Научно-технический журнал

H&ES RESEARCH

Scientific and Technical Journal

НАУКОЕМКИЕ ТЕХНОЛОГИИ В КОСМИЧЕСКИХ ИССЛЕДОВАНИЯХ ЗЕМЛИ

HIGH TECHNOLOGIES IN EARTH SPACE RESEARCH

Журнал **H&ES Research** издается с 2009 года, освещает достижения и проблемы российских инфокоммуникаций, внедрение последних достижений отрасли в автоматизированных системах управления, развитие технологий в информационной безопасности, исследования космоса, развитие спутникового телевидения и навигации, исследование Арктики. Особое место в издании уделено результатам научных исследований молодых ученых в области создания новых средств и технологий космических исследований Земли.

Журнал H&ES Research входит в перечень изданий, публикации в которых учитываются Высшей аттестационной комиссиейРоссии (ВАК РФ), в систему российского индекса научного цитирования (РИНЦ), а также включен в Международный классификатор периодических изданий.

Тематика публикуемых статей в соответствии с перечнем группспециальностей научных работников по Номенклатуре специальностей:

- 2.2.15 Системы, сети и устройства телекоммуникаций (техн. науки)
- 2.3.1 Системный анализ, управление и обработка информации (техн. науки)
- 2.3.5 Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (техн. науки)
- 2.3.6 Методы и системы защиты информации, информационная безопасность (техн. науки)
- 2.5.13 Проектирование, конструкция и производство летательныхъ аппаратов (техн. науки)
- 2.5.16 Динамика, баллистика, управление движением летательных аппаратов (техн. науки)

ИНДЕКСИРОВАНИЕ ЖУРНАЛА H&ES RESEARCH

H&ES Research is published since 2009. The journal covers achievements and problems of the Russian infocommunication, introduction of the last achievements of branch in automated control systems, development of technologies in information security, space researches, development of satellite television and navigation, research of the Arctic. The special place in the edition is given to results of scientific researches of young scientists in the field of creation of new means and technologies of space researches of Earth.

The journal H&ES Research is included in the list of scientific publications, recommended Higher Attestation Commission Russian Ministry of Education for the publication of scientific works, which reflect the basic scientific content of candidate and doctoral theses. IF of the Russian Science Citation Index.

Subject of published articles according to the list of branches of science and groups of scientific specialties in accordance with the specialties:

- 2.2.15 Telecommunication systems, networks and devices
- 2.3.1 System analysis, management and information processing
- 2.3.5 Mathematical and software support for computing systems, complexes and computer networks
- 2.3.6 Methods and systems of information security
- 2.5.13 Design, construction and production of aircraft
- 2.5.16 Dynamics, ballistics, aircraft motion control

JOURNAL H&ES RESEARCH INDEXING

• NEICON • CyberLenika (Open Science) • Google Scholar • OCLC WorldCat • Ulrich's Periodicals Directory • • Bielefeld Academic Search Engine (BASE) • eLIBRARY.RU • Registry of Open Access Repositories (ROAR)

Все номера журнала находятся в свободном доступе на сайте журнала **www.hes.ru** и библиотеке **elibrary.ru.**

Всем авторам, желающим разместить научную статью в журнале, необходимо оформить ее согласно требованиям и направить материалы на электронную почту: **HT-ESResearch@yandex.ru.**
С требованиями можно ознакомиться на сайте: **www.H-ES.ru.**

Язык публикаций: русский, английский.
Периодичность выхода – 6 номеров в год.
Свидетельство о регистрации СМИ ПИ № ФС 77-60899 от 02.03.2015
Территория распространения: Российская Федерация, зарубежные страны

Тираж 1000 экз. Цена 1000 руб.
Плата с аспирантов за публикацию рукописи не взимается.

All issues of the journal are in a free access on a site of the journal **www.hes.ru** and **elibrary.ru.**

All authors wishing to post a scientific article in the journal, you must register it according to the requirements and send the materials to your email: **HT-ESResearch@yandex.ru.**
The requirements are available on the website: **www.H-ES.ru.**

Language of publications: Russian, English.
Periodicity – 6 issues per year.
Media Registration Certificate PI No. FS77-60899. Date of issue: March 2, 2015.
Distribution Territory: Russian Federation, foreign countries

Circulation of 1000 copies. Price of 1000 Rur.
Postgraduate students for publication of the manuscript will not be charged

ISSN 2409-5419 (Print) ISSN 2412-1363 (Online) WWW.H-ES.RU Дата выхода: 25 февраля 2023

1 T.15 2023

Учредитель:
ООО "ИД Медиа Паблишер"

Издатель:
ДЫМКОВА С.С.

Главный редактор:
ЛЕГКОВ К.Е.

Адрес издателя:
111024, Россия, Москва,
ул. Авиамоторная, д. 8, корп. 1, офис 323.

Адрес редакции:
194044, Россия, Санкт-Петербург,
Лесной Проспект, 34-36, к. 1,
Тел.: +7(911) 194-12-42.

Адрес типографии:
Россия, Москва, ул. Складочная, д. 3, кор. 6.

СОДЕРЖАНИЕ

CONTENTS

Founder:
"Media Publisher", LLC

Publisher:
DYMKOVA S.S.

Address of publisher:
111024, Russia, Moscow,
st. Aviamotornaya, 8, bild. 1, office 323

Address of edition:
194044, Russia, St. Petersburg,
Lesnoy av., 34-36, h.1,
Phone: +7 (911) 194-12-42.

Address of printing house:
Russia, Moscow, st. Skladochnaya, 3, h. 6

doi: 10.36724/2409-5419-2023-15-1-4-9

ПРОБЛЕМЫ РЕАЛИЗАЦИИ ОБЛАЧНЫХ ТЕХНОЛОГИЙ В НИЗКООРБИТАЛЬНЫХ КОСМИЧЕСКИХ СИСТЕМАХ ПЕРСОНАЛЬНОЙ СПУТНИКОВОЙ СВЯЗИ И ПЕРЕДАЧИ ДАННЫХ И ТЕЛЕУПРАВЛЕНИЕ

КАРПЕНКО
Елена Анатольевна[1]

КРАВЧИНА
Максим Витальевич[2]

СЕРГИЕНКО
Алексей Викторович [3]

АНДРАШИТОВ
Дмитрий Сергеевич[4]

Сведения об авторах:

[1] Северо-Кавказский филиал ордена Трудового Красного Знамени ФГБОУ ВО "Московский технический университет связи и информатики", г. Ростов-на-Дону, Россия

[2] Донской государственный технический университет, г. Ростов-на-дону, Россия

[3] Донской государственный технический университет, г. Ростов-на-дону, Россия

[4] Военная академия Ракетных войск стратегического назначения имени Петра Великого Балашиха, Россия

Работа подготовлена в рамках научной темы "Разработка беспилотных технологий на основе комплексной поэтапной оптимизации с редукцией экстремальных задач и инструментов нейро-нечеткого моделирования" (FZNE-2022-0006).

АННОТАЦИЯ

Введение: В настоящее время в развитых странах мира сформировался устойчивый тренд на целенаправленную информатизацию разведывательного сообщества. Это объясняется, тем, что в современных условиях политической, экономической и технологической обстановки информационные технологии рассматриваются в качестве инновационного инструмента повышения возможностей и конкурентоспособности государства при одновременной экономии средств. **Цель исследования:** Актуальным в настоящее время является создание глобальной низкоорбитальной космической информационной системы (ГКНИС), основу которой составляет многоспутниковая орбитальная группировка малогабаритных космических аппаратов (МГКА) по принципу кластерной организации. Существенным отличием ГКНИС, как облачной сети, является то, что существующие облачные технологии адаптированы для наземных немобильных компьютеров, среди которых значительная часть играет роль не вычислителей, а просто накопителей и "держателей" данных. Таким образом, возникает проблема обеспечения управляемости ГКНИС при ограниченности вычислительных ресурсов каждого из отдельных КА, которые составляют космическую систему. Вряд ли будет эффективным путь решения этой проблемы путем включения в состав космической системы отдельных "больших" КА, масса и габариты которых позволят установить на борту мощную БЦВМ, чтобы позволить таким КА играть роль вычислительных узлов в рамках орбитальной системы. Очевидно, что экономическая состоятельность ГКНИС будет в первую очередь зависеть от возможности использовать в её составе КА именно малой массы и габаритов, унифицированных по платформе служебных систем и производимых поточно конвейерным способом. Разработка подобной технологии может оказаться ключевым моментом в создании экономически состоятельной ГКНИС. **Результаты:** В работе рассматриваются проблемы организации управления многофункциональной системой спутниковой связи при решении задач информационного обмена с использованием облачных технологий.

КЛЮЧЕВЫЕ СЛОВА: *спутниковая связь, облачные технологии, передача данных, глобальные навигационные спутниковые системы.*

Для цитирования: *Карпенко Е.А., Кравчина М.В., Сергиенко А.В., Андрашитов Д.С.* Проблемы реализации облачных технологий в низкоорбитальных космических системах персональной спутниковой связи и передачи данных и телеуправление // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 4-9. doi: 10.36724/2409-5419-2023-15-1-4-9

Введение

В области использования информации наблюдается тенденция расширения численности пользователей при одновременном понижении их ранга (от стратегического уровня к тактическому, вплоть до единичного потребителя). В настоящее время в развитых странах мира сформировался устойчивый тренд на целенаправленную информатизацию разведывательного сообщества. Это объясняется тем, что в современных условиях политической, экономической и технологической обстановки информационные технологии рассматриваются в качестве инновационного инструмента повышения возможностей и конкурентоспособности государства при одновременной экономии средств [1].

Такие пользователи отличаются высоким уровнем требований к детальности, обзорности, периодичности и оперативности доставки информации [2], которые не обеспечиваются современными наземными системами управления и мониторинга. Для удовлетворения этих требований все чаще используются системы аэрокосмического мониторинга регионального и глобального масштаба, включающие различные оптоэлектронные, радиолокационные и лазерные съемочные системы, глобальные навигационные спутниковые системы ГЛОНАСС и ГЛОНАСС/GPS, [6-9] а также современные спутниковые системы связи [10-15].

Постановка задачи

На данный момент отечественная космическая группировка по дистанционному зондированию Земли, включает всего лишь три спутника: гидрометеорологический спутник второго поколения «Электро-Л», гидрометеорологический и океанографический «Метеор-М» №1 и космический аппарат (КА) дистанционного зондирования Земли «Ресурс-ДК 1», а также спутник «Монитор-Э», который практически не эксплуатируется по целевому назначению [3]. Таким образом, объем требований к данным спутникам никак не сопоставим с их возможностями. Весьма важной остается и проблема отсутствия системы единых форматов хранения и представления данных по дистанционному зондированию земли и информационных продуктов в ведомственных информационных системах.

Решить указанные недостатки позволяет использование данных, полученных с иностранных космических аппаратов, в частности МЧС получает информацию с аппаратов *Terra*, *Aqua*, *SPOT* 4/5 и др. Однако оплата за предоставляемые услуги довольно высока.

Актуальным в настоящее время является создание глобальной низкоорбитальной космической информационной системы (ГКНИС), основу которой составляет многоспутниковая орбитальная группировка малогабаритных космических аппаратов (МГКА) по принципу кластерной организации.

Под кластером МГКА понимается иерархически структурированная группа совместно выполняющих целевую задачу космических аппаратов, воспринимаемая как потребителем, так и наземным комплексом управления, как единый объект. Иерархию кластера составляют малогабаритный космический аппарат (МКА)-лидер, решающий задачи организации межспутникового взаимодействия и информационного обмена с наземным комплексом управления и потребителем, и несколько КА-ведомых, в рамках кластера решающих целевые задачи. Ключевым фактором успешного функционирования кластера МГКА является его способность решать самостоятельно часть задач, которые в настоящее время по традиционной технологии управления КА решаются наземными службами.

Прежде всего это относится к возможности предварительной обработки полученной информации. Непрерывное (или высокопериодическое) наблюдение в режиме слежения, которое предполагается основным режимом работы ГКНИС на базе МГКА позволяет передавать потребителю не всю информацию, а лишь ту её часть, которая содержит существенные с точки зрения потребителя изменения с момента прошедшего наблюдения. Возможность резко снизить объем передаваемых данных является важнейшим условием реализуемости ГКНИС, так как канал связи «космос-земля» довольно энергоемок. Но осуществить такой режим работы можно лишь в случае способности бортовой аппаратуры отдельного космического аппарата произвести значительный объем вычислений в ходе предварительной обработки информации. Кроме того, такую обработку информации необходимо осуществлять крайне оперативно, фактически в режиме реального времени.

Это приводит к возникновению еще одной проблемы – ограниченности вычислительных ресурсов отдельно взятого МКА вне зависимости от его специализации в иерархии кластера. Очевидно, что как МКА-лидер, так и МКА-ведомый нуждаются в значительных вычислительных ресурсах. Трудности создания высокопроизводительного компьютера малых габаритов, массы и низкого уровня энергопотребления, способного стабильно работать в условиях космического полета, предопределяют трудность создания МКА, способного решить требуемые вычислительные задачи в ходе своего функционирования.

При этом трудность усугубляется тем, что если системные эффекты и позволяют снизить требования к массе и габаритам целевой аппаратуры, то необходимость их реализации в ходе многозвенного взаимодействия динамичных, удаленных друг от друга объектов наоборот предъявляет дополнительные требования к количеству и скорости обработки информации в каждом отдельном космическом аппарате. В результате при реализации ГКНИС её создателям придется иметь дело с ещё одним барьером – вычислительным, который обусловлен низкой производительностью вычислительных систем каждого отдельного МКА орбитальной группировки.

Важно отметить, что преодоление этого барьера не зависит от степени энерговооруженности КА – вычисления являются не энергоемким процессом, а производительность компьютера слабо зависит от мощности системы электропитания.

Таким образом, возникает проблема обеспечения управляемости ГКНИС при ограниченности вычислительных ресурсов каждого из отдельных КА, которые составляют космическую систему. Вряд ли будет эффективным путь решения этой проблемы путем включения в состав космической

системы отдельных «больших» КА, масса и габариты которых позволят установить на борту мощную БЦВМ, чтобы позволить таким КА играть роль вычислительных узлов в рамках орбитальной системы.

Очевидно, что экономическая состоятельность ГКНИС будет в первую очередь зависеть от возможности использовать в её составе КА именно малой массы и габаритов, унифицированных по платформе служебных систем и производимых поточно конвейерным способом.

Предпосылками решения этой проблемы при сохранении унификации МКА в орбитальной системе может быть учет двух важных особенностей ГКНИС.

Первая особенность ГКНИС заключается в том, что её МКА в силу своей многочисленности действуют в зоне прямой видимости друг друга в условиях отсутствия помех со стороны атмосферы Земли. Максимальная дальность взаимодействия двух КА низкоорбитальной группировки зависит от высоты орбиты и высоты верхней границы атмосферы, ниже которой помехи взаимной связи приходится принимать в расчет.

В простейшем случае максимальная дальность $D_{КА}$ между двумя КА ($КА_1$ – $КА_2$) на круговой орбите вычисляется по формуле, представлено на рисунке 1.

$$D_{КА} = \sqrt{\left(R_З + H_{КА}\right)^2 - \left(R_З + H_А\right)^2}, \qquad (1)$$

где $R_З$ = 6371 км – радиус Земли; $H_{КА}$ – высота орбиты КА, км; $H_А$ – высота верхней границы атмосферы, км.

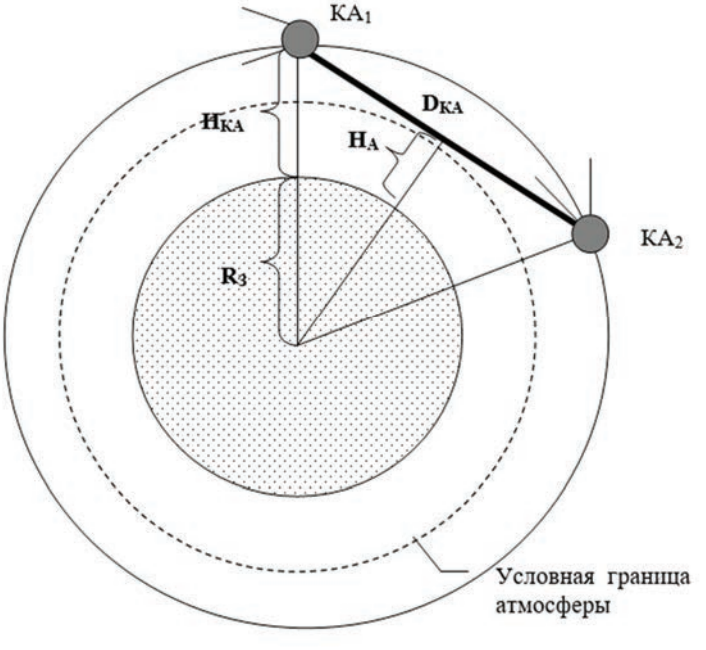


Рис. 1. Максимальная дальность взаимодействия между КА орбитальной группировки без помех со стороны атмосферы

Для случая $H_А$=150 км можно представить зависимость дальности $D_{КА}$ от высоты орбиты $H_{КА}$, в качестве которой взяты типичные высоты для низкоорбитальной системы мониторинга, представлено на рисунке 2.

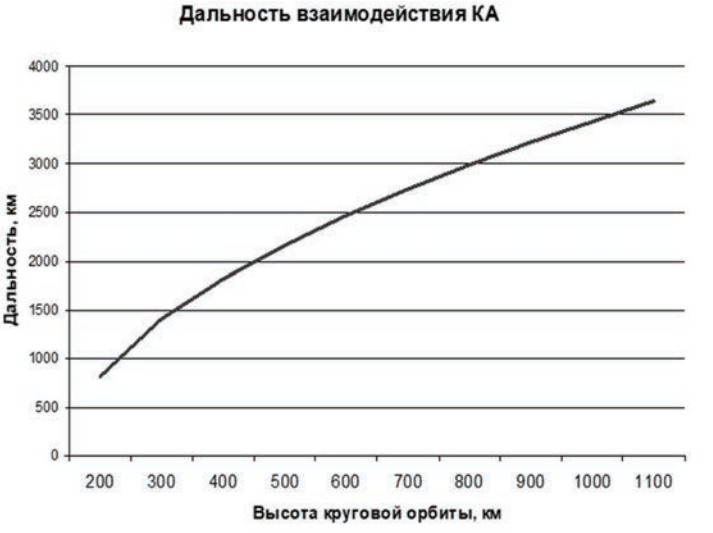


Рис. 2. Зависимость дальности взаимодействия КА при условной высоте верхней границы атмосферы $H_А$=150 км

Несмотря на значительные расстояния, условия космического полета позволяют организовать между космическими аппаратами обмен данными при незначительных энергетических затратах. Примером может быть канал межспутниковой связи космической системы Iridium.

В этой системе каждый КА орбитальной группировки имеет радиолинии связи с двумя соседними КА, находящимися в одной орбитальной плоскости с ним, и двумя КА в соседних (слева и справа) орбитальных плоскостях. Для поддержания межспутниковой связи на каждом КА имеются четыре щелевые антенные решетки с коэффициентом усиления 36 дБ.

Точность управления диаграммой направленности каждой антенны составляет ±5°. Используется полоса частот шириной 200 МГц в диапазоне 23,18-23,38 ГГц. Для исключения взаимных помех в межспутниковых каналах связи полоса частот шириной 200 МГц разбита на 8 отдельных частотных полос, которые образуют отдельные каналы связи. Скорость передачи информации в каждом канале 25 Мбит/с.

В радиолинии применяется временное разделение каналов. При помощи фазовой манипуляции ФМ-4 производится кодирование информации, которое обеспечивает сжатие речевой информации в цифровом виде. Информация о сжатии, а также сигналы циклической и тактовой синхронизации передаются по каналу управления, для чего в радиолинии «КА-абонент» задействовано 4 радиоканала.

Коэффициент сжатия. информации (2,2/1) позволяет обеспечить передачу в радиолинии «КА-абонент» 55 речевых каналов на 25 несущих частотах.

При передаче радиотелефонной информации вероятность ошибки на бит не выше 0,001, при передаче цифровых данных – 0,000001. Каждый канал межспутниковой линии связи поддерживает 600 телефонных каналов без сжатия (1300 каналов при коэффициенте сжатия информации 2,2/1).

Современные технологии связи, в том числе в области миллиметровых волн и особенно в оптическом диапазоне (по лазерному лучу) способны за пределами атмосферы обеспечить высокую пропускную способность обмена данными. Важно, что при этом энергоемкость типичного МКА является достаточной для организации межспутникового канала связи.

Вторая особенность заключается в том, что объекты наблюдения, которые могут интересовать потребителя ГКНИС, на поверхности Земли расположены довольно неравномерно, иллюстрируется на рисунке 3.

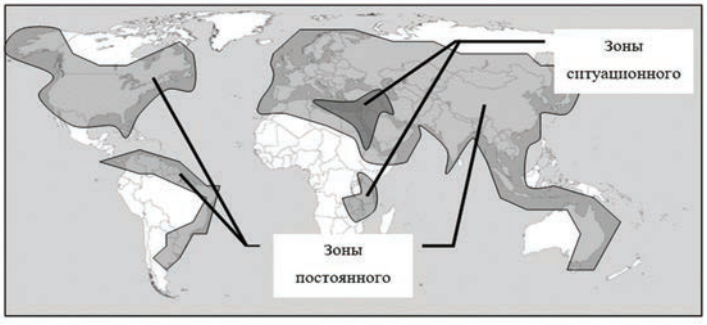


Рис. 3. Зоны интенсивности наблюдения ГКНИС (условно)

При этом зоны интереса потребителя могут иметь как постоянный характер, что позволяет планировать использование ресурсов ГКНИС заблаговременно, так и ситуационный характер, что вызывает необходимость рассчитывать технические возможности ГКНИС для случая оперативного перенацеливания внимания на район, ранее не наблюдаемый, либо для случая повышения интенсивности наблюдения в зоне постоянного интереса.

Неравномерность интенсивности использования целевой аппаратуры космических аппаратов ГКНИС в сочетании с технологиями межспутникового обмена данными предоставляют возможность преодолеть «вычислительный барьер» отдельных КА путем организации взаимодействия между ними в рамках всей системы.

Основой такого взаимодействия является превращение всей совокупности МКА ГКНИС в единую вычислительную сеть, в которой перегруженный вычислениями объект имеет возможность передать часть своих функций по обработке информации другому объекту, который в данный момент используется не в полной мере. Эта идея полностью соответствует сущности сетевым информационным технологиям.

Такие сетевые информационные технологии обеспечивают при развёртывании орбитальной группировки КА достижение потенциальных возможностей космических систем по глобальности и оперативности, по доступу к информации о любых пространственных объектах в космическом и воздушном пространствах, на суше и море.

Сегодня к таким технологиям, в первую очередь, относится получение и анализ изображений с низким, средним и высоким пространственным разрешением элементов земного рельефа, населенных пунктов и водной поверхности в различной цветовой гамме из космоса. Для задач оперативного контроля состояния природных ресурсов и экономически важных и/или опасных объектов РФ методы дистанционного зондирования Земли позволяют получать объективные данные в режиме реального времени с больших площадей, производить эффективную оценку ситуации на данной территории и оперативного принятия решения в чрезвычайной ситуации.

Эти методы дают хороший экономический результат в части рационального использования предоставляемых ресурсов, а также с более высокой достоверностью позволяют производить учет материального ущерба, причиненного различного рода катаклизмами. Технологии дистанционного зондирования Земли позволяют создавать аналитические модели суточного прогнозирования на основе физических законов, но для прогнозирования развития ситуации требуются также данные за большие периоды времени, плюс дополнительная гидрометеорологическая информация, наземные замеры.

Вся эта информация в комплексе даёт актуальный, объективный, прозрачный срез данных, необходимый для всех заинтересованных структур. Создание подобной сетевой архитектуры возможно только при широком использовании космических средств, при этом глобальность и оперативность могут обеспечить только низкоорбитальные КС с соответствующими задачам орбитальной группировки КА.

В настоящее время к такой сетевой технологии относится, так называемая, облачная или рассеянная технология [4]. Она представляет собой технологию обработки данных, в которой компьютерные ресурсы и мощности предоставляются «КА-абоненту» как Интернет-сервис. «КА-абонент» имеет доступ к собственным данным, но не может управлять и не должен заботиться об инфраструктуре, операционной системе и собственно программном обеспечении, с которым он работает.

С этой точки зрения всю систему ГКНИС, как совокупность взаимодействующих кластеров МГКА можно представить в виде единой сети. Ведь структурной основой реализации облачной технологии для наземных средств является кластерное построение сети компьютеров, которое в космосе может быть реализовано в виде кластерного построения многоспутниковой орбитальной группировки МКА, каждый из которых с точки зрения управления представляет собой периферийный компьютер ограниченной мощности в единой сети, предназначенной для решения ресурсоемких задач.

Такое описание кластеров МКА практически соответствует определению кластера в терминологии облачной технологий как типа параллельной или распределенной вычислительной системы, состоящей из набора соединенных между собой и работающих совместно однородных компьютеров, которые рассматриваются как «единый интегрированный вычислительный ресурс» (*Single System Image*, *SSI*).

Особенностью облачных технологий является обеспечение совместного использования ресурсов, распределенных по разным административным и географическим доменам, что применительно к орбитальной группировке, позволяет организовать информационный обмен не только между МКА одного кластера, но и между кластерами. Естественным для таких условий является решение передать часть ресурсоемких задач обработки информации от кластера, непосредственно

решающего задачу наблюдения, к «незагруженным» на данный момент кластерам МКА.

Такой подход требует представления всей орбитальной группировки МКА в виде единой иерархичной пространственно-распределенной сети разнородных вычислительных устройств, предназначенной для решения задач, вычислительный объем которых превышает не только возможности бортового цифрового вычислительного комплекса отдельного МКА, но и вычислительные возможности отдельного кластера МКА такой орбитальной группировки.

Существенным отличием ГКНИС, как облачной сети, является то, что существующие облачные технологии адаптированы для наземных немобильных компьютеров, среди которых значительная часть играет роль не вычислителей, а просто накопителей и «держателей» данных.

В терминологии облачных вычислений такие компьютеры относят к первому уровню иерархии, в которой нулевой уровень представлен датчиками получения первичной информации, а второй уровень – периферийные компьютеры, собственно, и решающие задачу обработки информации. В ГКНИС роль нулевого уровня облачной иерархии играют МКА-ведомые, решающие целевую задачу по получению первичной информации. Они же, вместе с МКА-лидерами кластеров могут играть роль объектов второго уровня иерархии, обрабатывающих полученную информацию.

Трудность возникает в том, что функции первого уровня иерархии, как хранилища накопленной информации, готовой для дальнейшей обработки, в ГКНИС возлагать на космические аппараты непродуктивно, а использование в качестве промежуточного хранилища наземной аппаратуры немедленно влечет за собой недопустимое увеличение объема передачи данных по линии «космос-земля-космос».

Заключение

В этой связи актуальным становится исследование вопроса о создании на базе мобильных космических носителей хранилища данных в виде управляемой «информационной волны». Предполагается, что первичные данные от целевых МКА (нулевого уровня облачной иерархии) будут переданы на некоторую совокупность МКА, территориально близких к КА нулевого уровня и временно играющих роль хранилища информации, к которому могут обращаться МКА второго уровня иерархии, решающие задачу обработки данных.

Поскольку КА-хранилища через определенный промежуток времени окажутся в зоне интересов потребителя и будут привлечены к выполнению целевых задач, они обязаны передать свою часть базовой информации следующим за ними свободным КА. Процесс должен напоминать волну, перетекающую от одного носителя к другому с целью удержать гребень этой волны над теми участками Земли, где основная масса МКА ГКНИС не задействована для решения целевых задач.

С учетом того, что ГКНИС должна быть рассчитана на обслуживание оперативно возникающих зон ситуативного интереса потребителя, возникает необходимость в разработке не только технологии формирования и поддержания «информационной волны», но и технологии её реконфигурации и переноса на другой участок «информационного затишья».

В рамках существующих облачных технологий подобные задачи пока не решены, поскольку специфика наземных компьютеров не вызывала необходимости в их разработке. В то же время разработка подобной технологии может оказаться ключевым моментом в создании экономически состоятельной ГКНИС.

Литература

1. *Кондратьев А., Затуливетер Ю.* Облачное будущее по-американски // Независимое военное обозрение. 2013. №4(745). С. 8-9.

2. *Макаренко Д. М., Потюпкин А. Ю.* Проблемы реализации GRID-технологий для решения задач информационного обмена в глобальной низкоорбитальной космической информационной системе // Вестник МАТИ. 2012. №19. С. 242-249.

3. *Галькевич А. И.* Концепция и перспектива создания и использования глобальной космической низкоорбитальной информационной системы "Космонет" для информационного обеспечения техносферной безопасности // Технологии техносферной безопасности. 2011. №4(38). С. 1-7.

4. *Широкова Е. А.* Облачные технологии // Современные тенденции технических наук. 2011. С. 30-33.

5. *Cherckesova L.V., Safaryan O.A., Trubchik I., Chumakov V., Yukhnov V.I., Yengibaryan I.A.* Modification and optimization of Miller – Rabin simplicity test algorithm implemented by parallel computation. В сборнике: IOP Conference Series: Materials Science and Engineering. Сер. "International Scientific and Practical Conference Environmental Risks and Safety in Mechanical Engineering, ERSME 2020" 2020. С. 012064.

6. *Platonov S.A., Platonov A.V., Postnikov M.E., Khadonova S.V., Dymkova S.S.* Using global navigation satellite systems to solve complex application problems // В сборнике: 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019. 2019. С. 8706807.

7. *Dymkova S.S.* Conjunction and synchronization methods of earth satellite images with local cartographic data // В сборнике: 2020 Systems of Signals Generating and Processing in the Field of on Board Communications. 2020. С. 9078561.

8. *Dymkova S.S.* Earth observation and global navigation satellite systems analitical report part i (aviation and space) // Synchroinfo Journal. 2022. Т. 8. № 1. С. 30-41.

9. *Дымкова С.С.* Облачные iot платформы и приложения для оптимизационного управления транспортом // REDS: Телекоммуникационные устройства и системы. 2020. Т. 10. № 4. С. 39-50.

10. *Алешин В.С.* Оценка реализуемости активной фазированной антенной решётки терминала системы спутниковой связи "Экспресс-РВ" // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 8. С. 13-21.

11. *Алешин В.С., Догаев С.Г.* Задержки распространения сигналов в сетях спутниковой связи // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 5. С. 4-11.

12. *Аджемов С.С., Рюмшин К.Ю., Чадов Т.А.* Декодер блочных турбокодов // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 9. С. 50-53.

13. *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.

14. *Шухардин А.Н., Шкорина А.В.* Методика оперативного оценивания вероятностей и сроков доставки сообщений в информационно-телекоммуникационных системах // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2021. № 1. С. 153-157.

15. *Алешин В.С.* Пять мифов о спутниковой связи // Телекоммуникации и информационные технологии. 2019. Т. 6. № 2. С. 5-11.

PROBLEMS OF CLOUD TECHNOLOGY IMPLEMENTATION IN LOW-ORBIT SPACE SYSTEMS OF PERSONAL SATELLITE COMMUNICATIONS AND DATA TRANSMISSION AND TELECONTROL

ELENA A. KARPENKO
Rostov-on-Don, Russia

MAKSIM V. KRAVCHINA
Rostov-on-Don, Russia

ALEXEY V. SERGIENKO
Rostov-on-Don, Russia

DMITRIY S. ANDRASHITOV
Balashikha, Russia

KEYWORDS: *satellite communication, cloud technology, data transfer, global low-orbit space information system*

ABSTRACT

Introduction: Currently, a steady trend has been formed for targeted informatization of the intelligence community. This is explained by the fact that in the current conditions of the political, economic and technological environment, information technologies are considered as an innovative tool for increasing the capabilities and competitiveness of the state while saving money. **Purpose of the study:** At present, the creation of a global low-orbit space information system, which is based on a multi-satellite orbital constellation of small-sized spacecraft based on the principle of cluster organization, is currently relevant. The essential difference between the global low-orbit space information systemas a cloud network is that the existing cloud technologies are adapted for terrestrial non-mobile computers, among which a significant part plays the role of not computers, but simply data storage devices and "holders". Thus, the problem arises of ensuring the controllability of the GKNIS with limited computing resources for each of the individual spacecraft that make up the space system. It is unlikely that there will be an effective way to solve this problem by including separate "large" spacecraft in the space system, the mass and dimensions of which will allow installing a powerful onboard computer on board to allow such spacecraft to play the role of computing nodes within the orbital system. It is obvious that the economic viability of the GKNIS will primarily depend on the ability to use in its composition spacecraft of precisely small mass and dimensions, unified by the platform of service systems and produced in-line by a conveyor method. The development of such a technology could be a key moment in the creation of an economically viable global low-orbit space information system. **Results:** The paper deals with the problems of managing the management of a multifunctional satellite communication system in solving problems of information exchange using cloud technologies.

REFERENCES

1. A. Kondrat'ev, Yu. Zatuliveter. Cloudy American future. *Independent military review.* 2013. No.4(745), pp. 8-9.
2. D.M. Makarenkoo, A.Yu. Potypkin. Problems of implementation of GRID-technologies for solving problems of information exchange in the global low-orbit space information system. *Vestnik MATI*. 2012. No.19, pp. 242-249.
3. A. I. Galkevich. The concept and prospects for the creation and use of global space low-orbit information system "Cosmonet" for information support of technosphere security. *Technospheric safety technologies.* 2011. No.4(38), pp. 1-7.
4. E. A. Shirokova. Cloud technologies. Modern trends in engineering sciences. 2011, pp. 30-33.
5. L.V. Cherckesova, O.A. Safaryan, I. Trubchik, V. Chumakov, V.I. Yukhnov, I.A. Yengibaryan. Modification and optimization of Miller - Rabin simplicity test algorithm implemented by parallel computation. *IOP Conference Series: Materials Science and Engineering. Cep. "International Scientific and Practical Conference Environmental Risks and Safety in Mechanical Engineering, ERSME 2020"* 2020. P. 012064.
6. S.A. Platonov, A.V. Platonov, M.E.Postnikov, S.V. Khadonova, S.S. Dymkova. Using global navigation satellite systems to solve complex application problems. *2019 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2019*. 2019. P. 8706807.
7. S.S. Dymkova. Conjunction and synchronization methods of earth satellite images with local cartographic data. *2020 Systems of Signals Generating and Processing in the Field of on Board Communications. 2020.* P. 9078561.
8. S.S. Dymkova. Earth observation and global navigation satellite systems analitical report part i (aviation and space). *Synchroinfo Journal*. 2022. Vol. 8. No. 1, pp. 30-41.
9. S.S. Dymkova. Облачные iot платформы и приложения для оптимизационного управления транспортом. *REDS*. 2020. Vol. 10. No. 4, pp. 39-50.
10. V. S. Aleshin. Estimation of an active phased antenna array feasibility of the satellite communication system "Express-RV" terminal. *T-Comm*. 2021. Vol. 15. No. 8, pp. 13-21.
11. V. S. Aleshin., S. G. Dogaev. Signal propagation delays in satellite networks. T-Comm. 2019. Vol. 13. No. 5, pp. 4-11.
12. S.S. Adzhemov, K.Yu. Ryumshin, T.A. Chadov. Block Turbo Code Decoder. *T-Comm*. 2018. Vol. 12. No. 9, pp. 50-53.
13. S.V. Pavlov, V.A. Dokuchaev, S.S.Mytenkov. Model of a fuzzy dynamic decision support system. *T-Comm*. 2020. Vol. 14. No. 9, pp. 43-47.
14. A.N. Shukhardin, A.V. Shkorina. Methods of operational estimation of probabilities and terms of message delivery in information and telecommunication systems. *Proceedings of the North Caucasian branch of MTUCI.* 2021. No. 1, pp. 153-157.
15. V.S. Aleshin. Five myths about satellite communications. Telecommunications and information technologies. 2019. Vol. 6. No. 2, pp. 5-11.

