Open Innovations Association (FRUCT), Moscow, Russia, 2021, pp. 96-102, URL: https://doi.org/10.23919/FRUCT50888.2021.9347629 (дата обращения: 18.01.2022).

19. *Ерохин С.Д., Борисенко Б.Б., Мартишин И.Д., Фадеев А.С.* Анализ существующих методов снижения размерности входных данных // T-Comm: Телекоммуникации и транспорт. 2022. Т.16. №1. с. 30-37.

20. CICFlowMeter, URL: https://github.com/CanadianInstituteForCybersecurity/CICFlowMeter (дата обращения: 18.01.2022

СИСТЕМА МОБИЛЬНОЙ СВЯЗИ КАК ОБЪЕКТ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

Косичкина Татьяна Павловна,
МТУСИ, зав. кафедрой, к.т.н., доц., Москва, Россия
t.p.kosichkina@mtuci.ru

Косичкин Георгий Романович,
МТУСИ, студент магистратуры, Москва, Россия
gosh-kos@yandex.ru

Аннотация
В данной работе на примере системы мобильной связи рассмотрена последовательность действий по анализу угроз безопасности объекта критической информационной инфраструктуры. Подробно анализируются источники угроз. Акцентируется внимание на возможностях предупреждения угроз.

Ключевые слова: *Критическая информационная инфраструктура, система мобильной связи, модель угроз, информационная безопасность, виды атак.*

Введение

За последние два десятилетия достижения в области информационных технологий и инфокоммуникаций оказали значительное влияние на функционирование и методы обеспечения безопасности правительств, коммерческих предприятий, исследовательских институтов, а также учреждений обороны и безопасности. Информационные системы быстро внедряются в такие важные сферы общества, как транспорт, связь, здравоохранение, сельское хозяйство, финансы и другие секторы экономики. В России в последние годы предприняты стратегические шаги по улучшению информационной инфраструктуры. Быстрый доступ к государственным услугам через цифровые платформы, предоставление образования через системы дистанционного обучения, электронная медицина и коммерция – это далеко не полный перечень того, без чего уже сложно представить нашу жизнь.

В связи с этим актуальной стала задача исследования объектов критической информационной инфраструктуры (КИИ). При этом под критической информационной инфраструктурой понимается такая инфраструктура, разрушение которой может привести к серьезным экономическим или социальным последствиям [1]. В этой трактовке к объектам КИИ в первую очередь относят объекты, связанные с производством и распределением электроэнергии, добычей нефти и газа, водоснабжением и водоотведением и т.п. Перечисленные объекты в значительной степени зависят от информационных технологий. Однако и компьютерная сеть, и информационно-коммуникационные сети сами по себе могут составлять критически важную инфраструктуру.

Из-за сложной природы систем мобильной связи большая часть усилий по обеспечению безопасности в этой области распределена по различным технологиям. Это привело к неясному представлению об общей безопасности систем мобильной связи. Попробуем проанализировать эту проблему, в том числе с помощью выявления угроз для данной предметной области.

Категорирование объектов критической информационной инфраструктуры

В некоторых источниках дается четкое разграничение между критической информационной инфраструктурой и информационной инфраструктурой. Информационные инфраструктуры – это технические, социальные и политические структуры, охватывающие людей, технологии, инструменты и услуги, используемые для облегчения распределенного совместного использования контента во времени и на расстоянии.

Критическая информационная инфраструктура представляет собой информационные и коммуникационные системы, техническое обслуживание, надежность и безопасность которых необходимы для безопасного функционирования предприятий, а в ряде случаев – для безопасности страны в целом.

Важный процесс для понимания, является ли организация субъектом КИИ – категорирование. Как следует из методических рекомендаций по категорированию объектов КИИ, принадлежащих субъектам КИИ, функционирующим в сфере связи [2], поскольку системы мобильной связи функционируют в одной из перечисленных в статье 2 Закона «О безопасности КИИ» сфер (связь), они однозначно являются субъектами КИИ. Формирование перечня объектов КИИ осуществляется оператором связи с использованием перечня типовых объектов. Этот этап достаточно подробно рассмотрен в [2].

Далее, согласно приказу ФСТЭК [3], должен быть проведен анализ угроз безопасности, который включает в себя:

– выявление источников угроз безопасности информации и оценку возможностей (потенциала) внешних и внутренних нарушителей;

– анализ возможных уязвимостей значимого объекта и его программных, программно-аппаратных средств;

– определение возможных способов (сценариев) реализации (возникновения) угроз безопасности информации;

– оценку возможных последствий от реализации (возникновения) угроз безопасности информации.

Поскольку данные вопросы не так подробно рассматриваются в [2], следует уделить им отдельное внимание.

Структура угроз в системах мобильной связи

Вопросам безопасности в сетях мобильной связи всегда уделялось много внимания. Угрозы, приводящие к отказам в обслуживании с использованием радиопомех всегда были проблемой радиосетей. В данной работе ограничимся рассмотрением архитектур и вопросов безопасности систем мобильной связи последних двух поколений (4G и 5G).

Цель разработки системы безопасности для LTE, или так называемой Evolved Packet System (EPS), состояла в том, чтобы поддерживать обратную совместимость и смягчать слабые места безопасности предыдущих поколений путем введения расширений в архитектуру безопасности EPS (рис. 1).

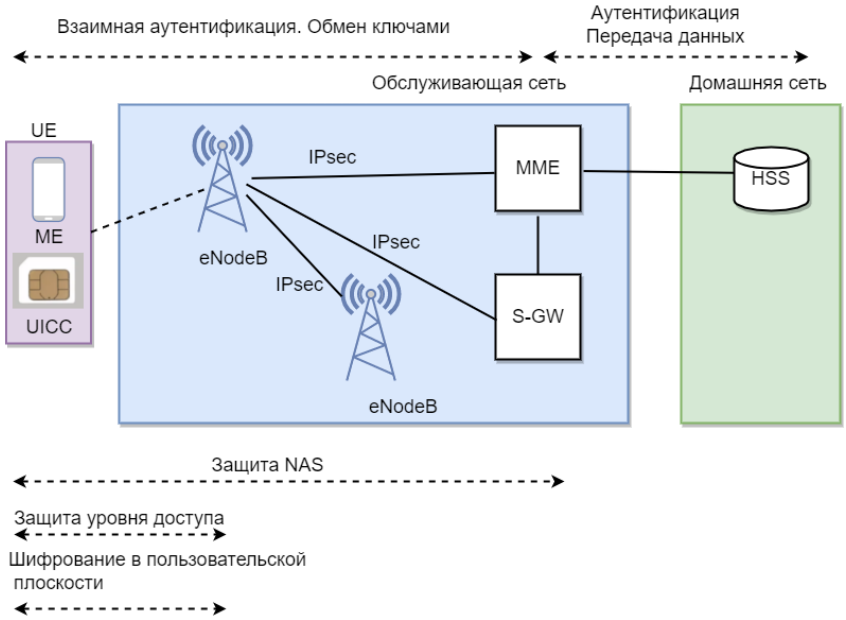

Рис. 1. Архитектура безопасности EPS

Шифрование данных сигнализации и защита целостности включают в себя защиту уровня без доступа (Non-Access Stratum, NAS) между пользовательским оборудованием (User Equipment, UE) и узлом управления мобильностью (Mobile Management Entity, MME) и уровня доступа (Access Stratum, AS) между UE и базовыми станциями (BS). Однако пользовательская плоскость, которая передает трафик данных, имеет только шифрование и не имеет защиты целостности. Кроме того, шифрование является необязательным как для сигнализации, так и для пользовательских данных между UE и BS.

Даже с учетом усовершенствований существуют некоторые угрозы EPS. Поскольку стандарт LTE должен поддерживать обратную совместимость, злоумышленник может использовать атаки с понижением версии, чтобы заставить UE подключиться к менее защищенной сети, в этом случае для атаки могут быть использованы протоколы предыдущих поколений (например, GTP, SS7, SIP).

Кроме того, поскольку LTE − это система на основе IP, сфера и методы атак хорошо известны злоумышленникам, к примеру, атаки на DNS-серверы.

В качестве усовершенствования по сравнению с SS7 протокол Diameter, включенный в стандарты 4G, представляет собой IP-протокол прикладного уровня, созданный на основе службы удаленной аутентификации пользователей с коммутируемым доступом (RADIUS) для обеспечения аутентификации, авторизации и учета. Однако протокол Diameter сохранил многие концепции SS7, унаследовав его уязвимые места. Более того, поскольку Diameter − это протокол на основе IP, он более доступен для злоумышленников, чем SS7. Diameter всегда отвечает на запрос, используя тот же путь, по которому был получен запрос, что позволяет злоумышленникам получать ответы на инициированные ими операции.

Другими уязвимостями являются отсутствие сквозной аутентификации, проверки целостности и шифрования, что делает невозможным проверку фактического отправителя и целостности сообщений Diameter.

Что касается мобильных сетей 5G, одной из особенностей их построения является мультивендорность, которая приводит к появлению большого количества программного обеспечения, построенного на базе открытых исходных кодов. Это дает больше шансов злоумышленникам для того, чтобы выявить уязвимости сети.

Кроме того, в сетях 5G, планируется подключение большого числа устройств Интернета вещей, которые обладают ограниченными криптографическими возможностями. Это может привести к появлению новых видов атак.

Архитектура безопасности 5G приведена на рисунке 2. На нем цифрами обозначены: 1 – безопасность доступа к сети; 2 – безопасность сетевого домена; 3 – безопасность пользовательского домена; 4 – безопасность домена приложения; 5 – безопасность домена сервисной архитектуры (SBA) [4].

Поскольку архитектуре 5G свойственна децентрализация, то есть расширение границ сети на периферийные узлы, которые могут выполнять обработку запросов и маршрутизацию пользовательского трафика, это предоставляет злоумышленникам дополнительные возможности атаки на локальные устройства. Однако в настоящее время продолжается работа над концепцией безопасности сетей 5G, затрагивающей разные стороны этого вопроса.

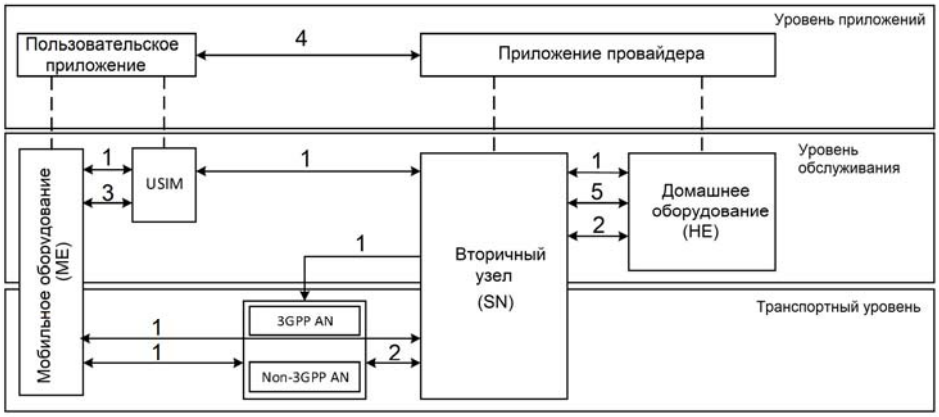

Рис. 2. Архитектура безопасности 5G

Следующим шагом в понимании защиты системы как объекта КИИ является определение потенциальных угроз. Специфика здесь заключается в том, чтобы среди возможных угроз выбрать те, которые могут повлечь за собой катастрофические последствия.

Среди возможных типов злоумышленников, например, в [5] отмечаются:

- взломщики радиоинтерфейса;
- мобильные операторы;

- инсайдеры;
- производители оборудования и SIM-карт;
- поставщики программного обеспечения (ПО) и операционных систем (ОС);
- мобильные пользователи со злыми намерениями.

Рассматривая систему мобильной связи как объект КИИ, следует особо выделить инсайдеров, производителей оборудования и поставщиков программного обеспечения.

Небрежность и отсутствие образования в области безопасности могут сделать инсайдеров легковерными, например, в отношении социальной инженерии, неправильной конфигурации узлов или игнорирования операционной безопасности. В любом случае инсайдеры - люди с прямым доступом к критической инфраструктуре являются самыми опасными злоумышленниками.

Ошибки на аппаратном уровне сложно отследить, однако они представляют реальные угрозы, создавая, например, каналы утечки информации, как было в случае с компанией Amazon. В заказанном в Китае оборудовании компанией были обнаружены микросхемы, не являвшиеся частью оригинального дизайна. Эти микросхемы при дальнейших исследованиях были признаны шпионскими [6].

Как уже отмечалось, современные мобильные сети включают в себя большое количество программных компонентов с открытым исходным кодом. Как и в случае с аппаратным обеспечением, цепочка поставок программного обеспечения также может содержать преднамеренные или случайные уязвимости. Из-за своей сложной и закрытой природы проникновение в основную сеть требует глубоких знаний и навыков, в то время как с помощью общедоступных криминалистических инструментов злоумышленники могут использовать распространенные уязвимости в стеке программного обеспечения маршрутизаторы и другие сетевые устройства.

В стратегии национальной безопасности России отдельное внимание уделяется экономической безопасности с точки зрения цепочки поставок и самодостаточности. Это означает, что Россия должна осуществлять рациональное импортозамещение, а также снижение критической зависимости от иностранных технологий и промышленной продукции. Однако, как стало известно 14 мая 2021 года, переход на отечественное оборудование и ПО на объектах КИИ, планировавшийся к 1 января 2023 года придется отложить. Министерство экономики России считает, что такие сжатые сроки приведут к дополнительным необоснованным расходам бюджета и бизнеса при импортозамещении в КИИ.