INFORMATION ABOUT AUTHORS:

Karpenko E.A., North Caucasus branch of Moscow Technical University of Communications and Informatics, Rostov-on-Don, Russia
Kravchina M.V., Don State Technical University, Rostov-on-Don, Russia
Sergienko A.V., Don State Technical University, Rostov-on-Don, Russia
Andrashitov D.S., Peter the Great Military Academy of Strategic Missile Forces, Balashikha, Russia

For citation: Karpenko E.A., Kravchina M.V., Sergienko A.V., Andrashitov D.S. Problems of cloud technology implementation in low-orbit space systems of personal satellite communications and data transmission and telecontrol. *H&ES Reserch*. 2023. Vol. 15. No 1. P. 4-9.
doi: 10.36724/2409-5419-2023-15-1-4-9 (In Rus)

doi: 10.36724/2409-5419-2022-15-1-10-20

МЕТОД УДАЛЕННОГО МОНИТОРИНГА ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ СРЕДСТВ СВЯЗИ И НАВИГАЦИОННОГО ОБОРУДОВАНИЯ РОСМОРРЕЧФЛОТА

АЛЛАКИН
Владимир Васильевич[1]

БУДКО
Никита Павлович [2]

ГОЛЮНОВ
Михаил Валерьевич [3]

КАРЕТНИКОВ
Владимир Владимирович [4]

АННОТАЦИЯ

Актуальность: наиболее действенным подходом к снижению аварийности и обеспечению безопасного судоходства на внутренних водных путях России зарекомендовал себя переход от лоцманского к инструментальному судовождению, а также внедрение в процессе управления движением судов и их удаленного мониторинга более совершенных информационно-телекоммуникационных технологий, что ведет к созданию в Росморречфлоте иерархических систем на основе ситуационных центров, отвечающих за безопасность судовождения. **Цель работы:** разработка метода удаленного мониторинга функционального состояния средств связи и навигационного оборудования. **Используемые методы:** для реализации метода удаленного мониторинга наиболее применимы средства мониторинга с использованием безэкипажных судов. **Новизна:** предложенный метод позволяет использовать в качестве объектов контроля и мониторинга широко используемые на берегу и на водных судах радиотехнические и радионавигационные средства, излучающие в радио или оптическом диапазоне волн, а в качестве средства мониторинга – маломерные безэкипажные водные суда различного класса, с размещением на них бортовых автоматизированных измерительных комплексов, имеющих под решаемые задачи сменное контрольно-измерительное оборудование. **Полученный результат:** в разработанном методе процесс мониторинга проводится по следующим этапам: предварительно для осуществления процедуры телеизмерений производят подготовку и ввод в береговой и бортовой автоматизированные измерительные комплексы исходных данных; на первом этапе на борту средства мониторинга производятся измерения доступных дистанционно параметров дальней и ближней зоны подконтрольных объектов с передачей аварийных сигналов на берег; на втором этапе – собранная измерительная информация обрабатывается береговым автоматизированным измерительным комплексом с установлением классов технического состояния объектов мониторинга; в завершении производится подготовка отчётов в интересах ситуационного центра управления движением судов.

КЛЮЧЕВЫЕ СЛОВА: *автоматизированный измерительный комплекс, измерительная информация, мониторинг, средства радиосвязи, безэкипажное водное судно*

Сведения об авторах:

[1] аспирант, ФГБОУ ВО "Государственный университет морского и речного флота имени адмирала С.О. Макарова". г. Санкт-Петербург, Россия
vladimir@duduh.ru

[2] аспирант, ФГБОУ ВО "Государственный университет морского и речного флота имени адмирала С.О. Макарова". г. Санкт-Петербург, Россия
budko62@mail.ru

[3] адъюнкт Военной академии связи, г. Санкт-Петербург, Россия
belka1213@mail.ru

[4] заведующий кафедрой судоходства на внутренних водных путях ФГБОУ ВО "Государственный университет морского и речного флота имени адмирала С.О. Макарова", доктор технических наук, доцент, г. Санкт-Петербург, Россия
kaf_svvp@gumrf.ru

Для цитирования: *Аллакин В. В., Будко Н. П., Голюнов М. В., Каретников В. В.* Метод удаленного мониторинга функционального состояния средств связи и навигационного оборудования Росморречфлота // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 10-20. doi: 10.36724/2409-5419-2023-15-1-10-20

Введение

Российская Федерация в настоящее время располагает самыми протяжёнными водными (речными, озерными и морскими) границами. Протяженность доступных для судоходства внутренних водных путей (ВВП) составляет более 100 тыс. км, что является самой разветвленной сетью водных коммуникаций в мире, причем более 60 тыс. км из которых безальтернативны [1].

Из [2, 3] известно, что одним из действенных подходов к сокращению аварийных ситуаций и повышению безопасности судоходства на ВВП РФ является замена лоцманского судовождения на инструментальную проводку судов, что требует активного внедрения автоматизированных систем управления движением судов (АСУ ДС) и совершенствования процедур удаленного мониторинга на основе новых информационно-телекоммуникационных технологий.

Для реализации данной концепции в Федеральном агентстве морского и речного транспорта Минтранса РФ (Росморречфлоте), на акваториях и водных коммуникациях (бассейнах рек) необходимо создание действенных иерархических систем, поддерживающих безопасность судовождения, например: речная информационная служба (РИС), корпоративная речная информационная система (КРИС), АСУ ДС и пр. В зависимости от размещения своих элементов (на берегу, на борту судна, на воде) данные системы могут реализовываться с использованием различных родов связи: от проводных (волоконно-оптических), спутниковых (Гонец, Инмарсат, Глобарстар и пр.) и сотовой связи стандарта *GSM*. Многообразие используемых телекоммуникационных систем и радиоканалов *GMDSS* (морской подвижной службы) позволяет оптимальным образом объединять бортовые, надводные и наземные радиоэлектронные средства (средства радиосвязи – СРС), а также средства навигационного оборудования (СНО) с ведомственными ситуационными центрами (СЦ) и региональными АСУ ДС на бассейнах ВВП РФ.

В последнее десятилетие все больше появляется публикаций, направленных на активное внедрение технологий робототехники и безэкипажного судовождения, а, следовательно, удаленного контроля (мониторинга) не только ТС безэкипажных водных судов (БЭВС), но и телеметрии географически распределенных береговых и бортовых СРС, а также стационарной и плавучей навигационной обстановки (СНО) в прибрежных морских районах (акваториях) и на ВВП РФ [1, 4-6].

При этом для дистанционного съема измерительной информации (ИИ) с СРС и СНО наиболее подходят средства мониторинга (СРМ), с применением БЭВС и размещением на них бортовых автоматизированных измерительных комплексов (АИК) со сменными измерительными модулями под решаемые задачи мониторинга сигналов оптического либо иных радиодиапазонов волн объектов контроля (ОК). Причем, из всего многообразия электромагнитного излучения (низкочастотное, радиоволновое, инфракрасное, видимое, ультрафиолетовое, рентгеновское, гамма-лучи) для дистанционного контроля СРС и СНО наиболее применимы оптический и радиоволновой методы. Поэтому в данной работе этим видам мониторинга уделено особое внимание.

Одним из наиболее близких по своей сущности к предлагаемому подходу для реализации удаленного мониторинга ТС ОК можно отнести способы проведения лётных проверок и настроек средств радиотехнического обеспечения (РТО) [7, 8], известные из авиационной отрасли. Однако к основному недостатку таких способов можно отнести задействование для него специального воздушного судна-лаборатории или беспилотного летательного аппарата (БПЛА) [7].

При этом отмечаются высокие затраты на проведение такого контроля выходных характеристик средств РТО и его низкая оперативность (один раз в год при проведении плановых проверок) [8]; чрезмерный расход канального ресурса (каналов воздушной радиосвязи), поскольку предполагает транслировать в направлении «борт-земля» всю получаемую по каждому параметру ОК ИИ независимо от наличия на нем аварийного режима.

Технической проблемой применимости такого подхода в интересах Росморречфлота является: неполный охват мониторингом всей географически распределенной системы стационарного и плавучего телекоммуникационного оборудования (СРС) и СНО на ВВП и в прибрежной морской зоне РФ; малая оперативность в проведении измерений, особенно в дальней зоне действия СРС и СНО; потребность в обратном канале связи «берег-борт» для настроек аварийных РЭС и СНО в ходе проведения цикла одного мониторинга в режиме *on-line*.

Также важно понимать, что способы контроля наземных и бортовых СРС и РТО, используемые в авиационной отрасли не всегда приемлемы для мониторинга береговых и бортовых (судовых) СРС и СНО, что связано, в том числе и с высотой подъема СРМ. В рассматриваемом случае использование БПЛА из-за высотности его применения может приводить к ошибкам при оценке пространственных параметров измерительных комплексов в условиях изменчивости высотного профиля береговой линии.

Для внедрения инструментальных методов навигации современный судоводитель должен получать навигационную информацию от нескольких надёжных и независимых источников, что будет способствовать безопасному судовождению и повышению эффективности АСУ ДС, включающей подсистему мониторинга технического состояния СРС и СНО.

Комплексное использование радиоканалов морской подвижной службы *GMDSS* позволит оптимизировать число используемых СРС и СНО для повышения безопасности на водном транспорте особенно в районах интенсивного движения акваторий портов, а также в «стеснённых» водах фарватеров, проливов, шлюзов, каналов, проходов и пр.

Разработка метода удаленного мониторинга функционального состояния средств связи и навигационного оборудования Росморречфлота

Предлагаемый метод удаленного мониторинга функционального состояния СРС и СНО в данной работе представлен блок-схемой алгоритма, показанным на рисунке 1.

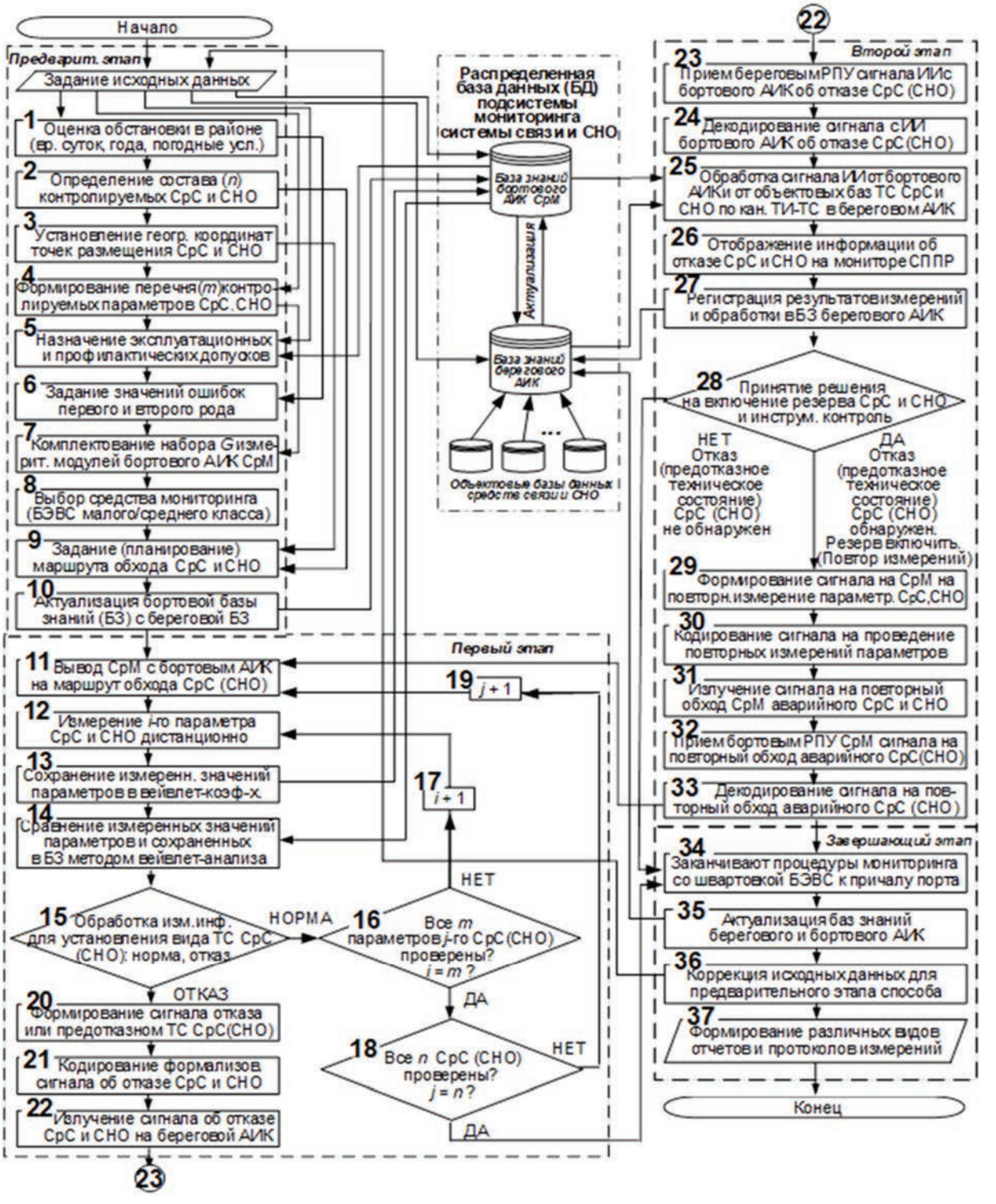


Рис. 1. Алгоритм метода удаленного мониторинга функционального состояния средств связи и навигационного оборудования Росморречфлота

В данном алгоритме в качестве СРМ используются – маломерные БЭВС или БПЛА среднего и малого класса, на борту которых должны быть размещены АИК, имеющие под решаемые задачи сменное контрольно-измерительное оборудование.

Процесс мониторинга в данном методе осуществляется поэтапно: предварительно, для осуществления процедуры телеизмерений производят подготовку и ввод исходных данных в береговой АИК и бортовой АИК; на первом этапе, на борту СРМ осуществляют измерения доступных дистанционно

параметров дальней и ближней зоны ОК, причем на берег передают только сигналы об аварийной ситуации; на втором этапе – собранная ИИ обрабатывается береговым АИК с установлением классов технического состояния ОК; в завершении производится подготовка отчётов в интересах СЦ ведомства (АСУ ДС).

Рассмотрим более подробно предложенный метод по его алгоритму (см. рис. 1).

Предварительный этап. Перед началом цикла мониторинга выполняют следующие действия в соответствие с блок-схемой алгоритма по шагам 1-10.

На шаге 1 определяют *условия проведения мониторинга* путем оценки обстановки в прибрежной морской зоне или на акватории бассейна ВВП РФ. При этом определяют период навигации (по времени года), климатические факторы, время суток, внешним воздействиям на СРС и СНО, интенсивность судоходства и пр. Оценивают обстановку визуально и с помощью инструментального метода: время суток – по фотоэлементам (по часам), осадки – гигрометру, время года – термометру, уровень радиопомех – по устройствам анализа помеховой обстановки и пр.

На шаге 2 устанавливают необходимый *объем (программу) мониторинга*, при этом с помощью базы данных (БД) АСУ ДС определяют состав n подвергаемых мониторингу СРС и СНО, влияющих на безопасность судовождения на данном участке бассейна ВВП РФ, с минимальным отклонением от оси хода водного судна днем и ночью, в условиях ограниченной видимости из-за метеоусловий, с учетов возможных воздействий помех и иных дестабилизирующих факторов (ДФ). Состав контролируемых СРС и СНО предложенным способом может меняться из-за категории ВВП, сроков работы СНО и судоходных гидротехнических сооружений, а также перечней судового хода.

На шаге 3 устанавливают *зону мониторинга* по географическим координатам точек размещения каждого j-го (j = 1, 2, …, n) стационарного или плавучего СРС и СНО с учётом их зон излучения и геопространственной информации. Причем к выбору точек расположения базовых станций (БС) АИС, а также СНО привлекаются специалисты лоцманской службы портов и представители инженерно-технических служб связи и СНО. Координаты развёртывания стационарных и плавучих СРС и СНО определяются условиями судовождения, разрешается совместное использование одной позиции несколькими средствами СНО с соблюдением требований по обеспечению их электромагнитной совместимости.

Если по условиям берегового ландшафта и водному бассейну типовое размещение средств СНО невозможно, допускают отступление от типового размещения с расчётом обеспечения их устойчивой работы в секторах с наибольшей интенсивностью движения судов (в том числе на встречных курсах). Несоответствие типовому расположению СРС и СНО может быть компенсировано эквивалентными мерами, обеспечивающими безопасность судовождения на ВВП.

На шаге 4 формируют состав m эксплуатационных параметров СРС и СНО, которые определяются в соответствии с тактико-техническими характеристиками (ТТХ) на них при разных режимах функционирования (*глубина мониторинга*). При этом состав параметров (i = 1, 2, …, m), входящих в процесс мониторинга для j-го СНО и СРС строго индивидуален. Одним из подходов при его формировании может быть использован коэффициент тяжести последствий при возникновении отказа (аварии) СРС (СНО) и «вклад» этих параметров в повышение надёжности ОК [1].

Этот коэффициент тяжести последствий будем называть коэффициентом значимости ($K_{зн}$) выбранного параметра, для определения которого (при включении параметра в процедуру мониторинга СРС и СНО) из всего множества параметров нормативно-технической документации (НТД) на объекте мониторинга будет применяться общий коэффициент значимости $K^{\Sigma}_{зн.i}$ контролируемого параметра в виде суммы назначаемых весов всех принятых по лингвистической шкале оценки показателей значимости $K_{зн}$, [9, 10]. Причём данный общий коэффициент значимости $K^{\Sigma}_{зн.i}$, влияющий на применение конкретных параметров в процессе мониторинга, получают путем суммирования коэффициентов значимости $K_{зн}$ рассматриваемых параметров в соответствие со шкалой лингвистической оценки по формуле [9, 10]:

$$K^{\Sigma}_{зн.i} = \sum_{i=1}^{m} K_{зн.i}\,.$$

Градацию степени значимости (ранжировку) контролируемых параметров, вошедших в программу мониторинга, производят на различных уровнях разукрупнения СРС (СНО) по узлам, агрегатам, стойкам, комплексам, на основе анализа их структурного взаимодействия. Чем выше вес коэффициента значимости элемента, тем больший вклад он вносит в обеспечение надёжности эксплуатации СРС (СНО), а также безопасного судовождения на ВВП РФ.

Формирование окончательного перечня наблюдаемых параметров по программе мониторинга проводится методом выстраивания вариационного ряда предпочтений по значениям, полученных с помощью таблицы суммы коэффициентов значимости для каждого i-го параметра наблюдаемого СРС (СНО) для последующего его включения в программу мониторинга из всего множества параметров НТД ОК по выражению:

$$K^{\Sigma}_{зн.l} \succ K^{\Sigma}_{зн.r} \succ K^{\Sigma}_{зн.k} \succ K^{\Sigma}_{зн.i} \succ K^{\Sigma}_{зн.u} \succ K^{\Sigma}_{зн.m}\,,$$

где i = 1, 2, …, k, …, l, …, u, …, r, …, m-1, m – параметры СРС и СНО по НТД. Причём в состав параметров, входящих в программу мониторинга, включают имеющие максимальное значение суммы коэффициентов (см. выражение выше). Количество наблюдаемых параметров зависит как от времени проведения программы мониторинга, так и от её глубины и применяемых технологий.

На шаге 5 задают *точность мониторинга* путем назначения эксплуатационных и профилактических допусков с учётом нормативно-технической документации (НТД) на СРС (СНО), а также актуализированной по результатам проведения последнего инструментального контроля (технического обслуживания) базы знаний (БЗ) берегового АИК АСУ ДС. Процедура формирования эксплуатационных и профилактических допусков параметров СРС и СНО с учётом частотного

диапазона, состояния среды распространения радиосигнала и условий функционирования объекта контроля приведена в [11].

На шаге 6 предварительно, на основе анализа условий функционирования СРС и СНО (день/ночь, осадки, помехи и пр.) задают *достоверность мониторинга* путем определения вероятности наступления ошибок контроля: «ложная тревога» (ошибка первого рода) и «пропуск отказа» (ошибка второго рода). Для повышения достоверности процедуры мониторинга проводят минимизацию данных ошибок [12].

На шаге 7 определяют *полезную нагрузку* СРМ путём комплектования набора G плат измерения $g_{ij} \in G$ бортового АИК БЭВС исходя из состава контролируемых береговых и надводных СрС и СНО, а также перечня их наблюдаемых параметров.

На шаге 8 подбирают *средство мониторинга* исходя из полезной нагрузки, глубины мониторинга, условий его проведения (погодных условий): – маломерное БЭВС с бортовым АИК. Варианты СРМ приведены на рисунке 2.

На шаге 9 задают (планируют или корректируют) маршрут обхода и точки наблюдения ближней и дальней зон излучения СРС (СНО) при проведении процедур дистанционного мониторинга значений параметров их технического состояния.

На шаге 10 реплицируют (актуализируют) БЗ бортового АИК измерительной информацией с БЗ берегового АИК для обеспечения режима мониторинга в режиме времени, близком к реальному (on-line). При актуализации (переносе) ИИ баз данных могут использоваться как проводные каналы, так и беспроводные каналы связи (Wi-Fi), либо отчуждаемый носитель (USB-накопитель).

Первым этапом метода является проведение программы телеизмерений с помощью бортового АИК и трансляция этого сигнала в береговой АИК при выявлении отказа (предотказного ТС) СРС (СНО). Этап описан шагами 11-22 (рис. 1).

На шаге 11 выводят СРМ (БПЛА, БЭВС) с бортовым АИК, на маршрут обхода (облёта) СРС (СНО) по заранее установленной программе мониторинга.

На шаге 12 удаленно, с помощью приёмников радионавигационных измерительных сигналов gij бортового АИК, замеряют доступные мониторингу i-е параметры радиоизлучающих j-х СРС (СНО), или, используя видеокамеру высокого разрешения, распознают излучающие в оптическом диапазоне навигационные огни.

а) б) в) г)

Рис. 2. Виды средств мониторинга: а) и в) на базе БЭВС среднего класса (безэкипажный катер «Искатель», АО «НПП «АМЭ»»); б) и г) малого класса с установленным бортовым АИК и его антенно-фидерной подсистемой (проектные решения под конкретные задачи процедур мониторинга СРС и СНО на ВВП РФ)

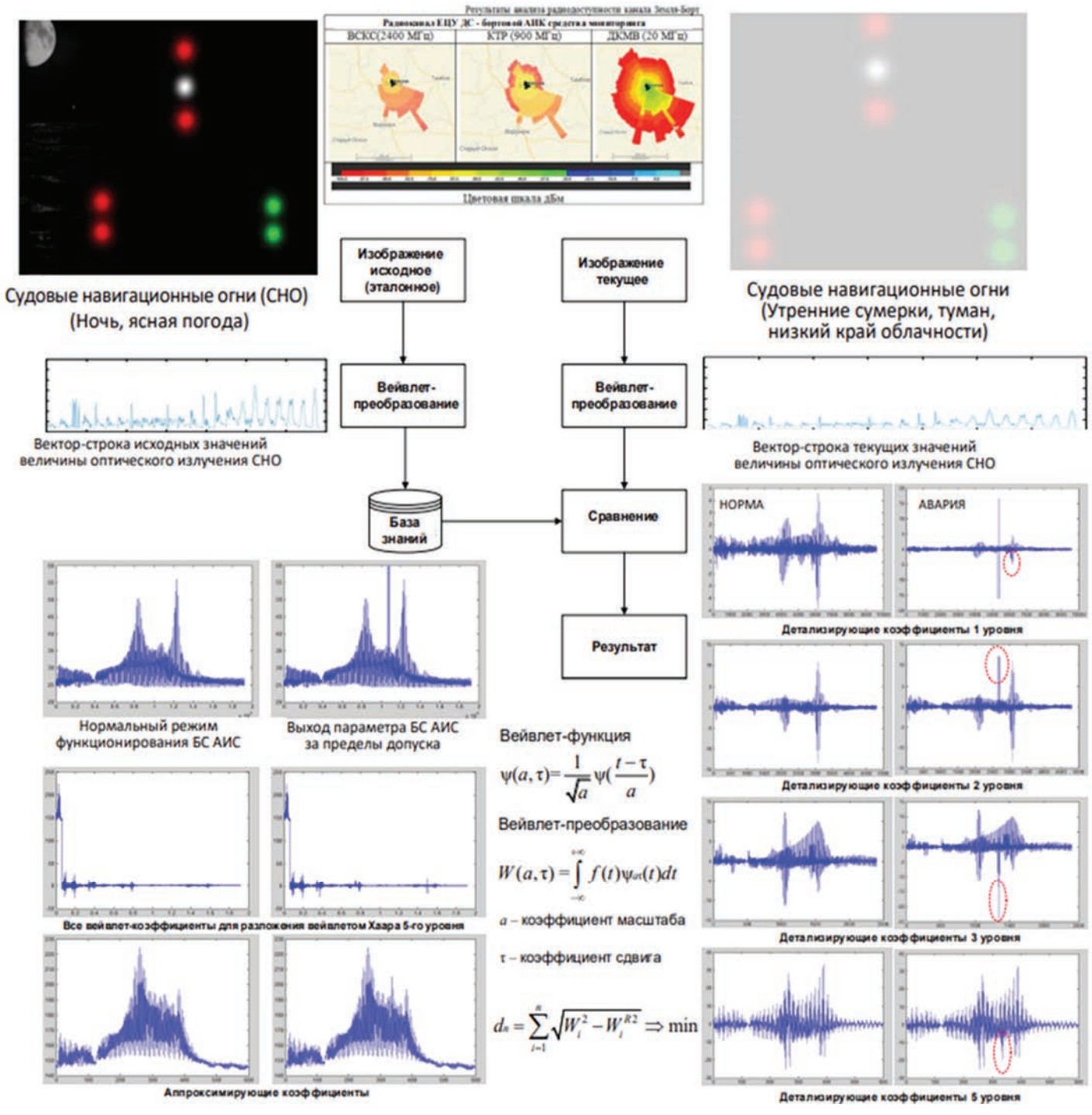


Рис. 3. Последовательность применения процедуры вейвлет-анализа в методе удаленного мониторинга функционального состояния СРС и СНО

На шаге 13 осуществляют запись значений измеренных бортовым АИК параметров в его БЗ. При этом для сокращения объёма сохраняемой ИИ, а также для повышения оперативности её обработки в дальнейшем хранят лишь вейвлет-коэффициенты измеренных параметров [13].

На рисунке 3 представлена последовательность использования процедуры вейвлет-анализа при мониторинге навигационных огней и средств управления БЭВС различных диапазонов: высокоскоростного канала связи (ВСКС), командно-телеметрической радиолинии (КТР) и СРС декаметровых волн (ДКМВ).

На шаге 14 в бортовом АИК сравнивают величины измеренных параметров СРС (СНО) с записанными в БЗ номинальными их значениями с учетом эксплуатационных и

профилактических допусков на них, путём проведения вейвлет-анализа [13].

На шаге 15 осуществляют обработку ИИ *i*-го параметра наблюдаемого СРС (СНО) с учётом заданной достоверности экспресс-контроля, а также эксплуатационных и профилактических допусков, минимизируя ошибки контроля при определении класса ТС ОК. При этом процедура идентификации отказа, реализуемая в ходе многоуровневого контроля, описывается вероятностным графом оценки класса ТС СРС и СНО с учётом ошибок первого и второго рода (рис. 4), на котором практически реализуется программа экспресс-контроля по этапам, когда сначала бортовым АИК обнаруживают отказ СРС или СНО, а затем в береговом АИК осуществляют его распознавание с применением объектовых АИК по наземных (проводным) каналам телеизмерения-телесигнализации (ТИ-ТС).

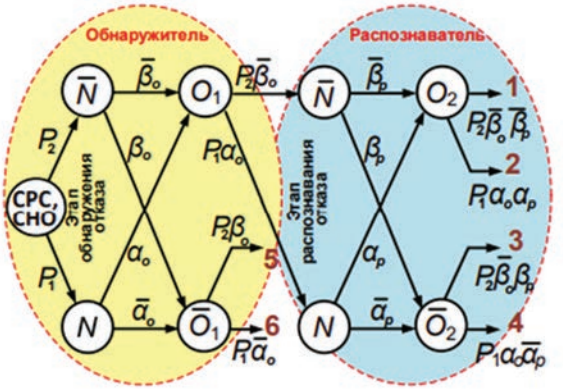


Рис. 4. Вероятностный граф идентификации класса ТС СРС и СНО

При этом на рисунке 4 показаны следующие классы ТС СРС и СНО [14]:

класс 1 – СРС (СНО) работоспособно, обнаружен ложный отказ и не распознан;

класс 2 – отказ СРС (СНО), при этом обнаруженный отказ не распознан;

класс 3 – СРС (СНО) работоспособно, отказ обнаружен ложно и не распознан;

класс 4 – отказ СРС (СНО) обнаружен и распознан, включение резерва;

класс 5 – СРС (СНО) работоспособно и правильно идентифицировано;

класс 6 – отказ СРС (СНО), который не обнаружен подсистемой мониторинга.

На шагах 16 и 17 осуществляется проверка всех доступных измерениям *m* параметров на *j*-м средстве связи (СНО).

На шагах 18 и 19 осуществляется проверка всех *n* объектов мониторинга. В случае отсутствия выявления отказа (предотказного ТС) на объекте мониторинга (СРС, СНО) – переход к шагу 34. В случае обнаружения отказа (предотказного ТС) СРС (СНО), когда значения наблюдаемых параметров выходят за пределы заданных эксплуатационных или профилактических допусков (см. состояние «ОТКАЗ» на рисунке 1), осуществляют переход к шагу 20.

На шаге 20 вырабатывают формализованный сигнал о наступлении отказа (аварийного, предаварийного ТС) СРС и СНО. Причём, в отличие от методов, опубликованных в [7, 8], из бортового АИК в сторону берегового АИК передаётся только класс ТС в формализованном виде, а не вся получаемая в ходе мониторинга ИИ, чем достигается выигрыш в сокращении объёмов передаваемой информации.

С позиции повышения достоверности процедуры мониторинга наиболее предпочтительным являются подтверждённые статусы результатов измерений об исправности ОК (зоны мониторинга), что соответствует классу 5 ТС – СРС (СНО) работоспособна и признана таковой (рис. 4). При идентификации класса 4 ТС (отказ СРС (СНО), который обнаружен и распознан) подсистема мониторинга АСУ ДС должна в автоматическом режиме выдать рекомендации должностному лицу системы поддержки принятия решения (СППР) о включении резервного комплекта СРС (СНО).

Передача полученной в ходе цикла мониторинга ИИ на береговой АИК производится только при идентификации класса ТС, отличного от статуса «подтверждено» нормальное функционирование или авария, к примеру, предотказное состояние, которое характеризуется статусом «недостоверный», или «ориентирующий», или «экстраполированный» [10]. На рисунке 4 такие статусы состояния на вероятностном графе показаны как классы 1, 2, 3, 6, которые характеризуются опасным ТС объекта мониторинга и требуют вмешательства оператора СППР (подсистемы мониторинга) СНО или АСУ ДС для ситуационного управления [6].

В отдельных случаях, при идентификации аварийного ТС наблюдаемого СРС или СНО после перевода АСУ ДС его на резерв, оператор подсистемы мониторинга должен по запросу получить от АИК СРМ доступную ИИ для более тщательной диагностики места отказа объекта мониторинга, поскольку для выявления отказа и проведения регулировок и тестовых проверок СРС и СНО только формализованного сигнала с классом его ТС недостаточно.

На шаге 21 формализованный сигнал о предотказном или неработоспособном техническом состоянии кодируют алгоритмами, используемыми в радиосвязи.

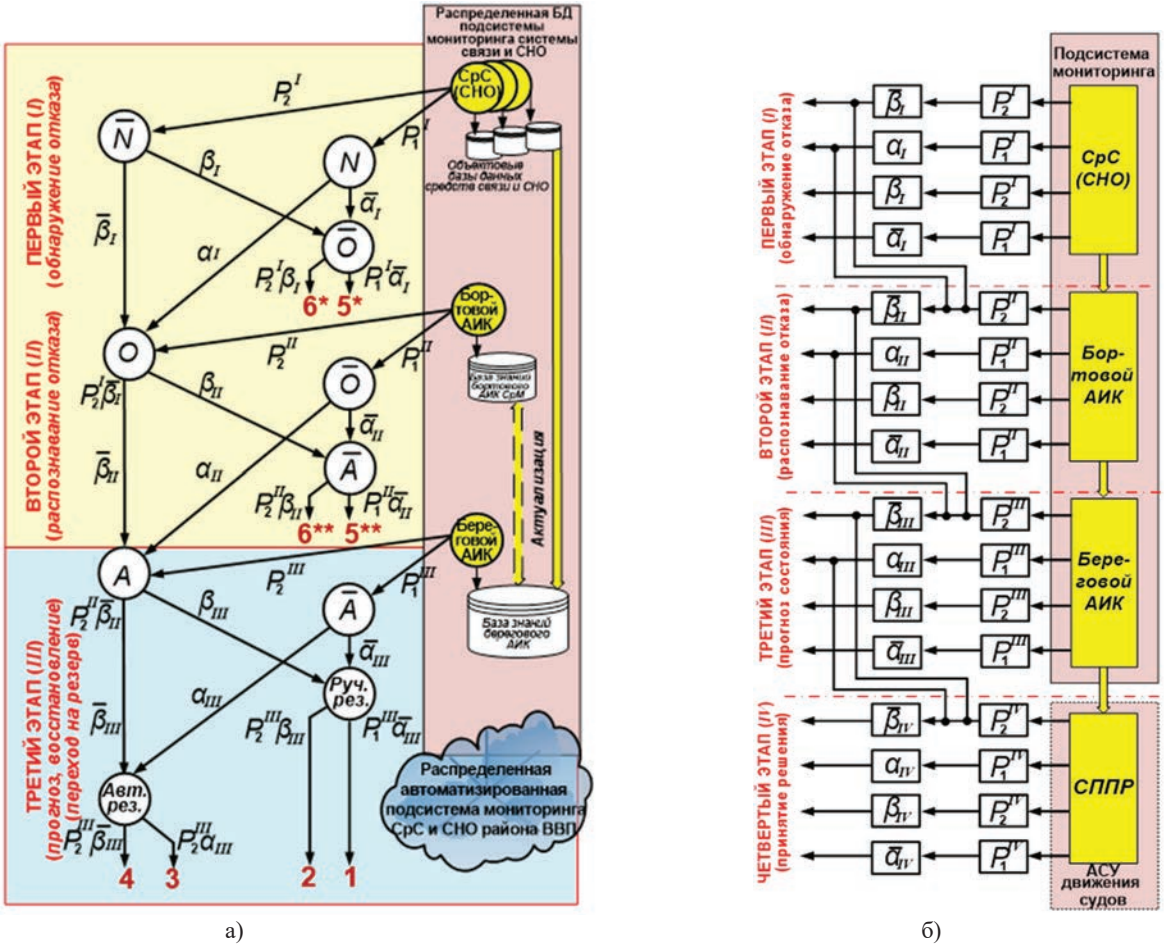
На шаге 22 формализованный сигнал о неработоспособном или предотказном ТС наблюдаемого СРС (СНО) излучают в свободное пространство.

Вторым этапом метода является процедура идентификации береговым АИК отказа, обнаруженного АИК на борту СРМ. Этап соответствует шагам 23-33.

На шаге 23 формализованный сигнал о неработоспособном (предотказном) ТС ОК от бортового АИК принимают радиоприёмным устройством (РПУ) берегового АИК.

На шаге 24 декодируют сигнал о предотказном или неработоспособном ТС СРС или СНО алгоритмами, которые используются системами радиосвязи.

На шаге 25 обрабатывают принятый от бортового АИК декодированный сигнал совместно с ИИ, хранящейся в БЗ берегового АИК и с показателями инструментального контроля, получаемым по каналам ТИ-ТС от объектовых АИК аварийных СРС и СНО, а также с ИИ предыдущих процедур технического обслуживания с записью результата в БЗ берегового АИК. В [14] подробно описан процесс идентификации отказа на втором этапе предложенного метода.



а) б)

Рис. 5. Вероятностный граф функционирования распределенной автоматизированной системы мониторинга СРС и СНО при определении финальной вероятности нормального функционирования системы $P_{нф}$ и финальной вероятности отказа $P_{отк}$ (а) и надёжностная схема замещения вероятностного графа автоматизированной системы мониторинга СРС и СНО при определении $P_{нф}$ и $P_{отк}$ через ошибки контроля первого α и второго β рода (б)

Поскольку подсистема мониторинга СНО работает как обеспечивающая подсистема распределённой АСУ ДС (т. е. в виде элемента СППР), а ошибки типа «пропуска отказа» и «ложной тревоги» могут происходить как непосредственно в АИК, так и при трансляции аварийных сигналов по каналам телеметрической системы, то процедуру определения вида ТС СРС и СНО, можно представить многоэтапным вероятностным графом, показанном на рис. 4, описывающий процесс функционирования всей АСУ ДС. При этом процедура определения финальной вероятности нормального функционирования системы $P_{нф}$ и финальной вероятности отказа $P_{отк}$ поясняется на рисунке 5.