Следует особо отметить, что если в отношении операционных систем и программного обеспечения российские разработки имеют какую-то перспективу, то в отношении аппаратного обеспечения сложно прогнозировать скорый прогресс. Известные попытки создать российские процессоры, мобильные телефоны и. т. п. несколько раз проваливались в основном ввиду отсутствия в России собственных производств.

Далее рассмотрим возможные виды атак. С точки зрения проникновения в сеть мобильной связи можно выделить несколько видов атак (см., например, [5]): атаки с целью доступа в сеть; атаки с целью поддержания постоянного доступа в течение длительного периода; атаки с целью сбора информации; уклонение от систем защиты. С технической точки зрения интерес представляет анализ точек входа в сети. Здесь возможны следующие виды атак:

- атаки со стороны пользовательского оборудования (UE), включающие использование программного или аппаратного обеспечения UE для отправки вредоносного трафика в мобильную сеть;
- атаки на основе SIM-карты, которые включают компрометацию или получение любых физических смарт-карт;
- атаки из сети радиодоступа, когда злоумышленник может выдать себя за мобильную сеть, установив мошенническую базовую станцию;
- атаки из других мобильных сетей, когда злоумышленник, имеющий доступ или получающий доступ к другим мобильным сетям, может запускать атаки на целевую мобильную сеть;
- атаки с физическим доступом к транспортной сети, когда злоумышленник может запустить атаку в опорной сети;
- атаки из сети на основе IP, которые включают отправку вредоносного трафика в сеть оператора через сеть услуг и приложений или связь по другому протоколу на основе IP;
- внутренние атаки и человеческие ошибки, которые включают преднамеренную атаку или непреднамеренные ошибки, совершенные кем-либо, имеющим доступ к любому элементу мобильной сети.

Имея представление о возможных источниках угроз и возможных уязвимостей значимого объекта и его программных/программно-аппаратных средств следует перейти к определению возможных сценариев возникновения угроз, а также к оценке возможных последствий от их возникновения. Рассмотрение возможных сценариев возникновения угроз представляет отдельную и достаточно интересную научную задачу, она может быть смоделирована, например, методом теории графов [7] или с помощью вероятностного моделирования.

Возникновение ущерба или его предупреждение?

В [2] приведена исчерпывающая методика оценки ущерба, которая позволяет присвоить субъекту КИИ определенную категорию. В ней расчет ущерба производится исходя из количества абонентов, зоны обслуживания и налогов оператора.

В частности, зная допустимый период простоя (Δt_{MAO} , календарных дней), закрепленный в локальных актах оператора связи, а также налог на прибыль организации n_f можно вычислить значение потенциально возможного ущерба бюджетам Российской Федерации, исходя из худшего сценария:

$$U_p = n_f \Delta t_{MAO} / 365, тыс.руб$$

Вычисленное значение сравнивается с размером прогнозируемого годового дохода бюджета для каждого конкретного объекта КИИ с учетом его зоны обслуживания и количества абонентов. Общее количество абонентов определяется на основании количества заключенных договоров на момент проведения категорирования объектов КИИ, принадлежащих оператору связи.

Последние годы свидетельствуют о все большем внимании, которое в России уделяется оборонительной кибербезопасности. Уязвимости КИИ понимаются как системные угрозы из-за горизонтальной важности сектора критической инфраструктуры в целом. Однако в течение последних лет в публикациях на тему КИИ происходит смещение акцента с защиты КИ на устойчивость [7]. Полная защита никогда не может быть гарантирована, и достижение желаемого гарантированного уровня защиты, как правило, экономически невыгодно по сравнению с реальными угрозами. Поэтому в последнее время все большее внимание уделяется адаптивным мерам и быстрому восстановлению.

Двумя понятиями, которые используются в документе [8] и впоследствии повторяются во многих постановлениях и других нормативных актах, являются «предупреждение» и «профилактика». В стратегическом документе по КИИ от 2011 г. [9] основными понятиями, касающимися оборонительных мероприятий, являются «минимизация» рисков и повышение уровня «защиты». Стратегия национальной безопасности [10] предусматривает противодействие угрозам КИИ за счет совершенствования и развития единой государственной системы «предупреждения» и «ликвидации» чрезвычайных ситуаций. То же самое касается большинства официальных стратегических документов и законодательства.

С точки зрения восстановления системы мобильной связи обладают большим потенциалом благодаря такому встроенному функционалу, как самоорганизация (SON).

Изначально SON создавались для экономии капитальных и эксплуатационных затрат. Самоорганизация систем мобильной связи включает в себя три уровня: самоконфигурация; самооптимизация; самовосстановление. В контексте КИИ интерес представляет, в первую очередь, самовосстановление. Оно состоит из нескольких функций, таких как: самовосстановление после программных сбоев сетевого элемента, самовосстановление неисправностей, полное отключение сотовой связи. Функция самовосстановления анализирует и отслеживает аварийные сигналы, уведомления и результаты самотестирования. Корректирующие действия с затронутым сбоем сетевым элементом запускаются автоматически. Поскольку цель применения SON в том, чтобы уменьшить вмешательство человека в сеть, ее наличие позволяет уменьшить риски, связанные с инсайдерскими действиями.

Кроме того, большой потенциал заложен в системах шестого поколения, в которых изначально заложена привязка мобильной связи к критической информационной инфраструктуре [11]. Концепция Network2030 предлагает новые возможности для обеспечения безопасного спасения людей. субъектов в любом месте в любое время в случае любого рода чрезвычайной ситуации. Выявление этих возможностей также является одной из основных целей Network 2030. Критически важные операции по обеспечению безопасности должны учитывать все особенности граждан, находящихся в зоне чрезвычайной ситуации. Например, предусматривается возможность знать местоположение пострадавшего до тех пор, пока его не спасут.

Это будет происходить с помощью его пользовательского оборудования со ссылкой на карту местности, доступ к которой осуществляется с помощью возможности навигации по безопасному пути. Данные продекларированные возможности не затрагивают поднятый вопрос об отнесении систем мобильной связи к объектам КИИ. Они лишь вселяют надежду на то, что в шестом поколении вопросы безопасности самих сетей будут проработаны не менее тщательно.

Заключение

Кибербезопасность в последние годы стала основным объектом интереса в России, вокруг которой была создана система регулирования с ее разнообразием законов и институтов.

Большинство критически важных инфраструктур либо основаны на уязвимых системах ИКТ, либо контролируются ими, что делает информационную инфраструктуру центром политики защиты КИ.

Система мобильной связи сложна, ее сеть работает на основе доверия между партнерами. Растущее использование IP-протоколов в системе увеличило вероятность атак. Существуют отдельные работы по вопросам обеспечения безопасности систем мобильной связи, однако в настоящее время отсутствует общая концептуальная основа для обзора состояния безопасности всей системы. Для дальнейших исследований в данной области необходим сбор информации об угрозах, а также обмен этой информацией между операторами и заинтересованными лицами.

Уязвимости представляют угрозу безопасности для систем, которые считаются критическими, и поэтому организациям рекомендуется вкладывать средства в инструменты оценки уязвимостей. Это поможет им обнаруживать, устранять, отслеживать и оценивать уязвимости в КИИ.

Литература

1. *Ерохин С.Д., Петухов А.Н., Пилюгин П.Л.* Управление безопасностью критических информационных инфраструктур. – М.: Горячая линия - Телеком, 2021. – 240с.

2. Методические рекомендации по категорированию объектов критической информационной инфраструктуры, принадлежащих субъектам критической информационной инфраструктуры, функционирующим в сфере связи. https://www.rans.ru/activity/news/316-2-jl-2019-goda (дата обращения 12.01.2022).

3. Приказ ФСТЭК России от 26 марта 2019 г. № 60.

4. ETSI TS 133 501 V15.6.0 (2019-10) 5G; Security architecture and procedures for 5G System (3GPP TS 33.501 version 15.6.0 Release 15)

5. *Rao, Siddharth Prakash & Holtmanns, Silke & Aura, Tuomas.* (2020). Threat modeling framework for mobile communication systems.

6. *Jordan Robertson and Michael Riley*. 2018. .e big hack: how China used a tiny chip to infiltrate US companies. Bloomberg Businessweek 4 (2018).

7. *Ерохин С.Д., Петухов А.Н., Пилюгин П.Л.* Принципы и задачи асимптотического управления безопасностью критических информационных инфраструктур // T-Comm: Телекоммуникации и транспорт. 2019. Том 13. №12. С. 29-35.

8. Методика отнесения объектов государственной и негосударственной собственности и критически важных объектов для национальной безопасности Российской Федерации. Утверждено МЧС 17.10.2012, г. Москва. Отменено постановлением МЧС России от 30 декабря 2019 г., N 43–7134–11.

9. Основы государственной политики в области обеспечения безопасности населения Российской Федерации и защищенности критически важных и потенциально опасных объектов от угроз природного, техногенного характера и террористических действий на период до 2020 года. Утверждено Президентом РФ 15 ноября 2011 г. № Пр-3400. Далее: Указ Президента РФ от 17.02.2017 № Пр-3400.

10. Указ Президента Российской Федерации от 31 декабря 2015 года N683 "О стратегии национальной" безопасности Российской Федерации».

11. *Li, Richard*. (2019). Network 2030 A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond.

РАЗВИТИЕ ИНТЕРВАЛЬНОГО МЕТОДА АНАЛИЗА ОЧЕРЕДЕЙ В СИСТЕМАХ МАССОВОГО ОБСЛУЖИВАНИЯ

Лихтциндер Борис Яковлевич,
ПГУТИ, профессор кафедры ССС, д.т.н, профессор. г. Самара, Россия
lixt@psuti.ru

Аннотация

В статье рассматриваются особенности трафика мультисервисных сетей связи, его обработки в системах массового обслуживания (СМО) и интервальные методы анализа трафика. В качестве анализируемой случайной величины рассматриваются количества заявок, поступающих в систему за постоянные промежутки времени, равные математическому ожиданию времен обработки заявок. Приведена обобщенная формула Хинчина-Поллачека для средних значений очередей, возникающих при обработке потоков заявок общего вида. Обобщенная формула устанавливает зависимость средних значений очередей от дисперсии и корреляционных свойств указанной случайной величины, при различных значениях коэффициента загрузки системы. В качестве примера потока, для которого корреляционная составляющая отсутствует, приведены групповые пуассоновские потоки. Рассмотрена формула для групповых пуассоновских потоков заявок. Для таких потоков, среднее значение очереди, при заданной загрузке системы, полностью определяется дисперсией чисел заявок на интервалах их обслуживания.

Показано, что процесс обработки заявок цикличен, и на последнем интервале активной части каждого цикла всегда отсутствует очередь. Для стационарного потока заявок с конечным математическим ожиданием и дисперсией приведены формулы, из которых следует, что при размещении поступающих заявок на интервалах времени удаленных от начала каждого цикла, их вклад в среднее значение очереди уменьшается. Очередь достигает своего максимального значения, если все заявки сосредоточены в началах циклов. Развитием интервального метода анализа очередей стало определение количества заявок в промежутках времени между соседними обслуженными заявками, покидающими очередь. Полученные соотношения являются фундаментальными, так как показывают, что при выполнении условий стационарности, эргодичности и сходимости ковариационной функции, условные средние значения очередей в одноканальной СМО полностью определяются дисперсией и суммой ковариаций количества заявок, поступивших в промежутках времени между соседними обслуженными заявками.

Ключевые слова*: системы массового обслуживания, потоки, заявки, очереди, интервальные методы, цикличность, случайные величины, ковариации.*

Введение

В мультисервисных сетях (МСС) с пакетной коммутацией поток пакетов существенно отличается от пуассоновского, Любой пакетный трафик является продуктом компьютерной обработки, выполняемой процессором при решении задач приложений. [1, 15-17], при этом, поток пакетов имеет явно выраженный пачечный характер. В большинстве случаев, такой поток характеризуется функцией распределения временных интервалов между соседними пакетами. Имеется также множество работ, в которых потоки в системах массового обслуживания характеризуются функцией распределения числа пакетов, поступающих за условную единицу времени.

1.Интервальный метод

В одном из предложенных нами методах анализа трафика в качестве указанной единицы времени, рассматривался постоянный (средний) интервал времени τ обработки одного пакета (заявки) [2, 3].

На рисунке 1, в верхней части, показаны заявки $m_i(\tau)$, поступающие в течение одинаковых интервалов времени τ, равных среднему времени обработки одной заявки. Ниже, показаны размеры возникающих очередей $q_i(\tau)$.

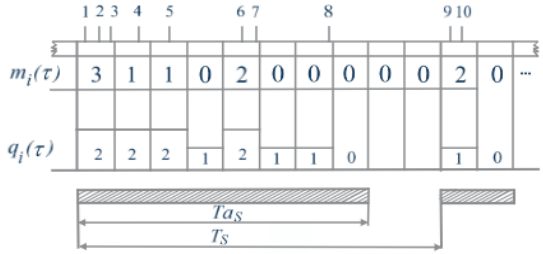

Рис. 1. Фрагмент потока заявок в СМО

Интервал времени обработки τ не может превышать значения $1/\lambda$, где λ – средняя интенсивность поступления заявок в потоке. Это равносильно тому, что значение суммарного коэффициента загрузки $\rho = \lambda\tau$ не должно превышать единицу. Именно в указанных пределах рассматривается изменение интервала τ.