Вероятностный граф АСУ ДС с учётом дополнительного этапа принятия решения в СППР на переход к резерву (восстановление оказавшего СРС и СНО) также можно представить надёжностной схемой замещения. Процесс расчёта финальной вероятности нормального функционирования системы $P_{нф}$ с учётом ошибок контроля (α и β) на разных этапах и её оценка имеет вид:

$$P_{нф}=1-\left(1-\bar{\beta}_{IV}\left\langle 1-\left[\bar{\beta}_{III}\left\{1-\left(1-P_1^{I}\alpha_I\bar{\beta}_{II}\right)\left(1-P_1^{II}\alpha_{II}\right)\right\}\right]\left[1-P_1^{III}\alpha_{III}\right]\right\rangle\right)\left(1-P_1^{IV}\alpha_{IV}\right),$$

Вычисление финальной вероятности отказа СРС (СНО) проводим по формуле:

$$P_{отк}=\left(1-\left\langle 1-\bar{\beta}_{III}\left[1-\bar{\beta}_{II}\left\{1-\left(1-P_2^{I}\alpha_I\bar{\beta}_I\right)\left(1-P_2^{II}\right)\right\}\right]\left[1-P_2^{III}\right]\right\rangle\left\langle 1-P_2^{IV}\right\rangle\right)\bar{\beta}_{IV}.$$

При этом в ходе этапа обнаружения предотказного и неработоспособного ТС осуществляют контроль комплексного показателя СРС и СНО $\Lambda(a)$ по установленному порогу параметра a_0. При соблюдении этого условия (например, $\Lambda(a) > a_0$) формируют сигнал о работоспособном ТС СРС (СНО). Допуски на значения параметров x_0^1, y_0^2, ..., γ_0^K эксплуатационных СРС и СНО, назначают на разных k уровнях их функционирования ($k = 1, 2, \ldots, K$), с учётом доступа для съёма ИИ дистанционно. Порядок определение эксплуатационных допусков на параметры СРС и СНО представлен в [15].

При несоблюдении этого условия и выходе текущего значения комплексного показателя ТС за значения установленного допуска проводят измерение параметров локального уровня ИТКС (их текущих показателей ТС) объектовым АИК (на объекте размещения СРС или СНО), которые далее также сравнивают с значением допуска на конкретный

эксплуатационный параметр $\Lambda(a) > a_0$ для последующей идентификации отказа. В результате такого сравнения определяют работоспособное ТС СРС (СНО) (N) с вероятностью $P_1 = P(N)$, либо его неработоспособное (предотказное) ТС ($\overline{N}$) с вероятностью $P_2 = 1 - P_1 = P(\overline{N})$.

Таким же образом осуществляют идентификации нарушения работоспособности системы связи и СНО – её переход в предотказное и неработоспособное ТС (аварийное) – (А) на вышестоящих уровнях управления и контроля сети связи. Переход на очередной уровень выявления отказа производят если выполнено условие обнаружения аварийного ТС на предыдущем уровне, а также если наблюдаемые метрики эксплуатационных параметров на рассматриваемом уровне превысили пределы допусков.

На шаге 26 отображают на мониторе пульта оператора или коллективном табло отображения подсистемы мониторинга (АСУ ДС) ИИ о классе ТС аварийного СРС и СНО и критичном параметре.

На шаге 27 сохраняют в виде вейвлет-коэффициентов результат измерения параметров ОК и его совместной обработки в БЗ АСУ ДС (береговом АИК) [13, 14].

На шаге 28 выносят решение на ситуационное управление (включение резерва) СРС и СНО с последующим инструментальным контролем аварийного средства связи (регулировку параметров, повышение энергетики, ориентация антенны и пр.). Если по результатам обработки всей доступной ИИ из берегового, бортового и объектового АИК отказ объекта мониторинга не определен, то переход к шагу 34.

Если по результатам обработки всей доступной ИИ из берегового, бортового и объектового АИК отказ объекта мониторинга определен, то принимается решение на ситуационное управление СРС (СНО), для чего необходимо дать команду на повторную процедуру мониторинга включённого резервного полукомплекта СРС (СНО), т. е. переход к шагу 29. И только после этого могут быть начаты процедуры диагностики (определения места отказа) на аварийном комплекте.

На шаге 29 подают команду на СРМ для проведения повторного цикла мониторинга включённого резерва ОК по программе первого этапа метода с обходом (облётом) его ближней или дальней зоны. При обратном переходе с резервного на основной комплект СНО (СРС) также формируют сигнал на СРМ для повторного обхода ОК после завершения на нем процедуры диагностики.

На шаге 30 производят кодирование команды на СРМ для повторного обхода аварийного СРС и СНО методами, используемыми системами радиосвязи.

На шаге 31 передают в свободное пространство сигнал с формализованной командой на повторный обход (облёт) СРМ аварийного СРС или СНО.

На шаге 32 осуществляют приём бортовым радиоприёмным устройством СРМ сигнала с формализованной командой на повторный обход (облёт) СРС (СНО).

На шаге 33 производят декодирование сигнала на повторный обход (облёт) СРМ аварийного СРС (СНО), после чего повторяют процедуры первого этапа, шаги 11-22.

Завершающий этап включает репликацию (актуализацию) БЗ берегового АИК подсистемы мониторинга АСУ ДС полученной ИИ бортового АИК по шагам 34-37.

На шаге 34 останавливают процедуру мониторинга и возвращают СРМ на пункт постоянной дислокации (швартуют БЭВС к причалу порта, приземляют БПЛА).

На шаге 35 реплицируют (актуализируют) БЗ берегового АИК подсистемы мониторинга АСУ ДС с учётом выполненной программы мониторинга с помощью ИИ БЗ из бортового АИК СРМ. Для актуализации (переноса) информации баз данных могут использоваться как проводные каналы, так и беспроводные каналы связи (*Wi-Fi*), либо отчуждаемый носитель (*USB*-накопитель).

На шаге 36 обновляют исходные данные процедуры предварительного этапа метода удалённого мониторинга функционального состояния СРС и СНО для процедур имитационного моделирования элементов системы связи и СНО и использования в учебно-тренировочных средствах для обучения судоводителей.

На шаге 37 формируют отчёты и протоколы измерений о ТС СРС и СНО по результатам экспресс-контроля, проведённого при процедуре цикла мониторинга.

Заключение

Новизна предложенного метода удаленного мониторинга функционального состояния средств связи и навигационного оборудования Росморречфлота отличается поэтапной процедурой идентификации класса ТС СРС и СНО, когда сначала происходит обнаружение нарушения функционирования (аварийной ситуации), а в последующем – распознавание вида отказа и идентификация класса ТС ОК с применением процедуры вейвлет-анализа. Это позволяет расширить перечень наблюдаемых стационарных и плавучих СНО (СРС), а также повысить оперативность процесса оценки их ТС, доведя до режима, близкому к реальному времени.

Практическая значимость представленного метода определяется доработкой теоретических положений прикладной теории надёжности до наглядных инструментов по внедрению на АСУ ДС районов водных путей и судоходства бассейнов ВВП РФ эффективной подсистемы удалённого мониторинга.

Предлагаемый метод может быть применен:

– пр вводе средств связи и СНО в эксплуатацию;

– при периодических плановых поверках СрС и СНО и в ходе их технического обслуживания (инструментальном контроле);

– при неплановых поверках СрС и СНО, связанных с замечаниями (жалобами) капитанов судов на некорректное функционирование СНО, а также при воздействиях на систему связи и СНО искусственного и естественного характера (техногенных катастроф, ураганов и пр.);

– при разработке и испытаниях новых образцов средств связи и СНО, систем посадки и радионавигационных систем гидродромов и посадочных площадок морских буровых платформ, а также для пилотируемой и беспилотной авиации наземного и морского базирования;

– при выполнении научно-исследовательских работ в области совершенствования средств связи и СНО и повышения их эффективности функционирования;

– при повышении квалификации специалистов речного и морского флота.

Литература

1. *Бекряшев В.А., Каретников В.В., Яснов А.П.* Система мониторинга плавучей навигационной обстановки на внутренних водных путях Российской Федерации // Морская радиоэлектроника. 2016. № 2. С. 20-23.

2. *Красников В.В., Сикарев А.А.* Создание современной инфраструктуры управления движением судов в Карском море с использованием автоматизированных идентификационных систем // Морская радиоэлектроника. 2014. № 4. С. 34-37.

3. *Федотов А.А., Емелин В.И.* Обоснование сетецентрических систем радиоэлектронного мониторинга // Морская радиоэлектроника. 2018. № 1. С. 10-15.

4. *Аллакин В.В., Будко Н.П., Васильев Н.В.* Общий подход к построению перспективных систем мониторинга распределенных информационно-телекоммуникационных сетей // Системы управления, связи и безопасности. 2021. № 4. С. 125-227.

5. *Мирошников В.И., Будко П.А., Виноградено А.М., Меженов А.В.* Комплексный подход в работе автоматизированной системы контроля в телеметрии технического состояния объектов связи морского базирования // Морская радиоэлектроника. 2018. № 4. С. 8-14.

6. *Каретников В.В., Будко Н.П., Аллакин В.В.* Синтез подсистемы интеллектуального мониторинга информационно-телекоммуникационной сети ведомственного ситуационного центра // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2021. № 3. С. 64-81.

7. *Войтович Н.И., Жданов Б.В.* Способ летных проверок наземных средств радиотехнического обеспечения полетов и устройства для его применения // Патент на изобретение RU 2501031 C2, опубл. 10.12.2013, бюл. № 34.

8. Приказ Минтранса РФ № 1 от 18.01.2005. Об утверждении Федеральных авиационных правил «Летные проверки наземных средств радиотехнического обеспечения полетов, авиационной электросвязи и систем светосигнального оборудования аэродромов гражданской авиации» // РГ № 3733. 31.03.2005.

9. *Виноградено А.М., Меженов А.В., Будко Н.П.* К вопросу обоснования понятийного аппарата неразрушающего экспресс-контроля технического состояния оборудования системы связи и радиотехнического обеспечения аэродрома // Наукоемкие технологии в космических исследованиях Земли. 2019. Т. 11. № 6. С. 30-44.

10. *Клюев В.В., Соснин Ф.Р.* Неразрушающий контроль и диагностика: справочник. М.: Машиностроение, 2005. 656 с.

11. *Аллакин В.В., Голюнов М.В.* Анализ научно-методического аппарата удаленного мониторинга технического состояния информационно-телекоммуникационных сетей и систем // Техника средств связи. 2020. № 4 (152). С. 17-37.

12. *Будко П.А.* Управление ресурсами информационно-телекоммуникационных систем. Методы оптимизации. Санкт-Петербург: ВАС, 2012. 512 с.

13. *Будко П.А., Жуков Г.А., Виноградено А.М., Гойденко В.К.* Определение аварийного состояния морского робототехнического комплекса по многоэтапной процедуре контроля на основе использования вейвлет-преобразований // Морская радиоэлектроника. 2016. № 4 (58). С. 20-23.

14. *Karetnikov V.V., Allakin V.V., Budko P.N., Butsanets A.A.* Monitoring of the technical state of communication and navigation equipment used for the inland waterways. – DOI 10.1088/1742-6596/2032/1/012083 // Journal of Physics: Conference Series. – Novosibirsk: International Conference on IT in Business and Industry (ITBI 2021) 12-14 May 2021, 2021. Vol. 2032. № 012083. Pp. 1-14.

15. *Абрамов О.В.* Планирование профилактических коррекций параметров технических устройств и систем // Информатика и системы управления. 2017. № 3. С. 55-66.

METHOD OF REMOTE MONITORING OF THE FUNCTIONAL STATE OF COMMUNICATIONS AND NAVIGATION EQUIPMENT OF ROSMORRECHFLOT

VLADIMIR V. ALLAKIN
St. Petersburg, Russia, vladimir@duduh.ru

NIKITA P. BUDKO
St. Petersburg, Russia, budko62@mail.ru

MIKHAIL V. GOLYUNOV
St. Petersburg, Russia, belka1213@mail.ru

VLADIMIR V. KARETNIKOV
St. Petersburg, Russia, kaf_svvp@gumrf.ru

KEYWORDS: *automated measuring complex, measuring information, monitoring, radio communications, unmanned watercraft.*

ABSTRACT

Relevance: the most effective approach to reducing accidents and ensuring safe navigation on the inland waterways of Russia has proven to be the transition from pilotage to instrumental navigation, as well as the introduction of more advanced information and telecommunication technologies in the process of vessel traffic management and remote monitoring, which leads to the creation of hierarchical systems in Rosmorrechflot based on situational centers responsible for safety of navigation. **Purpose of the work:** development of a method for remote monitoring of the functional state of communications and navigation equipment. **Methods used:** to implement the remote monitoring method, monitoring tools using unmanned vessels are most applicable. **Novelty:** the proposed method makes it possible to use radio engineering and radio navigation means widely used on shore and on water vessels, emitting in the radio or optical wave range, as objects of control and monitoring, and as a means of monitoring – small unmanned water vessels of various classes, with the placement of on-board automated measuring complexes on them, which have replaceable equipment for the tasks to be solved control and measuring equipment. **The result obtained:** in the developed method, the monitoring process is carried out according to the following stages: preliminary to the implementation of the procedure of tele-measurements, preparation and input of initial data into the onshore and onboard automated measuring complexes are carried out; at the first stage, measurements of remotely accessible parameters of the far and near zone of controlled objects are carried out on board the monitoring equipment with the transmission of emergency signals to the shore; at the second stage, the collected measuring information is processed by an onshore automated measuring complex with the establishment of classes of the technical condition of monitoring objects; at the end, reports are prepared in the interests of the situational vessel traffic control center.

REFERENCES

1. V. A. Bekryashev, V. V. Karetnikov, A. P. Yasnov. Monitoring system for floating navigation conditions on the inland waterways of the Russian Federation. *Marine radioelectronics*. 2016. No. 2. Pp. 20-23. (in Russian)

2. V. V. Krasnikov, A. A. Sikarev. Creation of a Modern Infrastructure for Vessel Traffic Control in the Kara Sea Using Automated Identification Systems. *Marine Radioelectronics*. 2014. No. 4. Pp. 34-37. (in Russian)

3. A. A. Fedotov, V. I. Emelin. Substantiation of network-centric systems of radio-electronic monitoring. *Marine radioelectronics*. 2018. No. 1. Pp. 10-15. (in Russian)

4. V. V. Allakin, N. P. Budko, N. V. Vasiliev. A general approach to the construction of advanced monitoring systems for distributed information and telecommunications networks. *Systems of Control, Communication and Security*, 2021, no. 4, pp. 125-227. DOI: 10.24412/2410-9916-2021-4-125-227 (in Russian)

5. V. I. Miroshnikov, P. A. Budko, A. M. Vinogradenko, A. V. Mezhenov. An integrated approach to the operation of an automated control system in the telemetry of the technical condition of sea-based communication facilities. *Marine radioelectronics*. 2018. No. 4. Pp. 8-14. (in Russian)

6. V. V. Karetnikov, N. P. Budko, V. V. Allakin. Synthesis of subsystem of intelligent monitoring of information and telecommunication network of departmental situational center. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics.* 2021, no. 3, pp. 64-81. (In Russ.) DOI: 10.24143/2072-9502-2021-3-64-81.

7. N. I. Voitovich, B. V. Zhdanov. The method of flight checks of ground-based means of radio-technical flight support and devices for its application. Patent for invention RU 2501031 C2, publ. 12/10/2013, bul. No. 34.

8. Order of the Ministry of Transport of the Russian Federation No. 1 dated January 18, 2005. On approval of the Federal Aviation Rules "Flight checks of ground-based means of radio-technical flight support, aviation telecommunications and lighting systems of civil aviation airfields". RG No. 3733. 31.03.2005. (In Russian)

9. A. M. Vinogradenko, A. V. Mezhenov, N.P. Budko. To the question of substantiation of the conceptual apparatus nondestructive express control of technical condition equipment of communication system and aerodrome radio engineering support. *H&ES Research*. 2019. Vol. 11. No. 6. Pp. 30-44. doi: 10.24411/2409-5419-2018-10293 (In Russian)

10. V. V. Klyuev, F. R. Sosnin. Non-destructive testing and diagnostics: a reference book. Moscow. Mashinostroenie, 2005. 656 p. (In Russian)

11. V. V. Allakin, M. V. Golyunov. Analysis of the scientific and methodological apparatus for remote monitoring of the technical condition of information and telecommunication networks and systems. Means of communication equipment. 2020. No. 4 (152). C. 17-37. (In Russian)

12. P. A. Budko. Resource management of information and telecommunication systems. Optimization methods. St. Petersburg: VAS, 2012. 512 p. (In Russian)

13. P. A. Budko, G. A. Zhukov, A. M. Vinogradenko, V. K. Goydenko. Determination of the emergency state of the marine robotic complex by a multi-stage control procedure based on the use of wavelet transforms. *Marine radioelectronics*. 2016. No. 4 (58). Pp. 20-23. (In Russian)

14. V. V. Karetnikov, V. V. Allakin, P. N. Budko, A. A. Butsanets. Monitoring of the technical state of communication and navigation equipment used for the inland waterways. DOI 10.1088/1742-6596/2032/1/012083. *Journal of Physics: Conference Series.* Novosibirsk: International Conference on IT in Business and Industry (ITBI 2021) 12-14 May 2021, 2021. Vol. 2032. № 012083. Pp. 1-14.

15. O. V. Abramov. [Planning of preventive corrections of parameters of technical devices and systems. *Informatika i sistemy upravleniya* [Informatics and control systems]. 2017. No. 3. Pp. 55-66. (In Russian)

INFORMATION ABOUT AUTHORS:

Vladimir V. Allakin, postgraduate student, Federal State Budgetary Educational Institution of Higher Education "Admiral S. O. Makarov State University of the Sea and River Fleet", St. Petersburg, Russia.

Nikita P. Budko, postgraduate student, Federal State Budgetary Educational Institution of Higher Education "Admiral S. O. Makarov State University of the Sea and River Fleet", St. Petersburg, Russia

Mikhail V. Golyunov, Postgraduate of the Military Academy of Communications. St. Petersburg, Russia

Vladimir V. Karetnikov, Head of the Department of Navigation on Inland Waterways of the Admiral S.O. Makarov State University of Marine and River Fleet, Doctor of Technical Sciences, Associate Professor, St. Petersburg, Russia

For citation: Allakin V. V., Budko N. P., Golyunov M. V., Karetnikov V. V. Method of remote monitoring of the functional state of communications and navigation equipment of Rosmorrechflot. H&ES Reserch. 2023. Vol. 15. No. 1. P. 10-20. doi: 10.36724/2409-5419-2023-15-1-10-20 (In Rus)

doi: 10.36724/2409-5419-2023-15-1-21-26

МЕТОДИКА ОПЕРАТИВНОГО ВЫБОРА ПУТЕЙ ДОВЕДЕНИЯ ИНФОРМАЦИИ В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

ШУХАРДИН
Александр Николаевич [1]

ШКОРИНА
Александр Васильевич [2]

АННОТАЦИЯ

Введение. Применение положений теории раскрашенных иерархических сетей Петри позволяет создавать модели современных сложных информационно-телекоммуникационных систем (ИТС), позволяющие с относительно невысокими вычислительными затратами проводить моделирование функционирования таких систем, проведение натурных исследований в которых в процессе эксплуатации либо невозможно, либо не целесообразно по различным причинам. **Методика проведения исследования:** В статье рассматривается методика, позволяющая выбрать пути доведения информации в информационно-телекоммуникационных системах, а также оценить выполнение предъявляемых требований к системе доведения информации до всех узлов системы при изменениях характеристик системы в процессе эксплуатации. Структурными элементами данной методики являются методика оценивания и математическая модель функционирования информационно-телекоммуникационных систем при доведении сообщений. **Результаты исследования:** при применении разработанной методики формируется три группы узлов: множество узлов, имеющих путь доведения информации, характеризуемый максимальным значением вероятности; множество узлов, имеющих совокупность путей доведения информации, выполнение которых позволит обеспечить выполнение требуемых значений вероятности и времени доведения информации; множество узлов, для которых отсутствуют пути, позволяющие обеспечить выполнение требуемых значений вероятности и времени доведения информации. Разработанная методика оперативного выбора путей доведения информации в ИТС позволяет для каждого узла рассматриваемой системы выбрать путь (или совокупность путей) доведения информации, позволяющий обеспечить выполнение требуемых значений вероятности и времени доведения информации, или выявить узлы, для которых таких путей не существует, для своевременного принятия компенсационных мер.

Сведения об авторах:

[1] Северо-Кавказский филиал ордена Трудового Красного Знамени ФГБОУ ВО "Московский технический университет связи и информатики", г. Ростов-на-Дону, Россия

[2] Военная академия Ракетных войск стратегического назначения им. Петра Великого, Московская обл., г. Балашиха, Россия

КЛЮЧЕВЫЕ СЛОВА: *оперативный выбор, информационно-телекоммуникационная система, пути доведения информации.*

Для цитирования: Шухардин А.Н., Шкорина А.В. Методика оперативного выбора путей доведения информации в информационно-телекоммуникационных системах // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 21-26. doi: 10.36724/2409-5419-2023-15-1-21-26

Введение

В работах [1, 2] было показано, что применение положений теории раскрашенных иерархических сетей Петри [3, 4] позволяет создавать модели современных сложных информационно-телекоммуникационных систем (ИТС), позволяющие с относительно невысокими вычислительными затратами проводить моделирование функционирования таких систем, проведение натурных исследований в которых в процессе эксплуатации либо невозможно, либо не целесообразно по различным причинам.

В работах [5, 6] представлена методика, позволяющая эксплуатирующему персоналу на основе построенных моделей таких систем при изменении её структуры, характеристик её элементов оперативно находить все возможные существующие в текущий момент времени пути доведения информации от пункта-источника до всех узлов в ИТС, а также рассчитать вероятность и сроки доставки сообщений для каждого пути.

Однако вопросы оценки результатов проведенного моделирования и принятия решения по результатам моделирования в этих работах не изложены. Требуется разработать методику оперативного выбора путей доведения информации в ИТС, которая позволяла бы эксплуатирующему персоналу оперативно для каждого узла системы выбрать путь (или совокупность путей), обеспечивающий выполнение требуемых значений вероятности и времени доведения информации, или выявить узлы, для которых таких путей не существует.

Структурными элементами данной методики являются методика оценивания [6] и математическая модель функционирования ИТС [2] при доведении сообщений.

Выполнение разработанной методики производится в четыре этапа. Порядок выполнения методики представлен на рисунке 1.

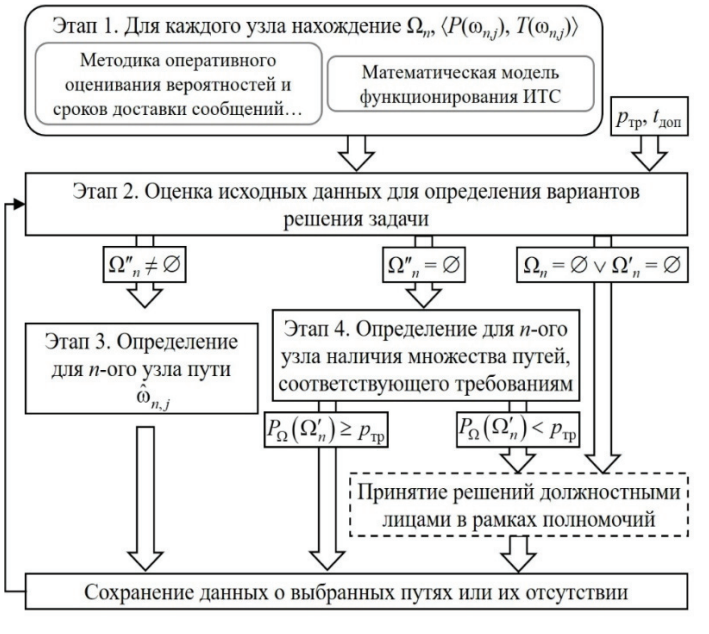


Рис. 1. Порядок выполнения методики оперативного выбора путей доведения информации в ИТС

На первом этапе методики с целью формирования данных для второго этапа производится выполнение методики оперативного оценивания вероятностей и времён доведения информации до каждого узла системы, описанной выше. В результате выполнения первого этапа формируются множества Ω_n существующих в расчётный момент времени путей $\omega_{n,j}$ доведения информации до каждого n-ого узла системы и кортежи значений вероятности и времени доведения информации по каждому из путей $\langle P(\omega_{n,j}), T(\omega_{n,j})\rangle$. Кроме того, исходными данными для выполнения второго этапа методики являются требуемая вероятность доведения информации $p_{тр}$ и допустимое время доведения информации $t_{доп}$ до узлов системы, задаваемые требованиями к системе.

На втором этапе для каждого узла системы проводится оценка полученных на первом этапе данных с целью определения возможных вариантов решения задачи.

Если определено, что $\Omega_n = \varnothing$, представляются данные об отсутствии путей доведения информации до n-ого узла системы, соответствующих предъявляемым требованиям, так как для него в данный расчетный момент времени в ИТС не существует путей доведения информации. На этом для n-ого узла выполнение методики завершается и осуществляется переход ко второму этапу для оценки данных следующих узлов.

Иначе, определяется подмножество Ω'_n путей $\omega'_{n,q}$ доведения информации до n-ого узла ИТС, характеризуемых значением времени, не превышающим $t_{доп}$:

$$\Omega'_n \subset \Omega_n,\ \left|\Omega'_n\right| = Q_n,\ \Omega'_n = \left\{\omega'_{n,q} \mid T\left(\omega'_{n,q}\right) \leq t_{доп},\ q = \overline{1, Q_n}\right\}, \quad (1)$$

где $\omega'_{n,q}$ – q-й путь доведения информации до n-ого узла.

Если определено, что $\Omega'_n = \varnothing$, представляются данные об отсутствии путей доведения информации до n-ого узла системы, соответствующих предъявляемым требованиям, так как для него в данный расчетный момент времени в ИТС не существует путей доведения информации. На этом для n-ого узла выполнение методики завершается и осуществляется переход ко второму этапу для оценки данных следующих узлов.

Если определено, что $\Omega'_n \neq \varnothing$, то в подмножестве Ω'_n определяется подмножество Ω''_n путей доведения информации до n-ого узла системы, характеризуемых значением вероятности, не ниже $p_{тр}$:

$$\Omega''_n \subset \Omega'_n,\ \left|\Omega''_n\right| = Q'_n,\ \Omega''_n = \left\{\omega''_{n,q'} \mid P\left(\omega''_{n,q'}\right) \geq p_{тр},\ q' = \overline{1, Q'_n}\right\}, \quad (2)$$

где $\omega''_{n,q'}$ – q'-й путь доведения информации до n-ого узла.

Если $\Omega''_n \neq \varnothing$, то осуществляется переход к третьему этапу методики, в ходе выполнения которого методом прямого перебора находится путь, соответствующий критерию

$$\hat{\omega}''_{n,q'} = \arg \max_{\omega''_{n,q'} \in \Omega''_n} P\left(\omega''_{n,q}\right). \quad (3)$$

Найденный путь $\hat{\omega}''_{n,q'}$ и является искомым $\hat{\omega}_{n,j}$. Данные о нём представляются эксплуатирующему персоналу.

Для отображения найденного пути $\hat{\omega}_{n,j}$ на схеме последовательность сработавших переходов представляется как последовательность пройденных пакетом информации узлов системы и каналов связи между ними до n-ого узла системы. На этом для n-ого узла выполнение методики завершается и осуществляется переход ко второму этапу для оценки данных следующих узлов.

Если $\Omega''_n = \varnothing$, то для n-ого узла единственного пути $\hat{\omega}_{n,j}$, соответствующего предъявляемым требованиям, не существует, и выполняется четвёртый этап методики, на котором производится оценка совокупности путей Ω'_n на соответствие установленным требованиям. Для подмножества путей вычисляется значение вероятности $P_\Omega(\Omega'_n)$ доведения информации до этого узла. Доведение информации по подмножеству путей Ω'_n есть наступление хотя бы одного из событий «доведение информации по отдельному пути $\omega'_{n,q} \in \Omega'_n$». Так как события «доведение информации по отдельному пути $\omega'_{n,q}$» являются независимыми в совокупности, то вероятность наступления хотя бы одного из событий, независимых в совокупности, определяется выражением:

$$P_\Omega\left(\Omega'_n\right) = 1 - \prod_{q=1}^{Q_n}\left(1 - P\left(\omega'_{n,q}\right)\right), \tag{4}$$

где $P(\omega'_{n,q})$ – вероятность доведения информации до n-ого узла системы по пути $\omega'_{n,q}$.

Вместе с тем, из-за особенностей функционирования ИТС (доведение информации по всем возможным каналам, распараллеливанием, дублированием, избыточностью структуры ИТС и т.п.) пути подмножества Ω'_n могут иметь общие элементы (узлы, ЛС). При этом значение вероятности, вычисленное в соответствии с выражением (4), будет некорректно и завышено.

Для получения более корректного результата предложен подход с использованием функции поглощения, описанный в ГОСТ Р 53111–2008 [7]. При определении вероятности $P_\Omega(\Omega'_n)$ доведения информации до n-ого узла по подмножеству путей Ω'_n в соответствии с выражением (4) вероятность $P(\omega'_{n,q})$ представляется как произведение значений вероятностей функционирования узлов системы и значений вероятностей передачи информации по линиям связи, и после раскрытия скобок у всех членов выражения заменяются на единицу значения показателей степени, имеющие значения больше единицы.

Таким образом, исключается возможность многократного учета вероятности существования узла или передачи информации по ЛС. В соответствии с этим выражение для вычисления вероятности доведения информации до n-ого узла по подмножеству путей Ω'_n принимает следующий вид:

$$P_\Omega\left(\Omega'_n\right) = E\left\{1 - \prod_{q=1}^{Q_n}\left(1 - \prod_{z=1}^{Z_q} p_{u,z}\right)\right\}, \tag{5}$$

где $p_{u,z}$ – значение вероятности срабатывания z-го помеченного перехода в q-ом пути до n-ого узла;

Z_q – количество помеченных переходов в q-ом пути до n-ого узла; E – функция поглощения.

Если $P_\Omega(\Omega'_n)$ не ниже $p_{\text{тр}}$, т.е. удовлетворяет критерию

$$P_\Omega\left(\Omega'_n\right) \geq p_{\text{тр}}, \tag{6}$$

то выполнение всего подмножества путей Ω'_n соответствует предъявляемым требованиям. В этом случае при выполнении доведении информации каждый путь подмножества Ω'_n является обязательным для реализации. Для отображения на схеме в подмножестве Ω'_n методом прямого перебора определяется путь, характеризующийся максимальным значением вероятности, по критерию

$$\hat{\omega}'_{n,q} = \arg \max_{\omega'_{n,q} \in \Omega'_n} P\left(\omega'_{n,q}\right). \tag{7}$$

Полученные данные представляются для отображения на схеме. На этом для n-ого узла выполнение методики завершается и осуществляется переход ко второму этапу для оценки данных следующих узлов.

Если $P_\Omega(\Omega'_n)$ не удовлетворяет критерию (6), представляются данные об отсутствии путей доведения информации до n-ого узла, соответствующих предъявляемым требованиям. На этом для n-ого узла выполнение методики завершается и осуществляется переход ко второму этапу для оценки данных следующих узлов.

Во всех случаях представления данных об отсутствии до n-ого узла ИТС путей доведения информации, соответствующих предъявляемым требованиям, должностные лица в рамках своих полномочий должны принять решение о порядке доведения информации до n-ого узла системы с использованием других средств или принять компенсационные меры, приводящие ИТС в состояние, позволяющее получить пути доведения информации до указанного узла, соответствующие требованиям.

Выполнение 2-4 этапов методики описано алгоритмом выбора пути доведения информации до каждого узла ИТС, изображённым на рисунках 2, 3, и начинается с загрузки исходных данных (рис. 2, блок 1) – данных, найденных на 1-ом этапе, а также значений $p_{\text{тр}}$ и $t_{\text{доп}}$.

Все три этапа методики проводятся в теле цикла «Перебор узлов» (блоки 3-17), в котором последовательно для каждого узла ИТС осуществляется выбор путей, обеспечивающих выполнение требуемых значений вероятности и времени доведения информации.

Второй этап методики реализован блоками 4-8. В случае если до n-ого узла не существует путей доведения информации, множество Ω_n пустое (блок 4, решение «Да»), выполнение методики для этого узла завершается.

Производится сохранение информации об отсутствии путей доведения информации до n-ого узла в ИТС, позволяющих обеспечить выполнение требуемых значений вероятности и времени доведения информации (блок 15) в данный расчетный момент времени. После этого счётчик n увеличивается на 1 (блок 16), и цикл может повториться для нового значения n до перебора всех N узлов рассматриваемой системы (блок 17).

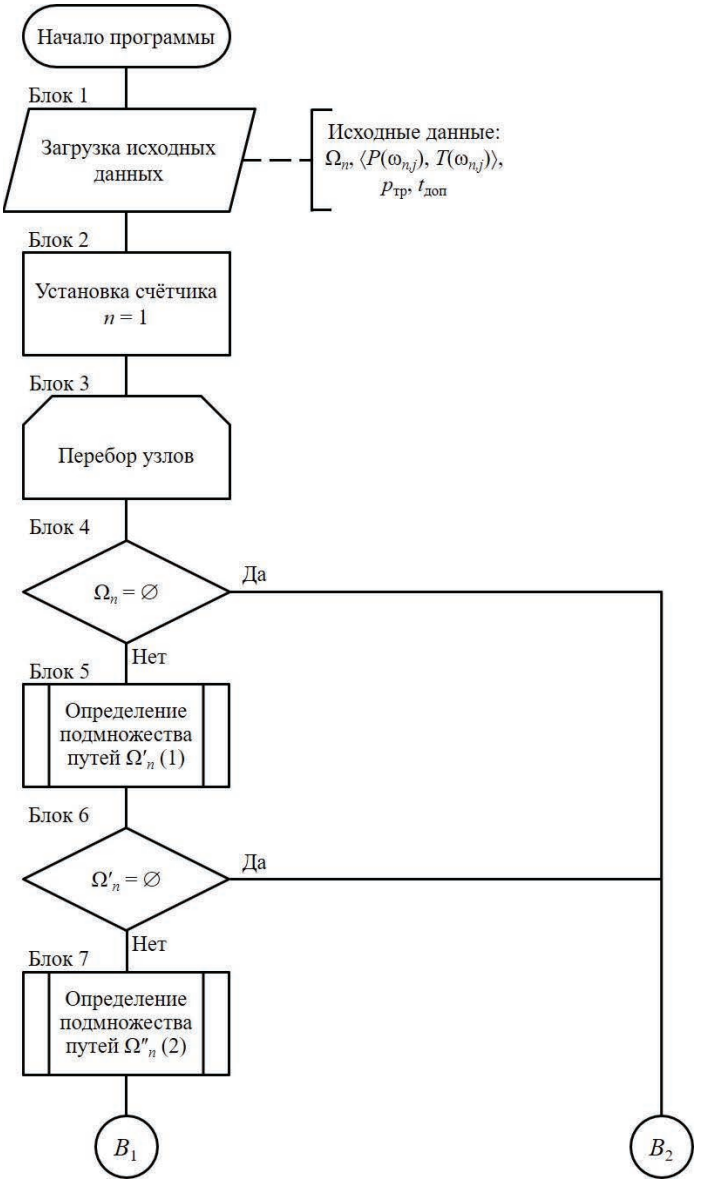


Рис. 2. Алгоритм выбора пути доведения информации до каждого узла ИТС (часть 1)

Если до n-ого узла существуют пути доведения информации (блок 4, решение «Нет»), то для него в блоке 5 методом прямого перебора определяется подмножество Ω'_n (1) путей, учитывающее ограничение по значению времени.

В случае, если в подмножестве Ω'_n не существует путей до n-ого узла (блок 6, решение «Да»), выполнение методики для этого узла завершается. Производится сохранение информации об отсутствии путей доведения информации до n-ого узла в ИТС, позволяющих обеспечить выполнение требуемых значений вероятности и времени доведения информации (блок 15) в данный расчетный момент времени. После этого счётчик n увеличивается на 1 (блок 16), и цикл может повториться для нового значения n до перебора всех N узлов рассматриваемой системы (блок 17).

Если подмножество Ω'_n не пустое (блок 6, решение «Нет»), то для неё в блоке 7 методом прямого перебора определяется подмножество Ω''_n (2) путей, учитывающее ограничение по значению вероятности. Второй этап методики заканчивается анализом подмножества Ω''_n.