Процесс обработки заявок в такой СМО всегда состоит из последовательности чередующихся периодов занятости обслуживающего прибора (прибор обрабатывает заявки) и периодов простоя, в течение которых заявки в обслуживающем приборе отсутствуют. На Рисунке 1 показан фрагмент, соответствующий S -му циклу Z_S с длительностью цикла T_S и периодом занятости T_{aS}. Следует отметить, что очередь на последнем интервале периода занятости (активной части цикла) всегда отсутствует. Однако, отсутствие очереди на активном интервале цикла еще не свидетельствует, о том, что этот интервал – последний. Для стационарного потока заявок $m_i(\tau)$ с конечным математическим ожиданием и дисперсией $D_m(\tau)$ нами были получены [4] формулы (1).

$$\overline{q(\tau)} = \lim_{N\to\infty}\frac{1}{N}\sum_{Z_s\subset[1,N]}\sum_{j=0}^{j_{s+1}-j_s}(N_s - j)(m_{j_s+j} - 1) \text{ и } \overline{q(\tau)} = \lim_{N\to\infty}\frac{1}{N}\sum_{Z_s\subset[1,N]}\sum_{j=0}^{j_{s+1}-j_s} j(1 - m_{j_s+j}). \quad (1)$$

Здесь, N -общее число заявок, $N_s = j_{s+1} - j_s + 1$ - число заявок в цикле Z_S, j_s - номер первой заявки цикла Z_S, а j - порядковый номер заявки, считая от начала цикла. $\sum_{Z_S\subset[1,N]}$ -означает суммирование по всем циклам. Из первой формулы непосредственно следует, что с увеличением числа j, т.е. удалением поступающих заявок от начала цикла, их вклад в среднее значение очереди уменьшается. Максимального размера средняя очередь в течение цикла Z_S достигает в случае, когда все N_S заявок сосредоточены в начале цикла, на первом интервале обработки:

$$\overline{q_{S\max}(\tau)} = \frac{N_s(N_s - 1)}{2}.$$

Применение постоянного интервала обслуживания τ позволило получить обобщение формулы Хинчина – Поллачека [5]-[11].

$$\overline{q(\tau)} = \frac{D_m(\tau) + 2\mu_{q_{i-1}m_i}(\tau)}{2(1-\rho)} - \frac{\rho}{2}, \quad (2)$$

где $D_m(\tau)$ - дисперсия $m_i(\tau)$, а $\mu_{q_{i-1}m_i}(\tau)$ второй взаимный центральный корреляционный момент последовательностей $q_{i-1}(\tau)$ и $m_i(\tau)$, называемый ковариацией.

Соотношение (2) справедливо для стационарных потоков заявок весьма общего вида, при постоянном времени обслуживания τ. Оно учитывает корреляционные свойства потока заявок. В приве-

денных выше публикациях нами показано, что для пачечных потоков, корреляционная составляющая вносит основной вклад в средние значения возникающих очередей.

Для пуассоновских потоков $D_m(\tau) = \lambda\tau = \rho$, а корреляция между заявками отсутствует, и, в соответствии с (2), мы приходим к формуле Хинчина Поллачека в ее обычном виде:

$$\overline{q(\tau)} = \frac{\rho^2}{2(1-\rho)}. \quad (3)$$

Корреляционная составляющая также отсутствует в групповых пуассоновских потоках.

Это пуассоновские потоки независимых событий с параметром λ. Каждое событие заключается в одновременном появлении в момент t_k «пачки» из μ_k независимых случайно распределенных чисел заявок. Нами показано [14], что дисперсия чисел заявок на интервалах обслуживания таких потоков линейно зависит от коэффициента загрузки ρ:

$$D_m(\rho) = \rho\bar{k}(1+\nu_k^2),$$

где $\bar{k}$ – среднее число заявок в «пачке». а ν_k^2 – квадрат коэффициента вариации чисел заявок в пачках. С учетом сказанного, формула (2) существенно упростится:

$$\overline{q(\rho)} = \frac{\rho\bar{k}(1+\nu_k^2)}{2(1-\rho)} - \frac{\rho}{2}.$$

Для обычного пуассоновского потока $\bar{k} = 1$, $\nu_k^2 = 0$ и вновь. получаем формулу (3).

2. Случайные интервалы времени обработки

Если интервалы времени обслуживания заявок не постоянны, то значения чисел заявок на указанных интервалах определяются последовательным расположением реальных интервалов обработки, как это показано на рисунке 2.

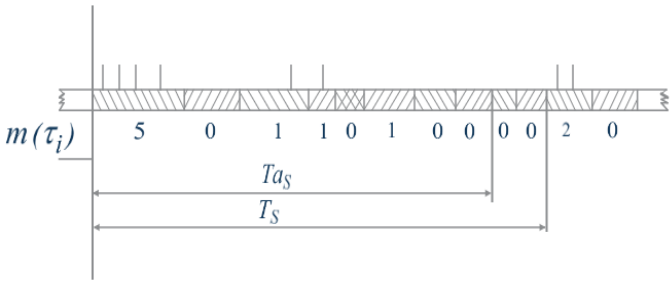

Рис. 2. Последовательное расположение интервалов обработки

На пассивных участках цикла, где отсутствует обработка заявок, размещаются случайно выбранные интервалы, а числа обрабатываемых на них заявок принимаются равными нулю. Очевидно, что отношение числа интервалов активной обработки заявок к общему числу интервалов представляет собой коэффициент загрузки ρ. Для полученного таким образом ряда значений $m(\tau_i)$ справедлива приведенная ранее формула (2).

3. Интервалы времени между соседними обслуженными заявками

Развитием интервального метода анализа очередей в СМО явилось определение чисел заявок на интервалах времени между соседними обслуженными заявками, покидающими очередь. Алгоритм такого определения показан на рисунке 3.

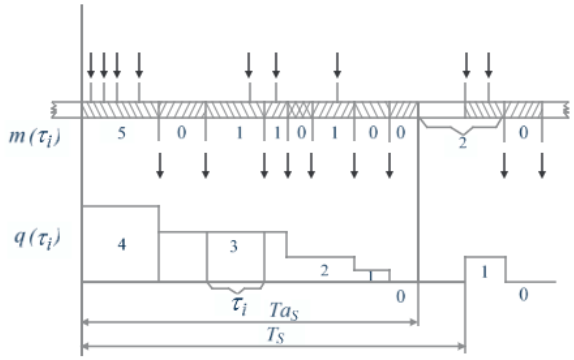

Рис. 3. Поступившие заявки на интервалах времени между соседними обслуженными заявками

На верхней части рисунка стрелками показаны поступающие и покидающие очередь заявки.

Пассивная часть каждого цикла объединяется с начальным интервалом обработки последующего цикла, образуя общий интервал времени между двумя соседними ушедшими из очереди заявками.

На рисунке 3 на указанный интервал поступили две заявки. Общее число полученных интервалов всегда равно общему числу поступивших заявок.

Произведем замену переменной интервалов τ_i на которых рассматривается поступление заявок, на другую переменную – θi, которая представляет собой интервал между двумя соседними заявками, покидающими очередь. Число заявок, поступивших на указанных интервалах, обозначим через $mi(\theta)$. Это и есть новый рассматриваемый случайный процесс.

Полученный таким образом ряд значений $m_i(\theta)$ полностью определяет процесс формирования очередей, и исключает из рассмотрения пассивные интервалы простоя, на которых очереди всегда отсутствуют. Нами показано [12], что средние значения очередей определяются соотношением:

$\overline{q(\rho)} = \frac{\rho}{2}(D_m(\theta) + 2\sum_{k=1}^{M-1}\frac{M-k}{M}\nu_m(k,\theta,M)$ где M – число членов соответствующей реализации процесса $m_i(\theta)$, $D_m(\theta)$ -дисперсия указанной реализации, а $\nu_m(k,\theta,M) = \frac{1}{M-k}\sum_{i=k+1}^{M} m_i(\theta)(m_{i-k}(\theta) - \overline{m_i(\theta)})$ – безусловная выборочная ковариация указанного процесса. Следует отметить, что здесь $\overline{m_i(\theta)} = 1$.

Если теоретическая ковариация $\nu_m(k,\theta) = \lim_{M\longrightarrow\infty}\nu_m(k,\theta,M)$ достаточно быстро убывает так, что ряд $\sum_{k=1}^{\infty}\nu_m(k,\theta)$ достаточно быстро сходится, то выражение для среднего значения очереди (3) примет вид:

$$\overline{q(\rho)} = \frac{\rho}{2}(D_m(\theta) + 2\sum_{k=1}^{\infty}\nu_m(k,\theta)) \quad . \tag{4}$$

В работе [13] нами было рассмотрено понятие условного среднего значения очереди

$$\overline{Q(\rho)} = \frac{\overline{q(\rho)}}{\rho} \quad ,$$

которое определяет среднее значение очереди только на активных участках циклов и не учитывает наличие интервалов простоя. С учетом (4), получим:

$$\overline{Q(\rho)} = D_m(\theta) + 2\sum_{k=1}^{\infty}\nu_m(k,\theta) \quad . \tag{5}$$

Условное среднее значение очереди в одноканальной СМО полностью определяется дисперсией и суммой ковариаций рассмотренного выше процесса.

Заключение

В большинстве случаев, при анализе очередей в СМО применяются две случайные величины: это интервалы времен между соседними заявками и интервалы времен обслуживания заявок. При интервальных методах обе указанных случайных величины заменяются одной, что существенно облегчает анализ возникающих очередей и позволяет получить результаты для СМО с потоками весьма общего вида.

Соотношение (5) имеет теоретическое значение и носит фундаментальный характер, ибо оно показывает, что условное среднее значение очереди в одноканальной СМО, при выполнении условий стационарности, эргодичности, и сходимости ковариационной функции, полностью определяется дисперсией и суммой ковариаций чисел заявок на интервалах времени между соседними заявками, покидающими очередь.

Литература

1. *Степанов С.Н.* Теория телетрафика. Концепции, модели, приложения. М.: Горячая линия-Телеком. 2015. 808 с.
2. *Лихтциндер Б.Я.* Интервальный метод анализа трафика мультисервисных сетей доступа. ПГУТИ. Самара, 2015. 121 с.
3. *Лихтциндер Б.Я.* Интервальный метод анализа мультисервисного трафика сетей доступа Электросвязь. №12. 2015. С. 52–54.
4. *Blatov I.A. Likhttsinder B. Ja,* On Estimate of Queue Lengths in QS with Certain Correlative Correspondence. Journal of Physics: Conference Series. 2018. Vol. 1096. P. 1-8. doi:10.1088/1742-6596/1096/1/012173.
5. *Лихтциндер Б. Я., Макаров И.С.* Алгоритм определения средней длины очереди СМО через обобщенную формулу Хинчина-Поллачека // Вестник самарского государственного технического университета. №1. 2012. С. 41-45.
6. *Лихтциндер Б.Я.* Интервальный метод анализа трафика мультисервисных сетей // Модели инфокоммуникационных систем: разработка и применение. Приложение к журналу ИКТ. Вып. 8, 2011. С. 101-152.
7. *Лихтциндер Б.Я.* Интервальный метод анализа трафика мультисервисных сетей доступа. Самара: ПГУТИ, 2015. 121с.
8. *Лихтциндер Б.Я.* Корреляционные свойства длин очередей в системах массового обслуживания с потоками общего вида // ИКТ. Т.13. №3. 2015. С. 276-280.
9. *Лихтциндер Б. Я.* О некоторых обобщениях формулы Хинчина-Поллачека // ИКТ. Т.5, №4. 2007. С. 253-258.
10. *Лихтциндер Б.Я.* Корреляционные связи в пачечных потоках систем массового обслуживания // Телекоммуникации № 9. 2015. С. 8-12.
11. *Лихтциндер Б.Я.* Интервальный метод анализа мультисервисного трафика сетей доступа // Электросвязь. №12. 2015. С. 52-54.
12. *Лихтциндер Б.Я.,* Блатов И.А., Китаева Е.В. Средняя длина очереди для систем массового обслуживания с коррелированными входными потоками // T-Comm: Телекоммуникации и транспорт. 2020. Том 14. №8. С. 13-20.
13. *Лихтциндер Б. Я.* Трафик мультисервисных сетей доступа (интервальный анализ и проектирование). – М.: Горячая линия – Телеком, 2018. 200 с
14. *Лихтциндер Б. Я.* Интервальные характеристики групповых пуассоновских моделей трафика телекоммуникационных систем. Инфокоммуникационные технологии. 2020. Т. 18. №3. С 302-311.
15. *Лихтциндер Б.Я.* Интервальный метод анализа очередей в системах массового обслуживания с пачечными потоками заявок // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 3. С. 17-23.
16. *Лихтциндер Б.Я., Блатов И.А.* Мощность обработки потоков заявок и размеры очередей в системах массового обслуживания // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 9. С. 10-16.
17. *Likhttsinder B.Ya., Bakai Yu.O.* Delays in queues of queuing systems with stationary requests flows // T-Comm. 2021. Т. 15. № 2. С. 54-58.

РАЗРАБОТКА СИСТЕМЫ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ ДЛЯ ЛОКАЛЬНОЙ СЕТИ ПРЕДПРИЯТИЯ

Поликарпова Анастасия Борисовна,
Московский технический университет связи и информатики, магистр, Москва, Россия
p-ab.97@mail.ru

Воронова Лилия Ивановна,
Московский технический университет связи и информатики, зав. каф. ИСУиА, д.ф.-м.н., профессор, Москва, Россия
voronova.lilia@ya.ru

Аннотация

Время повсеместной цифровизации в России не могло не затронуть малый бизнес, развивающиеся компании используют для своей работы локальные сети, но несмотря на их активное применение, нередко без внимания остаются вопросы, связанные с информационной безопасностью. В рамках данной статьи будет разработана криптографическая система защиты информации для локальной сети организации, которая имеет доступ к конфиденциальным данным и коммерческую тайну, которые должны быть надежно защищены.

Ключевые слова: *Информационная безопасность, локальная сеть, конфиденциальность данных, криптографическая защиты информации, целостность информации.*

Введение

Методы решения задач обеспечения безопасности очень тесно связаны с уровнем развития науки и техники и, особенно, с уровнем технологического обеспечения. А характерной тенденцией развития современных технологий является процесс тотальной интеграции.

По направлению, связанному с разработкой систем криптографической защиты информации [14-18] для локальной сети предприятия ведутся обширные исследования. Например, в [8] происходит анализ блочного шифра MARS с симметричным ключом, который был модернизирован метаморфической функцией.

Авторы [9] рассматривают создание нового ультра-легкого блочного шифра с целью поддержания конкуренции с существующими ведущими компактными потоковыми шифрами.

В работе [10] рассматривается разработка оригинального протокола шифрования в мессенджере Telegram.

Анализ нового семейства блочного шифра Камелия, с помощью которого можно проводить улучшенные атаки с уменьшенным количеством раунда проводится в работе [11]. В работе [12] исследуется безопасность of Neural Network Architectures for Cardio-Analysis in the Driver Support System, а в работе [13] строится безопасная сеть путем клстеризации of Wireless Sensor Network.

Основные тенденции организации информационной безопасности на предприятии

Был проведен анализ ряда кейсов, показывающих, как система криптографической защиты информации внедряется в российские компаниях. Данный вопрос решается одним из двух способов: собственными силами; с привлечением подрядчиков.

В России существует тенденция по переманиванию ведущих специалистов по кибербезопасности из компаний, специализирующихся на борьбе с киберугрозами, в крупные российские корпорации с целью выстраивания собственной модели ИБ, избегая рисков утечки данных, ожидаемых при работе с подрядчиками.

С точки зрения выбора программных средств для защиты информационной безопасности государственная политика поддержки национального производителя побуждает российские корпорации с госучастием или активно работающих с госконтрактами переходить на российское ПО. Небольшой бизнес предпочитает его по причинам доступности в цене, более высокой надежности, обеспечения информационной безопасности в узких секторах ИС.

Правовое обеспечение системы в области информационной безопасности и защиты информации

Подбор нормативных актов по защите информации регламентирует как аппаратно-программные, инженерно-технические аспекты системы обеспечения информационной безопасности защиты информации, так и вопросы их содержания.

Основным законам по защите информации является Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информационных технологиях и о защите информации». Данным законом регулируются три группы взаимосвязанных отношений:

- Право на поиск, получение, передачу, производство и распространение информации;
- Применение информационных технологий;
- Обеспечение защиты информации [5].

Имеется правовой акт устанавливающих требования по защите персональных данных Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных», который описывает следующие положения:

- Регламентирует основные понятия, связанные с персональными данными;
- Принципы и условия обработки персональных данных;
- Контроль и надзор за обработкой персональных данных;
- Ответственность за нарушения [6].

Так же существует Федеральный закон от 29 июля 2004 г. N 98-ФЗ «О коммерческой тайне», который содержит требования по защите коммерческой тайны и описывает следующие положения [7]:

- Регламентирует основные понятия, связанные с коммерческой тайной;
- Основные способы и методы хранения коммерческой тайны;
- Контроль и надзор за соблюдением коммерческой тайны;
- Ответственность за нарушения.

Данные правовые акты будут браться за основу и устанавливать требования при организации комплексной системы информационной безопасности, предъявляя требования к документации, аппаратным и программным средствам.

Выбор комплекса задач по повышению уровня криптографической защиты информации

Для модернизации существующей системы информационной безопасности в рассматриваемой компании необходимо оттолкнуться от имеющейся на предприятии аппаратно-программной архитектуры, усовершенствовав ее современными криптографическими средствами защиты информации, который организуют безопасность активов. Данные действия затронут локальную сеть предприятия, базы данных, пользователей, расширят должностную инструкцию системного администратора и внесут изменения в существующую Политику информационной безопасности.

Термины и определения разрабатываемой системы отражены в ГОСТ Р 50922-2006 «Защита информации. Основные термины и определения» [1].

После внедрения всех предложений по модернизации, компания получит следующие преимущества: Стабильная работа локальной сети; Хранение конфиденциальных данных в зашифрованном виде; Использование современных средств для защиты информации; Возможность восстановления данных с помощью системы резервного копирования; Повышение конкурентоспособности предприятия.

Оценка уровня информационной безопасности организации определяется с помощью групповых показателей информационной безопасности, позволяющих оценить степень выполнения требований информационной безопасности, которые высчитываются по методике, представленной в ГОСТ Р ИСО/МЭК 19794-7-2017.

$$R = P(t) * P(\upsilon) * S \tag{1}$$

где R – значение риска; P(t) – вероятность реализации угрозы ИБ; P(v) – вероятность наличия уязвимости; S – ценность актива.

В качестве примера значений вероятностей P(t) и P(v) приведена качественная шкала с тремя уровнями: низкая, средняя и высокая. Для оценки значения ценности актива S представлены число-

вые значения в интервале от 0 до 4. Сопоставление им качественных значений было проведено в организации, в которой производится оценка рисков информационной безопасности [3].

После внедрения разрабатываемого комплекса по обеспечению информационной безопасности показатели информационной безопасности улучшатся. Расчет перечня результатных показателей по формуле 1, рассчитываемых на базе использования совокупности исходных показателей в процессе выполнения этих функций и сравнение с имеющейся системой информационной безопасности представлен в таблице 1.

Таблица 1

Показатели информационной безопасности

Наименование группового показателя информационной безопасности	Показатели до модернизации системы обеспечения информационной безопасности
Обеспечение ИБ при использовании средств криптографической защиты информации	0
Общие требования по обработке персональных данных в организации и хранении их в защищенном виде	0

Обоснование выбора программных средств обеспечения криптографической защиты информации

В рассматриваемой организации информация хранится в незащищенном виде, именно поэтому необходима установка программного средства, которое будет осуществлять шифрования данных. При выборе программного средства важно учесть особенности шифрования данных и скрытые уязвимости.

На данном предприятии используется база данных, использующая язык запросов SQL, а он свою очередь содержит встроенные алгоритмы шифрования данных. Для этого системному администратору необходимо будет воспользоваться при администрировании базы функциями SQL AES_ENCRYPT() или DES_ENCRYPT(). Существует несколько видов встроенного шифрования данных на уровне SQL Server.

1. Шифрование на уровне ячеек (Преимущества – детальный уровень шифрования, можно зашифровать отдельную ячейку; данные не расшифровываются до тех пор, пока не приходит время их использования. Недостатки - зашифрованные данные должны быть сохранены с использованием типа данных varbinary, снижение работоспособности БД из доп. обработки файлов при шифровании и расшифровки файлов);
2. Прозрачное шифрование (Преимущества – не меняются использованные приложения; соблюдение требований нормативных актов и правил корпоративной безопасности. Недостатки – необходимо иметь копии сертификата и закрытого ключа);
3. Шифрование на уровне файлов через Windows (Преимущества – способ шифрования данных, сохраненных на дисках NTFS. Недостатки – автоматического шифрования резервных копий не происходи).

Из представленных типов шифрования данных прозрачное шифрование лучше всего подходит для решения поставленных задач. Данный вид шифрования позволит зашифровать базу данных без изменения других приложений. Он позволит сохранить выполнения требований нормативных актов и существующей Политики информационной безопасности. Данные будут шифроваться в режиме реального времени: по мере внесения записей в файлы (*.mdf) базы данных SQL Server и файлы (*.ldf) журнала транзакций.

Шифрование данных будет производиться перед записью на диск и расшифровываться перед извлечением. При начале использования прозрачного шифрования необходимо создать резервную копию сертификата и закрытого ключа, связанного с этим сертификатом, так как в случаи недоступности сертификата должны быть его копии и копии закрытого ключа, иначе невозможно будет открыть базу данных. Для шифрования будет использован ключ шифрования базы данных. Этот ассиметричный ключ хранится в загрузочной записи базы данных и потому всегда доступен при восстановлении.

Ключ шифрования базы данных шифруется с использованием сертификата сервера, который шифруется с помощью DMK базы данных master. DMK базы данных master шифруется с применени-

ем SMK. SMK – асимметричный ключ, зашифрованный с помощью Windows DPAPI. Ключ SMK автоматически формируется при первом шифровании любого объекта и привязан к учетной записи SQL Server Service.

Анализ локальной сети предприятия

В рамках данной статьи был проведен анализ активов, их уязвимостей и оценка рисков.

Активы представляют собой информацию, несущую ценность для фирмы и находящаяся в ее распоряжении, при этом форма представления информации не важна.

Информационные активы включают в себя основные базы данных и документы, участвующие в оптовых продажах лекарственных средств, а также отражающих основную деятельность предприятия, имеющие высокую размерность оценки. Основные информационные активы: Партнерская база; Документы бухгалтерского учета; Персональные данные сотрудников; Клиентская база; Производственная документация; База иностранных поставщиков.

Активы программного обеспечения состоят из программного обеспечения, использующегося в компании и представляющие ценность. Имеют среднюю размерность оценки и состоят из: Windows Server 2016; Microsoft SQL Server 2016; Windows 8; Windows 10; Microsoft Exchange Server 2016; Microsoft Outlook 2016; Radmin 3.5; с MS Office 2016; Dr. Web Security Space 11; TeamViewer 14; 1С: Предприятие 8.3; 1С: Бухгалтерия 8.3.

Физические активы имеют высокую размерность оценки и включают в себя: Персональные компьютеры; МФУ устройства; Сетевое оборудование; Проекторы; Кондиционеры; ЖК панели.

После анализа активов организации была проведена идентификация уязвимостей окружающей среды, организации, процедур, персонала, менеджмента, администрации, аппаратных средств, программного обеспечения и аппаратуры связи, которые могли бы быть использованы источником угроз для нанесения ущерба активам и деловой деятельности организации, осуществляемой с их использованием.

При проведении процедуры оценки были учтены требования стандарта ГОСТ Р ИСО/МЭК ТО 13335-3-2007 (Приложение D) [2].

Оценке подлежат следующие виды уязвимостей: Программные уязвимости; Аппаратные и сетевые уязвимости; Уязвимости, связанные с персоналом; Уязвимости, связанные с организацией; Уязвимости, связанные с расположением организации.

В процессе выявление рисков организации были рассмотрены угрозы критического характера, которые могут вывести из работоспособности сеть или ключевые элементы сети.

Все рассматриваемые риски приводят к нарушению целостности локальной сети, но в зависимости от ранга риска у них градируется вероятность появления, а также величина последствий. Список рисков включает в себя самые частые причины сбоев в локальных сетях, основанный на опыте аналогичных компаний существующий на рынке.

При анализе рисков были выделены следующие возможные риски: Утечка конфиденциальной информации; Потеря или недоступность данных, составляющих коммерческую тайну; Использование искаженных данных; Распространение во внешней среде информации, угрожающей репутации компании.

Стратегия повышения уровня криптографической защиты информации

Шифрование данных с помощью встроенных средств языка SQL обеспечит целостность и защиту данных от злоумышленников, защиту от копирования и распространения и минимизирует риски:

- Утечки конфиденциальной информации;
- Потери данных, составляющих коммерческую тайну.

Для обеспечения прозрачного шифрования данных системный администратор будет использовать при работе с базой SQL-функции AES_ENCRYPT() или DES_ENCRYPT(). Для этого необходимо будет создать сертификат DMK в базе данных master. Далее создать резервную копию сертификата и ключа DMK, связанного с сертификатом, после этого следует создание ключ шифрования базы данных, которая будет защищаться и включить прозрачное шифрование данных. Прозрачное шифрование будет активировано с помощью программного псевкода:

USE [master] GO // Создание DMK.

CREATE MASTER KEY ENCRYPTION BY PASSWORD = '$tr0ngPa$$w0rd1' GO // Создание сертификата сервера.

CREATE CERTIFICATE EncryptedDBCert WITH SUBJECT = 'Certificate to encrypt EncryptedDB'; // Создание ключа шифрования данных и указании базы данных, которая будет шифроваться

GO USE [master] GO CREATE DATABASE ENCRYPTION KEY WITH ALGORITHM = AES_256

ENCRYPTION BY SERVER CERTIFICATE [EncryptedDBCert] GO ALTER DATABASE [EncryptedDB] SET ENCRYPTION ON GO //Проверка состояния шифрования всех баз данных на сервере

USE [master] GO SELECT db.[name], db.[is_encrypted], dm.[encryption_state], dm.[percent_complete], dm.[key_algorithm], dm.[key_length] FROM [sys].[databases] db LEFT OUTER JOIN [sys].[dm_database_encryption_keys] dm ON db.[database_id] = dm.[database_id]; GO

Проведение оценки повышения уровня криптографической защиты информации

После внедрения разрабатываемого комплекса по обеспечению криптографической защиты информации, показатели информационной безопасности значительно улучшились, если сравнивать с показателями, представленными в таблице 1.

Описание перечня результатных показателей модернизированной системы, рассчитываемых на базе использования совокупности исходных показателей в процессе выполнения этих функций и сравнение с имеющейся системой информационной безопасности представлено в таблице 2. Расчет данных значений произведен по формуле (1).

Таблица 2

Показатели информационной безопасности

Наименование группового показателя информационной безопасности	Показатели после модернизации системы обеспечения информационной безопасности
Обеспечение ИБ при использовании средств криптографической защиты информации	0,75
Общие требования по обработке персональных данных в организации и хранении их в защищенном виде	0,9

Оценка эффективности разработанной системы по криптографической защите информации

В целях оценки эффективности разрабатываемой архитектуры системы обеспечения криптографической защиты информации применяются модели оценки защищенности локальной сети предприятия от несанкционированного доступа, утечки информации и нарушения ее целостности.

Защищенность локальной сети обеспечивается в течение заданного периода времени $T_{зад}$ в случае, если в течение заданного периода обеспечена ее целостность и доступность. В рамках оценки учитывается, что в отношении локальной сети предприятия проводятся мероприятия по периодическому мониторингу криптографической защиты информации (профилактическая диагностика). Архитектура локальной сети предусматривает включение в состав каждого устройства локальной сети решений, реализующих диагностику и восстановление необходимой целостности и доступности сети при выявлении источников нарушения конфиденциальности информации или следов их воздействия. При активизации источника опасности защищенность локальной сети считается нарушенной.