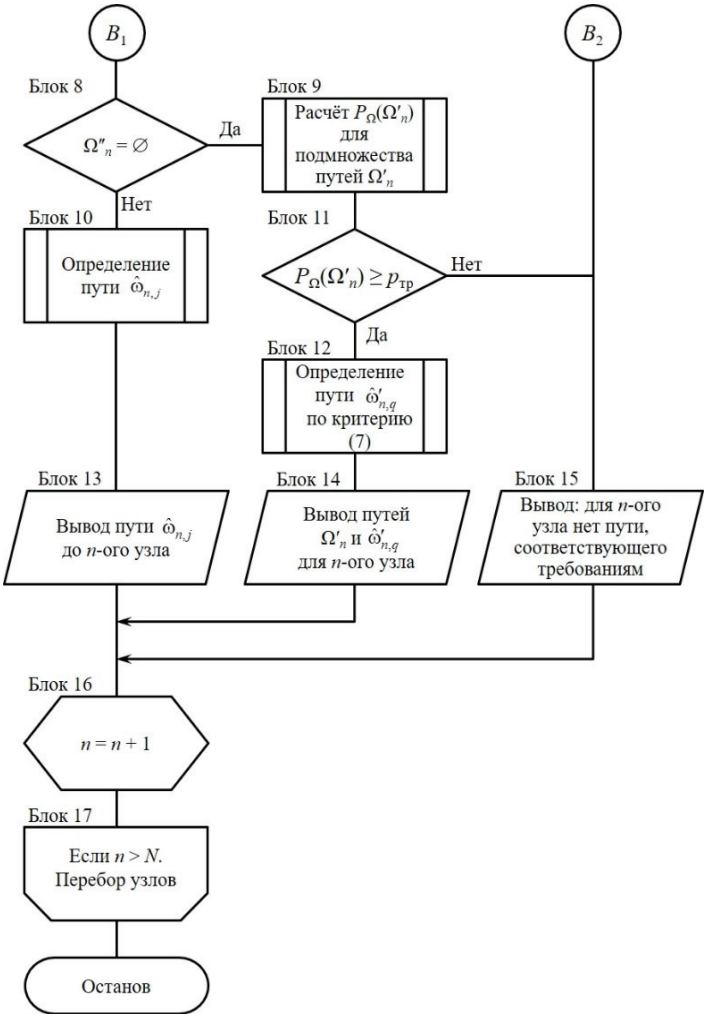


Рис. 3. Алгоритм выбора пути доведения информации до каждого узла ИТС (часть 2)

Если подмножество Ω''_n не пустое (блок 8, решение «Нет»), то начинается третий этап методики (блоки 10, 13). В блоке 10 методом прямого перебора в подмножестве Ω''_n определяется путь, характеризующийся максимальным значением вероятности по критерию (3). Далее производится его сохранение (блок 13).

На этом выполнение методики для n-ого узла завершается. После этого счётчик n увеличивается на 1 (блок 16), и цикл может повториться для нового значения n до перебора всех N узлов рассматриваемой системы (блок 17).

Если подмножество Ω''_n пустое (блок 8, решение «Да»), то начинается четвёртый этап методики (блоки 9, 11, 12, 14). Для подмножества путей доведения информации до n-ого узла Ω'_n, характеризуемого значением времени, не превышающим $t_{доп}$, вычисляется значение вероятности $P_\Omega(\Omega'_n)$ (5) доведения информации до этого узла по всему подмножеству путей (блок 9).

Если $P_{\Omega}(\Omega'_n)$ не ниже $p_{тр}$, т.е. удовлетворяет критерию (6), (блок 11, решение «Да»), то выполнение всего подмножества путей Ω'_n позволит удовлетворить заданные требования к вероятности и времени доведения информации, и, соответственно, при доведении информации каждый его путь является обязательным для реализации.

Для отображения на схеме в подмножестве Ω'_n методом прямого перебора определяется путь, характеризующийся максимальным значением вероятности по критерию (7) (блок 12). На этом выполнение методики для n-ого узла завершается, производится сохранение полученных данных для отображения на схеме (блок 14). После этого счётчик n увеличивается на 1 (блок 16), и цикл может повториться для нового значения n до перебора всех N узлов рассматриваемой системы (блок 17).

Если $P_{\Omega}(\Omega'_n)$ не удовлетворяет критерию (6) (блок 10, решение «Нет»), производится сохранение информации об отсутствии путей доведения информации до n-ого узла, позволяющих обеспечить выполнение требуемых значений вероятности и времени доведения информации (блок 11), выполнение методики для n-ого узла завершается. После этого счётчик n увеличивается на 1 (блок 16), и цикл может повториться для нового значения n до перебора всех N узлов рассматриваемой системы (блок 17).

Цикл завершается при окончании последовательного перебора всех узлов рассматриваемой системы (блок 17). На этом завершается выполнение методики.

Сходимость описанного алгоритма достигается конечностью цикла «Перебор узлов» (блоки 3-17), обеспечиваемой последовательным изменением значения счётчика и заданием его максимального значения.

В результате выполнения разработанной методики формируется три группы узлов:

a. множество узлов, имеющих путь доведения информации, характеризуемый максимальным значением вероятности, при этом не ниже $p_{тр}$, и значением времени, не превышающим $t_{доп}$;

b. множество узлов, имеющих совокупность Ω'_n путей доведения информации, выполнение которых позволит обеспечить выполнение требуемых значений вероятности и времени доведения информации;

c. множество узлов, для которых отсутствуют пути, позволяющие обеспечить выполнение требуемых значений вероятности и времени доведения информации.

Таким образом, разработанная методика оперативного выбора путей доведения информации в ИТС позволяет для каждого узла рассматриваемой системы выбрать путь (или совокупность путей) доведения информации, позволяющий обеспечить выполнение требуемых значений вероятности и времени доведения информации, или выявить узлы, для которых таких путей не существует, для своевременного принятия компенсационных мер.

Для одного из частных случаев информационно-телекоммуникационных систем, а, именно, автоматизированной системы управления, данная методика доведена до функционирующего программного продукта, она реализована в среде Delphi 10.3.3 Community Edition, получено свидетельство о регистрации программы для ЭВМ [8].

Литература

1. *Шкорина А.В., Шухардин А.Н.* Модель территориально-распределенной иерархической автоматизированной системы управления // Информация и космос. 2020. № 3. С. 94-99.

2. *Шухардин А.Н., Шкорина А.В.* Модель информационно-телекоммуникационной системы на базе сетей Петри // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2021. № 1. С. 158-161.

3. *Питерсон Дж.* Теория сетей Петри и моделирование систем. перевод с английского под ред. В. А. Горбатова. М.: Мир, 1984. 264 с.

4. *Тронин В. Г.* Применение раскрашенных сетей Петри в моделировании вычислительной сети // Автоматизация процессов управления. 2007. № 2. С. 97-102.

5. *Шухардин А.Н., Шкорина А.В.* Оценка вероятностно-временных характеристик доведения информации в автоматизированной системе управления войсками и оружием // Вестник Ярославского высшего военного училища противовоздушной обороны. 2019. № 4(7). С. 162-169.

6. *Шухардин А.Н., Шкорина А.В.* Методика оперативного оценивания вероятностей и сроков доставки сообщений в информационно-телекоммуникационных системах // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2021. № 1. С. 153-157.

7. ГОСТ Р 53111–2008 Устойчивость функционирования сети связи общего пользования. Требования и методы проверки: национальный стандарт Российской Федерации. М.: Стандартинформ, 2008.

8. *Шкорина А.В.* Методика выбора допустимого варианта доведения информации в автоматизированной системе управления // Свидетельство о регистрации программы для ЭВМ №2020616029. 08.06.2020.

METHODOLOGY FOR THE OPERATIONAL SELECTION OF WAYS TO COMMUNICATE INFORMATION IN INFORMATION AND TELECOMMUNICATION SYSTEMS

ALEXANDER N. SHUKHARDIN
Rostov-on-Don, Russia

ALEXANDER V. SHKORINA
Balashikha, Russia

KEYWORDS: *operational selection, information and telecommunication system, ways of communicating information.*

ABSTRACT

Introduction. The application of the provisions of colored hierarchical Petri theory nets makes it possible to create models of modern complex information and telecommunication systems (ITS), which allow, with relatively low computational costs, to simulate the functioning of such systems, in which field studies are either impossible or not advisable during operation for various reasons. **Practical relevance:** The article discusses a methodology that allows you to choose ways to communicate information in information and telecommunication systems, as well as evaluate the fulfillment of the requirements for a system for communicating information to all nodes of the system with changes in the characteristics of the system during operation. The structural elements of this methodology are the evaluation methodology and mathematical model of the information and telecommunication systems unctioning when delivering messages. **Discussion:** The developed methodology for the rapid choice of ways to deliver information to the ITS allows for each node of the system under consideration to choose a path (or a set of ways) for delivering information that makes it possible to ensure that the required values of the probability and time of delivering information are met, or to identify nodes for which such paths do not exist, for timely acceptance compensatory measures.

REFERENCES

1. A. V. Shkorina, A. N. Shukhardinu Model of a geographically distributed hierarchical automated control system. *Information and space*. 2020. No. 3, pp. 94-99.

2. A. N. Shukhardin, A. V. Shkorin. Model of information and telecommunication system based on Petri nets. *Proceedings of the North Caucasian branch of the Moscow Technical University of Communications and Informatics*. 2021. No. 1, pp. 158-161.

3. J. Peterson. Theory of Petri nets and system modeling: translation from English, ed. V. A. Gorbatov. Moscow: Mir, 1984. 264 p.

4. V. G. Troniny. Application of colored Petri nets in computer network modeling. *Automation of control processes*. 2007. No. 2, pp. 97-102.

5. A. N. Shukhardin, A. V. Shkorin. Evaluation of the probabilistic-temporal characteristics of information delivery in an automated control system for troops and weapons. *Bulletin of the Yaroslavl Higher Military School of Air Defense*. 2019. No. 4(7), pp. 162-169.

6. A. N. Shukhardin, A. V. Shkorin. A technique for operative estimation of the probabilities and terms of message delivery in information and telecommunication systems. *Proceedings of the North Caucasian branch of the Moscow Technical University of Communications and Informatics*. 2021. No. 1. S. 153-157.

7. GOST R 53111–2008 Stability of functioning of a public communication network. Requirements and verification methods: national standard of the Russian Federation. Moscow: Standartinform, 2008.

8. A. V. Shkorin. Methodology for choosing a valid option for communicating information in an automated control system. Certificate of registration of the computer program No. 2020616029. 06/08/2020.

INFORMATION ABOUT AUTHORS:

Shukhardin A.N., North Caucasus branch of Moscow Technical University of Communications and Informatics, Rostov-on-Don, Russia
Shkorina A.V., Military Academy of Strategic Missile Forces, Moscow region, Balashikha, Russia

For citation: Shukhardin A.N., Shkorina A.V. Methodology for the operational selection of ways to communicate information in information and telecommunication systems. *H&ES Reserch*. 2023. Vol. 15. No 1. P. 21-26. doi: 10.36724/2409-5419-2023-15-1-21-26 (In Rus)

doi: 10.36724/2409-5419-2023-15-1-27-36

РАЗРАБОТКА АЛГОРИТМА ВЫЯВЛЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ ДЛЯ ПЛАТФОРМЫ ANDROID ПУТЕМ ПРОВЕДЕНИЯ АНАЛИЗА ФАЙЛА МАНИФЕСТА

БАЙРАШНЫЙ
Алексей Олегович[1]

БОЛЬШАКОВ
Александр Сергеевич[2]

АННОТАЦИЯ

Введение: Введение: имеющиеся алгоритмы выявления вредоносного программного обеспечения могут недостаточно эффективно обнаруживать модифицированные виды вирусов в приложениях смартфонов. В связи с этим рассмотрены вопросы, связанные с использованием анализа файла манифеста приложения операционной системы Android. Предложено проведение комплексного анализа кода и манифеста мобильного приложения с целью формирования атрибутов вредоносного программного обеспечения и создания алгоритмов выявления для операционной системы Android. **Цель исследования:** показать актуальность реверс-инжиниринга и разработать алгоритм эвристического анализа, позволяющий обнаруживать наличие скрытого вредоносного кода в мобильных приложениях посредством анализа метаданных кода манифеста приложения. **Методы:** в статье использован метод реверс-инжиниринга на примере актуальных троянов для формирования атрибутов, характеризующих возможность наличия скрытого вредоносного кода в мобильном приложении. Предложенный алгоритм обеспечивает принятие совокупного решения о возможности наличия скрытой вредоносной нагрузки с учетом рангов и логических значений анализируемых атрибутов. **Результаты:** предложенный подход протестирован с использованием базы данных VirusTotal для оценки ошибок первого и второго рода и показал 78% эффективность по выявлению вредоносного программного обеспечения по метаданным кода манифеста с использованием предложенных атрибутов, не требует сложных вычислений и трудоемких затрат и может быть основанием использования искусственного интеллекта в сочетании с проведением статического и динамического анализа для выявления вирусов операционной системы Android. **Практическая значимость:** разработанный алгоритм может быть использован в виде дополнительного элемента в комплексной системе антивирусной защиты смартфона от заражения вредоносом.

Сведения об авторах:

[1] бакалавр МТУСИ, Москва, Россия, bayir678@gmail.com

[2] к.т.н., доцент кафедры ИБ МТУСИ, Москва, Россия, alexbol57@mail.ru

КЛЮЧЕВЫЕ СЛОВА: *манифест приложения; антивирусная защита; информационная безопасность; вредоносное ПО; Android-Banking.*

Для цитирования: Байрашный А.О., Большаков А.С. Разработка алгоритма выявления вредоносных программ для платформы Android путем проведения анализа файла манифеста // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 27-36. doi: 10.36724/2409-5419-2023-15-1-27-36

Введение

Вредоносное программное обеспечение (ВПО или вредонос), нацеленное на пользователей устройств на базе операционной системы (ОС) Android, развивается быстрее многих других видов вирусного программного обеспечения, в основном из-за того, что исходный код этой платформы является открытым и устройства этой ОС пользуются наибольшей популярностью. При этом одним из источников распространения вредоносов такого типа является официальный магазин Play Маркет.

Функциональные особенности многих вредоносов для Android-банкинга позволяют осуществлять кражу банковских данных жертв с целью совершения несанкционированных покупок, а также перевода денежных средств на сторонние счета.

С каждым годом появляется огромное число новых вирусов, которым удаётся обходить автоматизированные средства защиты. И именно поэтому специалистам информационной безопасности так необходимо исследовать новые вирусы, чтобы улучшать алгоритмы выявления угроз. Под исследованием понимается реверс-инжиниринг вредоноса с целью получения кода ВПО для анализа его методов скрытия от антивирусов, нанесения ущерба или кражи информации.

Так, например, в феврале 2022 года специалисты ThreatFabric обнаружили новый банковский троян для Android, который они назвали Xenomorph-ом [1]. Он скрывался в общеизвестном магазине приложений Play market под видом приложения «Fast Cleaner», предназначенного для ускорения работы устройства и очистки лишних данных.

Произведя анализ кода, специалисты выявили его особенности, цели и методы. Однако это не было бы возможно без реверс-инжиниринга, так как вредоносная нагрузка зашифрована, а динамический анализ не даёт полной картины его возможностей, что обусловливает актуальность проведения реверс-инжиниринга.

В ряде статей [2, 3, 4, 5] было отмечено, что для автоматизированного обнаружения вирусов используется эвристический анализ кода, который представляется в виде решения задачи с использованием методов искусственного интеллекта, поэтому в данной статье была предпринята попытка сформировать атрибуты, которые бы легли в основу машинного обучения, и разработать соответствующий алгоритм, позволяющий выявлять наличие ВПО.

К сожалению, практика показывает [6, 7], что при скачивании мобильных приложений сигнатурный и поведенческий анализы не всегда позволяют определить наличие «полезной вредоносной» нагрузки. В связи с этим предложено использовать метод реверс-инжиниринга для формирования атрибутов, характеризующих возможность наличия скрытого вредоносного кода в мобильном приложении с целью идентификации вредоносности приложения.

Результаты диссертации (https://ugatu.su/assets/files/documents/dissov/07/2017/GavrilovGN/Dissert-GavrilovGN.pdf) показали высокую эффективность выявления вредоносов (93%) с помощью нейронной сети, в то время как большинство антивирусов выявляли вредоносы с вероятностью обнаружения от 43% до 80%. В этой диссертации рассматривалось принятие решения нейросети по системным вызовам, манифесту и коду. Что позволяет сделать вывод о том, что эффективность данной нейронной сети может быть улучшена за счёт добавления новых атрибутов. Именно поэтому авторы статьи предлагают произвести реверс-инжиниринг вредоносов для формирования новых атрибутов и произвести тестирование их эффективности на разработанном алгоритме с их использованием.

В статье основной акцент при проведении реверс-инжиниринга смещён в сторону анализа манифеста (файл, в котором содержатся точки входа, метаданные и разрешения, запрашиваемые приложением) поскольку автоматизация его анализа – гораздо проще, чем автоматизация анализа обфусцированного (запутанного) кода.

В данной работе авторы статьи предлагают использование реверс-инжиниринга на примере актуального по версии ThreatFabric [8] вредоноса Hydra для демонстрации структуры ВПО и формирования атрибутов, указывающих на признаки ВПО. С использованием данного вредоноса продемонстрировано практическое использование результатов реверс-инжиниринга в сочетании с разработанным эвристическим алгоритмом и исследована эффективность такого алгоритма, которая подтвердила актуальность применения реверс-инжиниринга.

1. Формирование атрибутов наличия ВПО

Поскольку вредоносы пытаются обойти проверки антивирусов, они зачастую прячут и шифруют вредоносную нагрузку от них. Поэтому проводят реверс-инжиниринг, чтобы понять каким образом они обходят проверки антивирусов и как улучшить выявление вредоносов.

1.1. Формирование атрибута A

Проводя реверс инжиниринг самых активных вредоносов по версии ThreatFabric [8], была выявлена следующая закономерность: отсутствие имён классов, заявленных в манифесте, в открытой части кода, но присутствующих в скрытой части кода, содержащей вредоносную нагрузку. В статье выполнен поиск такой закономерности на примере банковского трояна Hydra.

Все файлы, полученные в результате реверс-инжиниринга вредоноса Hydra, были открыты в среде разработки IntelliJ Idea как проект для удобства анализа его кода. Алгоритм проведения анализа ВПО представлен на рисунке 1.

После открытия файла манифеста среда разработки при помощи статического анализа определила отсутствующие классы, которые она выделила красным цветом. Фрагмент манифеста после статического анализа показан на рисунке 2. Определить отсутствующие классы можно и вручную, проводя поиск каждого в коде, но это не рациональный путь.

Рис. 1. Схема проведения анализа ВПО с использованием реверс-инжиниринга

Рис. 2. Фрагмент манифеста Hydra

На рисунке 2 визуализирован фрагмент манифеста с семью отсутствующими классами кода (они выделены красным цветом) в открытой его части, хотя при полном проведении анализа манифеста их насчитывается в количестве 26-ти:

1. com.sdktools.android.MainActivity
2. com.sdktools.android.MainActivity2
3. com.sdktools.android.bot.HelperAdmin$MyHomeReceiver
4. com.sdktools.android.bot.PermissionsActivity
5. com.sdktools.android.bot.components.commands.NLService
6. com.sdktools.android.bot.components.injects.system.InjAccessibilityService
7. com.sdktools.android.bot.components.locker.LockerActivity
8. com.sdktools.android.bot.components.locker.LockerActivity$DummyActivity
9. com.sdktools.android.bot.components.screencast.ScreencastService
10. com.sdktools.android.bot.components.screencast.ScreencastStartActivity
11. com.sdktools.android.bot.components.screencast.UnlockActivity
12. com.sdktools.android.bot.components.socks5.Socks5ProxyService
13. com.sdktools.android.bot.receivers.MainReceiver
14. com.sdktools.android.bot.sms.ComposeSmsActivity
15. com.sdktools.android.bot.sms.HeadlessSmsSendService
16. com.sdktools.android.bot.sms.MmsReceiver
17. com.sdktools.android.bot.sms.SmsReceiver
18. com.sdktools.android.core.PeriodicJobReceiver
19. com.sdktools.android.core.PeriodicJobService
20. com.sdktools.android.core.injects_core.CHandler
21. com.sdktools.android.core.injects_core.Screen

22. com.sdktools.android.core.injects_core.Worker
23. info.pluggabletransports.dispatch.service.DispatchReceiver
24. info.pluggabletransports.dispatch.service.DispatchService
25. info.pluggabletransports.dispatch.service.DispatchVPN
26. org.torproject.android.service.OrbotService

С целью определения назначения выявленных 26-и классов в виртуальной среде был активирован данный вредонос посредством дешифрации скрытой части кода, хранившейся в папке активов (assets) в файле с названием «rfrNI.json». После выполнения реверс-инжиниринга дешифрованного кода файла определено назначение отсутствующих классов, представленных в таблице 1.

Таблица 1

Назначение отсутствующих классов вредоноса Hydra

№	Имя класса	Назначение
1	com.sdktools.android.MainActivity	класс запуска
2	com.sdktools.android.MainActivity2	класс запуска
3	com.sdktools.android.bot.HelperAdmin$MyHomeReceiver	класс для управления службой администрирования устройства
4	com.sdktools.android.bot.PermissionsActivity	класс получения разрешений
5	com.sdktools.android.bot.components.commands. NLService	прослушка уведомлений
6	com.sdktools.android.bot.components.injects.system.InjAccessibilityService	класс, используемый для эксплуатации сервиса доступности
7	com.sdktools.android.bot.components.locker.LockerActivity	управление спящим режимом
8	com.sdktools.android.bot.components.locker.LockerActivity$DummyActivity	пустая активити с перезапуском
9	com.sdktools.android.bot.components.screencast.ScreencastService	запись экрана
10	com.sdktools.android.bot.components.screencast.ScreencastStartActivity	запроса разрешения на запись экрана
11	com.sdktools.android.bot.components.screencast.UnlockActivity	удержание экрана включённым
12	com.sdktools.android.bot.components.socks5.Socks5ProxyService	соединение через ssl
13	com.sdktools.android.bot.receivers.MainReceiver	прослушивает событие перезагрузки устройства
14	com.sdktools.android.bot.sms.ComposeSmsActivity	Не реализован
15	com.sdktools.android.bot.sms.HeadlessSmsSendService	Не реализован
16	com.sdktools.android.bot.sms.MmsReceiver	Не реализован
17	com.sdktools.android.bot.sms.SmsReceiver	Не реализован
18	com.sdktools.android.core.PeriodicJobReceiver	периодический запуск приёмников
19	com.sdktools.android.core.PeriodicJobService	периодический запуск сервисов
20	com.sdktools.android.core.injects_core.CHandler	логирование событий
21	com.sdktools.android.core.injects_core.Screen	определение параметров атаки наложения
22	com.sdktools.android.core.injects_core.Worker	выполнение переданных действий
23	info.pluggabletransports.dispatch.service.DispatchReceiver	запуск сервиса DispatchService
24	info.pluggabletransports.dispatch.service.DispatchService	управление сервисом DispatchVPN
25	info.pluggabletransports.dispatch.service.DispatchVPN	Не реализован
26	org.torproject.android.service.OrbotService	соединение с сервером управления, получение команд

Как видно таблицы 1, отсутствующие классы содержат вредоносные действия (№ =3;5;6;7;9) или запускающие их вспомогательные.

Таким образом отсутствие некоторых классов в открытой части кода, но при этом указанных в манифесте, может указывать на наличие скрытой вредоносной нагрузки. А значит можно выделить этот факт в виде атрибута для выявления ВПО.

1.2. Формирование атрибута Б

Для выполнения вредоносных действий необходимы опасные разрешения, такие как разрешения показа системных диалоговых окон, записи системных настроек, записи звука, получения, отправки и чтения смс сообщений, так как без них невозможно выполнять соответствующие действия, приводящие к утечке конфиденциальной информации пользователя смартфона.

```
<uses-permission android:name="android.permission.READ_PHONE_NUMBERS"/> <!-- чтение телефонных номеров устройства -->
<uses-permission android:name="android.permission.WRITE_SETTINGS"/> <!-- редактирование системных настроек -->
<uses-permission android:name="android.permission.WAKE_LOCK"/>
<uses-permission android:name="android.permission.REORDER_TASKS"/>
<uses-permission android:name="android.permission.WRITE_SMS"/>
<uses-permission android:name="android.permission.FOREGROUND_SERVICE"/>
<uses-permission android:name="android.permission.MODIFY_AUDIO_SETTINGS"/> <!-- редактирование глобальных настроек звука -->
<uses-permission android:name="android.permission.BLUETOOTH"/>
<uses-permission android:name="android.permission.CALL_PHONE"/> <!-- звонок без подтверждения -->
<uses-permission android:name="android.permission.CHANGE_WIFI_STATE"/>
<uses-permission android:name="android.permission.GET_TASKS"/>
<uses-permission android:name="android.permission.REQUEST_DELETE_PACKAGES"/>   <!-- удаление приложений -->
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE"/>
<uses-permission android:name="android.permission.ACCESS_WIFI_STATE"/>
<uses-permission android:name="android.permission.VIBRATE"/>
<uses-permission android:name="android.permission.SYSTEM_ALERT_WINDOW"/>   <!-- диалоговое окно поверх приложений -->
<uses-permission android:name="android.permission.CAPTURE_VIDEO_OUTPUT"/>
<uses-permission android:name="android.permission.REQUEST_INSTALL_PACKAGES"/>   <!-- установка приложений -->
<uses-permission android:name="android.permission.ACCESS_COARSE_LOCATION"/>
<uses-permission android:name="com.google.android.gms.permission.ACTIVITY_RECOGNITION"/>
<uses-permission android:name="android.permission.USE_FINGERPRINT"/>
<uses-permission android:name="android.permission.REQUEST_IGNORE_BATTERY_OPTIMIZATIONS"/>
<uses-permission android:name="android.permission.RECEIVE_BOOT_COMPLETED"/>
<uses-permission android:name="android.permission.ACCESS_BACKGROUND_LOCATION"/> <!-- доступ к местоположение -->
<uses-permission android:name="android.permission.QUICKBOOT_POWERON"/>
<uses-permission android:name="android.permission.INTERNET"/>   <!-- интернет -->
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE"/>
<uses-permission android:name="android.permission.DISABLE_KEYGUARD"/>
<uses-permission android:name="android.permission.CHANGE_WIFI_STATE"/>
<uses-permission android:name="android.permission.READ_CONTACTS"/>  <!-- чтение контактов -->
<uses-permission android:name="android.permission.RECEIVE_SMS"/>   <!-- получение sms -->
<uses-permission android:name="android.permission.RECEIVE_LAUNCH_BROADCASTS"/>
<uses-permission android:name="android.permission.QUERY_ALL_PACKAGES"/>
<uses-permission android:name="android.permission.ACTION_MANAGE_OVERLAY_PERMISSION"/>
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE"/>
<uses-permission android:name="android.permission.ACCESS_FINE_LOCATION"/>
<uses-permission android:name="android.permission.ACCESS_NOTIFICATION_POLICY"/>
<uses-permission android:name="android.permission.SEND_SMS"/>   <!-- отправка sms -->
<uses-permission android:name="android.permission.RECORD_AUDIO"/>   <!-- запись аудио -->
<uses-permission android:name="android.permission.READ_EXTERNAL_STORAGE"/>
<uses-permission android:name="android.permission.READ_SMS"/>   <!-- чтение sms -->
```

Рис. 3. Часть манифеста Hydra с разрешениями

В зависимости от содержания целевой атаки, направленной на смартфон, ВПО требуется получить определенный набор опасных разрешений. Как видно из рисунка 3, для перехвата конфиденциальных данных вредонос должен иметь опасные разрешения (*чтение телефонных номеров; редактирование системных настроек и т.д).* В связи с этим, можно сформировать следующий атрибут: если количество опасных разрешений превышает определённый порог, то это приложение является потенциально опасным.

1.3. Формирование атрибута B

В связи с тем, что целью активации вредоноса, как правило, является установление контроля над пользовательскими приложениями, который позволяет выполнить нелегитимные действия мошеннического характера, связанные с получением пользовательской конфиденциальной информацией, то интерес представляет возможность эксплуатации сервиса специальных возможностей (AccessibilityService) [9, 10, 11]. Данный сервис предназначен для лиц с ограниченными возможностями, но его эксплуатация может позволить злоумышленнику реализовать утечку конфиденциальной информации. Поскольку большинство пользователей и приложений в подобном сервисе не нуждается, то получение разрешения на его эксплуатацию указывает на вредоносный характер приложения в ОС Android. Поэтому целесообразно сформировать атрибут B. Наличие строки «AccessibilityService» в манифесте будет означать, что данное приложение является вредоносом.

На рисунке 2 в манифесте вредоноса Hydra видны строки с подстрокой «AccessibilityService», относящиеся к описанию сервиса com.sdktools.android.bot.components.injects.system.InjAccessibilityService, который эксплуатирует возможности AccessibilityService.

2. Разработка алгоритмов определения значений атрибутов ВПО

В основе алгоритмов определения атрибутов лежит сравнение манифеста Android приложения и его открытой части кода.

Таким образом для реализации контроля присутствия ВПО в мобильном приложении авторы предлагают проводить комплексный анализ по сформированным атрибутам:

А) анализ манифеста приложения и классов кода приложения с точки зрения их несогласованности;

Пусть конечное множество имён классов из манифеста – **M**, а конечное множество имён классов из кода (не включая скрытую часть) – **K.** Тогда атрибут A может принимать следующее логическое значение:

«Истина» – если $\mathbf{M} \not\subset \mathbf{K}$, что указывает на наличие скрытого кода, в том числе вредоносного кода;

«Ложь» – если $\mathbf{M} \subset \mathbf{K}$, что указывает на отсутствие скрытого кода.

На рисунке 4 приведена блок-схема алгоритма определения значения атрибута A.

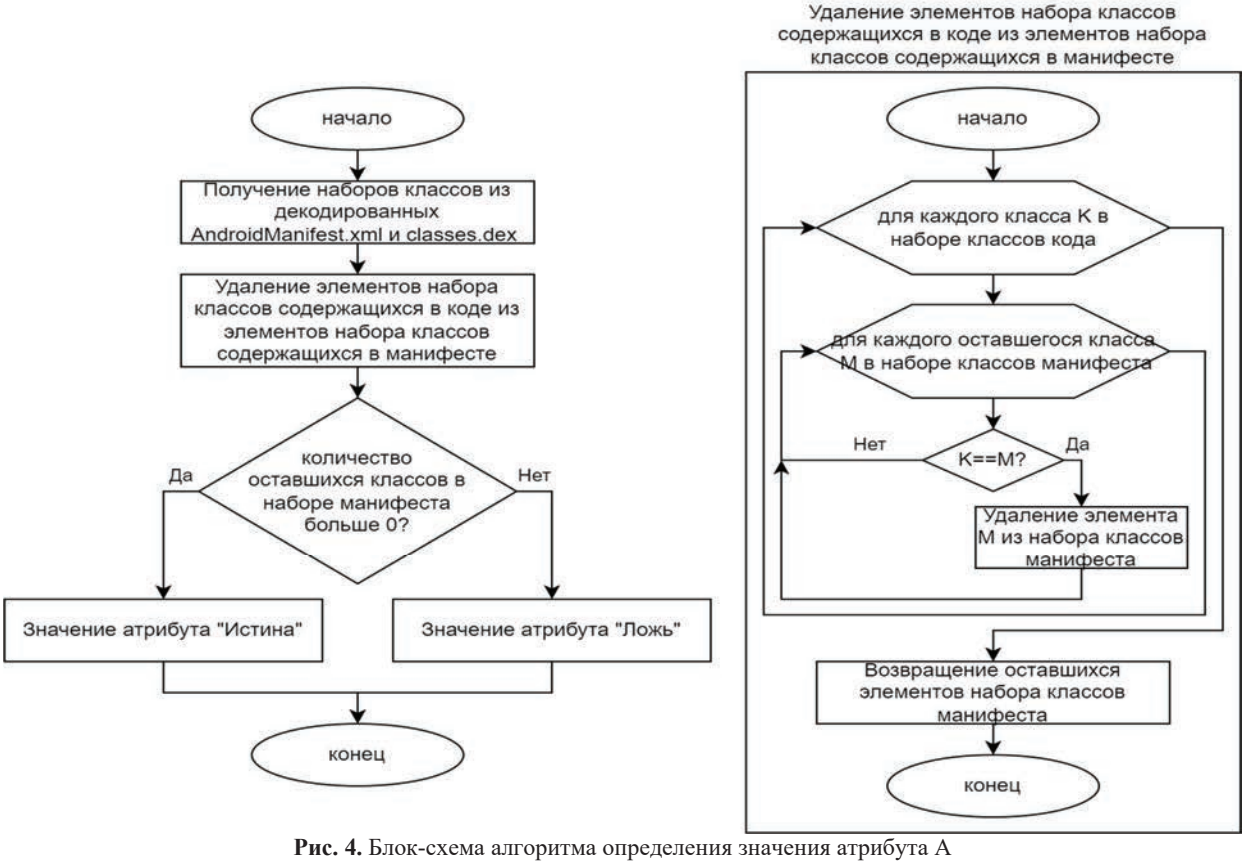


Рис. 4. Блок-схема алгоритма определения значения атрибута A

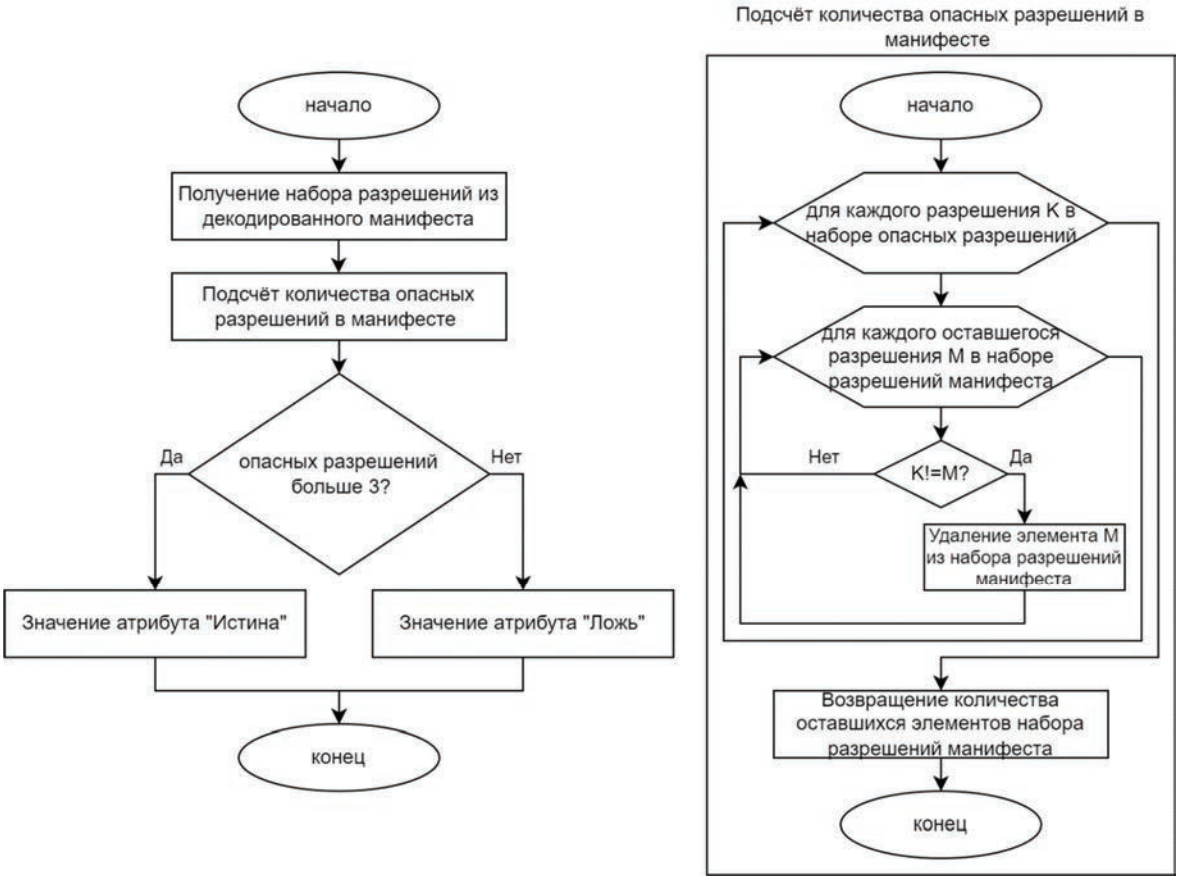


Рис. 5. Блок-схема алгоритма определения значения атрибута Б

Б) анализ опасных разрешений в манифесте приложения;

Пусть конечное множество имён разрешений из манифеста – **P**, конечное множество имён разрешений из опасных разрешений – **D**, функция μ является мерой множества, а n – число максимально допустимого количества опасных разрешений. Тогда атрибут Б может принимать следующее логическое значение:

«Истина» – если $\mu(\mathbf{P} \cap \mathbf{D}) > n$, то количество опасных разрешений превышает заданный порог и указывает на вредоностность кода;

«Ложь» – если $\mu(\mathbf{P} \cap \mathbf{D}) \leq n$, то предполагается отсутствие вредоносного кода.

Априори принимается, что ранги атрибутов А и Б равны друг другу.

На рисунке 5 приведена блок-схема алгоритма определения значения атрибута Б.

В) анализ манифеста на наличие строки «AccessibilityService».

Значение атрибута В может принимать следующие значения:

«Истина», если в манифесте содержится строка «AccessibilityService», что означает наличие вредоносного кода в приложении;

«Ложь» – в противном случае.

На рисунке 6 приведена блок-схема алгоритма определения значения атрибута В.

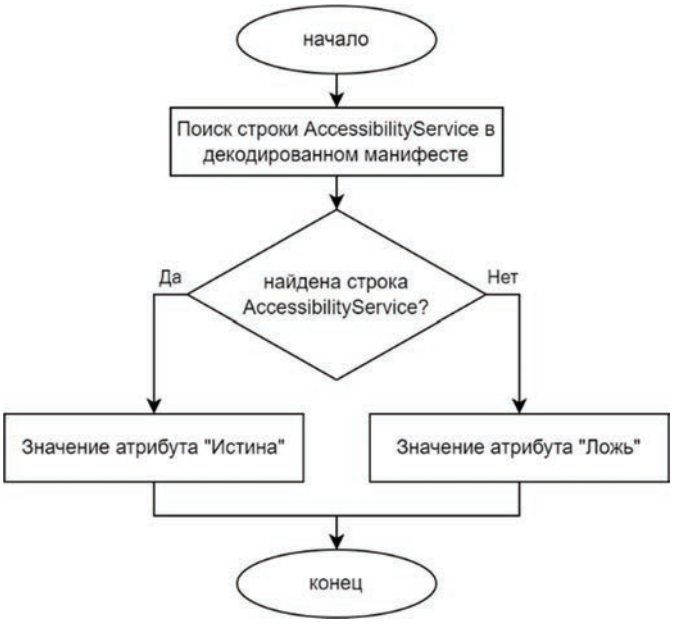


Рис. 6. Блок-схема алгоритма определения значения атрибута В

Алгоритмы, представленные на рисунках 4-6, применимы для Android приложений, поскольку используют особенности архитектуры Android платформы.

3. Разработка алгоритма выявления ВПО

Принимая во внимание рисунок 1 и на основании предложенных алгоритмов определения значений атрибутов А, Б, В разработан алгоритм принятия решения о наличии вредоносной нагрузки в мобильном приложении. Выявление наличия ВПО в приложении смартфона базируется на определении наличия строки «AccessibilityService» в манифесте и на возможность скрытого кода при достаточном количестве разрешений на запрос опасных разрешений. В случае выполнения одного из этих условий файл приложения считается вредоносным.

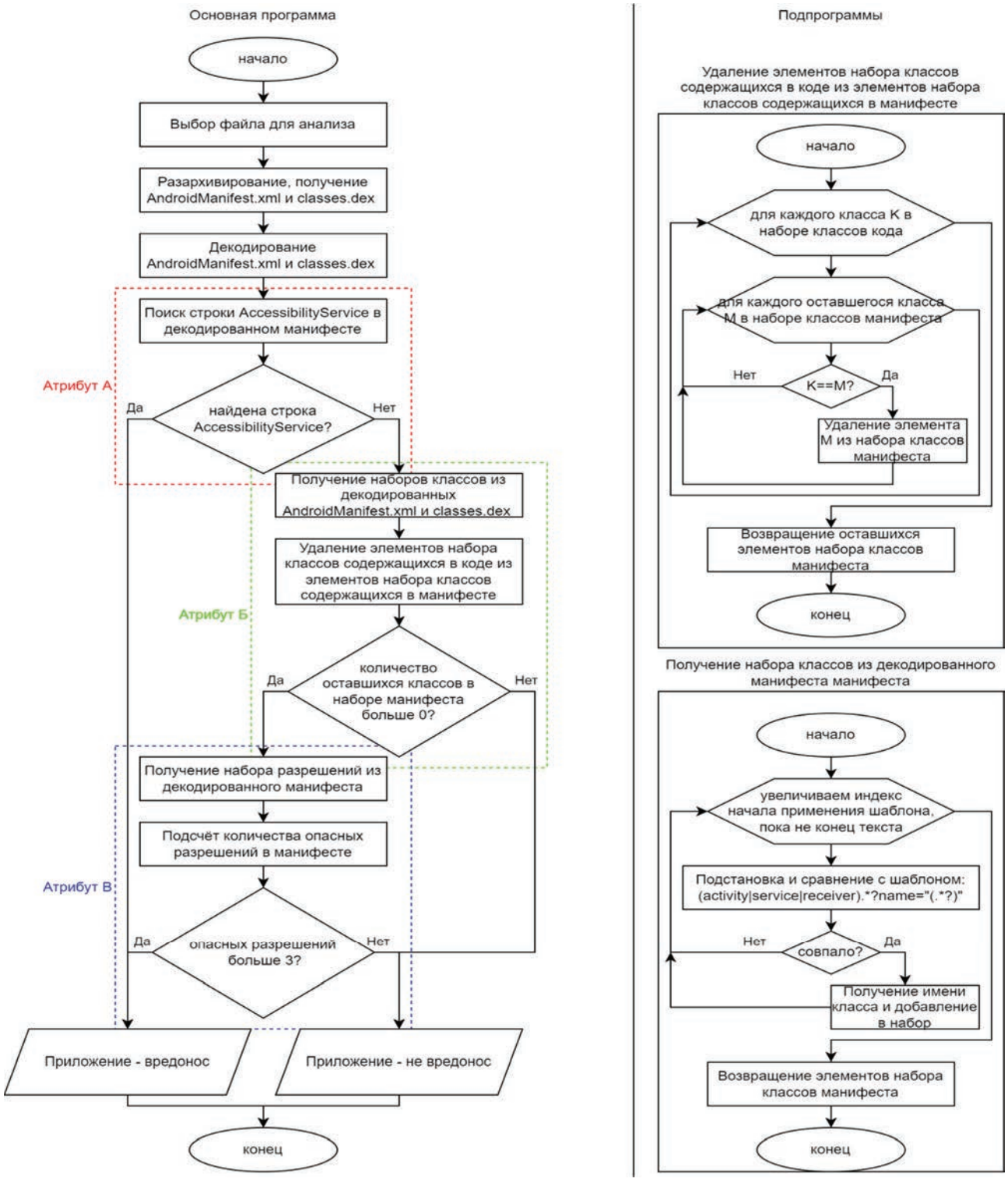


Рис. 7. Алгоритм выявления ВПО

Согласно рис. 7 из установочного файла выгружаются файл манифеста (AndroidManifect.xml) и файл кода (classes.dex), а затем производится их декодирование.

Посредством поиска через регуляторные выражения ищется строка «AccessibilityService», и если она присутствует, то файл считается вредоносным.

В случае, если не найдена строка «AccessibilityService», то путём поиска строк при помощи регуляторных выражений определяются классы, которые указаны в манифесте. Из этого набора классов алгоритм исключает классы, встречаемые в наборе классов кода (этот набор классов получен в процессе декодирования файла кода). Если после этого исключения остаётся хоть один класс, то это указывает на возможное наличие скрытой нагрузки или использование других приложений и сервисов (например: для использования сервисов оплаты банковской картой или получение ответа от переводчика слов).

Поскольку эта возможность неоднозначно характеризует вредонос, то проверяется количество опасных разрешений. Опасные разрешения – это разрешения на запрос разрешений показа системных диалоговых окон, записи системных настроек, записи звука, получения, отправки и чтения смс сообщений. Если таких разрешений больше 3-х (для обычных приложений их обычно не больше двух, поэтому на основе тестирования выбрано это число, характеризующее максимальное количество допустимых разрешений) и установлена возможность наличия скрытой нагрузки, то приложение считается вредоносом.

4. Тестирование алгоритма выявления ВПО

В данной работе авторы тестировали эффективность разработанного алгоритма с использованием открытой базы данных VirusTotal [12], базы данных ВПО (https://bazaar.abuse.ch) и базы данных (https://d.apkpure.com) приложений ОС Android, априори не содержащих ВПО. Под эффективностью разработанного алгоритма понимается оценка матрицы ошибок. Ошибками первого рода являются ситуации, когда разработанный алгоритм не определяет файл как вредонос, а ответ от VirusTotal не подтверждает эту гипотезу. Ошибками второго рода будут являются ситуации, когда разработанный алгоритм определят файл как вредонос, а ответ от VirusTotal не подтверждает эту гипотезу.

Диаграмма последовательности тестирования представлена на рисунке 8.

Согласно разработанному алгоритму, вычислялись хеши анализируемых файлов, которые сравнивались с содержимым базы данных VirusTotal. На основании сравнения производился расчет ошибок первого и второго рода. Результаты оценки матрицы ошибок отображены на рисунках 9 и 10, где:

V-VirusTotal;
A-Algorithm (разработанный алгоритм);
T-True F-False;
E1 – Количество ошибок первого рода (количество ситуаций, когда разработанный алгоритм дал ответ False, а ответ от VirusTotal – True);

E2 – Количество ошибок второго рода (количество ситуаций, когда разработанный алгоритм дал ответ True, а ответ от VirusTotal – False).

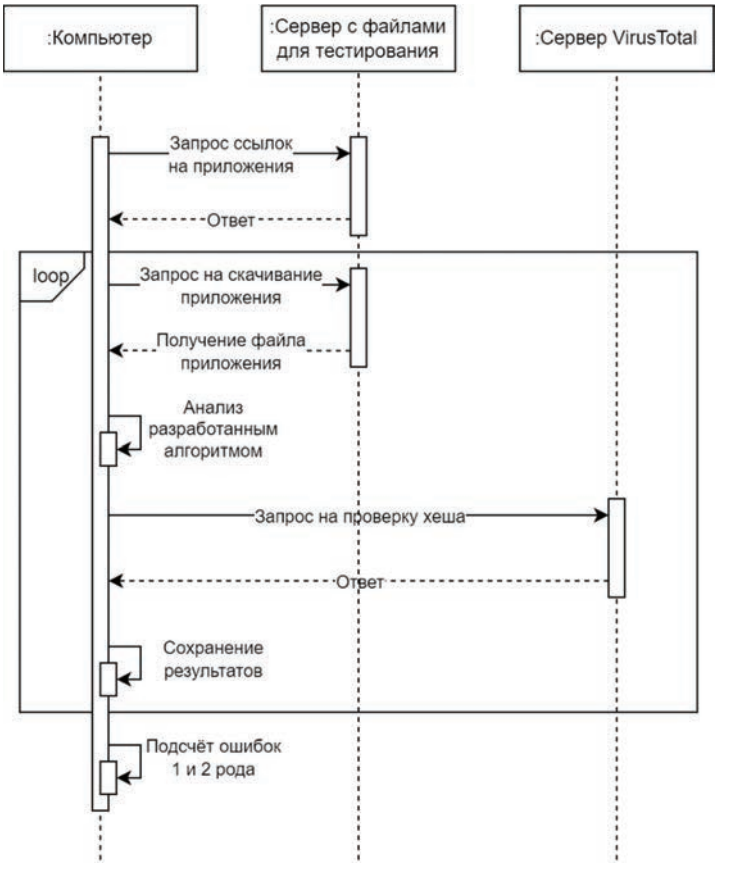


Рис. 8. Диаграмма последовательности тестирования алгоритма

https://bazaar.abuse.ch – сайт с базой данных вредоносов
https://d.apkpure.com – сайт мобильных приложений (обычных приложений)

Рис. 9. Результаты выявления ВПО для 1000 файлов

Для наглядности результатов, на рисунке 10 приведены диаграммы количества ошибок и правильных ответов.

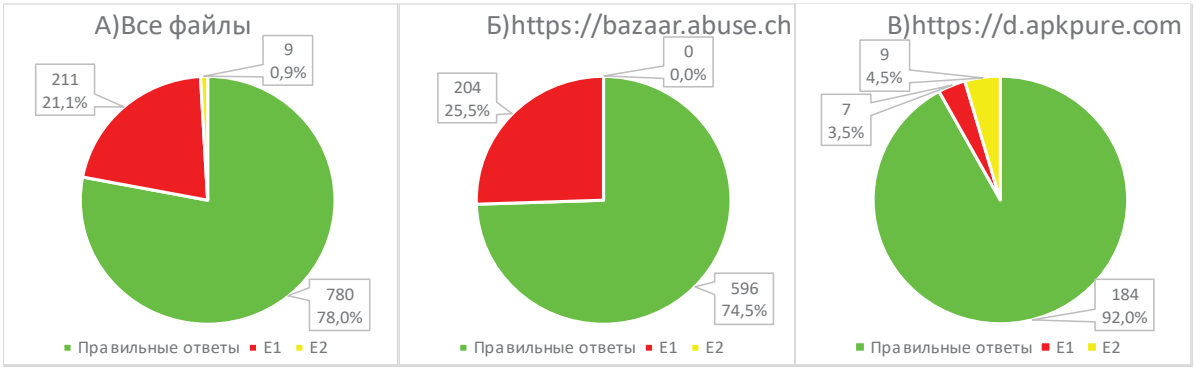


Рис. 10. Диаграммы количества ошибок первого и второго рода

По результатам тестирования, представленных на рисунке 10б, видно, что при анализе файлов из базы данных ВПО ошибок второго рода не наблюдается, а доля ошибок первого рода довольно большая (25,5%). Это говорит о том, что эти файлы не используют AccessibilityService, а используют малое количество опасных разрешений, связанных с выполнением узконаправленных задач ВПО, которые не фиксируются алгоритмом. Проводя анализ содержания файлов приложения, не содержащих ВПО, из рисунка 10в видно, что предложенный алгоритм допускает ошибки первого и второго рода в размере 3,5% и 4,5% соответственно.

На диаграмме рисунка 10а видно, что тестирование работы алгоритма при анализе файлов двух баз данных показало малый процент ошибок второго рода (0,9%) и довольно высокий процент обнаружения ВПО (78%).

Данный результат указывает на целесообразность применения алгоритма по выявления вредоносов для операционной системы Android в качестве дополнительного метода по выявлению ВПО. На наш взгляд, предложенный алгоритм с использованием анализа манифеста мобильного приложения и вышеуказанных атрибутов может повысить эффективность обнаружения модифицированного ВПО, а также его целесообразно использовать в качестве дополнительной метки для машинного обучения антивирусов, что позволит эффективней обнаруживать любое вредоносное программное обеспечение.

Заключение

На основе структур ВПО, рассмотренных при проведении реверс-инжиниринга вредоноса Hydra, разработан эвристический алгоритм по выявлению ВПО для операционной системы Android. Проведено его тестирование и показана эффективность его использования. Результаты предложенного алгоритма указали на целесообразность его использования в качестве дополнительного метода по выявлению модифицированного ВПО средствами антивирусной защиты.

Литература

1. Xenomorph: A newly hatched Banking Trojan [Электронный ресурс]. URL: Режим доступа: https://www.threatfabric.com/blogs/xenomorph-a-newly-hatched-banking-trojan.html (Дата обращения 12.11.2022)
2. Deceive the Heavens to Cross the sea [Электронный ресурс]. URL: Режим доступа: https://www.threatfabric.com/blogs/deceive-the-heavens-to-cross-the-sea.html (Дата обращения 12.11.2022)
3. Хеш четкий и хеш нечеткий. Как средства защиты ловят и классифицируют малварь [Электронный ресурс]. URL: Режим доступа: https://xakep.ru/2020/09/28/clear-hash/ (Дата обращения 12.11.2022)
4. Bypass Antivirus Dynamic Analysis [Электронный ресурс]. URL: Режим доступа: https://wikileaks.org/ciav7p1/cms/files/BypassAVDynamics.pdf (Дата обращения 12.11.2022)
5. *Лысенко А.* В. Анализ методов обнаружения вредоносных программ // Молодой ученый. 2016. № 21 (125). С. 758-761.
6. Deceive the Heavens to Cross the sea [Электронный ресурс]. URL: Режим доступа: https://www.threatfabric.com/blogs/deceive-the-heavens-to-cross-the-sea.html (Дата обращения 12.11.2022)
7. Уязвимости и угрозы мобильных банков sea [Электронный ресурс]. URL: Режим доступа: https://www.ptsecurity.com/ru-ru/research/analytics/vulnerabilities-mobile-banks-2020/#id3 (Дата обращения 12.11.2022)
8. 2022 mobile threat landscape update [Электронный ресурс]. URL: Режим доступа: https://www.threatfabric.com/blogs/h1-2022-mobile-threat-landscape.html (Дата обращения 08.11.2022)
9. Accessibility. Как сделать приложение доступным для пользователей с ограниченными возможностями [Электронный ресурс]. URL: Режим доступа: https://habr.com/ru/company/arcadia/blog/498476/ (Дата обращения 28.11.2022)
10. A11y Attacks: Exploiting Accessibility in Operating Systems [Электронный ресурс]. URL: Режим доступа: https://www.makeuseof.com/tag/android-accessibility-services-can-used-hack-phone (Дата обращения 28.11.2022)
11. How Android Accessibility Services Can Be Used to Hack Your Phone [Электронный ресурс]. URL: Режим доступа: https://www.makeuseof.com/tag/android-accessibility-services-can-used-hack-phone/ (Дата обращения 28.11.2022)
12. From zero to Zanubis [Электронный ресурс]. URL: Режим доступа: https://blog.virustotal.com/2022/11/from-zero-to-zanubis.html (Дата обращения 28.11.2022)

DEVELOPMENT OF AN ALGORITHM FOR DETECTING MALWARE FOR THE ANDROID PLATFORM BY ANALYZING THE MANIFEST FILE

ALEXEY O. BAYRASHNY
Moscow, Russia, bayir678@gmail.com

ALEXANDER S. BOLSHAKOV
Moscow, Russia, alexbol57@mail.ru

KEYWORDS: *reverse engineering, antivirus protection, information security, malware, Android-Banking, Android.*

ABSTRACT

Introduction: the available algorithms for detecting malicious software are often not enough to detect new types of viruses. In this regard, the issues related to the use of reverse engineering of malicious software codes intended for the Android operating system are considered. It is proposed to conduct a comprehensive analysis of the code and manifest of the mobile application in order to form the attributes of malicious software and create detection algorithms for a specific operating system. **Research objective:** to show the relevance of reverse engineering and to develop a heuristic analysis algorithm that allows detecting the presence of hidden malicious code in mobile applications by analyzing code metadata. **Methods:** the article uses the reverse engineering method on the example of actual Trojans to form attributes that characterize the possibility of hidden malicious code in a mobile application, according to the logical values of which the algorithm makes a cumulative decision about the possibility of hidden malicious load. **Results:** the proposed approach was tested using the VirusTotal database to evaluate errors of the first and second kind and showed 78% efficiency in detecting malicious software only from code metadata using the proposed attributes, does not require complex calculations and time-consuming costs and can be the basis for using artificial intelligence in combination with static and dynamic analysis to identify viruses of the Android operating system. Practical significance: the developed algorithm can be used as an additional element in the integrated antivirus protection system of a smartphone against malware infection.

REFERENCES

1. Xenomorph: A newly hatched Banking Trojan. URL: https://www.threatfabric.com/blogs/xenomorph-a-newly-hatched-banking-trojan.html (Date of access 12.11.2022).

2. Deceive the Heavens to Cross the sea URL: https://www.threatfabric.com/blogs/deceive-the-heavens-to-cross-the-sea.html (Date of access 12.11.2022).

3. Hash is crisp and hash is fuzzy. How security tools catch and classify malware. URL: https://xakep.ru/2020/09/28/clear-hash/ (Date of access 12.11.2022).

4. Bypass Antivirus Dynamic Analysis. URL: https://wikileaks.org/ciav7p1/cms/files/BypassAVDynamics.pdf (Date of access 12.11.2022).

5. A. V. Lysenko. Analysis of malware detection methods. Young scientist. 2016. No. 21 (125). pp. 758-761.

6. Deceive the Heavens to Cross the sea URL: https://www.threatfabric.com/blogs/deceive-the-heavens-to-cross-the-sea.html (Date of access 12.11.2022).

7. Vulnerabilities and threats of sea mobile banks. URL: https://www.ptsecurity.com/ru-ru/research/analytics/vulnerabilities-mobile-banks-2020/#id3 (Date of access 12.11.2022).

8. 2022 mobile threat landscape update. URL: https://www.threatfabric.com/blogs/h1-2022-mobile-threat-landscape.html (Date of access 08.11.2022).

9. Accessibility. How to make an application accessible to users with disabilities. URL: https://habr.com/ru/company/arcadia/blog/498476/ (Date of access 28.11.2022).

10. A11y Attacks: Exploiting Accessibility in Operating Systems. URL: https://wenke.gtisc.gatech.edu/ papers/a11y.pdf (Date of access 28.11.2022).

11. How Android Accessibility Services Can Be Used to Hack Your Phone. URL: https://www.makeuseof.com/tag/android-accessibility-services-can-used-hack-phone/ (Date of access 28.11.2022).

12. From zero to Zanubis. URL: https://blog.virustotal.com/2022/11/from-zero-to-zanubis.html (Date of access 28.11.2022).

INFORMATION ABOUT AUTHORS:

Alexey O. Bayrashny, bachelor MTUCI, Moscow, Russia
Alexander S. Bolshakov, Doctor of Technical Sciences, assistant professor of the Departament of IS of MTUCI, Moscow, Russia

For citation: Bayrashny A.O., Bolshakov A.S. Development of an algorithm for detecting malware for the android platform by analyzing the manifest file. H&ES Reserch. 2023. Vol. 15. No 1. P. 27-36. doi: 10.36724/2409-5419-2023-15-1-27-36 (In Rus)

doi: 10.36724/2409-5419-2023-15-1-37-47

МОДЕЛИ КОМПЬЮТЕРНЫХ АТАК НА ПРОГРАММНО-КОНФИГУРИРУЕМЫЕ СЕТИ

САЕНКО
Игорь Борисович[1]

КОТЕНКО
Игорь Витальевич[2]

ЛАУТА
Олег Сергеевич[3]

СКОРОБОГАТОВ
Сергей Юрьевич[4]

АННОТАЦИЯ

Введение: Технология SDN в ближайшем будущем позволит внедрить аспекты открытости кода сетевой составляющей облачной инфраструктуры, которая считается наиболее благоприятной основой для разработки и внедрения широкого спектра приложений. Она основана на реализации сетевых устройств и выполняемых ими функций не в отдельном сетевом оборудовании, а на любом хосте сети при помощи программного коммутатора OpenvSwitch. Такой подход позволяет использовать практически любое вычислительное устройство в сети в качестве коммутатора/маршрутизатора. В то же время применение новых технологий влечет за собой появление новых дестабилизирующих факторов на них, особое место в которых занимают кибератаки. Возможными результатами воздействия кибератак на SDN являются блокирование контроллера SDN, канала управления и мониторинга Open Flow, внедрение ложной информации о пользователе сети SDN, получающем сетевые услуги, нарушение установленных регламентов сбора, обработки и передачи информации в SDN, отказы и сбои в работе SDN, а также компрометация передаваемой или получаемой информации. **Цель работы** заключается в применении комплексного подхода при исследовании новых подходов построения сетей передачи данных в условиях компьютерных атак, получения вероятностно-временных характеристик компьютерных атак, характерных для программно-конфигурируемых сетей, что в свою очередь позволяет определить наиболее опасные компьютерные атаки, подверженные им элементы, а также задать требования для системы, обеспечивающей защиту от наиболее вероятных воздействий. **Используемые методы:** используемый при моделировании метод топологического преобразования стохастических сетей позволяет получить параметры, обладающие заданной полнотой и достоверностью. Разница не более 5% между значениями, полученными в имитационной и аналитической моделях, подтверждают их адекватность. Реализация в имитационной модели большого числа реальных устройств, а также использование нескольких средств сбора сетевой статистики позволяет достичь требуемой полноты моделирования. Значительная детализация этапов компьютерных атак в методе топологического преобразования стохастических сетей способствует получению расчетных выражений, достаточно точно описывающих вероятностно-временные характеристики системы. Научная новизна полученных результатов определяется использованием метода топологического преобразования стохастических сетей (ТПСС) для аналитического моделирования кибератак на SDN. **Результат:** представленная модель позволяет получить вероятностно-временные характеристики компьютерных атак, характерных для программно-конфигурируемой сети. Основу предложенной модели представляет имитационная модель основной целью которой является наиболее точное воспроизведение процессов реальной системы. Полученные значения вероятностно-временных характеристик компьютерных атак позволяют достаточно точно определить наиболее опасные их типы, а также наиболее вероятные места проявления. **Практическая значимость:** представленный метод является универсальным и может быть применен в государственной системе обнаружения, предупреждения и ликвидации последствий компьютерных атак при решении задач, связанных с обеспечением общественной безопасности.

Сведения об авторах:

[1] ведущий научный сотрудник, Санкт-Петербургский федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), доктор технических наук, профессор, Санкт-Петербург, Россия, ibsaen@comsec.spb.ru

[2] главный научный сотрудник, СПб ФИЦ РАН, доктор технических наук, профессор, Санкт-Петербург, Россия, ivkote@comsec.spb.ru

[3] профессор кафедры Государственного университета морского и речного флота им. адмирала С.О. Макарова (ГУМРФ), доктор технических наук, Санкт-Петербург, Россия, laos-82@yandex.ru

[4] адъюнкт 32 кафедры Военной академии связи имени маршала Советского союза С.М. Буденного, Санкт-Петербург, Россия, skorobogatovsu-vas@yandex.ru

Работа выполнена при частичной финансовой поддержке бюджетной темы FFZF-2022-0007

КЛЮЧЕВЫЕ СЛОВА: *компьютерные атаки, метод топологического преобразования стохастических сетей, программно-конфигурируемые сети, моделирование.*

Для цитирования: Саенко И.Б., Котенко И.В., Лаута О.С., Скоробогатов С.Ю. Модели компьютерных атак на программно-конфигурируемые сети // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 37-47. doi: 10.36724/2409-5419-2023-15-1-37-47

Введение

Резкий рост объемов трафика и изменение предоставляемых услуг большому числу пользователей, формирование высокопроизводительных кластеров для обработки больших данных и хорошо масштабируемых виртуализированных сред для предоставления облачных сервисов серьезно изменило структуру и требования, предъявляемые к современным сетям передачи данных [1]. Одной из концепций для построения сети передачи данных различных корпоративных структур является программно-конфигурируемая сеть (англ. – Software-defined network, SDN), которая работает, начиная с сетевого уровня [2, 3].

Программно-конфигурируемые сети помогают решить целый ряд имеющихся проблем, а также способствуют созданию автоматизированных, программируемых, гибких и экономичных сетевых инфраструктур [4-6]. Они помогают системно решить большинство накопившихся проблем, в том числе связанных с обеспечением сетевой и информационной безопасности [7, 8].

SDN – это открытая сетевая архитектура, предложенная в последние годы для устранения некоторых ключевых недостатков традиционных сетей передачи данных. Сторонники SDN утверждают, что логики управления сетью и сетевыми функциями являются двумя отдельными понятиями и поэтому должны быть разделены на разные уровни. С этой целью в SDN были введены понятия плоскости управления и плоскости данных: централизованная плоскость управления (иначе называемая *контроллером*) управляет логикой сети, контролирует функции инжиниринга траффика с плоскости данных (называемой *коммутаторами*), которые просто заботятся о пересылке пакетов между сетями.

Таким образом, SDN можно рассматривать как физически распределенную структуру коммутации с логически централизованным управлением, предназначенную для обеспечения высоко динамичного управления и качества обслуживания / политик безопасности.

Однако, как во многих новых решениях, в SDN есть ряд недостатков [9]:

полностью или практически полностью программное решение влечет за собой тысячи строк программного кода, который, в свою очередь, влечет за собой наличие непреднамеренных ошибок;

значительная часть уязвимостей перешла в технологию из стека протоколов TCP/IP;

наличие устройства, полностью управляющего сетью и владеющего всей информацией о сети, требует дополнительных средств и механизмов защиты;

новая технология подразумевает под собой интенсивное появление новых уязвимостей, характерных для новых решений данной технологии [10].

Таким образом, применение новых технологий влечет за собой появление новых сетевых дестабилизирующих факторов, особое место среди которых занимают кибератаки. Возможными результатами воздействия кибератак на SDN являются: блокирование контроллера SDN, канала управления и мониторинга Open Flow; внедрение ложной информации о пользователе сети SDN, получающем сетевые услуги; нарушение установленных регламентов сбора, обработки и передачи информации в SDN; отказы и сбои в работе SDN; компрометация передаваемой или получаемой информации.

Это позволяет считать, что кибератаки и способность противодействовать их реализации являются ключевыми факторами, определяющими устойчивость SDN. По этой причине в настоящей статье мы акцентируем свое внимание именно на кибератаках как наименее изученной, по нашему мнению, а также наиболее важной группе дестабилизирующих факторов. Более того, мы будем трактовать термин «устойчивость SDN» как способность сети передачи данных, в которой уровень управления сетью отделен от устройств передачи данных и реализуется программным образом, как одну из форм виртуализации сети, позволяющую реализовывать свои функции и процессы в условиях кибератак.

Рассматриваемый подход предполагает разработку вербальных моделей кибератак и построение их аналитических моделей при реализации. С целью построения аналитических моделей кибератак применяется метод топологического преобразования стохастических сетей. Для получения исходных данных при моделировании обоснован и развернут имитационной стенд SDN на виртуальной среде EVE-NG. Результатом моделирования является функция распределения времени и среднее время реализации кибератаки. Эти результаты используются затем для оценки показателей устойчивости SDN, нахождение которых осуществляется с использованием методов теории Марковских процессов [11]. Этот подход отличается более высокой точностью и сходимостью получаемых решений и хорошо зарекомендовал себя при моделировании многошаговых стохастических процессов различной природы.

Результаты исследований

Рассмотренный подход в настоящей статье получил экспериментальную проверку для некоторых наиболее известных и популярных видов атак. Атаки «Подмена сетевой топологии» являются характерными примерами атак пассивного типа, которые не наносят разрушений в SDN, но выявляют важную информацию, которую впоследствии злоумышленник может использовать для проведения более серьезных атак. Атака «Взлом, сбой контроллера» является характерным примером активных атак, которые существенно нарушают работоспособность SDN. Указанные типы атак будут рассмотрены в настоящей статье в качестве объектов аналитического моделирования.

Основные способы выполнения этапов реализации кибератак на SDN представлен в таблице 1.

В целях обеспечения основных требований, предъявляемых к модели с заданным уровнем точности, и получения наиболее достоверных вероятностно-временных характеристик, требуется разработать комплексную модель КА на SDN, состоящую из вербальной модели КА на SDN, математической модели и имитационной. С этой целью предлагается использовать эталонные модели КА и метод топологического преобразования стохастических сетей (ТПСС).

Таблица 1

Процесс кибератак на SDN

Этапы реализации воздействия	Основные способы выполнения
Сбор информации	Первый этап реализации атак — это сбор информации об атакуемой системе или узле. Он включает такие действия, как определение сетевой топологии, типа и версии операционной системы атакуемого узла, а также доступных сетевых и иных сервисов и т.п. Эти действия реализуются различными методами.
Изучение окружения	На этом этапе нападающий исследует сетевое окружение вокруг предполагаемой цели атаки. К таким областям, например, относятся узлы Internet-провайдера "жертвы" или узлы удаленного офиса атакуемой компании. На этом этапе злоумышленник может пытаться определить адреса "доверенных" систем (например, сеть партнера) и узлов, которые напрямую соединены с целью атаки (например, маршрутизатор ISP) и т.д. Такие действия достаточно трудно обнаружить, поскольку они выполняются в течение достаточно длительного периода времени и снаружи области, контролируемой средствами защиты (межсетевыми экранами, системами обнаружения атак и т.п.).
Идентификация топологии сети	Существует два основных метода определения топологии сети, используемых злоумышленниками: 1) изменение TTL (TTL modulation); 2) запись маршрута (record route).
Идентификация узлов	Идентификация узла, как правило, осуществляется путем посылки при помощи утилиты ping команды ECHO_REQUEST протокола ICMP. Ответное сообщение ECHO_REPLY говорит о том, что узел доступен. Существуют свободно распространяемые программы, которые автоматизируют и ускоряют процесс параллельной идентификации большого числа узлов, например, fping или nmap. Опасность данного метода заключается в том, что стандартными средствами узла запросы ECHO_REQUEST не фиксируются. Для этого необходимо применять средства анализа трафика, межсетевые экраны или системы обнаружения атак.
Идентификация сервисов или сканирование портов	Идентификация сервисов, как правило, осуществляется путем обнаружения открытых портов (port scanning). Такие порты очень часто связаны с сервисами, основанными на протоколах TCP или UDP. Например: • открытый 80-й порт подразумевает наличие Web-сервера; • 25-й порт - почтового SMTP-сервера; • 31337-й - серверной части троянского коня BackOrifice; • 12345-й или 12346-й - серверной части троянского коня NetBus и т.д.
Идентификация операционной системы	Основной механизм удаленного определения ОС - анализ ответов на запросы, учитывающие различные реализации TCP/IP-стека в различных операционных системах. В каждой ОС по-своему реализован стек протоколов TCP/IP, что позволяет при помощи специальных запросов и ответов на них определить, какая ОС установлена на удаленном узле. Другой, менее эффективный и крайне ограниченный способ идентификации ОС узлов - анализ сетевых сервисов, обнаруженных на предыдущем этапе. Например, открытый 139-й порт позволяет сделать вывод, что удаленный узел, вероятнее всего, работает под управлением ОС семейства Windows. Для определения ОС могут быть использованы различные программы, например, nmap или queso.
Определение роли узла	Предпоследним шагом на этапе сбора информации об атакуемом узле является определение его роли, например, выполнение функций межсетевого экрана или Web-сервера. Выполняется этот шаг на основе уже собранной информации об активных сервисах, именах узлов, топологии сети и т.п. Например, открытый 80-й порт может указывать на наличие Web-сервера, блокировка ICMP-пакета указывает на потенциальное наличие межсетевого экрана, а DNS-имя узла proxy.domain.ru или fw.domain.ru полностью раскрывает роль узла.
Определение уязвимостей узла	Последний шаг - поиск уязвимостей. На этом шаге злоумышленник при помощи различных автоматизированных средств или вручную определяет уязвимости, которые могут быть использованы для реализации атаки. В качестве таких автоматизированных средств могут быть использованы ShadowSecurityScanner, nmap, Retina и т.д.
Реализация атаки	С этого момента начинается попытка доступа к атакуемому узлу. Доступ может быть как непосредственный, т.е. проникновение на узел, так и опосредованный, например, при реализации атаки типа "Отказ в обслуживании". Реализация атак в случае непосредственного доступа также может быть разделена на два этапа: • проникновение; • установление контроля.
Цели реализации атак	Необходимо отметить, что злоумышленник на втором этапе может преследовать две цели. Во-первых, это может быть получение несанкционированного доступа к самому узлу и содержащейся на нем информации. Во-вторых, это может быть получение несанкционированного доступа к узлу для осуществления дальнейших атак на другие узлы. Первая цель, как правило, осуществляется только после реализации второй. Иными словами, сначала злоумышленник создает себе базу для дальнейших атак и только после этого проникает на другие узлы. Это необходимо для того, чтобы скрыть или существенно затруднить нахождение источника атаки.
Завершение атаки	Этапом завершения атаки является "заметание следов" со стороны злоумышленника. Обычно это реализуется путем удаления соответствующих записей из журналов регистрации узла и других действий, возвращающих атакованную систему в исходное, "предъатакованное" состояние.

Таблица 2

Классификация атак, специфичных для SDN

Плоскость SDN	Угроза/атака	Описание
1. анные	1.1. Синхронная атака (Flooding attacks)	Таблицы потоков коммутаторов содержат только ограниченное количество правил потока.
	1.2. Атака «человек посередине»	Активное прослушивание, при котором злоумышленник устанавливает независимые связи, поскольку TLS – это вариант дополнения, а не стандарт.
	1.3. Взлом или сбой контроллера	Взлом контроллера увеличивает риск для плоскости данных.
2. Управление	2.1. Вмешательство в цепочку обслуживания (Service Chain Interference)	Данная атака может привести к двум последствиям: 1) вредоносное приложение может участвовать в цепочке и удалить управляющее сообщение до того, как другие приложения получат необходимую информацию; 2) вредоносное приложение может попасть в бесконечный цикл, чтобы остановить цепное выполнение приложений.
	2.2. Злоупотребление внутренней памятью	Использование внутренней памяти контролера
	2.3. Манипуляции с управляющими сообщениями	Манипуляции управляющими сообщениями.
	2.4. Злоупотребление северным API	Приложение SDN может манипулировать поведением других приложений, используя плохо спроектированный северный API.
	2.5. Манипуляции с системными переменными	Манипуляция системными переменными.
	2.6. Отравление топологии сети	Изменение топологии сети.
	2.7. DoS-атаки	Нет значительной аутентификации.
	2.8. Несанкционированный доступ к контроллеру	Нет достоверных прав доступа пользователей.
	2.9. Масштабируемость и доступность	Увеличение размера и сдвига сети создает проблемы.
3. Приложен	3.1. Отсутствие аутентификации и авторизации	Для приложений не используются никакие средства аутентификации.
	3.2. Вставка мошеннических правил потока	Подключенные вредоносные приложения могут вставлять ложные правила в таблицы потоков.
	3.3. Отсутствие контроля доступа	Сложно реализовать контроль доступа.