В случае, если заданный период безопасного функционирования меньше периода проведения диагностики локальной сети (периода между окончаниями соседних диагностик), вероятность отсутствия опасного воздействия в течение периода $T_{зад}$ вычисляется по формуле:

$$P_{imp(1)} = \{(\sigma - \beta^{-1})^{-1} * (\sigma * e^{\frac{T_{set}}{\beta}} - \beta^{-1} * \sigma^{\sigma T_{set}}),\ if\ \sigma \neq \beta^{-1} e^{-\frac{T_{set}}{\beta}} * (1 + \sigma T_{set}), if\ \sigma = \beta^{-1} \quad (2)$$

где σ – частота воздействия источника опасности на локальную сеть; β – среднее время, в течение которого источник опасности, проникший в локальную сеть, активизируется; $T_{зад}$ – задаваемый период, в течение которого должна быть обеспечена непрерывная работа локальной сети.

В случае, если заданный период безопасного функционирования локальной сети превышает или равен периоду проведения периодического мониторинга, вероятностная оценка защищенности локальной сети определяется следующим образом:

$$P_{imp} = P_{av} + P_{cert} \tag{3}$$

где $P_{average}$ - вероятность отсутствия опасного воздействия в течение всего времени между проведением периодического мониторинга в рамках заданного периода.

С учетом доли периодов в пределах заданного периода T_{set} расчет осуществляется по формуле:

$$P_{av} = \frac{N(T_{\text{int}} + T_{diag})}{T_{set}} P^{N}_{imp(1)}(\sigma, \beta, T_{\text{int}}, T_{diag}, T_{\text{int}} + T_{diag}) \tag{4}$$

где T_{int} – время между окончанием предыдущей и началом очередной проверки защищенности локальной сети; T_{diag} – длительность проверки защищенности локальной сети, включающая время, необходимое на восстановление локальной сети; N – число периодов между диагностиками, которые целиком вошли в пределы T_{set}, с округлением до целого числа, $N = \frac{T_{set}}{T_{\text{int}} + T_{diag}}$ – целая часть;

$P^{N}_{imp(1)}(\sigma, \beta, T_{\text{int}}, T_{diag}, T_{\text{int}} + T_{diag})$ – вероятность отсутствия негативного воздействия, оказываемого источниками опасности в течение периода между проведением мониторинга, в пределах заданного периода функционирования локальной сети, (расчет осуществляется в соответствии с (2).

P_{cert} – вероятность отсутствия опасного воздействия после последнего проведения мониторинга (в конце заданного периода функционирования локальной сети). С учетом доли остатка $T_{rem} = T_{set} - N(T_{av} + T_{diag})$ в общем заданном периоде T_{set} и независимости исходных характеристик расчет осуществляется по формуле:

$$P_{cert} = \frac{T_{rem}}{T_{set}} P_{imp(1)}(\sigma, \beta, T_{av}, T_{diag}, T_{rem}) \tag{5}$$

Значение $P_{imp(1)}(\sigma, \beta, T_{av}, T_{diag}, T_{rem})$ вычисляется по формуле (2).

Вероятность обеспечения защищенности локальной сети от утечки информации P_{imp} в течение T_{set} определяется расчетами по (2) и (3) в зависимости от исходных данных (соотношения T_{diag} и T_{set}).

Исходные значения устанавливаются в ТЗ или в постановках функциональных задач с учетом сценариев возможных угроз, а также определяется в результате практических экспериментов и включается в эксплуатационную документацию.

Проведенный аудит разработанной криптографической системы безопасности показал, что внедренные средства криптографической защиты удовлетворяют требованиям, предъявляемым к защите информации в локальной сети организации.

Принятые меры в достаточной мере обеспечивают защиту важных активов и минимизируют риски. Модернизированная Политика информационной безопасности описывает сценарий поведения системного администратора при управлении системы комплексной защиты информации.

Оценка уровня информационной безопасности организации определяется с помощью групповых показателей информационной безопасности, позволяющих оценить степень выполнения требований информационной безопасности.

Разработанная система соответствует ГОСТ Р ИСО/МЭК 13335-1-2006 «Информационная технология (ИТ). Методы и средства обеспечения безопасности» и Международному стандарту безопасности информационных систем ISO 17799 [4].

Заключение

В рамках статьи проведен анализ актуальности темы по информационной безопасности в локальных сетях, исследованы стандарты в области информационной безопасности и защиты информации. Проведено исследование, выявляющие основные активы организации и их ранжирование, построена модель угроз информационной безопасности, описаны риски. Предложено проектное предложение по разработке системы криптографической защиты информации в рамках локальной сети. Описанное решение может носить типовой характер и применимо в локальных сетях малых и средних организаций, имеющие проблемы с передачей незащищенных данных.

Литература

1. ГОСТ Р 50922-2006. Защита информации. Основные термины и определения;
2. ГОСТ Р ИСО/МЭК ТО 13335-3-2007. Информационная технология (ИТ). Методы и средства обеспечения безопасности;
3. ГОСТ Р ИСО/МЭК ТО 7. Национальный стандарт Российской Федерации. Методы и средства обеспечения безопасности. Часть 3. Методы менеджмента безопасности информационных технологий. Москва. 20с.
4. Международный стандарт ISO/IEC 17799 – Информационные технологии – практические правила управления информационной безопасностью;
5. Федеральный закон от 27 июля 2006 г. N 149-ФЗ Об информации, информационных технологиях и о защите информации
6. Федеральный закон от 27 июля 2006 г. N 152-ФЗ "О персональных данных"
7. Федеральный закон от 29 июля 2004 г. N 98-ФЗ "О коммерческой тайне"
8. *Ahmed Helmy, Magdy Saeb,A. Baith Mohamed.* A Metamorphic-Enhanced MARS Block Cipher // The International Journal of Computer Science and Communication Security (IJCSCS),July, 2013.
9. *Bogdanov A., Knudsen L.R., Leander G., Paar C., Poschmann A., Robshaw M.J.B., Seurin Y., Vikkelsoe C.* PRESENT: An Ultra-Lightweight Block Cipher.
10. *Jakobsen J.B., Orlandi C.* A practical cryptanalysis of the Telegram messaging protocol. Aarhus university, September 2015.
11. *Aoki K., Ichikawa T., Kanda M., Matsui M., Moriai S., Nakajima J., Tokita T. Camelia:* Improved Attacks on Reduced-Round Camellia-128/192/256. CT-RSA 20, April 2015
12. *Brus V. R., Voronova L. I.* The Research of Neural Network Architectures for Cardio-Analysis in the Driver Support System // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, pp. 01-04, doi: 10.1109/IEEECONF51389.2021.9416024.
13. *Lilia I. Voronova, Vyacheslav I. Voronov, Nawar Mohammad.* Modeling the Clustering of Wireless Sensor Networks Using the K-means Method // 2021 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), Yaroslavl, Russia, 2021.
14. *Воронов В.И., Быков А.Д., Воронова Л.И*. Проектирование подсистемы детектирования лиц и интерфейса работы с базой данных в программно-аппаратном комплексе биометрической идентификации на основе нейросетевого распознавания лиц // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 4. С. 31-38.
15. *Voronov V.I., Zharov I.A., Bykov A.D., Trunov A.S., Voronova L.I.* Designing a neural network identification subsystem in the hardware-software complex of face recognition // T-Comm. 2020. Т. 14. № 5. С. 69-76.
16. *Шишкин А.О., Воронова Л.И.* Проектирование IOT системы "умный дом" с криптографической защитой данных // Телекоммуникации и информационные технологии. 2018. Т. 5. № 2. С. 66-71.
17. *Врагова Е.В., Воронова Л.И.* Методы и средства защиты от botnet's (зомби-сетей) // Телекоммуникации и информационные технологии. 2018. Т. 5. № 1. С. 112-116.
18. *Воронов В.И., Воронова Л.И., Скрябин В.И., Лукманов К.Д.* Разработка лабораторного практикума "подсистема управления безопасностью в программно-аппаратном комплексе умный дом" // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2020. Т. 9. № 2. С. 20-26.

РОЛЬ И РЕЗУЛЬТАТЫ АНАЛИЗА ВКЛАДОВ НА СОВЕТ МСЭ ДЛЯ ВЫЯВЛЕНИЯ ПРИОРИТЕТОВ В ПОЗИЦИЯХ ГОСУДАРСТВ-ЧЛЕНОВ ПРИ ПОДГОТОВКЕ К ПК-22

Резникова Наталья Петровна,
ФГУП «НИИР», главный научный сотрудник, д.э.н., профессор, Москва, Россия
reznikova.natalya1946@yandex.ru

Артемьева Галина Станиславовна,
Московский технический университет связи и информатики, к.э.н. доцент кафедры «Цифровая экономика, управление и бизнес-технологии», Москва, Россия
artemieva-g-s@yandex.ru

Калюга Дарья Викторовна,
ФГУП «НИИР», начальник лаборатории НТЦА ЭМС
kdv@niir.ru

Аннотация

В статье дана характеристика вопросов, поднятых во вкладах Государств-Членов (ГЧ) МСЭ и обсуждавшихся на сессиях Совета МСЭ в период после ПК-18. Структурирована основная тематика вкладов, выявлены связи вкладов с резолюциями и решениями ПК-18, приведена статистика вкладов по годам и активности ГЧ, показаны возможные направления развития тем и дискуссий на ПК-22.

Ключевые слова: *Международный союз электросвязи (МСЭ), Государства-Члены, электросвязь/ИКТ (ЭИКТ), вклады, Совет МСЭ; Полномочная конференция МСЭ (ПК); политика и стратегия; информационное общество; цели устойчивого развития, резолюции, решения*

Введение

Международный Союз Электросвязи (МСЭ, Союз) является ведущей международной организацией в области ЭИКТ [1] и до сих пор не теряет своей значимости для решения многих вопросов, затрагивающих интересы всех его членов [2]: суверенных Государств-Членов (ГЧ) и Членов секторов (ЧС). МСЭ входит в систему Организации Объединенных Наций (ООН) в качестве специализированного агентства и, если это подтверждают ГЧ, принимает к исполнению решения и рекомендации ГА ООН и других органов ООН в рамках своего мандата, зафиксированного в Уставе и Конвенции МСЭ и Административных регламентах, являющихся международными договорами. Высшим органом власти Союза является Полномочная Конференция (ПК) – собрание делегаций всех ГЧ Союза. На ПК принимается основная организационная политика и стратегия, определяются инструменты и виды деятельности.

Только ПК, определяя направления деятельности Союза, принимает решения относительно изменения текстов действующих Устава и Конвенции МСЭ и утверждает решения и резолюции, определяющие работу МСЭ на четыре года. Решения на ПК принимаются на основе консенсуса, поэтому очень важно анализировать тенденции в отношении обсуждаемых вопросов и проблем, которые возникают в периоды между ПК, чтобы заранее примерно представлять, какие резолюции могут пересматриваться, а какие появятся вновь, чтобы занять соответствующую позицию.

Роль Совета МСЭ в осуществлении решений и резолюций ПК. Цели и задачи статьи

В периоды между ПК проводятся ежегодные сессии Совета МСЭ (ITU Council, Совет), действующего в качестве руководящего органа по реализации решений ПК и управления текущей деятельностью Союза.

Совет МСЭ, учрежденный в 1947 г., объединяет 25% от общего числа ГЧ, которые избираются в него на ПК в соответствии с требованием объективного распределения мест в Совете между пятью регионами мира (Америкой (Северной и Южной); Западной Европой; Восточной Европой; Африкой (включая некоторые арабские государства); Азией (Азиатско-тихоокеанский регион) и Австралией).

Деятельность Совета направлена на рассмотрение широкого круга вопросов в области ЭИКТ и обеспечивает соответствие между направлениями деятельности Союза, заданными на ПК, и быстро происходящими изменениями современного общества и сферы ЭИКТ (внешней среды).

Совет в качестве законодательного и контрольного органа санкционирует бюджет, несет общую ответственность за результаты стабильной текущей работы МСЭ, которую, в свою очередь выполняет Генеральный секретариат и директораты бюро секторов; контролирует финансы Союза. По решению Совета создаются его рабочие группы (РГС), которые отвечают за определенные направления работы. В частности, в обязанности Совета входит подготовка, силами соответствующих РГС, проектов документов по стратегическому планированию Союза: проекта Стратегического плана МСЭ (СП) и проекта Финансового плана МСЭ на четырехлетний период, которые должны быть подготовлены к последнему перед ПК заседанию Совета (в данном периоде к Совету-22, который запланирован на март 2022 г.). ПК-22, которая должна утвердить эти планы, запланирована на период с 26 сентября по 14 октября 2022 г. [3, 4].

Работа Совета идет в соответствии с Общим регламентом конференций, ассамблей и собраний Союза [1], Правилами процедуры Совета [1] и повесткой дня, в которой перечислены все документы, подлежащие рассмотрению, и которые каждая администрация связи ГЧ может заблаговременно анализировать с целью формирования соответствующей позиции. Можно назвать несколько групп документов. Во-первых, это так называемые «вклады» ГЧ, в которых администрация связи соответствующего ГЧ выражает свое мнение по тем или иным вопросам, которые считает необходимым обсудить и/или принять по ним решение. Как правило, ГЧ сначала выносит свой вклад на обсуждение в соответствующую РГС. Во-вторых, это документы, представляемые Секретариатом на рассмотрение Совета, носящие характер отчетных документов, документов в помощь ГЧ или документов для сведения. По этим двум видам документов Советники могут принимать соответствующие решения: отклонять, направлять на доработку, принимать к сведению или утверждать на их основе решение или резолюцию Совета, обязательную для исполнения, прежде всего Секретариатом, в интересах членства, для совершенствования текущих процессов работы, повышения подотчётности и прозрачности, управления рисками, внедрения современных информационных систем управления организацией, выполнения одобренных советниками рекомендаций надзорных органов, в том числе органов ООН, и т.п.