Вербальная модель кибератак на SDN. Для SDN существует ряд специфичных КА, нацеленных на плоскости, реализуемые в рамках данной технологии построения сети. Классификация атак, специфичных для SDN, представлена в таблице 5.

Большинство сетевых атак относятся к следующим типам:

Спуфинг-атаки (Spoofing);

«Атака человек посередине» (Man-in-the-Middle);

Несанкционированный доступ (Tampering);

Отказ (Repudiation);

Утечка информации (Information Disclosure);

Отказ в обслуживании (Deny-of-Service).

Кибератаки на элементы SDN реализуются в виде целенаправленных программно-аппаратных воздействий, приводящих к нарушению или снижению эффективности выполнения технологических циклов.

Модель КА на SDN «Подмена сетевой топологии». Для построения математической модели кибератаки типа «Подмена сетевой топологии» представим сценарий ее реализации в виде последовательности действий, указанных в таблице 3. Итогом кибератаки «Подмена сетевой топологии» становится работа злоумышленника, выдаваемого за доверенное сетевое устройство с использованием идентификационных данных, полученных от средств технической компьютерной разведки.

Описанный выше процесс реализации КА представим в виде стохастической сети (рис. 1).

Таблица 3

Сценарий кибератаки типа «Подмена сетевой топологии»

№ п/п	Описание этапа проведения КА	Условное обозначение этапа
1	Проверка канала подключения к атакуемой сети	*w(s)*
2	Обмен с сетевым контроллером по протоколу управления Open Flow, выдача своего устройства за легитимное устройство сети	*m(s)*
3	Отправка данных сетевой статистики на контроллер по протоколу управления, проверка отклика контроллера	*l(s)*
4	Построение топологии сети путем отправки ложной сетевой статистики	*z(s)*
5	Управление работой сети методом обмана сетевого контроллера ложными сообщениями Open Flow протокола	*d(s)*

Для определения эквивалентной функции вводится понятие замкнутой стохастической сети, а также петель первого и *k*-го порядков [12-14].

Эквивалентная функция петли *k*-го порядка определяется как

$$Q_k(s) = \prod_{i=1}^{k} Q_i(s), \quad (1)$$

где $Q_i(s)$ – эквивалентная функция i-й петли первого порядка, определяемая как произведение эквивалентных функций ветвей, входящих в эту петлю.

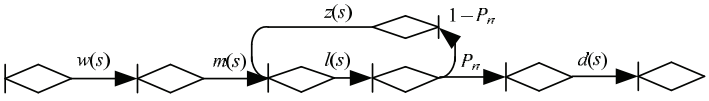


Рис. 1. Стохастическая сеть КА типа «Подмена сетевой топологии»

Преобразуем исходную стохастическую сеть в замкнутую (рис. 2).

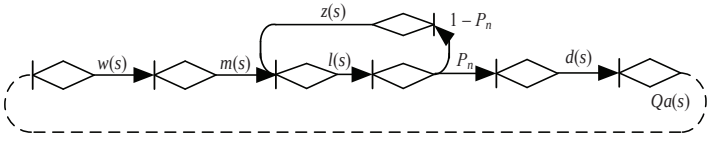


Рис. 2. Замкнутая стохастическая сеть КА типа «Подмена сетевой топологии»

Замкнутая сеть позволяет использовать для определения эквивалентной функции исходной сети топологическое уравнение Мейсона для замкнутых графов

$$H = 1 + \sum_{k=1}^{K}(-1)^k\, Q_k(s) = 0, \tag{2}$$

где K – максимальный порядок петель, входящих в стохастическую сеть.

После того как стохастическую сеть замкнули фиктивной ветвью $Q_a(s) = 1/h(s)$, где $h(s)$ – искомая эквивалентная функция, определим все петли.

Петли первого порядка:

$$w(s)\cdot m(s)\cdot l(s)\cdot P_п\cdot d(s)/h(s);$$
$$(1-P_п)\cdot z(s)\cdot l(s). \tag{3}$$

Петель второго и более высоких порядков нет.

Тогда уравнение (3) можно записать так

$$1 - w(s)\cdot m(s)\cdot l(s)\cdot P_п\cdot d(s)/h(s) - (1-P_п)\cdot z(s)\cdot l(s) = 0. \tag{4}$$

Эквивалентная функция в этом случае имеет вид

$$h(s) = \frac{w(s)\cdot m(s)\cdot l(s)\cdot P_п\cdot d(s)}{1-(1-P_п)\cdot z(s)\cdot l(s)}. \tag{5}$$

Для определения расчетного соотношения функции распределения допустим, что

$$\begin{cases} W(t) = 1 - exp[-wt]; \\ M(t) = 1 - exp[-mt]; \\ L(t) = 1 - exp[-lt]; \\ D(t) = 1 - exp[-dt]; \\ Z(t) = 1 - exp[-zt], \end{cases} \tag{6}$$

где $w = 1/\bar{t}_{зап}$, $m = 1/\bar{t}_{пар}$, $l = 1/\bar{t}_{перех}$, $d = 1/\bar{t}_{отчет}$, $z = 1/\bar{t}_{повт}$, $\bar{t}_{зап}$, $\bar{t}_{пар}$, $\bar{t}_{перех}$, $\bar{t}_{отчет}$, $\bar{t}_{повт}$ – средние времена реализации отдельных процессов КА.

С использованием преобразования Лапласа находим изображения функций плотности распределения времени выполнения k-го процесса КА:

$$l(s) = \int_0^\infty exp(-st)d[L(t)] = \frac{l}{l+s};$$
$$d(s) = \int_0^\infty exp(-st)d[D(t)] = \frac{d}{d+s};$$
$$z(s) = \int_0^\infty exp(-st)d[Z(t)] = \frac{z}{z+s};$$
$$w(s) = \int_0^\infty exp(-st)d[W(t)] = \frac{w}{w+s};$$
$$m(s) = \int_0^\infty exp(-st)d[M(t)] = \frac{m}{m+s}. \tag{7}$$

После подстановки выражения (6) в (7), а полученных результатов в (5), получим [15]

$$h(s) = \frac{w\cdot m\cdot l\cdot P_п\cdot d\cdot(z+s)}{(w+s)\cdot(d+s)\cdot(m+s)\cdot[(l+s)\cdot(z+s)-(1-P_п)\cdot z\cdot l]} \tag{8}$$

Для упрощения расчетов определим:

$$A = d + l + m + w + z;$$
$$B = [l\cdot(d+m+z) + w\cdot(d+l+m+z) + d\cdot(m+z) + m\cdot z - (1-P_n)\cdot m\cdot z];$$
$$C = [w\cdot[l\cdot(d+m+z) + d\cdot(m+z) + m\cdot z - (1-P_n)\cdot m\cdot z] +$$
$$+d\cdot(m\cdot z - (1-P_n)\cdot m\cdot z) + l\cdot[d\cdot(m+z) + m\cdot z - (1-P_n)\cdot m\cdot z]];$$

$$D = \Big[w\cdot\big[d\cdot(m\cdot z - (1-P_n)\cdot m\cdot z) + l\cdot[d\cdot(m+z) + m\cdot z - (1-P_n)\cdot m\cdot z]\big] +$$
$$+d\cdot l\cdot\big(m\cdot z - (1-P_n)\cdot m\cdot z\big)\Big];$$
$$E = d\cdot l\cdot w\big(m\cdot z - (1-P_n)\cdot m\cdot z\big). \tag{9}$$

С целью определения оригинала эквивалентной функции (8) используем разложение Хэвисайда:

$$h(s) = \sum_{k=1}^{n}\frac{f(s_k)}{\varphi'(s_k)}\cdot\frac{1}{s-s_k} = \sum_{k=1}^{5}\frac{w\cdot m\cdot l\cdot P_п\cdot d\cdot(z+s_k)}{5s_k^4+4A\cdot s_k^3+3B\cdot s_k^2+2C\cdot s_k+D}\cdot\frac{1}{s-s_k}. \tag{10}$$

Тогда имеем:

$$h(t) = L^{-1}\{h(s)\} = \sum_{k=1}^{5}\frac{w\cdot m\cdot l\cdot P_п\cdot d\cdot(z+s_k)}{5s_k^4+4A\cdot s_k^3+3B\cdot s_k^2+2C\cdot s_k+D}\cdot exp[s_k t]. \tag{11}$$

Полученное выражение является функцией плотности вероятностей. Поэтому искомая интегральная функция распределения вероятностей определяется следующим образом:

$$F(t) = \int_0^t h(t)\,dt = \sum_{k=1}^{5}\frac{w\cdot m\cdot l\cdot P_п\cdot d\cdot(z+s_k)}{5s_k^4+4A\cdot s_k^3+3B\cdot s_k^2+2C\cdot s_k+D}\cdot\frac{1-exp[s_k t]}{-s_k}. \tag{12}$$

Среднее время $\bar{t}_{КА}$, затрачиваемое на реализацию КА, определяется так:

$$\bar{t}_{КА} = \int_0^\infty t\cdot h(t)\,dt = \sum_{k=1}^{5}\frac{w\cdot m\cdot l\cdot P_п\cdot d\cdot(z+s_k)}{5s_k^4+4A\cdot s_k^3+3B\cdot s_k^2+2C\cdot s_k+D}\cdot\frac{1}{(-s_k)^2}. \tag{13}$$

Таким образом, определена интегральная функции распределения и среднее время $\bar{t}_{КА}$ реализации КА. Понятно, что указанные расчеты производятся для каждого этапа КА.

С целью получения нормированных значений для ВВХ кибератак разработана имитационная модель *SDN*.

Экспериментальные результаты

Описание имитационного стенда. С целью получения исходных данных по реализации КА была разработана комплексная компьютерная имитационная модель сети передачи данных с применением SDN (рис. 3).

Модель разработана в целях исследования устойчивости программно-конфигурируемой сети в условиях КА с учетом особенностей ее функционирования (информационного обмена, настройки сетевого оборудования, использования средств мониторинга и т.д.). Особенностью данной модели является ее комплексность. Во-первых, при построении SDN в виртуальной среде были использованы образы реальных сетевых устройств операторов связи, а также образы устройств Open Flow коммутаторов и контроллера Runos 2.

Во-вторых, для имитации нагрузки вместо программ-генераторов пакетов были использованы программы, предназначенные для информационного обмена на реальном оборудовании. В-третьих, использовалось несколько параллельно собирающих и анализирующих трафик программных продуктов, что позволяет оценить достоверность полученной сетевой статистики. Наконец, при планировании и проведении этапов КА учитывалось среднее время их выполнения, которое вычислялось при проведении заданного числа экспериментов.

Модель разработана на основе виртуальной сетевой лаборатории EVE-NG, в которую добавлены образы устройств реального оборудования, используемого в корпоративных сетях передачи данных: Dionis-NX, Juniper, Cisco, Huawei и др. Кроме того, два узла распределенной корпоративной сети было реализовано при помощи технологии SDN. Для моделирования SDN были использованы образы Linux (Ubuntu) с установленными программными маршрутизаторами OpenvSwitch, а в качестве образа контроллера использовалась модель контроллера Runos 2.0.

Модель сети передачи данных состоит из сети связи общего пользования, опорной сети передачи данных и оконечных узлов. Осуществляется подключение внешних устройств к виртуальной модели сети передачи данных. Внешними устройствами служат виртуальные или реальные устройства генерации трафика.

В модели используются рабочие места на операционных системах Linux Ubuntu, с помощью которых организуется видеоконференцсвязь, а также образ автоматической телефонной станции Протей-СП, с помощью которого имитируется обмен информацией между пользователями по средствам IP-телефонной связи и передачи сообщений с помощью почтового сервера.

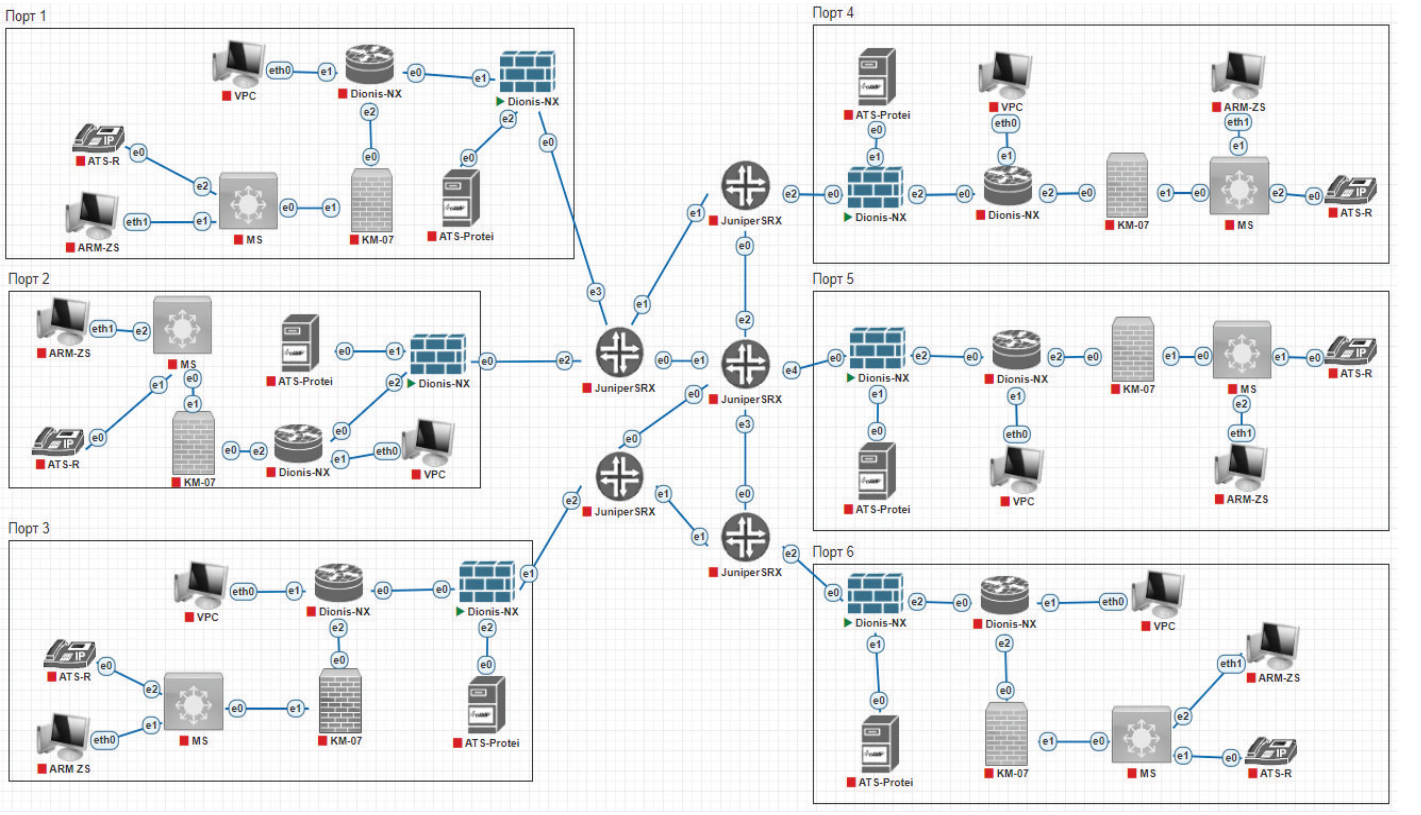


Рис. 3. Компьютерная имитационная модель сети передачи данных с применением SDN

В таблице 4 приведен перечень используемых программных продуктов, образов устройств и реального оборудования, используемого в имитационной компьютерной модели сети передачи данных с использованием технологии SDN.

Таблица 4

Перечень используемых программных продуктов

№ п/п	Название устройства (программного продукта)	Примечание
Коммутаторы, маршрутизаторы		
1	JuniperSRX-240 (QEMU)	Сетевое устройство, выполняющее роль граничного маршрутизатора
2	Dionis-NX (QEMU)	Сетевое устройство, выполняющее роль межсетевого экрана
3	Cisco3845 (QEMU)	Предназначено для имитации работы СС ОП
4	OpenvSwitch (Linux Ubuntu)	Программный маршрутизатор SDN
Средства имитации информационного обмена		
5	Linux Ubuntu	Операционная система
6	Lifesize	Программное средство для ВКС
7	Протей_СП (iso)	Телефонная станция IP телефонии
8	SIP-T22R	IP-телефон
9	Runos 2.0	Контроллер SDN
Среда моделирования, вспомогательные средства		
10	EVE-NG	Среда моделирования СПД
11	NFsen	Средство сбора информации с сетевых устройств об информационных потоках
12	Zabbix 3.4	Средство мониторинга
13	Wireshark	Средство перехвата трафика в СПД
14	VMware	Среда виртуализации для работы гостевых операционных систем

Исследованию устойчивости SDN к различным кибератакам посвящено достаточно много работ, однако в них для построения имитационной компьютерной модели используется программа моделирования виртуальной сетевой среды *Mininet*, которая не позволяет, в отличии от *EVE-NG*, создавать комплексную систему с внедрением реального оборудования и программ для имитации работы реальной SDN.

Виртуальная сетевая среда *EVE-NG* позволяет создавать различные топологии исследуемых сетей, проводить на них сценарии КА и параллельно собирать сетевую статистику при различных вариантах работы. В данную среду можно добавлять различные образы устройств, имеющих свою операционную систему, а также объединять между собой с помощью виртуальной сети придачи данных реальные устройства, применяемые для обеспечения SDN. При этом и виртуальные, и реальные устройства, работающие в виртуальной среде или подключенные к ней, полностью выполняют заложенный в них функционал, что позволяет осуществить их настройку согласно принятой методики.

Для сбора и обработки сетевой статистики на отдельном компьютере были установлены программные продукты *Nfsen2* и *Wireshark*, а также средство мониторинга сети *Zabbix* 3.4.

Создание имитационной модели информационного обмена при помощи программ, работающих на реальной СПД, и подключение средств мониторинга, которые позволяют отслеживать сгенерированный трафик, образует лабораторный стенд, который позволяет исследовать особенности изменения свойств СПД при различных условиях функционирования. Так как в данной работе рассматривается SDN, следовательно, в модели два взаимодействующих узла реализованы с помощью данной технологии.

Пример имитационной модели КА на SDN

После формирования имитационного стенда, описания логической структуры SDN и характерных для нее КА, необходимо перейти к формированию имитационной модели кибератак. Рассмотрим пример имитационной модели процесса КА на SDN. Для этого каждый этап процесса реализации атаки характеризуется среднем временем выполнения. Для проведения экспериментов и получения параметров среднего времени t для этапов КА предлагается воспользоваться средствами виртуального моделирования компьютерных сетей *EVE-NG*.

Для осуществления удаленной кибератаки злоумышленник может находится в любой точке, имеющей связь с ССОП. Этапы КА можно обобщенно представить в виде схемы, представленной на рисунке 4.

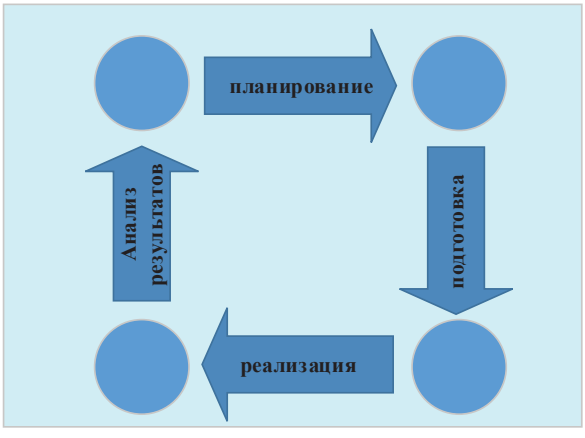


Рис. 4. Обобщенная схема этапов КА на SDN

На выбор типа КА влияют два основных параметра:

1) цели и задачи предстоящей КА;
2) данные ТКР о СПД, на которую предполагается реализация КА.

Таким образом первыми параметрами построения имитационной модели КА на SDN являются цели и задачи атаки. Допустим, целью КА является получение управления или частичного управления СПД для сбора и анализа передаваемого трафика. Одним из типов КА в SDN для достижения такой цели является атака типа «Подмена сетевой топологии».

Для осуществления данной КА ТКР должна обладать рядом данных о SDN-устройствах. Для SDN это могут быть стандартные IP/MAC – адреса сетевых устройств, и свойственные для данной технологии *UUID* – уникальные номера программного коммутатора/контроллера, SSL-ключ и т.д.

Получив необходимые данные, злоумышленник может приступать к проведению кибератаки, осуществив все необходимые настройки устройства/устройств, с которого предполагается осуществлять КА типа «Подмена сетевой топологии» (рис. 5). Нумерация этапов КА взята из таблицы 3.

Таким образом, для успешной реализации атаки злоумышленнику необходимо реализовать этапы, приведенные в таблице 3, и реализованные в имитационной компьютерной сети.

В процессе проведения экспериментов были получены данные сетевого трафика. Итогом кибератаки стала работа выданного за доверенное сетевое устройство злоумышленника, используя идентификационные данные, полученные ТКР.

Для получения адекватных данных в имитационной модели необходимо определить модельное время и количество экспериментов. При проведении экспериментов на имитационной модели модельным временем будет являться время проведения исследуемой КА на SDN. Логично, что с каждым *i*-ым экспериментом время его реализации будет меньше предыдущего, пока не достигнет значения минимального времени выполнения. Однако стоит учесть и неопределенность, которая влияет на дисперсию времени этапа. Количество экспериментов определим как необходимое для выполнения неравенства $T_{i-1} - T_i \leq 5$ сек., T_{i-1}-$T_i \geq 5$сект.е. время выполнения процесса КА следующего эксперимента не меньше, чем время выполнения предыдущего, на 5 сек.

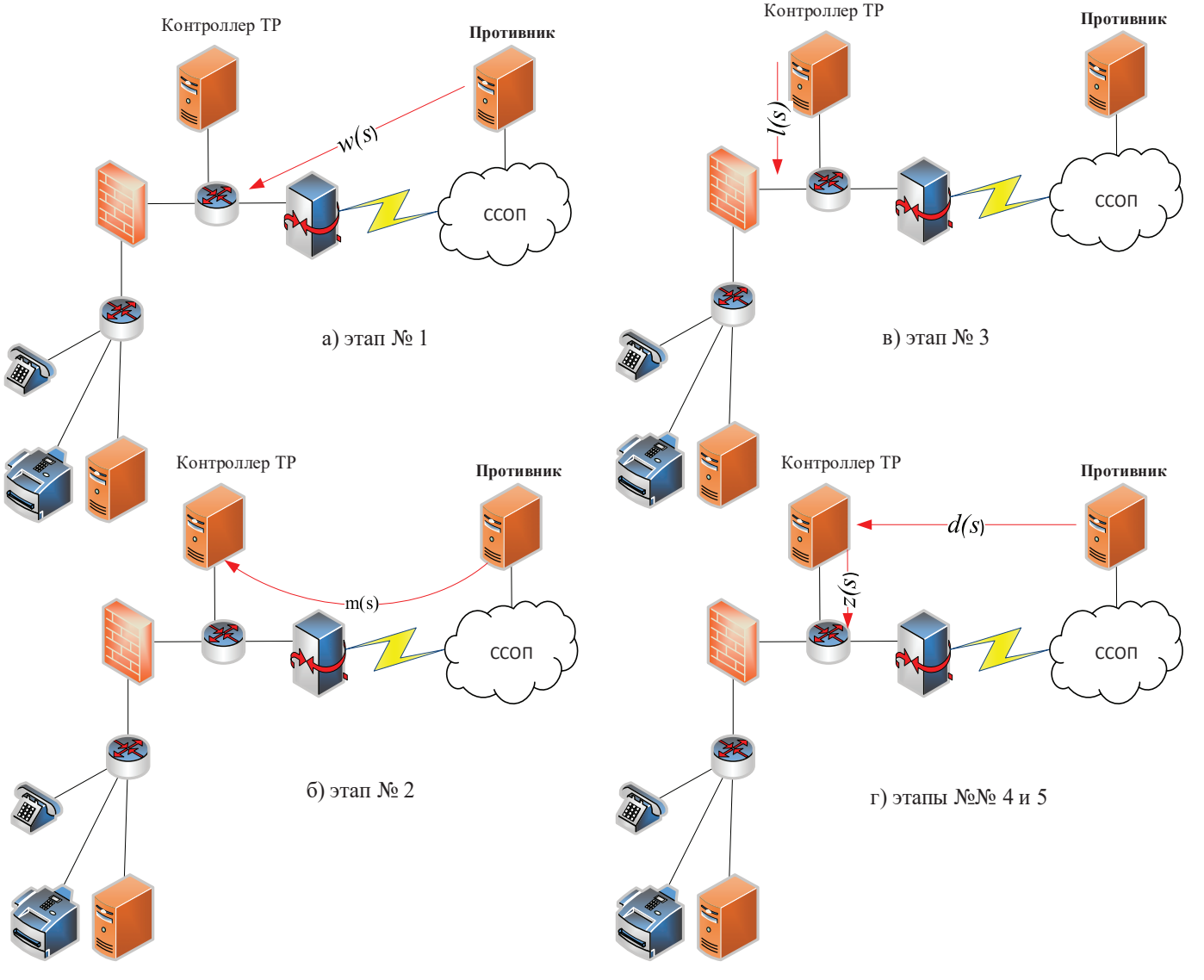


Рис. 5. Проведение атаки типа «Подмена сетевой топологии»

Также исключим из формулы определения среднего времени проведения КА время первого и последнего экспериментов, т.к. первый эксперимент является самым долгим, потому что проводится поэтапно с параллельным тестированием всех элементов модели, а последний практически равен предшествовавшему, и последнее время не более чем на 5 секунд отличается от предыдущего. Обозначим N – количество экспериментов. Тогда среднее время i-го этапа КА вычисляется по следующей формуле:

$$T_{cpi} = \frac{\sum_2^{i-1} T_i}{N} \mathrm{T_{cpi}} = \frac{\sum_2^{i-1} \mathrm{T_i}}{\mathrm{N}} \quad (17)$$

Аналогичным образом были реализованы все характерные для SDN атаки, приведенные в таблице 2. Результаты имитационного моделирования использовались в расчетах ВВХ с использованием ТПСС.

В таблице 5 приведены экспериментальные значения времени проведения этапов КА, характерных для технологии SDN.

Таблица 5

Экспериментальные значения времен проведения этапов КА

№ этапа КА	Время i-го эксперимента каждого эксперимента								T_{cpi}
	1	2	3	4	5	6	7	8	
1.1 Синхронная атака (Flooding attacks)									
1	4,34	4,22	4,10	3,59	3,45	3,36	3,27	3,10	3,57
2	2,55	2,44	2,47	2,35	2,17	1,58	1,56	1,37	2,25
3	5,02	4,56	4,44	4,30	4,27	3,67	3,58	3,53	4,32
4	5,58	5,44	5,45	5,24	5,24	4,56	4,45	4,38	5,25
5	3,45	3,30	3,20	3,18	3,14	2,43	2,44	2,25	3,00
1.2 «Атака человек по середине» (Man-in-the-middle)									
1	2,50	2,51	2,29	2,20	2,10	1,59	1,45	1,41	2,13
2	3,54	3,48	3,32	3,27	3,24	2,59	2,43	2,46	3,21
3	3,33	3,27	3,32	2,10	2,44	2,26	2,16	2,28	2,55
4	5,30	5,13	5,33	4,53	4,51	4,28	4,25	4,32	4,59
1.3 Взлом или сбой контроллера									
1	2,02	1,55	1,49	1,42	1,21	0,58	0,49	0,38	1,35
2	2,38	2,21	2,05	1,63	1,57	1,35	1,19	1,15	1,55
3	3,57	3,50	3,33	3,21	3,08	2,48	2,37	2,29	3,14
4	4,46	3,54	3,52	3,36	3,31	3,12	2,55	2,40	3,41

Используя полученные значения в качестве исходных данных, были получены зависимости $F(t)$ и $\bar{t}_{КА}$, представленные на рисунках 6-8. В качестве исходных данных использовались следующие значения времени и вероятности, соответствующие профильной модели КА: среднее время для каждого этапа КА – значения из таблицы 5; P_n = 0,2; 0,6; 0,8.

Представленные зависимости позволяют определить: вероятность $P_n(t \le T_з)$ реализации каждой из КА за время, не более заданного $T_з$; среднее время их реализации; время, соответствующее заданному уровню угрозы их реализации. Так, например, при вероятности нарушения работоспособности сети P_n= 0,8, определяющей ее доступность средствам КА злоумышленника, через 17 минут функционирования сеть будет не работоспособна с вероятностью F (t =17) = 0,8. При этом среднее время реализации КА составляет около 10 минут.

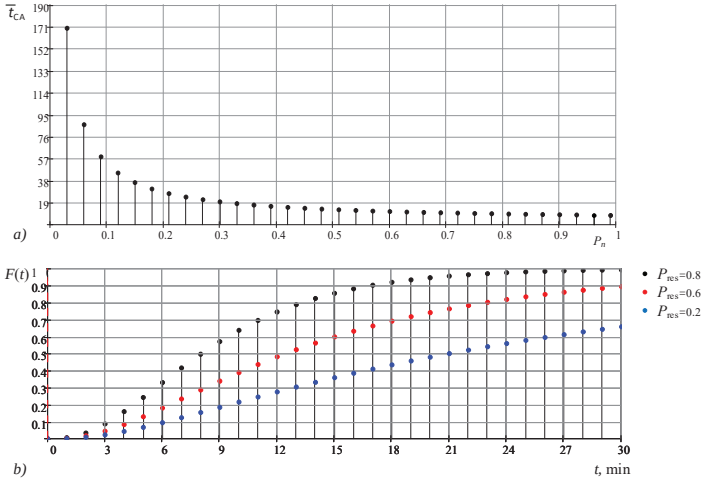


Рис. 6. Вероятностно-временные характеристики КА типа «Человек посредине» (*а* – зависимость среднего времени от вероятности реализации КА; *б* – зависимость интегральной функции распределения вероятности от времени реализации КА)

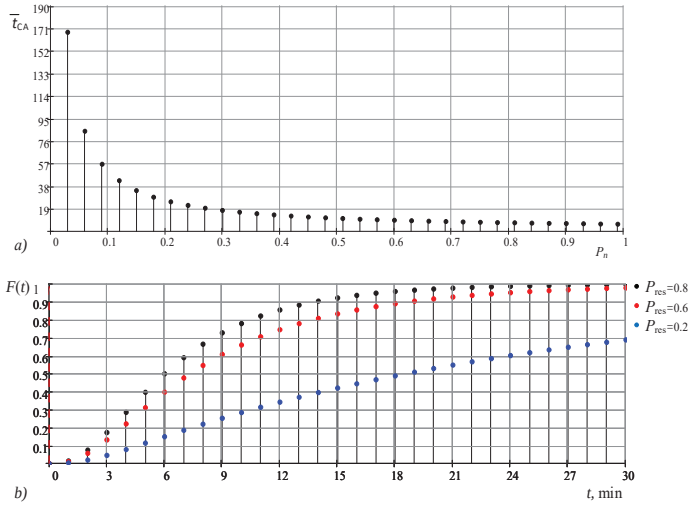


Рис. 7. Вероятностно-временные характеристики КА типа «Взлом или сбой» контроллера

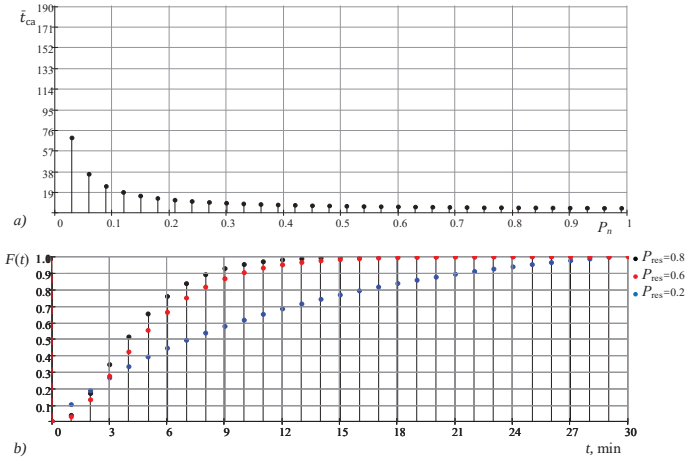


Рис. 8. Вероятностно-временные характеристики КА типа «DoS-атака»

Анализ полученных результатов показывает, что при одновременном воздействии нескольких способов реализации КА среднее время их реализации сокращается в разы, что подтверждает целесообразность их комплексирования.

В свою очередь становится актуальным развитие таких средств защиты плоскости передачи в программно-конфигурируемой сети, которые способны не только нейтрализовать попытки КА, но и предупреждать о совершении таких попыток. В этом случае возможно подготовиться к отражению новых атак другими способами, что увеличивает устойчивость сети.

Заключение

В статье предлагается новый подход к аналитическому моделированию КА, основанный на методе топологического преобразования стохастических сетей. Сущность данного метода заключается в замене множества элементарных ветвей стохастической сети одной эквивалентной ветвью с последующим определением эквивалентной функции сети, а также начальных моментов и функции распределения времени реализации КА.

Проверка предложенного подхода была проведена для моделирования КА типов «Подмена сетевой топологии» и «Взлом или сбой контроллера», которые являются одними из наиболее распространенных и опасных для SDN.

Определяя дальнейшие направления исследований, следует отметить, что в предложенном подходе было принято ограничение, согласно которому новая кибератака начинается через некоторое время после того, как была обнаружена предыдущая, и были устранены последствия ее реализации. Однако такое событие следует рассматривать как частный случай, при котором на компьютерную сеть воздействует только один злоумышленник. В реальности одновременно злоумышленников может быть достаточно много, и компьютерные атаки, активируемые ими, могут накладываться друг на друга. Поэтому сценарий проведения массированных кибератак следует считать одним из направлений дальнейших исследований.

Другое ограничение рассмотренного подхода связано с тем, что сценарии возможных атак заранее считаются известными, а сценарии реализации мер противодействия атак не рассматриваются. В то же время множество возможных сценариев противодействия кибератакам является конечным. По этой причине возможно построить аналитические модели для реализации контрмер и интегрировать их с аналитическими моделями кибератак.

В результате получится интегрированная аналитическая модель поведения компьютерной сети в условиях кибервоздействий, позволяющая оценивать и выбирать наиболее эффективные меры противодействия. Это направление также следует считать достаточно перспективным для дальнейших исследований.

Литература

1. *Ананченко И.В., Щербович-Вечер А.В.* Проектирование современных инфокоммуникационных сетей с использованием технологии программно-конфигурируемых сетей и виртуализации сетевых функций // Символ науки: международный научный журнал, 2015. № 12-1. С. 11-13.

2. *Калмыков Н.С., Докучаев В.А.* Применение концепции программно-конфигурируемых сетей для построения территориально-распределенных сетей // Телекоммуникации и информационные технологии, 2020. Т. 7. № 2. С. 51-56.

3. *Евглевская Н.В., Хмелляр Н.А., Шинкарев С.А.* Построение сети передачи данных как программно-конфигурируемой сети // Известия Тульского государственного университета. Технические науки, 2021. № 11. С. 272-278.

4. *Мурадова А.А.* Анализ работы основных элементов программно-конфигурируемых сетей, используемых на транспортном уровне инфокоммуникационных сетей // Перспективы развития информационных технологий, 2016. № 32. С. 136-144.

5. *Атея А.А., Мутханна А.С., Кучерявый А.Е.* Интеллектуальное ядро для сетей связи 5G и тактильного интернета на базе программно-конфигурируемых сетей // Электросвязь, 2019. № 3. С. 34-40.

6. *Локтионов О.В.* Применение программно-конфигурируемых сетей для построения сетей связи специального назначения // Вестник Санкт-Петербургского университета МВД России, 2017. № 4 (76). С. 127-130.

7. *Степанов М.Д., Павленко Е.Ю., Лаврова Д.С.* Обнаружение сетевых атак в программно-конфигурируемых сетях с использованием алгоритма изолирующего леса // Проблемы информационной безопасности. Компьютерные системы, 2021. № 1. С. 62-78.

8. *Пименова А.А., Никитин Д.Д., Никишин К.И.* Моделирование сценариев безопасности в программно-конфигурируемых сетях // Вестник Рязанского государственного радиотехнического университета, 2022. № 82. С. 60-72.

9. *Kotenko I., Doynikova E.* Security Assessment of Computer Networks based on Attack Graphs and Security Events // Lecture Notes in Computer Science, Vol.8407, 2014, pp.462-471.

10. *Чернов И.В., Орлов В.Г.* Особенности программно-конфигурируемых сетей // Телекоммуникации и информационные технологии, 2018. Т. 5. № 1. С. 21-25.

11. *Малафеев О.А., Зайцева И.В., Шлаев Д.В., Шматко С.Г., Брейдер Н.А.* Моделирование процесса взаимодействия в информационно-вычислительной сети как системе с марковскими процессами // Известия высших учебных заведений. Приборостроение, 2021. Т. 64. № 6. С. 444-451.

12. *Kotenko I., Saenko I., Lauta O.* Analytical modeling and assessment of cyber resilience on the base of stochastic networks conversion // Proceedings of 2018 10th International Workshop on Resilient Networks Design and Modeling, RNDM 2018. 10, 2018. С. 8489830.

13. *Kotenko I., Saenko I., Lauta O., Kocinyak M.* Assessment of computer network resilience under impact of cyber attacks on the basis of stochastic networks conversion // Communications in Computer and Information Science, 2018. Т. 797. pp. 107-117.

14. *Kotenko, I., Saenko, I., Lauta, O.*; Analytical Modeling and Assessment of Cyber Resilience on the base of Stochastic Networks Conversion. 2018 10th International Workshop on Resilient Networks Design and Modeling (RNDM 2018), Longyearbyen, 2018, Norway, pp. 1-8.

15. *Котенко В.И., Саенко И.Б., Коцыняк М.А., Лаута О.С.* Оценка киберустойчивости компьютерных сетей на основе моделирования кибератак методом преобразования стохастических сетей // Труды СПИИРАН, 2017. № 6(55). С.160-184.

COMPUTER ATTACK MODELS ON SOFTWARE-CONFIGURABLE NETWORKS

VASILIY A. GOLOVSKOY
St. Petersburg, Russia

VASILIY A. GOLOVSKOY
St. Petersburg, Russia

VASILIY A. GOLOVSKOY
St. Petersburg, Russia

VASILIY A. GOLOVSKOY
St. Petersburg, Russia

KEYWORDS: *computer attacks, method of topological transformation of stochastic networks, software-configurable networks, modeling.*

ABSTRACT

Introduction: SDN technology in the near future will allow to introduce aspects of the openness of the code of the network component of the cloud infrastructure, which is considered the most favorable basis for the development and implementa-tion of a wide range of applications. It is based on the implementation of network devices and their functions not in separate network equipment, but on any network host using the OpenvSwitch software switch. This approach allows you to use almost any computing device in the network as a switch/router. At the same time, the use of new technologies entails the emergence of new destabilizing factors on them, in which cyber attacks occupy a special place. Possible results of cyberattacks on SDN are blocking of the SDN controller, Open Flow control and monitoring channel, introduction of false information about the SDN network user receiving network services, violation of established regulations for collecting, processing and transmitting information to SDN, failures and failures in SDN operation, as well as compromise of transmitted or received information. **The purpose of the work** is to apply an integrated approach to the study of new approaches to building data transmission networks in the conditions of computer attacks, obtaining probabilistic and temporal characteristics of computer attacks characteristic of software-configurable networks, which in turn makes it possible to determine the most dangerous computer attacks, the elements subject to them, as well as to set re-quirements for a system that provides protection against the most likely impacts. **Methods used:** the method of topological transformation of stochastic networks used in modeling allows us to obtain parameters with a given completeness and reliability. The difference is no more 5% between the values obtained in the simulation and analytical models confirm their adequacy. The implementation of a large number of real devices in the simulation model, as well as the use of several means of collecting network statistics, allows achieving the required completeness of modeling. Significant detailing of the stages of computer attacks in the method of topological transformation of stochastic networks contributes to obtaining computational expressions that accurately describe the probabilistic-temporal characteristics of the system. **The scientific novelty** of the results obtained is determined by the use of the method of topological transformation of stochastic networks (TPSS) for analytical modeling of cyberattacks on SDN. **Result:** the presented model allows us to obtain probabilistic and temporal characteristics of computer attacks characteristic of a software-configurable network. The basis of the proposed model is a simulation model, the main purpose of which is the most accurate reproduction of the processes of a real system. The obtained values of the probabilistic-temporal characteristics of computer attacks allow us to accurately determine the most dangerous types of them, as well as the most likely places of manifestation. **Practical significance:** the presented method is universal and can be applied in the state system for detecting, preventing and eliminating the consequences of computer attacks when solving tasks related to ensuring public safety.

REFERENCES

1. Ananchenko I.V., Shcherbovich-Vecher A.V. Designing modern infocommunication networks using the technology of software-defined networks and virtualization of network functions. *Symbol of science: international scientific journal*, 2015. No. 12-1, pp. 11-13.
2. Kalmykov N.S., Dokuchaev V.A. Application of the concept of software-defined networks for building geographically distributed networks. *Telecommunications and information technologies*, 2020. Vol. 7. No. 2, pp. 51-56.
3. Evglevskaya N.V., Khmellyar N.A., Shinkarev S.A. Building a data transmission network as a software-defined network. *Izvestia of the Tula State University. Technical sciences*, 2021. No. 11, pp. 272-278.
4. Muradova A.A. Analysis of the operation of the main elements of software-defined networks used at the transport level of infocommunication networks. *Prospects for the development of information technologies*, 2016. No. 32, pp. 136-144.
5. Ateya A.A., Muthanna A.S., Kucheryavyi A.E. Intelligent core for 5G communication networks and tactile Internet based on software-defined networks. *Elektrosvyaz*, 2019. No. 3, pp. 34-40.
6. Loktionov O.V. The use of software-defined networks for building special-purpose communication networks. *Bulletin of the St. Petersburg University of the Ministry of Internal Affairs of Russia*, 2017. No. 4 (76), pp. 127-130.
7. Stepanov M.D., Pavlenko E.Yu., Lavrova D.S. Detection of network attacks in software-defined networks using the isolation forest algorithm. *Problems of information security. Computer Systems*, 2021. No. 1, pp. 62-78.
8. Pimenova A.A., Nikitin D.D., Nikishin K.I. Modeling security scenarios in software-defined networks. Bulletin of the Ryazan State Radio Engineering University, 2022. No. 82, pp. 60-72.
9. Kotenko I., Doynikova E. Security Assessment of Computer Networks based on Attack Graphs and Security Events. *Lecture Notes in Computer Science*, Vol.8407, 2014, pp. 462-471.
10. Chernov I.V., Orlov V.G. Features of software-defined networks. *Telecommunications and information technologies*, 2018. Vol. 5. No. 1, pp. 21-25.
11. Malafeev O.A., Zaitseva I.V., Shlaev D.V., Shmatko S.G., Breider N.A. Modeling the process of interaction in the information-computing network as a system with Markov processes. *Izvestia of higher educational institutions*. Instrumentation, 2021. Vol. 64. No. 6, pp. 444-451.
12. Kotenko I., Saenko I., Lauta O. Analytical modeling and assessment of cyber resilience on the base of stochastic networks conversion. *Proceedings of 2018 10th International Workshop on Resilient Networks Design and Modeling*, RNDM 2018. 10, 2018. P. 8489830.
13. Kotenko I., Saenko I., Lauta O., Kocinyak M. Assessment of computer network resilience under impact of cyber attacks on the basis of stochastic networks conversion. *Communications in Computer and Information Science*, 2018. Vol. 797, pp. 107-117.
14. Kotenko, I., Saenko, I., Lauta, O.; Analytical Modeling and Assessment of Cyber Resilience on the base of Stochastic Networks Conversion. *2018 10th International Workshop on Resilient Networks Design and Modeling (RNDM 2018)*, Longyearbyen, Norway, 2018, pp. 1-8.
15. Kotenko V.I., Saenko I.B., Kotsynyak M.A., Lauta O.S. Estimation of cyber stability of computer networks based on the simulation of cyber attacks by the method of transformation of stochastic networks. *Proceedings of SPIIRAS*, 2017. No. 6(55), pp.160-184.

INFORMATION ABOUT AUTHORS:

Igor B. Saenko, leading researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, doctor of technical sciences, professor, St. Petersburg, Russia
Igor V. Kotenko, chief researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, doctor of technical sciences, professor, St. Petersburg, Russia
Oleg S. Lauta, professor, Admiral Makarov State University of Maritime and inland shipping, doctor of technical sciences, St. Petersburg, Russia
Sergey Yu. Skorobogatov, graduate student, Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, St. Petersburg, Russia

For citation: Saenko I.B., Kotenko I.V., Lauta O.S., Skorobogatov S.Yu. Computer attack models on software-configurable networks. H&ES Reserch. 2023. Vol. 15. No 1. P. 37-47. doi: 10.36724/2409-5419-2023-15-1-37-47 (In Rus)

doi: 10.36724/2409-5419-2023-15-1-48-56

ВЛИЯНИЕ ПРОБЛЕМЫ МНОГОЗНАЧНОСТИ МЕТОК КЛАССОВ СИСТЕМНЫХ ЖУРНАЛОВ НА ЗАЩИЩЕННОСТЬ КОМПЬЮТЕРНЫХ СЕТЕЙ

РАКОВСКИЙ
Дмитрий Игоревич[1]

АННОТАЦИЯ

Введение: Защищенность информации, циркулирующей в компьютерной сети, связана с защищенностью поддерживающей инфраструктуры. Важной проблемой интеллектуальной обработки данных системных журналов является существование наборов данных, содержащих записи с несколькими ассоциациями меток классов. Работы, так или иначе исследующие проблемы многозначности, объединены термином: многозначное обучение, Multi-Label Learning. Отечественных работ, посвященных анализу наборов данных, порожденных компьютерными сетями, с многозначными метками классов, в настоящий момент не представлено, что актуализирует исследования в указанной области. **Цель исследования:** повысить защищенность компьютерных сетей за счет использования методов многозначного обучения при решении задачи классификации меток классов системных журналов. **Результаты:** Проведен сравнительный анализ однозначных и многозначных классификаторов в вычислительном эксперименте по метрике Mean accuracy. Обнаружена нелинейная зависимость между долей участков экспериментальных данных, содержащих многозначные метки классов, и точностью классификации данных. Несмотря на то, что многозначных участков в исследуемых экспериментальных данных всего 3%, выигрыш в точности достигает 23% по указанной метрике. По результатам проведенного анализа 80% однозначных классификаторов уступили в точности классификации по метрике Mean accuracy многозначным аналогам, что может сигнализировать о сильном влиянии многозначности меток классов на рассматриваемые модели. Показано, что рассматриваемая структура экспериментальных данных табличного вида подвержена влиянию проблемы многозначности гораздо сильнее, чем это может быть оценено стандартной частотной проверкой, что актуализирует дальнейшие исследования в данном направлении. **Практическая значимость:** Практическая значимость полученных результатов заключается в повышении защищенности компьютерных сетей за счет использования многозначного подхода в задаче классификации. Задачи информационной безопасности, решаемые многозначной классификацией, могут включать в себя: область мониторинга, обнаружения или предупреждения нарушениям и компьютерным атакам в компьютерных сетях. **Обсуждение:** Поскольку предсказательная способность частотной проверки влияния результатов многозначности меток классов на результаты классификации однозначных классификаторов невелика, планируются дальнейшие исследования на эту тему. Планируется расширение перечня метрик оценки качества классификации в дальнейших экспериментах.

Сведения об авторе:

[1] аспирант, ассистент кафедры информационная безопасность. Московский технический университет связи и информатики, Москва, Россия, Prophet_alpha@mail.ru

КЛЮЧЕВЫЕ СЛОВА: *обучение с учителем, multi-label classification, многозначная классификация, многоклассовая классификация, multiclass classification, информационная безопасность, multi-label learning.*

Для цитирования: Раковский Д.И. Влияние проблемы многозначности меток классов системных журналов на защищенность компьютерных сетей // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 48-56. doi: 10.36724/2409-5419-2023-15-1-48-56

Введение

Современные компьютерные сети (КС) обладают сложной инфраструктурой, требующей постоянного мониторинга для выявления аномальных состояний, которые могут вызвать сбои в работе, что недопустимо для масштабных распределённых КС [1]. Устойчивость КС к последствиям реализаций достигается за счет модернизации поддерживающей инфраструктуры КС; в том числе за счет повышения защищенности.

Защищенность КС может достигаться за счет применения классических мер для предотвращения перехвата траффика – установки программно-аппаратных средств защиты информации [2]; систем обнаружения и предотвращения вторжений [3, 4]; антивирусного программного обеспечения [5] и прочих решений [6].

Актуальны работы, посвященные разработке программных решений для детектирования, обнаружения и нивелирование киберугроз в КС [7]. Известны работы по оценивание и прогнозирование состояния сложных объектов: применение для информационной безопасности [8, 9].

Важной проблемой интеллектуальной обработки данных системных журналов является существование наборов данных, содержащих записи с несколькими ассоциациями меток классов. То есть *класс*, ассоциированный с объектом, характеризуется множеством меток.

Набор данных, пригодный к классификации, как правило содержит множество признаков и ассоциированное с ним множество меток класса. Целью классификации является обученная модель, способная присвоить соответствующий класс неизвестному объекту (записи в «исторических данных»).

В зарубежных публикациях данная проблема также известна как «*multi-label…*», где вместо троеточия может располагаться уточняющее слово или словосочетание. Работы, так или иначе исследующие проблемы многозначности, объединены термином: *многозначное обучение*, *Multi-Label Learning, MLL* [10].

Многозначное обучение обобщает понятие анализа данных на область задач, в которых каждому объекту может быть сопоставлено несколько меток. Среди данных статей выделяется кластер работ по анализу текстовых корпусов [11] и тональности сообщений в социальных сетях [12].

Отечественных работ, посвященных анализу наборов данных, порожденных КС, с многозначными метками классов в настоящий момент не представлено. Существующие работы, например, [13, 14], посвящены аспектам нечеткой классификации. Нечеткая классификация относится к области нечеткой логики (*Fuzzy logic*), являющейся частью методов многоклассового обучения.

Многозначное обучение косвенно связано с понятием «смешанное обучение» (смешанная классификация - *Misclassification*). Термин в настоящее время используется для маркировки работ, посвященных решению проблем неправильной разметки данных [15] и повышению точности классификации [16].

Информационная безопасность характеризует сохранение свойств конфиденциальности, целостности и доступности информации [17]. Анализ влияния многозначности меток классов на защищенность КС необходимо проводить в определенном терминологическом контексте. В качестве такого контекста избран ГОСТ Р ИСО/МЭК 27000, из которого проистекает вышеизложенное определение информационной безопасности, а также ГОСТ Р ИСО/МЭК12207[1]. Согласно упомянутому документу, п. 3.25, «Защищенность (*security*): Способность компьютерной системы защитить информацию и данные так, чтобы не допустить их несанкционированного прочтения или изменения другими системами и отдельными лицами, и для того, чтобы допущенные к ним системы и лица не получали отказов».

Конкретизируем защищенность информации, циркулирующей в КС: «Защищенность информации – поддержание на заданном уровне тех параметров находящейся в автоматизированной системе информации, которые характеризуют установленный статус ее хранения, обработки и использования» [18].

Из двух определений следует, что защищенность информации, циркулирующей в КС, связана с защищенностью поддерживающей инфраструктуры [19]. В рамках данной работы будет проведен анализ влияния многозначности на точность классификации состояний КС, непосредственно связанных с профилем нормального функционирования КС [20].

Целью работы является повышение защищенности компьютерных сетей за счет использования методов многозначного обучения при решении задачи классификации меток классов системных журналов.

Формализация задачи

КС можно представить в виде множества из M наборов значений дискретно изменяющихся атрибутов «исторических данных» КС:

$$\boldsymbol{A} \subseteq A_{first} \cup A_{second} = \\ = \{A_{first\,1} \times A_{first\,2} \times ... \times A_{first\,len_1}\} \cup \{A_{second\,1} \times A_{second\,2} \times ... \times A_{second\,len_2}\}; \quad (1)$$

где $A_m = \{a_{mn}; m = \overline{1, M}, n = \overline{1, N}\}$, $A_m \subset \boldsymbol{A}$, $M = len_1 + len_2$.

Атрибуты в записи (1), могут подразделяться на два типа: первичные $\{A_{first\,k_1}; k_1 = \overline{1, len_1}\}$ и вторичные $\{A_{second\,k_2}; k_2 = \overline{1, len_2}\}$.

Первичные атрибуты получают непосредственно с системных датчиков, установленных внутри КС. Вторичные атрибуты получают в результате обработки первичных атрибутов. Примерами вторичных атрибутов могут быть, например, среднее время задержки сигнала в КС, количество потерянных пакетов в КС для конкретного хоста и прочее.

[1] ГОСТ Р ИСО/МЭК 12207-2010. Национальный стандарт российской федерации. Информационная технология. Системная и программная инженерия. Процессы жизненного цикла программных средств. Information technology. System and software engineering. Software life cycle processes. Дата введения – 2012.03.01.

Введем для описания КС множество меток классов категориального типа – $\boldsymbol{S}$, которые будем называть «состояниями КС». Состояния КС могут вводиться также в виде множества:

$$\boldsymbol{S} = \{S_1, S_2, ..., S_M\} \cup \{s_{normal}\};\ S_m = \{s_i; i = \overline{1, I}\} \quad (2)$$

где S_m – m-тое подмножество состояний КС, ассоциированное с соответствующим A_m атрибутом КС. Мощность подмножества S_m имеет верхнюю границу, равную I. На практике подмножества, входящие в $\boldsymbol{S}$, могут иметь разную мощность.

Примером элементов с разной мощностью является неравенство $|S_1| \neq |S_2|$. В случае, если $\forall S_m = \varnothing$ вводится состояние s_{normal}, характеризующее нормальное функционирование КС.

Для автоматизации процесса определения состояний КС введем множество решающих правил

$$\begin{aligned} &\boldsymbol{METARULES} = \{RULE_1, RULE_2, ..., RULE_M\}, \\ &RULE_m = \{r_{mj}; j = \overline{1, |S_m|}\} \end{aligned} \quad (3)$$

Каждое подмножество – $RULE_m$ – ассоциировано с соответствующим подмножеством состояний КС по m-тому атрибуту – S_m. Мощность подмножества $RULE_m$ зависит от мощности соответствующего подмножества S_m. Итерационная переменная j введена для учета различия мощностей различных подмножеств S_m. В случае идентичности всех подмножеств S_m, $j = \overline{1, |S_m|} \equiv i = \overline{1, I}$, верхняя граница будет тождественна I.

Решающие правила предлагается выбирать на основании вводимой индивидуально характеристики уровня обслуживания *Service Level Objectives*, *SLO* исходя из технических и эксплуатационных характеристик КС.

Решающие правила предлагается выбирать на основании характеристики уровня обслуживания *SLO*, вводимой индивидуально исходя из технических и эксплуатационных характеристик КС.

Рассмотрим процесс маркировки множества атрибутов, соответствующего n-ному наблюдению исторических данных (n-ной строки в таблице исторических данных) - $\{a_{1n}, a_{2n}, ..., a_{Mn}\}$. Указанная строка является аргументом функции маркировки $mark(\{a_{1n}, a_{2n}, ..., a_{Mn}\})$, и формирует множество меток set_n, соответствующих n-ной строке.

Множество set_n формируется в ходе проверки каждого атрибута n-ной строки $\{a_{1n}, a_{2n}, ..., a_{Mn}\}$ – на соответствие правилам из соответствующего множества $RULE_m$ (3) – r_{mj}. Если правило r_{mj} выполняется, то в множество меток set_n добавляется элемент s_{mj}, где $j = \overline{1, |S_m|}$.

Процесс маркировки можно формализовать в виде:

$$\begin{aligned} &mark : \{a_{1n}, a_{2n}, ..., a_{Mn}\} \to set_n;\ set_n \subseteq \boldsymbol{S},\ где \\ &mark(\{a_{1n}, a_{2n}, ..., a_{Mn}\}) = \begin{cases} set_n,\ если\ set_n \neq \varnothing \\ s_{normal},\ иначе \end{cases}, \\ &где\ set_n = \left\{ \begin{matrix} s_{mj} \in S_m \mid r(a_{mn}, j) = 1, \\ j = \overline{1, |S_m|},\ m = \overline{1, M} \end{matrix} \right\}, \\ &где\ r(a_{mn}, j) = \begin{cases} 1,\ если\ выполняется \\ правило\ r_{mj} \in RULE_m \\ 0,\ иначе \end{cases} \end{aligned} \quad (4)$$

Если ни одно из правил не выполняется, то $set_n = \left\{ s_{mj} \in S_m \mid r(a_{mn}, j) = 1,\ j = \overline{1, |S_m|},\ m = \overline{1, M} \right\} = \varnothing$.

Это означает, что результатом маркировки будет являться заранее определенное состояние КС – s_{normal}.

Каждый элемент r_{mj}, является свободно задаваемым *вербально-логическим* правилом, вводимым для конкретной КС. Правила могут быть сопряжены с политикой безопасности, актуальной для КС: с моделью угроз; с показателями уровня обслуживания *SLO*; иными методиками оценки защищенности и качества предоставляемых услуг. При воздействии маркирующих правил на данные КС, каждой записи (строке – $\{a_{1n}, a_{2n}, ..., a_{Mn}\}$) присваивается либо множество состояний set_n, в соответствии с соотношением (4), либо состояние s_{normal}.

Маркировка «исторических данных» о поведении КС может быть представлена в виде таблицы размером M столбцов на N строк:

$$D_N = \{(\{a_{1n}, a_{2n}, ..., a_{Mn}\}, set_n); m = \overline{1, M}, n = \overline{1, N}\},$$

где n-ной строке значений атрибутов записи $\{a_{1n}, a_{2n}, ..., a_{Mn}\}$ ставится в соответствие состояние КС и множество меток set_n.

Хотя это не единственный способ разметки экспериментальных данных, однако маркировка является наиболее удобной с точки зрения организации обработки и анализа данных специализированными программными средствами.

Структура и описание исследуемой сетевой инфраструктуры

Исследования для оценки сетевых характеристик проводилось на КС, состоящей из 6 хостов, образующих кластер под управлением *Rancher* (рис. 1) [21]. Архитектура взаимодействия хостов исследуемой КС построена на принципе виртуализации и взаимодействия *Docker*-контейнеров; служб под управлением кластера *Apache Spark*; базы данных (*PostgreSQL*; *Apache Ignite; Apache Cassandra; Redis*); кластера *Apache Ignite*; программного обеспечения на основе микросервисной архитектуры и других вспомогательных модулей.

Технические характеристики хостовых машин распределенной КС приведены в таблице 1. Машины №1 – №3 формируют физическую топологию распределенной КС; машины №4 – 6 функционируют посредством виртуализации операционной системой *VMware ESXI* на базе машин №1 – 3. Для сбора данных на шесть машин КС использовались специальное программное обеспечения получения информации системных датчиков: *packetbeat* (агрегирует трафик протоколов *HTTP* и *DNS* запросов); *metricbeat* (агрегирует данные по использованию центрального процессора; диска; использованию памяти; сети; по процессам системы); *filebeat* (агрегирует данные журналов сообщений); *execbeat* (агрегирует выполнение специализированных скриптов и отправка результата их выполнения).

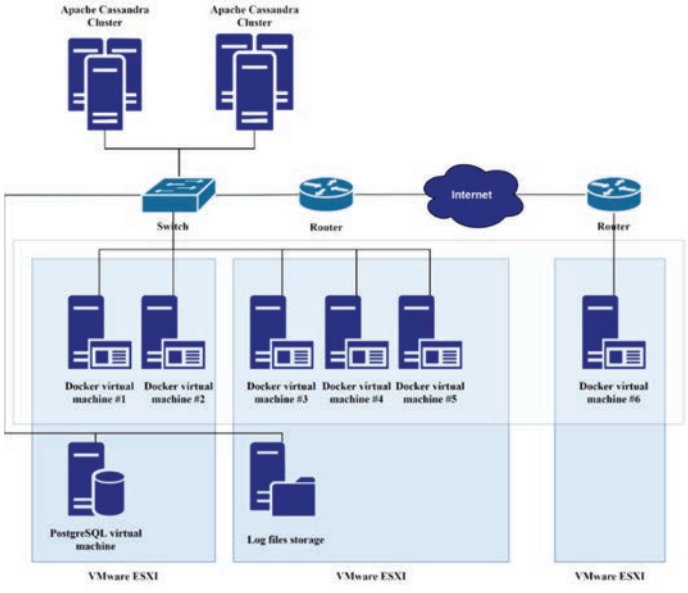


Рис. 1. Схема исследуемой сетевой инфраструктуры

Для сбора показателей, связанных с *SLO*, в рассматриваемой КС реализована система по синхронному мониторингу всех хостов. Схема сбора показателей приведена на рисунке 2. Полученные данные агрегируют в централизованном хранилище под управлением *Apache Cassandra*.

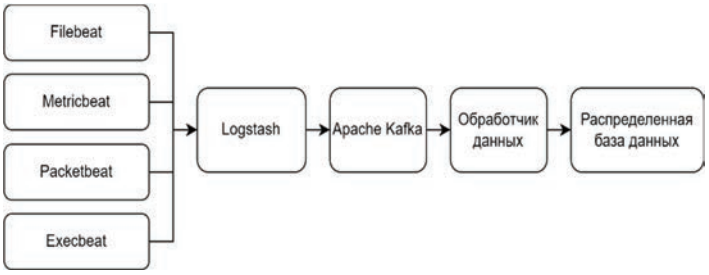


Рис. 2. Схема использования системы агрегации показателей, связанных с *SLO*

На хосты, приведенные в таблице 2, в соответствии с схемой на рисунке 2, установлены агрегаторы программно-аппаратных метрик: *Packetbeat* (агрегатор сетевой активности хоста - мониторинг трафика, протокола *HTTP* и *DNS* запросов)[2] [22]; *Metricbeat* – агрегатор показателей, ассоциированных с операционной системой и аппаратными устройствами хоста - использование *CPU*, памяти, дисков, запущенных процессов[3]; *Filebeat* – агрегатор системных журналов[4]; *Execbeat* – программное обеспечение для тестов КС методом формирования и отправки на исполнение произвольных скриптов[5]. *Execbeat* использовался для отправки *ICMP*-запросов (*ping*-запросов) с целью определения задержки в сети и отправки *GET*-запросов с использованием для определения времени реакции сервера на посланный запрос.

Таблица 1

Конфигурация КС

№	Соответствие хоста виртуальной машине	Операционная система (ОС)	Количество ядер	Оперативная память, Гб	Емкость жесткого диска (суммарная), Гб	Модель процессора
1	server3-20 (физическая машина из кластера Apache Cassandra)	CentOS Linux 7	4	64	1524	Intel(R) Xeon(R) CPU E3-1220 v6 @ 3.00GHz
2	server3-21 (физическая машина из кластера Apache Cassandra)	CentOS Linux 7	4	64	1524	Intel(R) Xeon(R) CPU E3-1220 v6 @ 3.00GHz
3	server3-22 (физическая машина из кластера Apache Cassandra)	CentOS Linux 7	4	64	1524	Intel(R) Xeon(R) CPU E3-1220 v6 @ 3.00GHz
4	server24- 384-1 (Docker virtual machine №1; Docker virtual machine №2)	Ubuntu 18.04.1 LTS	5	50,05	68	Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60GHz
5	server24- 384-2 (Docker virtual machine №3; Docker virtual machine №4; Docker virtual machine №5)	Ubuntu 18.04.1 LTS	6	48,61	265	Intel(R) Xeon(R) CPU E5-2420 v2 @ 2.20GHz
6	server24- 384-3 (Docker virtual machine №6)	Ubuntu 18.04.1 LTS	8	60	285	Intel(R) Xeon(R) CPU E5-2420 v2 @ 2.20GHz

[2] PACKETBEAT. Lightweight shipper for network data // Elastic [Электронный ресурс] URL: https://www.elastic.co/beats/packetbeat (дата обращения: 22.10.2022).

[3] METRICBEAT. Lightweight shipper for metrics. // Elastic [Электронный ресурс] URL: https://www.elastic.co/beats/metricbeat (дата обращения: 22.10.2022).

[4] FILEBEAT. Lightweight shipper for logs. // Elastic [Электронный ресурс] URL: https://www.elastic.co/beats/filebeat (дата обращения: 22.10.2022).

[5] Elastic beat to call commands in a regular interval and send the result to Logstash // Elasticsearch [Электронный ресурс] URL: https://github.com/christiangalsterer/execbeat (дата обращения: 22.10.2022).

Каждый из четырех типов агрегаторов отправляет данные в центральную точку КС – агрегатор логов *Logstash*, преобразующий всю поступающую информацию в файлы формата JSON. Выбор формата обусловлен общепринятой нотацией строения структуры JSON файлов. Описанный стек аггрегаторов широко используется при построении систем обработки информации в области информационной безопасности [23-26].

После преобразования, JSON файл отправляется в обработчик сообщений Apache Kafka[6], выполняющий буферизирующую функцию между большим потоком входных данных и распределённой базой данных.

Проблема первичных и вторичных атрибутов

Актуальной прикладной задачей является определение состояний КС без знания вторичных атрибутов. В этом случае метки классов *SLO* определяются только на основании первичных данных системных датчиков в условиях частичной неопределенности остальных параметров.

Рассмотрим два случая:

1. Полная априорная определенность как первичных, так и вторичных атрибутов КС в каждый момент времени;

2. Частичная неопределенность вторичных атрибутов КС, которые либо неизвестны, либо вычисляются с большой задержкой.

При наличии полной информации об атрибутах (A_{first} и A_{second},), в силу полной зависимости set_n от A_{second}, задача классификации состояния КС выполняется многозначным классификатором с точностью, близкой к идеальной, т.е. без ошибок. Препятствием к такой идеальной классификации является выявление непосредственных правил преобразования A_{second} в set_n ($A_{second} \rightarrow set_n$).

Если правила представлены тривиальными логическими условиями «если … то …», то точность классификации многими классификаторами, основанными на правилах (например, деревья решений или нейронные сети), будет близка к идеальной. Если вторичные атрибуты неизвестны, но известны первичные атрибуты и соответствующие состояния – вторичные атрибуты будут являться скрытой переменной. В случае отсутствия информации о вторичных атрибутах, однозначность отображения первичных атрибутов в состояния КС не гарантируется, поскольку вторичные атрибуты становятся скрытыми переменными. Однако принципиальная возможность отображения первичных атрибутов в состояния КС, все же, возможна.

Вычислительный эксперимент

Для сравнения двух способов классификации – «классического» однозначного – и многозначного – проведем вычислительный эксперимент на Python со следующими входными данными.

Однозначный подход к классификации рассмотрим на примере многоклассовых алгоритмов, отобранных по двум критериям:

– открытость исходного кода (библиотека, реализующая данный алгоритм, находится в открытом доступе);

– наличие многозначной реализации данного алгоритма.

По установленным критериям из открытой библиотеки scikit-learn языка программирования *Python*[7] [27] отобраны следующие алгоритмы:

– *Tree.DecisionTreeClassifier* – Классификатор, сформированный на основе алгоритма «Decision Tree» (непараметрический контролируемый метод обучения);

– *Tree.ExtraTreeClassifier* – Классификатор, сформированный на основе алгоритма «Extra Decision Tree» (непараметрический контролируемый метод обучения). При поиске наилучшего разделения для разделения выборок узла на две группы для каждой из случайно выбранных атрибутов выбирается наилучшее разделение по задаваемому критерию;

– *Ensemble.ExtraTreesClassifier* – Классификатор, сформированный на основе алгоритма «Extra Decision Tree» (ансамблевая реализация);

– *Neighbors.KNeighborsClassifier* – Классификатор, сформированный на основе алгоритма голосования «K-Neighbors»;

– *Ensemble.RandomForestClassifier* – Классификатор, сформированный на основе алгоритма «Random Forest» (ансамблевая реализация).

Получены показатели уровня обслуживания *SLO* (решающие правила) и соответствующие им состояния КС, ассоциированные со вторичными атрибутами сформированные в виде порогов, определяющих категориальные маркеры состояния КС:

– Если ни одна из целей уровня обслуживания не была нарушена, то состояние КС равно маркеру *normal*.

– Если время задержки сигнала к тестовому серверу (*ping_avg*) > 5 мс., то состояние КС равно маркеру *signal_delay*.

– Если время ответа тестового сервера (*server_response_timetotal*) > 1.5 с., то состояние КС равно маркеру *server_response_delay*.

– Если количество пакетов, потерянных при передаче к тестовому серверу (*network_outdropped*) > 0 шт., то состояние КС равно маркеру *packets_dropped*.

– Если время обработки запроса диском хостовой машины (*disk_ioreadmergespersec*) > 2 с., то состояние КС равно маркеру *disk_iowriteawait*.

При желании, количество решающих правил и затрагиваемых ими атрибутов КС может быть увеличено, однако для иллюстрации достаточно и 5 меток классов. Рассмотрим распределение экспериментальных данных по числу одновременного нарушаемых показателей уровня обслуживания. Изначальное распределение приведено в таблице 2.

[6] Apache Kafka. A distributed streaming platform. // Apache Kafka [Электронный ресурс] URL: https://kafka.apache.org/ (дата обращения: 22.10.2022).

[7] Multiclass and multioutput algorithms // scikit-learn URL: https://scikit-learn.org/stable/modules/multiclass.html (дата обращения: 26.12.2022).

Таблица 2

Распределение экспериментальных данных по числу одновременного нарушаемых показателей уровня обслуживания

Число одновременно нарушаемых показателей уровня обслуживания, *Anomaly*	Количество записей в экспериментальных данных, ед.	Количество записей в экспериментальных данных, %	
0	170931	71,870	
1	60447	25,416	0.28057
2	6282	2,641	
3	175	0,074	
4	0	0	

Как видно из представленной таблицы, более 71% экспериментальных данных занимает состояние нормального функционирования КС, что порождает проблему классового дисбаланса.

Для вычислительного эксперимента взяты первые 200 тысяч записей исходных экспериментальных данных [21]. Объем экспериментальных данных выбирался исходя из имеющихся вычислительных ресурсов.

Указанные атрибуты – *ping_Avg, network_outdropped, disk_ioreadmergespersec, server_response_timetotal* – преобразованы в соответствующие состояния КС и исключены из последующего анализа. Таким образом указанные вторичные атрибуты КС становятся скрытыми переменными.

В качестве первичных атрибутов, в иллюстративных целях, выбраны следующие атрибуты: *disk_await, disk_writebytes, network_outbytes, network_inbytes, ping_max*.

Поскольку одной записи может соотноситься несколько состояний КС одновременно, был выбран метод сведения многозначных меток классов к однозначному виду – *Label Powerset* (*LP,* [28]), порождает новый класс для каждой возможной комбинации меток посредством унитарного кодирования алфавита всевозможных комбинаций состояний КС, а затем решает задачу многозначного анализа как задачу однозначной многоклассового анализа.

Для повышения объективности классификации, точность оценивалась кросс-валидацией: выборка делилась на 10 равных частей; поочередно одна из частей становилась тестовой. Метрика оценки эффективности классификации – *Mean accuracy* (является стандартной метрикой для всех алгоритмов, представленных библиотекой scikit-learn.org).

Эксперимент проводился при стандартных гиперпараметрах, устанавливаемых для алгоритмов по умолчанию. Оптимизации гиперпараметров не проводилось. Для пар «однозначный алгоритм классификации X – многозначный алгоритм классификации X» устанавливались одинаковые гиперпараметры.

Результаты вычислительного эксперимента приведены в таблице 3. В таблице приведено название алгоритма, результаты для однозначного и многозначного случая. Светлым цветом выделена ячейка с наивысшим значением метрики *Mean accuracy* среди всех видов классификации.

Таблица 3

Сравнительный анализ однозначных и многозначных классификаторов в вычислительном эксперименте

Название алгоритма классификации	Значение метрики *Mean accuracy* для случая однозначной классификации	Значение метрики *Mean accuracy* для случая многозначной классификации
Tree.DecisionTreeClassifier	0,52	0,75
Tree.ExtraTreeClassifier	0,66	0,69
Ensemble.ExtraTreesClassifier	0,64	0,81
Neighbors.KNeighborsClassifier	0,64	0,91
Ensemble.RandomForestClassifier	0,70	0,13

Как видно из таблицы, 80% однозначных классификаторов уступили в точности классификации по метрике *Mean accuracy* многозначным аналогам, что может сигнализировать о сильном влиянии многозначности меток классов на рассматриваемые модели. Несмотря на то, что многозначных участков всего 3% (см. табл. 2), выигрыш в точности достигает 23% по метрике *Mean accuracy* для алгоритмов *MLL*.

Проведенный эксперимент позволяет сформировать следующие выводы.

Метод LP, используемый для разметки однозначных данных, приводит к высоким погрешностям классификации у бустинговых алгоритмов при кросс-валидации.