В процессе обсуждения первых двух категорий документов могут появляться, так называемые, временные документы, в которых специально созданные в ходе сессии Совета группы ad-hoc, отражают найденный консенсус по содержанию документов, которые были важными, но неоднозначно воспринятыми ГЧ первоначально: позиции не совпали полностью или частично. Дело в том, что в МСЭ документ считается одобренным только в том случае, если никто против него не возразил. В противном случае, он направляется на доработку в РГС, из которой он и поступил на сессию Совета, или возвращается в Секретариат. Наконец, четвертый тип документов – информационные, они могут сопровождать два основных типа документов или устную презентацию, если требуется что-то разъяснить особенно подробно, или необходимо представить советникам значительные объемы статистических данных. Как правило, третий и четвёртый типы документов не переводятся на шесть официальных языков Союза, для работы с ними используется английский язык. Устные презентации могут использоваться в отдельных случаях, если документ отсутствует в письменном виде или он слишком велик по объему.

В настоящее время наступила финишная прямая подготовки к ПК-22. Совет-22 должен принять квази-окончательную[1] версию двух стратегических планов: собственно, стратегического и финансового. Последние два года работа осложнялась тем, что все собрания РГС и самого Совета проходили в виртуальной форме, что было вызвано распространением по всему миру пандемии Covid-19 со всеми его модификациями и вытекающими последствиями. Ни в Уставе, ни в Конвенции, ни в других упомянутых выше регламентах нет рекомендаций по процедурам виртуальных собраний. Конечно, были осуществлены определенные договоренности, решения по важнейшим документам принимались после виртуальных консультаций по переписке между Советниками, что сказалось на активности ГЧ. Вместе с тем, именно активность ГЧ по вкладам на сессиях Совета дает возможность снять,

[1] Окончательная версия принимается на ПК. По опыту ПК-18, стратегический план МСЭ на 2020-2023 гг. был принят на ночном специальном заседании глав делегаций накануне последнего дня конференции

хотя бы частично, неопределенность для делегации Российской Федерации в отношении того, какие темы будут затрагивать и отстаивать те или иные государства (их группы) и как они будут вести себя на ПК-22, в каком бы формате она ни проходила. Именно это обстоятельство позволило сформулировать основную цель настоящей статьи – провести ретроспективный анализ вкладов администраций связи ГЧ, начиная с 2018 г. (после ПК-18), для выявления тенденций в изменении/сохранении позиций ГЧ на сессиях Совета МСЭ, степени влияния администраций на формирование повестки дня сессий Совета во исполнение решений Полномочной конференции 2018 г. (Дубай), роли, тематики и активности России по вкладам для подготовки позиции на ПК-22.

Для достижения указанной цели авторами были решены следующие задачи:

- проанализированы вклады ГЧ на сессиях Совета МСЭ в 2018 гг. и в период после ПК-18 (2019-2021 гг.) по следующим критериям: разработчики вкладов, содержание вкладов и связь с решениями и резолюциями ПК-18;
- выявлены определенные тенденции в отношении некоторых областей, в которых с большой вероятностью можно ожидать неоднозначных позиций.

Результаты анализа позиций Государств-Членов на Совете МСЭ в период 2018-2021 гг. Выявленные тенденции и рекомендации

На Совете МСЭ 2018 г., который проходил в Женеве (Швейцария) в апреле 2018 г. в обычном формате, а его заключительное заседание состоялось 27.10.2018 г. в Дубае (ОАЭ) в соответствии с повесткой дня Совета было рассмотрено 108 документов, из которых 26 документов – вклады по разным вопросам деятельности Союза, при этом 10 ГЧ представили двадцать четыре вклада от своего имени, 2 вклада были совместными. Наибольшую активность проявили Российская Федерация и Соединенные Штаты Америки, представившие по пять вкладов. На рис. 1 приведено общее количество вкладов ГЧ (как индивидуальных, так и совместных) за 2018-2021 гг.

Степень участия ГЧ МСЭ в формировании повестки дня путем постановки наиболее важных, по их мнению, вопросов для рассмотрения составила 24,1% (отношение числа вкладов к общему числу документов), значительная доля которых (75,9% в 2018 г.) постоянно формируется ее Генеральным секретариатом, т.е. исполнительным органом Союза.

На Совете МСЭ 2019 г., который проходил в Женеве (Швейцария) в июне 2019 г., в соответствии с повесткой дня было рассмотрено 119 документов, из которых 44 документа – вклады ГЧ по разным вопросам деятельности Союза, при этом 12 ГЧ представили двадцать пять вкладов от своего имени, 19 вкладов были совместными, см. рис.1.

Степень участия ГЧ в формировании повестки дня и выявлении наиболее значимых вопросов для рассмотрения составила 36,97%, а доля, как и прежде, значительная, хотя и меньшая (63,03% в 2019 г.) продолжала формироваться самой организацией (Секретариатм). При этом следует отметить существенный рост активности ГЧ по сравнению с сессией Совета-18 (увеличение на 18 вкладов (на 69,23%)), что можно объяснить необходимостью выполнять решения ПК-18.

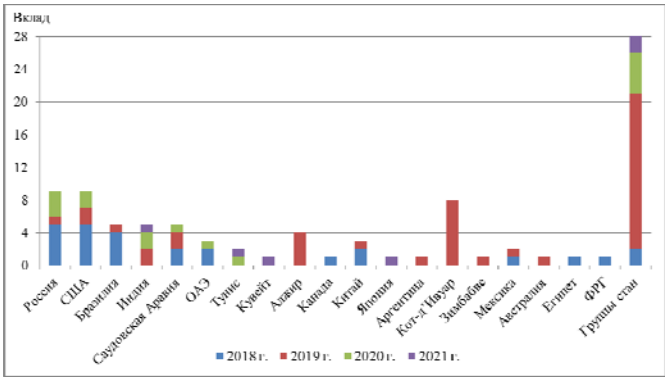

Рис. 1. ГЧ – участники сессий Совета МСЭ, сделавшие индивидуальные и совместные вклады, и их активность в 2018-2021 гг.

На Совете 2019 г. было представлено наибольшее количество вкладов за весь период анализа. При этом было представлено самое большое число совместных вкладов (19), что составило около 43,2%

от их общего числа. Это объяснялось тем, что стали формироваться группы стран, как правило, по региональному признаку, с целью обеспечения большей поддержки предложениям, содержащимся во вкладах, например, такая группа ГЧ, как: Соединенные Штаты Америки, Канада, Мексика, Аргентина. Россия от своего имени представила 1 вклад, но участвовала в представлении шести совместных вкладов, инициатором которых была Россия и к которым присоединились некоторые другие ГЧ (Республика Армения, Азербайджанская Республика, Киргизская Республика, Республика Таджикистан, Республика Узбекистан). Экстраординарную активность проявила делегация Республики Кот-д'Ивуар, представившая 8 вкладов, что составило 18,2% от общего их количества. Кроме того, на сессии Совета-19 вклады представили: Алжир (4 вклада); Индия, Саудовская Аравия и США − по два вклада; Бразилия, Китай, Аргентина, Зимбабве, Мексика и Австралия − по одному вкладу.

В 2020 г. из-за пандемии Covid-19 Совет МСЭ проходил дважды в форме виртуальных консультаций Советников: в июне (ВКС-1) и в ноябре (ВКС-2). На виртуальных консультациях Советников 2020 г. в соответствии с повесткой дня было представлено 84 документа, из которых 17 документов – вклады по разным вопросам деятельности Союза, при этом 8 администраций (ГЧ) представили 12 вкладов от своего имени, 5 вкладов были совместными (см. рис. 1). Особую активность проявила Российская Федерация, представившая три вклада и принявшая участие в трех совместных вкладах.

Степень участия ГЧ в формировании повестки дня и выявлении наиболее значимых вопросов для рассмотре-ния составила 20,2%. Как и в прошлые годы от имени Генерального секретариата поступило 79,8% документов. Следует отметить существенное снижение активности ГЧ по сравнению с Советом-19 (уменьшение на 61,4 %), что можно в значительной степени объяснить влиянием пандемии.

На виртуальных консультациях Советников 2021 г. в июне 2021 г., в соответствии с повесткой дня было рассмотрено 64 документа (а 13 документов было предложено отложить до следующего собрания). Шесть из рассмотренных документов – вклады по разным вопросам деятельности Союза. При этом 4 ГЧ представили по одному вкладу от своего имени (Индия, Кувейт, Тунис и Япония), два вклада были совместными (Соединенные Штаты Америки, Канада, Мексика, Аргентина; Австралия, Канада, Чешская Республика, Франция, Румыния, Соединенное Королевство Великобритании и Северной Ирландии). Обращает на себя внимание, что появились «блоки» не только по региональному признаку, но и по интересам.

Степень участия ГЧ в формировании повестки дня и выявлении наиболее значимых вопросов для рассмотре-ния составила 9,4%, а Секретариата − 90,6%. Следует отметить дальнейшее существенное снижение активности ГЧ. По сравнению с виртуальными консультациями Советников в 2020 г. число вкладов уменьшилось на 64,7%.

Делегации России и США за период 2018-2021 гг. представили по 9 индивидуальных вкладов, что составило 19,4% от общего количества вкладов. На Совете 2018 г. делегация от России, как и делегация от США, представили максимальное количество вкладов (по 5 вкладов), что составило 38,5% от общего количества вкладов, однако в дальнейшем как Россия, так и США значительно снизили свою активность, хотя и участвовали в совместных вкладах.

Всего за рассматриваемый период было представлено 93 вклада, которые охватывали достаточно широкий круг вопросов, о чем свидетельствуют данные таблицы 1, в которой приведены для примера из-за ограниченности объема статьи только несколько вкладов. Тем не менее, на основе полных данных о вкладах нами была сделана примерная группировка их тематики. Ниже приведены эти темы с указанием стран, которые представляли или присоединялись к вкладам, имеющим отношение к теме. Порядок тем представлен в соответствии с количеством присоединившихся ГЧ:

1. Регламент международной электросвязи (РМЭ), Всемирная конференция по международной электросвязи и Группа экспертов по РМЭ (США, Алжир, Зимбабве, Египет, Бразилия, Канада, Мексика, Парагвай, Российская Федерация, Армения, Азербайджан, Киргизская Республика, Республика Таджикистан, Республика Узбекистан, Аргентина, Содружество Багамских Островов, Австрия, Болгария, Дания, Эстония, Финляндия, Германия, Греция, Литва, Люксембург, Мальта, Молдова, Норвегия, Польша, Румыния, Словения, Испания, Швеция, Чешская Республика, Нидерланды, Словацкая Республика, Соединенное Королевство, Египет, Объединенные Арабские Эмираты, Саудовская Аравия). 40 стран.
2. Повышение эффективности открытых консультаций РГС-Интернет (Федеративная Республика Бразилия, Алжир, США, Саудовская Аравия, Албания, Австрия, Азербайджан, Бельгия, Босния и

Герцеговина, Болгария, Чешская Республика, Дания, Грузия, Германия, Греция, Венгрия, Италия, Латвия, Литва, Мальта, Молдова, Нидерланды, Норвегия, Польша, Румыния, Российская Федерация, Словацкая Республика, Испания, Швеция, Швейцария, Турция, Украина, Соединенное Королевство Великобритании и Северной Ирландии, Ватикан). 34 страны.

3. Предложения для Всемирного форума по политике в области электросвязи (Республика Кот-д'Ивуар, Алжир, Мексика, Аргентина, Содружество Багамских Островов, Федеративная Республика Бразилия, Канада, США, Австрия, Болгария, Чешская Республика, Дания, Эстонская Республика, Финляндия, Германия, Греция, Венгрия, Литовская Республика, Люксембург, Мальта, Республика Молдова, Королевство Нидерландов, Норвегия, Польша, Румыния, Словацкая Республика, Республика Словения, Испания, Швеция, Соединенное Королевство Великобритании и Северной Ирландии, Египет, Объединенные Арабские Эмираты, Саудовская Аравия). 33 страны.

4. Формат, тематика и место проведения конференций и ассамблей МСЭ (США, Индия, Канада, Германия, Болгария, Республика Кипр, Хорватия, Дания, Испания, Финляндия, Франция, Греция, Венгрия, Литовская Республика, Мальта, Норвегия, Королевство Нидерландов, Польша, Словакия, Чешская Республика, Румыния, Швеция, Соединенное Королевство Великобритании и Северной Ирландии, Объединенные Арабские Эмираты, Саудовская Аравия). 25 стран.

5. Предлагаемые изменения решений и резолюций ПК и Совета с точным указанием номера решения или резолюции и формулированием характера поправок (Кувейт, Тунис, США, Канада, Мексика, Аргентина, Республика Кот-д'Ивуар, Объединенные Арабские Эмираты, Алжир, Саудовская Аравия, Индия, Кения, Руанда, Южная Африка, Китай, Российская Федерация, Армения, Азербайджан, Киргизская Республика, Республика Таджикистан, Республика Узбекистан, Бразилия, Парагвай). 23 страны.

6. Строительство нового здания штаб-квартиры МСЭ в Женеве и связанные вопросы (США, Германия, Российская Федерация, Армения, Азербайджан, Киргизская Республика, Республика Таджикистан, Республика Узбекистан). 8 стран.

7. Влияние пандемии Covid-19 на различные стороны работы МСЭ (Индия, Китай, Азербайджанская Республика, Республика Беларусь, Россия, Республика Узбекистан). 6 стран.

8. Полномочия избираемых лиц на должности руководителей и их заместителей в РГ, исследовательских комиссиях и другие аспекты полномочности и этики, круг ведения (Республика Кот-д'Ивуар, Россия, Бразилия, США, Канада, Парагвай). 6 стран.

9. Руководящие указания по использованию Глобальной программы кибербезопасности (Австралия, Канада, Чешская Республика, Франция, Румыния, Соединенное Королевство Великобритании и Северной Ирландии). 6 стран.

10. Вопросы регионального присутствия МСЭ (Мексика, Индия, Народная Республика Бангладеш, Буркина-Фасо, Федеративная Республика Нигерия). 5 стран.

11. Роль МСЭ в выполнении решений Всемирной встречи на высшем уровне по вопросам информационного общества и повестки дня в области устойчивого развития на период до 2030 года (Россия, США, Саудовская Аравия, Китай). 4 страны.

12. Возмещение затрат на обработку заявок на регистрацию сложных НГСО спутниковых систем (Россия, США, Австралия, Канада). 4 страны.

13. Об участии МСЭ в деятельности Малых и средних предприятий (Республика Кот-д'Ивуар, ОАЭ, Аргентина). 3 страны.

14. Предложения по обеспечению принятия Полномочной конференцией 2018 г. реалистичных стратегического и финансового планов, основ бюджета и стратегического плана по управлению людскими ресурсами на 2020−2023 гг. (Россия, США). 2 страны.

15. Совершенствование менеджмента, отчетности и прозрачности в МСЭ (США, ОАЭ). 2 страны.

16. О Фонде развития информационно-коммуникационных технологий (ФРИКТ) и оценка мероприятий ITU TELECOM (Республика Кот-д'Ивуар, Япония). 2 страны.

17. Включение вопросов развития цифровой экономики в общие стратегические цели стратегического плана МСЭ (Китай). 1 страна.

18. Измерение информационного общества и статистические данные по ИКТ в МСЭ (Федеративная Республика Бразилия). 1 страна.

19. Разное. Организационные вопросы, защита ребенка в онлайновой среде [5], предложение о работе по видам деятельности, связанным с OTT и др.

Таблица 1

Названия вкладов ГЧ за 2018-2021 гг. (пример)

Вклады			
2018 г.	2019 г.	2020 г.	2021 г.
Документ C18/80. Вклад Российской Федерации − Предложения по обеспечению принятия Полномочной конференцией 2018 года реалистичных стратегического плана, основ бюджета и финансовых пределов МСЭ на 2020-2023 гг.	Документ C19/75. Вклад Российской Федерации – Предложения по пересмотру Резолюции 1299 «Разработка стратегического плана в области людских ресурсов»	Документ VC/2. Вклад от Российской Федерации – Предложения по повестке дня виртуальных консультаций Советников	Документ C21/78. Вклад от Индии – ВАСЭ-20: Возможные сценарии в свете пандемии COVID-19 и рекомендации
…			
Документ C18/88. Вклад от Соединенных Штатов Америки − Редакционные поправки Соединенных Штатов Америки к пересмотренному проекту Приложения 1 к Резолюции 71: Стратегический план МСЭ на 2020−2023 годы	Документ C19/93. Вклад от Республики Кот-д'Ивуар − Предлагаемая тема для Всемирного форума по политике в области электросвязи (ВФПЭ) 2021 года	Документ VC-2/2. Вклад от Республики Индии − Проведение Всемирной ассамблеи по стандартизации электросвязи (ВАСЭ) МСЭ в Индии	Документ C21/82. Вклад от Австралии, Канады, Чешской Республики, Франции, Румынии, Соединенного Королевства Великобритании и Северной Ирландии – Руководящие указания по использованию Глобальной программы кибербезопасности
…			

Какие выводы напрашиваются, исходя из полученных данных? Очевидно, что вопросы Регламента международной электросвязи, интернет и политики в области ЭИКТ привлекли наибольшее внимание. Много ГЧ обратились к резолюциям и решениям ПК и Совета с целью их улучшения в дальнейшем в соответствии с действующими процедурами. В то же время, вопросы развития цифровой экономики, их включение в общие стратегические цели стратегического плана МСЭ, также, как и измерение информационного общества и статистические данные по ЭИКТ в МСЭ, практически, не интересовали членство в рассматриваемом периоде.

Что же будут обсуждать на ПК-22? Скорее всего, будут подняты многие из названных выше тем в привязке к связанными с ними резолюциями и решениями ПК, что также следует из рассмотренных нами 93 вкладов.

Авторами был проведен анализ связи рассматриваемых вкладов с решениями и резолюциями ПК-18, что позволило провести анализ резолюций и решений на их актуальность с точки зрения проблем, которые волнуют государства, а также – с точки зрения уровня их исполнения в течение почти всего срока, прошедшего с ПК-18. Всего обращения к резолюциям ПК содержались в 27 вкладах (29% от 93 вкладов). Чаще всего делегаты обращались к следующим резолюциям [6]: Резолюция 2 (Пересм. Дубай, 2018 г.), 4 вклада; Резолюция 71 (Пересм. Дубай, 2018 г.); Резолюция 102 (Пересм. Дубай, 2018 г.), 5 вкладов; Резолюция 146 (Пересм. Дубай, 2018 г.), 10 вкладов. Остальные резолюции были использованы в 1-3 вкладах: Резолюция 77 (Пересм. Дубай, 2018 г.); Резолюция 91 (Пересм. Гвадалахара, 2010 г.); Резолюция 101 (Пересм. Дубай, 2018 г.), 2 вклада; Резолюция 123 (Пересм. Дубай, 2018 г.); Резолюция 130 (Пересм. Дубай, 2018 г.); Резолюция 133 (Пересм. Дубай, 2018 г.); Резолюция 140 (Пересм. Дубай, 2018 г.); Резолюция 144 (Пересм. Пусан, 2014 г.), 3 вклада; Резолюция 154 (Пересм. Дубай, 2018 г.); Резолюция 169 (Пересм. Дубай, 2018 г.); Резолюция 175 (Пересм. Дубай, 2018 г.), 2 вклада; Резолюция 179 (Пересм. Дубай, 2018 г.); Резолюция 180 (Пересм. Дубай, 2018 г.); Резолюция 203 (Пересм. Дубай, 2018 г.); Резолюция 206 (Дубай, 2018 г.), 3 вклада.

Естественно, что Резолюция 146 (Пересм. Дубай, 2018 г.) «Регулярное рассмотрение и пересмотр Регламента международной электросвязи» оказалась лидером в этом перечне. При этом также неоднократно упоминалась Резолюция 4 Всемирной конференции по международной электросвязи (Ду-

бай, 2012 г.). Резолюция 102 (Пересм. Дубай, 2018 г.) «Роль МСЭ в вопросах международной государственной политики, касающихся интернета и управления ресурсами интернета, включая наименования доменов и адреса», использованная в пяти вкладах, подписанных множеством ГЧ, также указывает на направление их интересов.

Заключение

1. Извлеченные уроки − это всегда возможности. Есть все основания полагать, что вопросы, поднятые ГЧ, в той или иной мере будут обсуждаться и на ПК-22. Сфокусированность на некоторых резолюциях свидетельствует о существенности для ГЧ тех вопросов, которые с ними связаны. Вместе с тем, в настоящее время в Совете МСЭ, силами его РГС-СФП и Генерального секретариата, ведется разработка Стратегического плана МСЭ на 2024-2027 гг. (Резолюция 71 (Пересм. Дубай, 2018 г.)). В нем определены два стратегических направления работы: первое, «Универсальная связь», оно фокусируется на обеспечении всеобщего доступа к недорогим, высококачественным и безопасным ЭИКТ и включает вопросы развития ЭИКТ и подключения всех неподключенных к сетям и услугам связи. Второе «Устойчивая цифровая трансформация», позволяет справедливо использовать ЭИКТ для расширения прав и возможностей людей и общества в области устойчивого развития.
2. Поскольку «инновации» присущи всей работе в сфере ЭИКТ, а также внутри МСЭ, то это нашло отражение во всех тематических приоритетах стратегического плана и было также включено в качестве фактора реализации, а именно: «эксплуатационная эффективность, действенность и инновации», то есть вопросы инновационной работы МСЭ могут быть подняты.
3. С учетом того, что вклады были представлены от имени примерно 21 процента общего количества ГЧ, то с большой вероятностью на ПК-22 могут появиться темы, не затронутые в ходе обсуждения на сессиях Совета в рассмотренном периоде.
4. Безусловно, на ПК будут затронуты болезненные вопросы кибербезопасности, поднятые в анализируемом периоде, и как уже показывает ход разработки СП МСЭ на 2024-2027 гг., вызывающие существенные разногласия.
5. Хотя поднятые во вкладах ГЧ за четыре года после ПК-18 темы, связанные с полутора-двумя десятками из 213 резолюций и решений ПК-18, т.е. − около 10%, очень важны, остальные резолюции остались как бы вне поля зрения членства, но они также имеют свой потенциал. Поэтому нужно использовать еще неиспользованные возможности МСЭ, отраженные в этих резолюциях, для развития информационного общества и достижения ЦУР. И здесь уместно привести слова Генерального секретаря МСЭ господина Хоулинь Чжао, предпосланные краткому отчету о ВКС-21: «Почти половина населения мира до сих пор не подключена к Интернету. Существует ложное представление о том, что индустрия информационных и коммуникационных технологий, в отличие от других отраслей, неплохо справляется с этой пандемией и что нам не нужно беспокоиться о развитии ИКТ. Если мы не изменим это мышление, к 2030 году подключить к сети другую половину мира будет очень сложно. Нам нужно использовать этот момент для поощрения инвестиций в ИКТ» [1].

Литература

1. МСЭ: Верен идее соединить мир – ITU. – URL: https://www.itu.int/ru/Pages/default.aspx.
2. *Резникова Н.П., Артемьева Г.С., Куликова К.Н.* Национальные интересы как основа подхода к обоснованию необходимости присутствия РФ в международных организациях связи и ИКТ // T-Comm: Телекоммуникации и транспорт. 2013. Том 7. № 12. С. 79-83.
3. *Резникова Н.П., Артемьева Г.С.* Подход к обоснованию стратегических направлений развития МСЭ на период 2024–2027 гг. с использованием SWOT-анализа // Электросвязь. 2022. № 2. С. 43-50.
4. *Резникова Н.П., Артемьева Г.С., Калюга Д.В.* Направления совершенствования стратегического плана Международного союза электросвязи на 2023-2027 гг // Электросвязь. 2021. № 10. С. 39-44.
5. *Резникова Н.П., Артемьева Г.С., Куликова К.Н.* От чего защищать ребенка в эру развитых инфокоммуникационных технологий и информационного общества // В сборнике: Телекоммуникационные и вычислительные системы. Труды конференции. 2015. С. 216-219.
6. Международный союз электросвязи. Заключительные акты Полномочной конференции (Дубай, 2018 г.). Решения, Резолюции и Рекомендация. RL: https://www.itu.int/dms_pub/itu-s/opb/conf/S-CONF-ACTF-2018-R1-PDF-R.pdf.

АСПЕКТЫ ПРОРАБОТКИ СИСТЕМЫ БЕЗОПАСНОСТИ УМНОГО ГОРОДА

Щербонос Егор Борисович,
МИРЭА, студент, Россия
shcherbonos.egor@yandex.ru

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

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

Аннотация

В рамках развития современных информационных технологий появилась возможность их внедрения, выражающаяся в построении системы «Умный город». Данное решение предполагает объединение отдельных городских систем жизнедеятельности и жизнеобеспечения в единую систему посредством информационно-телекоммуникационных технологий и разработки комплекса мер обеспечения информационной безопасности от возможных угроз.

Ключевые слова: *Интеллектуальные информационные системы, умный город (Smart city), рэнкинг, угроза информационной безопасности, обеспечение безопасности, носители угроз, злоумышленники, программно-аппаратное обеспечение, несанкционированный доступ, кибератака, киберинцидент, интернет вещей (IoT).*

Введение

В настоящее время существует тенденция урбанизации городского населения. Города разрастаются, что приводит к повышению уровня нагрузки на городскую инфраструктуру, усложняет функционирование и управление её подсистемами. Так количество подключенных к городскому IoT цифровых датчиков и оборудования в мире, по подсчетам консалтинговой компании Strategy Analytics к 2025 г. достигнет 38,6 млрд, а к 2030 г. – 50 млрд [1].

В рамках развития информационных технологий появилось решение этого вопроса, выражающаяся в создании Smart city. Концепция подхода Smart city предусматривает интеграцию отдельных городских систем жизнеобеспечения в единую систему посредством современных информационно-телекоммуникационных технологий.

Системы умного города связаны с объектами жизнедеятельности, поэтому любые сбои её функционирования представляют угрозу здоровью и жизни тысяч, а то и миллионов людей. В тесно завязанной технологической платформе умного города его общая надёжность определяется надёжностью самого слабого звена. Инфраструктура умных городов представляет собой совокупность новых и устаревших ИТ решений, программного и аппаратного обеспечения с недостаточным уровнем защиты и т.д. Обеспечение безопасности умного города, в том числе его информационно-коммуникационных систем, является одной из основных составляющих его концепции. Поэтому неудивительно, что объем рынка кибербезопасности для IoT растет и по расчетам экспертов GSMA достигнет $36,6 млрд к 2025 году [2].

Основная часть

Создание аппаратно-программного комплекса защиты информации умного города, в том числе его интеллектуальной составляющей, основывается на его архитектуре. Поскольку каждый умный город имеет собственную концепцию развития, для создания общей картины было рассмотрено несколько решений, реализованных в разных городах и странах.

Бизнес-школа IESE при университете Наварры ежегодно публикует Cities in Motion Index, анализирующий уровень развития умных городов по всему миру. Данная методика основана на 96 признаках, объединенных в группы (рис. 1).

Рис. 1. Методика развития Smart city бизнес-школы при университете Наварры

Согласно рэнкингу Cities in Motion Index по итогам 2020 года в топ-10 умных городов мира по показателям развития и привлечения талантов вошли Лондон, Нью-Йорк, Амстердам, Париж, Рейкьявик, Токио, Сингапур, Копенгаген, Берлин и Вена [3]. В итоговом рэнкинге Москва на 86-й позиции (рис. 2).

Ranking	City	Performance	CIMI
1	London - United Kingdom	A	100.00
2	New York - USA	A	95.73
3	Paris - France	RA	85.50
4	Tokyo - Japan	RA	81.95
5	Reykjavik - Iceland	RA	80.47
6	Copenhagen - Denmark	RA	78.51
7	Berlin - Germany	RA	77.46
8	Amsterdam - Netherlands	RA	77.31
9	Singapore - Singapore	RA	76.71
10	Hong Kong - China	RA	76.04

Рис. 2. Топ-10 самых умных городов мира 2020 года

Существуют другие методики оценки развития городов: EasyPark Smart Cities Index, «Индекс цифровой жизни российских городов», рейтинг НИИТС, IQ российских городов, а также разработки компаний Juniper Research, Price water house Coopers, Cisco, Bosch, Nokia, Huawei, Шанхайской академии социальных наук (SASS), PwC, Forbs и стандарты, в которых рассматриваются методики оценки развития Smart city (ISO 37120:2014 «Устойчивое развитие сообществ - показатели городских услуг и качества жизни», ISO 37122 «Устойчивые города и сообщества – показатели для умных городов», Стандарт ISO 37151:2015 «Интеллектуальные инфраструктуры коммунального хозяйства. Принципы и требования к системе рабочих показателей». ISO 37106:2018 «Устойчивые города и сообщества. Руководство по созданию операционных моделей умного города для устойчивых сообществ»).

После рассмотрения архитектуры топ-10 умных городов мира и в России (Москва, Санкт-Петербург, Казань, Новосибирск) можно выделить несколько основных подсистем, характерных для любого города, стремящегося к званию умного. К ним можно отнести: транспорт, здравоохранение, цифровое правительство, безопасность, образование, экология, культура.

Эти подсистемы можно считать ключевыми и на них нужно опираться при проектировании и разработке проектов Smart city. Так, в случае разработки комплексного решения обеспечения защиты информации, необходимо учитывать особенности каждой из представленных подсистем.

Угрозы безопасности системы Smart city представляют собой нарушения трёх основных свойств информации: целостности, доступности и конфиденциальности. Они могут быть преднамеренными или непреднамеренными, могут носить антропогенный или природный характер.

С учетом масштаба проектирования, разработки, внедрения и эксплуатации системы «Умный город» угрозы классифицируют по двум параметрам: по носителям угрозы и по основным причинам сбоя системы. Носителями угрозы могут выступать разработчики системы, обслуживающий персонал, поддерживающий её работоспособность, и пользователи, а также прочие внешние злоумышленники. Масштабы и сложности системы умного города делают её уязвимой к атакам и приводят опасным последствиям в случае сбоев. Среди основных причин сбоя системы выделяют ошибки в проектировании и разработке системы (в том числе аппаратные и программные), проблемы с эксплуатацией оборудования, проблемы, связанные с хранением, обработкой данных и их передачей.

Актуальной угрозой информационной безопасности умного города считается несанкционированный доступ, позволяющий злоумышленнику выводить из строя функционирующие системы жизнеобеспечения. Применяемые нарушителями сложные распределенные кибератаки, использование не по назначению сетей типа LPWAN (NB-IoT, LoRaWAN, SigFox, Стриж, Weightless, RPMA, Ingenu, LTE-M и другие), недостаток криптографических средств защиты, знаний о методах социальной инженерии и защиты от DDoS-атак представляют собой слабые места в системе защиты умного города, требующие пристального внимания.

Основными источниками угроз в цифровом технологическом ландшафте Smart city являются угрозы из интернета, угрозы при подключении съемных носителей, угрозы в почтовых вложениях, вредоносные программы-вымогатели [4].

Менее заметной в общегородских масштабах, но не менее критичной для успешной реализации концепции представляет собой угроза доступа к данным отдельных пользователей. Особенность данной угрозы заключается в том, что конечные устройства отдельных пользователей гораздо сложнее поддаются учёту и регулированию, особенно при рассмотрении нерезидентов конкретной системы умный город.

Методологии, стандарты и методы построения защищенных автоматизированных систем целесообразно рассматривать на основе трёх направлений: аутентификации (защите от несанкционированных воздействий на систему), целостности (защите существующих данных от модификаций) и конфиденциальности (недоступности критически важных данных злоумышленникам) [5].

При этом основополагающими решениями в области безопасности систем умный город и его интеллектуальных подсистем можно назвать [6]:

1) Стандартизация протоколов передачи данных для систем умный город. Существует несколько главных стандартов передачи данных, а именно:

– IP, самый простой и примитивный протокол, не обеспечивающий надёжность и целостность, а потому использующийся для маршрутизации;

– TCP/IP, надстроенный протокол TCP над IP контролирует надёжность передачи и отслеживает целостность;

– UDP, очень быстрый, но не обеспечивающий никакого контроля за целостностью, а потому использующийся только при недопустимости задержек, протокол;

– FTP, протокол передачи файлов, работающий по принципу клиент-серверной архитектуры, и использующийся при удалённом доступе;

– DNS, специализированный протокол, использующийся при запросе IP-адреса у DNS-сервера;

– HTTP, протокол клиент-серверного взаимодействия, не сохраняющий промежуточные состояния, имеющий расширение HTTPS, основанное на шифровании TLS;

– NTP, протокол синхронизации локальных часов с сетевым временем;

– SSH, протокол удалённого управления операционной системой с полным шифрованием трафика.

2) Использование гомоморфного шифрования. Такой алгоритм шифрования позволяет производить операции над данными без их расшифровки. Однако, его использование может привести к снижению производительности системы.

3) Защита содержимого зашифрованных пакетов в облачных хранилищах. Существуют компании, предоставляющие услуги по шифрованию данных в своих облачных хранилищах, однако они хранят ключи у себя, а, следовательно, могут ими воспользоваться. Другой подход предполагает заключение данных в криптоконтейнеры, но этот подход не позволяет синхронизировать данные без полной распаковки. Поэтому, самым совершенным методом на данный момент является использование специализированных программ или криптографических файловых систем.

4) Разделение собственно сети умного города и всемирной сети интернета. Это довольно очевидный и удобный подход, когда устройства и каналы передачи данных, относящихся к системам умного города, изолированы от общедоступной сети интернет. Фактически, некоторые системы не могут быть обособлены из-за необходимости обслуживания широкого круга лиц. В таком случае, рекомендуется уделять особое внимание защите каналов связи между внутренней частью системы и сетью интернет.

5) Автоматизация в управлении критически важных объектов инфраструктуры умного города - очевидная мера, предполагающая максимальную автоматизацию систем умного города. Это позволит минимизировать количество сотрудников и влияние человеческого фактора, но потребует значительной квалификации у администрирующего и обслуживающего персонала.

6) Предпочтение комплексным решениям, а не заказной разработке. В настоящее время наметилась растущая тенденция потребности в готовых сервисах при внедрении технологии Интернета Вещей. Частная разработка приводит к увеличению затрат, сложности внедрения IoT-решений, сложности построения системы защиты от угроз информационной безопасности умного города. Предпочтение комплексным предложениям внедряющих технологии Интернета Вещей (IoT) позволяет обеспечить наибольшую безопасность составляющих умного города.

7) Встроенная защита в процессе производства и внедрения продукции OEM-производителями. Использование компаниями - производителями оборудования встроенной защиты на этапе производства и внедрения продукции, позволит минимизировать возможные риски.

Также в качестве мер по снижению угроз информационной безопасности системы «Умный город» можно предложить:

– анализ, тестирование и тщательную подборку предлагаемых к использованию программных и аппаратных средств;

– должное обучение и инструктирование обслуживающего персонала, а также обеспечение контроля над его действиями, включающего в себя разделение прав и режимов доступа и ведение журналов, совершаемых персоналом действий;

– контроль над действиями пользователей Smart city;

– комплекс мероприятий по предотвращению несанкционированного доступа к оборудованию, хранилищам и базам данных, а также каналам связи и передачи информации и использование методов шифрования.

– использование интеллектуальных систем защиты информации.

Использование искусственного интеллекта в информационной безопасности обусловлено необходимостью оперативного реагирования при наступлении киберинцидента и нехваткой специалистов в области защиты информации. Недостаток опытных квалифицированных специалистов по защите, который сложился в настоящее время, с одной стороны и масштабные инциденты информационной безопасности, которые могут развиваться стремительно, с другой, могут привести к непоправимому урону. В компаниях, где отсутствует круглосуточная дежурная смена специалистов по ИБ, без системы оперативного автономного реагирования на инциденты ИБ будет проблематично обеспечить качественную защиту в нерабочее время. Кроме того, нарушители перед атакой могут выполнить отвлекающий манёвр - например, запустить DDoS-атаку или активное сетевое сканирование, отвлекая специалистов по ИБ. В таких ситуациях система реагирования на инциденты ИБ на основе искусственного интеллекта поможет обрабатывать большое количество инцидентов информационной безопасности, автоматизировать многие процессы для специалистов по безопасности и обеспечить своевременное реагирование на возможные киберинциденты.

При этом предпочтение должно отдаваться именно системам искусственного интеллекта. При применении традиционных систем анализа отклонений сравниваются данные, которые имеют место быть и некоторые, заранее заданные данные (которые должны быть), и на основе этих сравнений принимаются управленческие решения. Интеллектуальные же информационные системы относятся к системам поддержки принятия решения и работают не по принципу сравнения фактических и типичных показателей, а сами могут подсказать решение на основе выявленных закономерностей.

Для защиты данных пользователя рекомендуется использовать системы определения аномалий. В качестве примера можно привести банковский интернет-сервис, который собирает данные, проводит

анализ характерных признаков работы клиента, таким образом, определяя аномальные учетные записи.

Также активно используются когнитивные технологии для оценивания потенциальных заемщиков, определение финансовых рисков в банковских организациях.

Можно привести достаточно много примеров использования систем машинного обучения в кибербезопасности, например, работа с внутренними нарушителями [7]. Зная паттерны работы пользователя, системы искусственного интеллекта в случае типичного изменения шаблона сотрудника может отправить предупреждение специалистам по безопасности. Использование систем защиты снабженных функциями распознавания речи и машинным зрением помогает охранным структурам, сообщая им о несанкционированных попытках прохода посторонних лиц в помещения организации, позволяет классифицировать активность работников, анализировать работу менеджеров с клиентами и т.п.

Заключение

Тенденция урбанизации городского населения приводит к необходимости интеграции отдельных городских систем жизнеобеспечения в единую систему посредством современных информационно-телекоммуникационных технологий.

Конкретные решения умных технологий уже получают практическую реализацию и активно применяются в городских инфраструктурных проектах. Во многих городах мира разрабатываются собственные, с учетом специфики, проекты умных городов, каждый из которых имеет своё уникальное видение на реализацию концепции развития проектирования Smart city. Однако, на данный момент отдельные умные системы ещё не объединены в единые всеобъемлющие планы градостроительства. Поэтому для реализации концепции умного города на данном этапе необходимо аккумулирование лучших практик и решений, выработка актуальных требований и стандартов.

Проектирование и разработка Smart city предполагает широкую и глубокую интеграцию современных цифровых технологий с уже существующими платформами систем умного города. Сложность и многогранность взаимодействия элементов и частей систем умного города, огромный объем обмена данными приводит к необходимости тщательной проработки концепции информационной безопасности, основанной на комплексном решении с использованием интеллектуальных систем защиты информации от возможных современных угроз.

Литература

1. Strategy Analytics: Internet of Things Now Numbers 22 Billion Devices But Where Is The Revenue? // Strategy Analytics [Электронный ресурс]. URL: //https://news.strategyanalytics.com/press-release/iot-ecosystem/strategy-analytics-internet-thingsnownumbers-22-billion-devices-where (дата обращения: 10.01.2022).

2. IoT Security Market. [Электронный ресурс]. URL: https://www.marketsandmarkets.com/Market-Reports/iot-security-market-67064836.html (дата обращения: 10.01.2022).

3. IESE Cities in Motion Index 2020. [Электронный ресурс]. URL: https://blog.iese.edu/cities-challenges-and-management/2020/10/27/iese-cities-in-motion-index-2020 (дата обращения: 15.01.2022).

4. Ландшафт угроз для систем промышленной автоматизации. Первое полугодие 2021. [Электронный ресурс]. URL: https://ics-cert.kaspersky.ru/publications/reports/2021/09/09/threat-landscape-for-industrial-automation-systems-statistics-for-h1-2021 (дата обращения: 15.01.2022).

5. *Шукенбаев А.Б., Лозовецкий В.В.* Автоматизированные системы обработки информации: учебное пособие. Москва: МИГКУ. 2009. 125 с.

6. *Шукенбаев А.Б., Глазов А.С.* Подсистема защищенного удаленного доступа // Международная научно-практическая конференция «Современные проблемы и задачи обеспечения информационной безопасности» СИБ-2015. Москва. НОУ ВО МФЮА, 2015. С. 58-64.

7. *Панин Д.Н., Железнова П.В., Лапаева О.С.,* Новикова Д.Д. Цифровая безопасность умных городов // Международный научно-исследовательский журнал. 2019. № 11 (89) Ч. 1. С. 31-34.