Структура данных [21] подвержена влиянию проблемы многозначности гораздо сильнее, чем это может быть оценено стандартной частотной проверкой, выполненной в таблице 2. Одна из возможных причин возникновения столь сильного влияния использование в качестве аргументов первичных атрибутов, напрямую не связанных с классифицируемыми состояниями КС.

Поскольку предсказательная способность частотной проверки влияния результатов многозначности меток классов на результаты классификации однозначных классификаторов невелика, планируются дальнейшие исследования на эту тему. Проведение исследований в области многозначного анализа может привести к повышению точности как статического, так и динамического обнаружения неисправностей в КС и сетевых атак [29].

Заключение

Проанализированы результаты исследования оценки характеристик состояний распределенной компьютерной системы, состоящей из шести хостов при заданных показателях уровня обслуживания *SLO*.

Метки классов (состояния КС) порождаемые в результате функционирования КС, в общем случае, многозначны в следствие съема и анализа информации по нескольким атрибутам КС (с нескольких системных датчиков).

Природа многозначности состояний КС отлична от природы возникновения многозначности при анализе текстовых корпусов или данных социальных сетей.

Аномалии, связанные с нарушением установленных порогов *SLO*, регулярно возникают одновременно по нескольким анализируемым атрибутам.

Результаты проведенного вычислительного позволяют судить о нелинейной зависимости частотного распределения многозначных меток классов на степень влияния многозначности, оказываемую на результаты классификации, что, в свою очередь, непосредственно отражается на защищенности информации, циркулирующей в КС.

В связи с полученными результатами, в случае наличия приоритета в классификации определенных меток классов (что важно для задач информационной безопасности), предлагаются к использованию многозначные классификаторы.

Литература

1. *Kuznetsov A., Babenko V., Kuznetsova K., Kavun S., Smirnov O., Nakisko O.* Malware correlation monitoring in computer networks of promising smart grids // В сборнике: 2019 IEEE 6th International Conference on Energy Smart Systems, ESS 2019 - Proceedings. 6. 2019. С. 347-352. DOI: 10.1109/ESS.2019.8764228

2. *Большаков, А.С., Раковский Д.И.* Эффективный метод многокритериального анализа в области информационной безопасности // Правовая информатика. 2020. № 4. С. 55-66. DOI 10.21681/1994-1404-2020-4-55-66.

3. *Котенко, И.В., Хмыров С.С.* Анализ моделей и методик, используемых для атрибуции нарушителей кибербезопасности при реализации целевых атак // Вопросы кибербезопасности. 2022. Т. 50. № 4. С. 52-79. DOI 10.21681/2311-3456-2022-4-52-79.

4. *Гайфулина, Д.А. Котенко И.В.* Применение методов глубокого обучения в задачах кибербезопасности. Часть 1 // Вопросы кибербезопасности. 2020. № 3(37). С. 76-86. DOI 10.21681/2311-3456-2020-03-76-86.

5. *Alrammal M., Naveed M., Rihawi S.* Using heuristic approach to build anti-malware // В сборнике: ITT 2018 - Information Technology Trends: Emerging Technologies for Artificial Intelligence. 5, Emerging Technologies for Artificial Intelligence. 2019. С. 191-196. DOI: 10.1109/CTIT.2018.8649499

6. *Большаков А.С., Раковский Д.И.* Программное обеспечение моделирования угроз безопасности информации в информационных системах // Правовая информатика. 2020. № 1. С. 26-39. DOI: 10.21681/1994-1404-2020-1-26-39.

7. *Павленко Е.Ю., Гололобов Н.В., Лаврова Д.С., Козачок А.В.* Распознавание киберугроз на адаптивную сетевую топологию крупномасштабных систем на основе рекуррентной нейронной сети // Вопросы кибербезопасности. 2022. № 6 (52). С. 93-98. DOI:10.21681/2311-3456-2022-6-93-99

8. *Израилов К.Е., Буйневич М.В., Котенко И.В., Десницкий В.А.* Оценивание и прогнозирование состояния сложных объектов: применение для информационной безопасности // Вопросы кибербезопасности. 2022. Т. 52. № 6. С. 2 – 21. DOI:10.21681/23113456-6-2022-2-21

9. *Sheluhin O.I., Osin A.V., Rakovsky D.I.* New Algorithm for Predicting the States of a Computer Network Using Multivalued Dependencies // Automatic Control and Computer Sciences, 2023, Т. 57, №. 1. С. 48–60. DOI: 10.3103/S0146411623010091

10. *Gibaja E., Ventura S.* A Tutorial on Multi-Label Learning // ACM Computing Surveys. 2015. № 47. С. 1-40. DOI: 10.1145/2716262

11. *Lima A.C.E.S., de Castro L.N.* A multi-label, semi-supervised classification approach applied to personality prediction in social media // Neural Networks. 2014. Т. 58. С. 122-130.

12. *Карпович С.Н.* Многозначная классификация текстовых документов с использованием вероятностного тематического моделирования ML-PLSI // Труды СПИИРАН. 2016. Т. 47. № 4. С. 92-104 DOI: 10.15622/sp.47.5

13. *Котенко И.В., Саенко И.Б., Браницкий А.А., Паращук И.Б., Гайфулина Д.А.* Интеллектуальная система аналитической обработки цифрового сетевого контента для защиты от нежелательной информации // Информатика и автоматизация. 2021. Т. 20. № 4. С. 755-792. DOI 10.15622/ia.20.4.1

14. *Куликов Г.Г., Антонов В.В., Антонов Д.В.* Анализ возможности извлечения аналитических знаний из формальной модели информационной системы предметной области нейросетевыми методами // Нейрокомпьютеры: разработка, применение. 2013. № 3. С. 12-16.

15. *Azad, M. Moshkov M.* A Bi-criteria Optimization Model for Adjusting the Decision Tree Parameters // Kuwait Journal of Science. 2022. Т. 49. № 2. С. 1- 14. DOI 10.48129/kjs.10725

16. *Niemistö A., Yli-Harja O., Shmulevich I., Lukin V.V., Dolia A.N.* Correction of misclassifications using a proximity-based estimation method // Eurasip Journal on Applied Signal Processing. 2004. № 8. С. 1142-1155. DOI: 10.1155/S1110865704402145

17. *Марков А.С.* Кибербезопасность и информационная безопасность как бифуркация номенклатуры научных специальностей // Вопросы кибербезопасности. 2022. № 1(47). С. 2-9. DOI 10.21681/2311-3456-2022-1-2-9

18. *Ловцов, Д.А.* Принципы обеспечения защищённости информации в эргасистемах // Правовая информатика. 2021. № 1. С. 36-50. DOI 10.21681/1994-1404-2021-1-36-50

19. *Большаков, А.С., Осин А.В., Хусаинов Р.В.* Обнаружение аномалий трафика с использованием нейронной сети для обеспечения защиты информации // I-methods. 2021. Т. 13. № 4. С. 1 – 15.

20. *Шелухин О.И., Раковский Д.И.* Прогнозирование профиля функционирования компьютерной системы на основе многозначных закономерностей // Вопросы кибербезопасности. 2022. № 6. С. 28-45. DOI:10.21681/2311-3456-2022-6-53-70 DOI: 10.36724/2072-8735-2021-15-6-40-47

21. *Шелухин О.И., Раковский Д.И.* Выбор метрических атрибутов редких аномальных событий компьютерной системы методами интеллектуального анализа данных // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 6. С. 40-47. DOI: 10.36724/2072-8735-2021-15-6-40-47

22. *Raja B., Ravindranath K., Jayanag B.* Monitoring and analysing anomaly activities in a network using packetbeat // International Journal of Innovative Technology and Exploring Engineering. 2019. Т. 8. № 6. С. 45-49.

23. *Котенко, И.В., Кулешов А.А., Ушаков И.А.* Система сбора, хранения и обработки информации и событий безопасности на основе средств elastic stack // Труды СПИИРАН. 2017. Т. 54. № 5. С. 5-34. DOI 10.15622/sp.54.1

24. *Петров В.В., Брюханов К.В., Авксентьева Е. Ю.* Сетевой мониторинг: анализ сетевого трафика с помощью ELK // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2020. № 5. С. 102-105. DOI 10.37882/2223-2966.2020.05.34.

25. *Calderon G., Del Campo G., Saavedra E., Santamaria A.* Management and Monitoring IoT Networks through an Elastic Stack-based Platform // Proceedings - 2021 International Conference on Future Internet of Things and Cloud, FiCloud 2021: 8, Virtual, Online, 23-25 августа 2021 года. Virtual, Online, 2021. С. 184-191. DOI 10.1109/FiCloud49777.2021.00034.

26. *Kotenko I., Kuleshov A., Ushakov I.* Aggregation of elastic stack instruments for collecting, storing and processing of security information and events // 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), California, USA, 04-08 августа 2017 года. California, USA: Institute of Electrical and Electronics Engineers, 2017. DOI 10.1109/UIC-ATC.2017.8397627.

27. *Chaudhuri U., Dey S., Banerjee B., Bhattacharya A., Datcu M.* Interband Retrieval and Classification Using the Multilabeled Sentinel-2 BigEarthNet Archive // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2021. Vol. 14. Pp. 9884-9898. DOI 10.1109/JSTARS.2021.3112209

28. *Maltoudoglou L., Paisios A., Papadopoulos H., Lenc L., Martínek J., Král P.* Well-calibrated confidence measures for multi-label text classification with a large number of labels // Pattern Recognition. 2022. T. 122. C. 108271. DOI: 10.1016/j.patcog.2021.108271

29. *Шелухин О.И., Рыбаков С.Ю., Ванюшина А.В.* Модификация алгоритма обнаружения сетевых атак методом фиксации скачков фрактальной размерности в режиме online // Труды учебных заведений связи. 2022. Т. 8. № 3. С. 117-126. DOI 10.31854/1813-324X-2022-8-3-117-126.

INFLUENCE OF MULTI-LABEL CLASS PROBLEM OF SYSTEM LOGS ON THE SECURITY OF COMPUTER NETWORKS

DMITRIY I. RAKOVSKIY
Moscow, Russia

KEYWORDS: *supervised learning, multi-label classification, multiclass classification, information security, multi-label learning.*

ABSTRACT

Intoduction: The security of information circulating in a computer network is related to the security of the supporting infrastructure. An important problem in the intelligent processing of syslog data is the existence of multi-label datasets. Among the Russian-language scientific publications, the problem under consideration in the context of information security of computer networks is not presented. **Purpose:** increase the security of computer networks by using multi-label learning methods when solving the problem of classifying system logs class labels. **Results:** A comparative analysis of single-valued and multi-label classifiers was carried out in a computational experiment on the Mean accuracy metric. A non-linear relationship was found between the proportion of experimental data sections containing multi-label class labels and the overall accuracy of data classification. Despite the fact that multi-label plots in the studied experimental data are only 3%, the gain in accuracy reaches 23% according to the specified metric. According to the results of the analysis, 80% of unambiguous classifiers were inferior in classification accuracy according to the Mean accuracy multi-label metric to their analogues, which may signal a strong influence of multi-label class labels on the models under consideration. It is shown that the considered structure of experimental data in a tabular form is affected by the multi-label problem much more strongly than it can be estimated by a standard frequency check, which actualizes further research in this direction. **Practical relevance:** The practical significance of the results obtained lies in increasing the security of computer networks through the use of a multi-label approach in the classification problem. The tasks of information security solved by multi-label classification may include: the area of monitoring, detection or prevention of violations and computer attacks in computer networks. **Discussion:** Since the predictive power of frequency testing of the influence of multi-label class label results on the classification results of unambiguous classifiers is low, further research on this topic is planned. It is planned to expand the list of classification quality assessment metrics in future experiments.

REFERENCES

1. Kuznetsov A., Babenko V., Kuznetsova K., Kavun S., Smirnov O., Nakisko O. Malware correlation monitoring in computer networks of promising smart grids. *Proceedings of the IEEE 6th International Conference on Energy Smart Systems*, ESS 2019. 2019. Pp. 347-352. DOI: 10.1109/ESS.2019.8764228

2. Bol'shakov A.S., Rakovskii D.I. An efficient multiple-criteria decision analysismethod in the field of information security. *Pravovaya informatika* [Legal Informatics]. 2020. No 4. Pp. 55-66. DOI 10.21681/1994-1404-2020-4-55-66. (In Rus)

3. Kotenko I.V., Khmyrov S.S. Analysis of models and techniques used for attribution of cyber security violators in the implementation of targeted attacks. *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti]. 2022. Vol 50. No 4. Pp. 52-79. DOI 10.21681/2311-3456-2022-4-52-79. (In Rus)

4. Gaifulina D.A., Kotenko I.V. Application of deep learning methodsin cybersecurity tasks. *Voprosy kiberbezopasnosti* [Voprosy kiber-

bezopasnosti]. 2020. Vol 37. No 3. Pp. 76-86. DOI 10.21681/2311-3456-2020-03-76-86. (In Rus)

5. Alrammal M., Naveed M., Rihawi S. Using heuristic approach to build anti-malware. *Proceedings of the ITT 2018 - Information Technology Trends: Emerging Technologies for Artificial Intelligence.* 5, Emerging Technologies for Artificial Intelligence. 2019. Pp. 191-196. DOI: 10.1109/CTIT.2018.8649499.

6. Bol'shakov A.S., Rakovskii D.I. Software for modelling information security threats in information systems. *Pravovaya informatika* [Legal Informatics]. 2020. No 1. Pp. 26-39. DOI: 10.21681/1994-1404-2020-1-26-39. (In Rus)

7. Pavlenko E.Y., Gololobov N.V., Lavrova D.S., Kozachok A.V. Recognition of cyber threats on the adaptive network topology of large-scale systems based on a recurrent neural network. *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti]. 2022. Vol. 52. No 6. Pp. 93-98. DOI:10.21681/2311-3456-2022-6-93-99 (In Rus)

8. Izrailov K.E., Buinevich M.V., Kotenko I.V., Desnitsky V.A. Assessment and prediction of the complex objectsstate: applicatioin for information security. *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti]. 2022. Vol. 52. No 6. Pp. 2-21. DOI:10.21681/23113456-6-2022-2-21 (In Rus)

9. Sheluhin O.I., Osin A.V., Rakovsky D.I. New Algorithm for Predicting the States of a Computer Network Using Multivalued Dependencies. *Automatic Control and Computer Sciences.* 2023. Vol. 57. No. 1. pp. 48-60. DOI: 10.3103/S0146411623010091(In Rus)

10. Gibaja E., Ventura S. A Tutorial on Multi-Label Learning. *ACM Computing Surveys*. 2015. No 47. Pp. 1-40. DOI: 10.1145/2716262

11. Lima A.C.E.S., de Castro L.N. A multi-label, semi-supervised classification approach applied to personality prediction in social media. *Neural Networks*. 2014. vol. 58. Pp. 122-130.

12. Karpovich S.N. Multi-Label Classification of Text Documents using Probabilistic Topic Model ml-PLSI. *Trudy SPIIRAN* [SPIIRAS Proceedings]. 2016. vol 47. no 4. Pp. 92-104 DOI: 10.15622/sp.47.5 (In Rus)

13. Kotenko I.V., Saenko I.B., Branitsky A.A., Paraschuk I.B., Gayfulina D.A. Intelligent system of analytical processing of digital network content for its protection from unwanted information. *Informatics and automation* [Informatics and automation]. 2021. vol. 20, no 4. Pp. 755-784. (In Rus)

14. Kulikov G.G., Antonov V.V., Antonov D.V. Analysis of the possibility of analytical knowledge extraction of a formal model of subject domain information system by neural network methods. *Neurocomputers* [Neurocomputers]. 2013. No 3. Pp. 12-16. (In Rus)

15. Azad M., Moshkov M. A Bi-criteria Optimization Model for Adjusting the Decision Tree Parameters. *Kuwait Journal of Science*. 2022. Vol. 49. No 2. Pp. 1 -14. DOI 10.48129/kjs.10725

16. Niemisto A., Yli-Harja O., Shmulevich I., Lukin V.V., Dolia A.N. Correction of misclassifications using a proximity-based estimation method. *Eurasip Journal on Applied Signal Processing*. 2004. Vol. 2004. No 8. Pp. 1142-1155. DOI: 10.1155/S1110865704402145

17. Markov A.S. Cybersecurity and information security as nomenclature bifurcation scientific specialties (russian text). *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti]. 2022. Vol 47. No 1. Pp. 2-9. DOI 10.21681/2311-3456-2022-1-2-9 (In Rus)

18. Lovtsov D.A. Principles of ensuring information security in ergasystems. *Pravovaya informatika* [Legal Informatics]. 2021. No 1. Pp. 36-50. DOI 10.21681/1994-1404-2021-1-36-50

19. Bolshakov A.S., Khusainov R. V., Osin A.V. Traffic anomaly detection using a neural network to ensure information protection. *I-methods*. 2021. Vol. 13. No 4. Pp. 1-15. (In Rus)

20. Sheluhin O.I., Rakovskiy D.I. Prediction of the profile functioning of a computer system (network) based on multivalued patterns. *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti]. 2022. No 6. Pp. 28-45. DOI:10.21681/2311-3456-2022-6-53-70 (In Rus)

21. Sheluhin O.I., Rakovsky D.I. Selection of metric and categorical attributes of rare anomalous events in a computer system using data mining methods. *T-Comm*. 2021. Vol. 15. No. 6. Pp. 40-47. DOI: 10.36724/2072-8735-2021-15-6-40-47 (In Rus)

22. Raja B., Ravindranath K., Jayanag B. Monitoring and analysing anomaly activities in a network using packetbeat. *International Journal of Innovative Technology and Exploring Engineering*. 2019. Vol. 8. No. 6. Pp. 45-49.

23. Kotenko I.V., Kuleshov A.A., Ushakova I.A. System for collecting, storing and processingsecurity information and events based on elasticstack tools. *Informatika i avtomatizatsiya (Trudy SPIIRAN)* [Informatics and Automation (SPIIRAS Proceedings)]. 2017. Vol. 54. No. 5. Pp. 5-34. DOI 10.15622/sp.54.1(In Rus)

24. Petrov V.V., Bryukhanov K.V., Avksentieva E.Y. Network monitoring: network traffic analysis using ELK. *In Modern Science: actual problems of theory & practice*. 2020. No 5. Pp. 102-105. DOI 10.37882/2223-2966.2020.05.34. (In Rus)

25. Calderon G., Del Campo G., Saavedra E., Santamaria A. Management and Monitoring IoT Networks through an Elastic Stack-based Platform. *Proceedings of 2021 International Conference on Future Internet of Things and Cloud, FiCloud 2021*. Virtual, Online, 2021. Pp. 184-191. DOI 10.1109/FiCloud49777.2021.00034.

26. Kotenko I.V., Kuleshov A.A., Ushakov I.A. Aggregation of elastic stack instruments for collecting, storing and processing of security information and events. *Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI).* California, USA: Institute of Electrical and Electronics Engineers. 2017. Pp. 1-8. DOI 10.1109/UIC-ATC.2017.8397627.

27. Chaudhuri U., Dey S., Banerjee B., Bhattacharya A., Datcu M. Interband Retrieval and Classification Using the Multilabeled Sentinel-2 BigEarthNet Archive. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2021. Vol. 14. Pp. 9884-9898. DOI 10.1109/JSTARS.2021.3112209

28. Maltoudoglou L., Paisios A., Papadopoulos H., Lenc L., Martinek J., Kral P. Well-calibrated confidence measures for multi-label text classification with a large number of labels. *Pattern Recognition*. 2022. Vol. 122. Pp. 108271. DOI: 10.1016/j.patcog.2021.108271

29. Sheluhin O.I., Rybakov S.Yu., Vanyushina A.V. Modified Algorithm for Detecting Network Attacks Using the Fractal Dimension Jump Estimation Method in Online Mode. *Trudy uchebnykh zavedeniy svyazi* [Proceedings of Telecommunication Universities]. 2022. Vol. 8 No 3. Pp. 117-126. (In Rus) https://doi.org/10.31854/1813-324X-2022-8-3-117-126.

INFORMATION ABOUT AUTHOR:

Dmitriy I. Rakovskiy, Lecturer, Postgraduate at the Department of Information Security of Moscow Technical University of Communication andInformatics, Moscow, Russia, Prophet_alpha@mail.ru

For citation: Rakovskiy D.I Influence of multi-label class problem of system logs on the security of computer networks. *H&ES Reserch*. 2023. Vol. 15. No 1. P. 48-56. doi: 10.36724/2409-5419-2023-15-1-48-56 (In Rus)

doi: 10.36724/2409-5419-2023-15-1-57-64

ВЛИЯНИЕ ФРАКТАЛЬНОЙ РАЗМЕРНОСТИ НА КАЧЕСТВО КЛАССИФИКАЦИИ КОМПЬЮТЕРНЫХ АТАК МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

ШЕЛУХИН
Олег Иванович[1]

РЫБАКОВ
Сергей Юрьевич[2]

ВАНЮШИНА
Анна Вячеславовна [3]

АННОТАЦИЯ

Введение: Для построения эффективной системы сетевой защиты компьютерных сетей от атак перспективным направлением является совместное использование фрактального анализа и интеллектуального анализа данных. Предлагается повысить эффективность классификации сетевых атак путем введения дополнительной статистики фрактальной размерности (ФР) атак наряду с другими атрибутами. **Методы:** В отличие от известных работ предлагается дополнительно повысить эффективность классификации сетевых атак путем использования в качестве информационных признаков не только среднего значения, но и других статистических характеристик ФР атак и нормального трафика. Это могут быть дисперсия, коэффициенты асимметрии и эксцесса, характеризующие форму и параметры распределения ФР. Эффективность предлагаемого способа оценивается с помощью алгоритмов машинного обучения путем оценки качества бинарной классификации сетевых атак и нормального трафика на примере использования базы данных UNSW-NB15. Для классификации набора данных были использованы следующие алгоритмы классификации метод k-ближайших соседей (k-NN), множественная логистическая регрессия (LR), дерево решений (DTC), случайный лес (RF), ada boost. Для оценки эффективности построенных моделей использовались метрики: точность (precision), полнота (recall), F-мера (F-score), ROC-кривые, AUC-ROC. **Результаты исследования:** Показано, что использование в качестве дополнительных информационных признаков в виде среднего значения, дисперсии, коэффициентов асимметрии и эксцесса, характеризующих форму и параметры распределения статистических характеристик распределения ФР позволяет повысить эффективность классификации атак в среднем на 10% Наибольший эффект от учета дополнительных статистических параметров ФР заметен для алгоритмов классификации k-NN и LR. Для алгоритмов DTC и RF наибольший эффект от использования дополнительных атрибутов оказывается в сокращении времени обучения и тестирования и составляет около 3,5 раз для каждого из алгоритмов.

Сведения об авторах:

[1] Московский Технический Университет Связи и Информатики (МТУСИ), д.т.н., заведующий кафедрой "Информационная безопасность", Москва, Россия
sheluhin@mail.ru

[2] Московский Технический Университет Связи и Информатики (МТУСИ), аспирант кафедры "Информационная безопасность", Москва, Россия, s.i.rybakov@mtuci.ru

[3] Московский Технический Университет Связи и Информатики (МТУСИ), к.т.н., доцент кафедры "Информационная безопасность", Москва, Россия
a.v.vaniushina@mtuci.ru

КЛЮЧЕВЫЕ СЛОВА: *Фрактальная размерность, бинарная классификация, сетевые атаки, машинное обучение, показатель Херста.*

Для цитирования: *Шелухин О.И., Рыбаков С.Ю., Ванюшина А.В.* Влияние фрактальной размерности на качество классификации компьютерных атак методами машинного обучения // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 1. С. 57-64. doi: 10.36724/2409-5419-2023-15-1-57-64

Введение

Статистический анализ измерений сетевого трафика в компьютерной сети показывает четкое присутствие у него фрактальных или самоподобных свойств [1-4].

В работах [5-9] для решения задач информационной безопасности используются фрактальный анализ.

Для оценки степени самоподобия используются понятия фрактальной размерности множества (по Хаусдорфу) D и показатель Херста H, характеризующий степень самоподобия процесса, связанные между собой соотношением: $D = 2 - H$. В подавляющем большинстве работ в области телекоммуникаций для обнаружения аномалий сетевого трафика используется показатель Херста [3,4,5,16].

Для построения эффективной системы сетевой защиты перспективным направлением является совместное использование фрактального анализа и интеллектуального анализа данных.

В работе [10] на примере базы данных KDD Cup1999 [11,12] показано положительное влияние оценки фрактальных свойств сетевого трафика и атак на качество бинарной классификации. В качестве дополнительного признака нормального трафика и сетевых атак предложено использовать среднее значение показателя Херста H.

В отличие от [10] предлагается дополнительно повысить эффективность классификации сетевых атак путем использования в качестве информационных признаков не только среднего значения, но и других статистических характеристик ФР атак и нормального трафика. В частности, это могут быть дисперсия, коэффициенты асимметрии и эксцесса, характеризующие форму и параметры распределения ФР.

Эффективность предлагаемого подхода может быть оценена путем оценки качества бинарной классификации сетевых атак и нормального трафика на примере использования базы данных (например, UNSW-NB15[13,14]) с помощью широкого класса алгоритмов машинного обучения.

1. Набор данных

В таблице 1 представлена статистика набора данных UNSW-NB15, которая содержит в себе следующие данные: период моделирования, номера потоков, общее количество байтов от источника и получателя, количество пакетов источника, количество пакетов назначения, тип протоколов, количество нормальных и ненормальных записей и количество уникальных IP-адресов источника/назначения [13,14].

Составленные признаки на основе сырых данных представлены в таблице 2. Признаки с 1 по 35 представляют интегрированную собранную информацию из данных пакетов. Большинство признаков генерируется из заголовков пакетов, а дополнительные признаки 35-47 создаются на основе потока.

Основными метками набора данных UNSW-NB15 являются нормальные записи и атаки. В наборе данных представлены 9 типов атак.

Таблица 1

Статистика база данных

		1й день (16 часов)	2й день (15 часов)
No_of_flows		987627	976882
Src_bytes		4860168866	5940523728
Des_bytes		44743560943	44303195509
Src_Pkts		41168425	41129810
Des_Pkts		53402915	52585462
Типы протоколов	TCP	771488	720665
	UDP	301528	688616
	ICMP	150	374
	Others	150	374
Нормальная запись		1064987	1153774
Атака		22215	299068
Количество уникальных IP-адресов источника		40	41
Количество уникальных IP-адресов назначения		44	45

Таблица 2

Признаки набора данных UNSW-NB15

№	Признак	Описание
Потоковые признаки		
1	Scrip	IP адреса отправителя
2	Sport	Номер порта отправителя
3	Dstip	IP адреса получателя
4	Dsport	Номер порта получателя
5	Proto	Протокол связи
Базовые признаки		
6	State	Состояние и его соответствующий протокол, например, ACC, CLO, еще (-)
7	Dur	Общая продолжительность записи
8	Sbyte	Число байтов от отправителя к получателю
9	Dbyte	Число байтов от получателя к отправителю
10	Sttl	Время существования от отправителя к получателю
11	Dttl	Время существования от получателя к отправителю
12	Sloss	Пакеты отправителя ретранслированы или потеряны
13	Dloss	Пакеты получателя ретранслированы или потеряны
14	Service	http, ftp, ssh, dns...,else
15	Sload	Биты отправителя в секунду
16	Dload	Биты получателя в секунду
17	Spkts	Количество пакетов от отправителя к получателю
18	dpkts	Количество пакетов от получателя к отправителю
Содержательные признаки		
19	Swin	Окно подтверждения TCP отправителя
20	Dwin	Окно подтверждения TCP получателя
21	Stcpb	Номер очереди TCP отправителя
22	Dtcpb	Номер очереди TCP получателя
23	Smeansz	Среднее значение размера пакета, переданного с помощью src
24	Dmeansz	Среднее значение пакета, переданного с помощью dst

25	Trans_depth	Глубина подключения http транзакции запроса/ ответа
26	Res_bdy_len	Размер данных, переданных от http службы сервера
Временные признаки		
27	Sjit	Джиттер отправителя (мс)
28	Djit	Джиттер получателя (мс)
29	Stime	Начало времени записи
30	Ltime	Конец времени записи
31	Sintpkt	Время поступления inter-packet отпр отправителя
32	Dinpkt	Время поступления inter-packet получателя (мс)
33	Tcprtt	Сумма ‘synack’ и ‘ackdat’ TCP
34	Synack	Время между SYN и SYN и SYN_ACK пакетами TCP
35	Ackdat	Время между SYN_ACK и ACK пакетами TCP
Дополнительные признаки		
36	Is_sm_ips_port	Если отправитель (1) и получатель (3) имеют одинаковые IPадреса и номера портов (2) (4) равны, тогда эта переменная принимает значение 1, в противном случае 0
37	Ct_state_ttl	Число для каждого состояния (6), соответствующее определенному диапазону значений времени жизни отправителя/получате ля (10) (11).
38	Ct_flw_http_mthd	Число потоков, у которых есть такие методы, как Get и Post в http службе
39	Is_ftp_login	Если сеанс ftp инициирован пользователем и пароль правильный, тогда 1, в противном случае 0.
40	Ct_ftp_cmd	Число потоков, у которых есть команда в ftp сессии.
Признаки соединений		
41	Ct_srv_src	Число соединений, которые содержат одинаковые службы (14) и адреса отправителя в 100 соединениях, согласно последнему времени (26).
42	Ct_srv_dst	Число соединений, которые содержат одинаковые службы (14) и адреса получателя в 100 соединениях согласно последнему времени (26).
43	Ct_dst_ltm	Число соединений одного и того же адреса получателя (3) в каждых 100 соединениях согласно последнему времени (26)
44	Ct_src_ltm	Число соединений одного и того же адреса отправителя (1) в каждых 100 соединениях согласно последнему времени (26).
45	Ct_src_dport_ltm	Число соединений одного и того же адреса отправителя (1) и порта получателя (4) в 100 соединениях согласно последнему времени (26).
46	Ct_dst_sport_ltm	Число соединений одного и того же адреса получателя (1) и порта отправителя (4) в 100 соединениях согласно последнему времени (26).
47	Ct_dst_src_ltm	Число соединений одного и того же адреса отправителя (1) и адреса получателя (3) в 100 соединениях согласно последнему времени (26).
Признаки метки классов		
48	Attack_cat	Название каждого типа атаки. В этом наборе данных содержится 9 типов атак (Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode and Worms)
49	Label	0 для нормальной записи и 1 для записи атаки

В анализируемой базе данных UNSW-NB15 в наборе тестовых и обучающих выборок отсутствуют признаки 29-30, а также признаки 1-4. Всего выборки содержат 175341 и 82332 обучающих и тестовых записей соответственно. Распределение записей по категориям показано на рисунках 1 и 2.

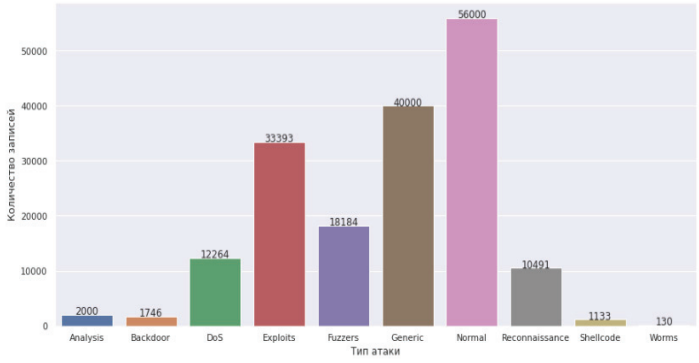


Рис. 1. Распределение записей в обучающей выборке для всех классов

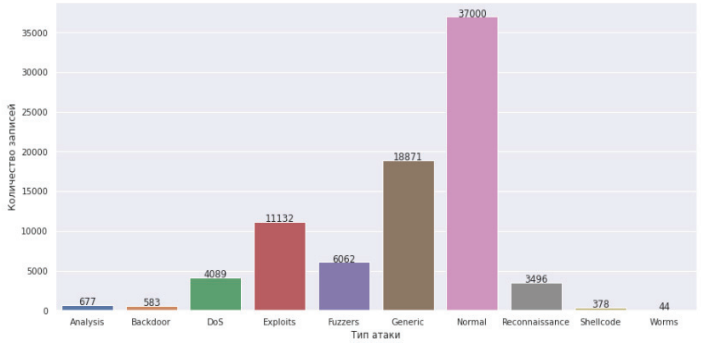


Рис. 2. Распределение записей в тестовой выборке для всех классов

Учитывая, что классы в наборе данных UNSW-NB15 не сбалансированы, т.е. количество записей с классами *normal* и *DoS* в несколько раз превосходит количество записей по другим классам необходимо использовать *нормализацию* набора данных.

Нормализация проводилась по принципу *мини-макс* в соответствии с *формулой* $x' = \frac{x - \min(X)}{\max(X) - \min(X)}$, где $\min(X)$ и $\max(X)$ – минимальное и максимальное значение поля из всего набора данных.

С помощью признаков 1-5 и 29, 30 из исходных данных были выделены потоковые данные с разделением по каждой категории. Фрагмент нормального трафика UNSW-NB15 представлен на рисунке 3.

Используя экспериментально снятые характеристики атак и нормального трафика, можно оценить статистические характеристики ФР атак и нормального трафика на этапе обучения и использовать их затем на этапе классификации.

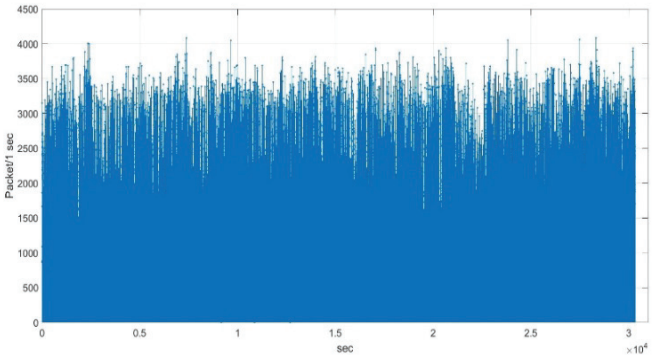


Рис. 3. Фрагмент нормального трафика UNSW-NB15

2. Алгоритмы и метри классификации

Для классификации набора данных были использованы следующие алгоритмы классификации [15]:

- ***Метод k-ближайших соседей*** (k-Nearest Neighbors, neighbors, *k-NN*). Использовался *нормализованный* набор данных.
- ***Множественная логистическая регрессия*** (Logictic Regression, *LR*). Для решения уравнения логистической регрессии использовался алгоритм SAGA. Использовался *нормализованный* набор данных.
- ***Мультиномиальный Наивный Баиес*** (*Multinomial Naive Bayes*, *NB*). Использовался *нормализованный* набор данных.
- ***Дерево решений (****Decision Tree Classifier, DTC*). В качестве оценочной функции использовался коэффициент неопределенности Gini. *Нормализация* данных *не требуется*. В ходе эмпирического анализа было выяснено, что лучший результат достигается при *количестве признаков 28* и при *глубине дерева 23*.
- ***Случайный лес (****Random Forest - RF)*. Из-за того, что основой алгоритма является дерево решений, *нормализация не требуется.* Наилучший результат для рассматриваемого набора данных был получен при разбивке данных на 100 подвыборок.
- ***Ada Boost (****AB)*.

Наилучший результат для рассматриваемого набора данных был получен при разбивке данных на 1000 подвыборок.

В задачах машинного обучения наиболее часто используются следующие метрики для оценки эффективности построенных моделей: точность (*precision*), полнота (*recall*), *F*-мера (*F-score*), *ROC*-кривые (*Receiver Operating Characteristic curve – кривая ошибок*), *AUC-ROC* и *AUC-PR* (*Area Under Curve -площадь под кривой ошибок и площадь под кривой pricison-recall*)

3. Д олнительные фрактальные признаки атак

Для повышения эффективности бинарной классификации анализируемого набора данных предлагается в отличие от работы [10] ввести дополнительные признаки (атрибуты) для каждого из типов обнаруживаемых атак.

При проведении численных расчетов учитывались только реализации атак, для которых количество наблюдений n>100. В этом случае погрешность оценки показателя Херста не превышала 5%.

В качестве совокупности фрактальных атрибутов предлагается использовать экспериментально полученные статистические характеристики ФР такие как: среднее значение ФР (показатель Херста) M_H, дисперсию показателя Херста D_H, коэффициенты асимметрии K_{acc}, и эксцесса $K_Э$, характеризующие форму плотности распределения вероятностей фрактальной размерности $w(H)$.

В таблице 3 представлены результаты оценки указанных выше статистических параметров показателя Херста H для атак всех категорий трафика при количестве реализаций атак равном *N*.

При вычислениях не были учтены атаки Shellcode и Worms, поскольку для них отсутствовали продолжительные интервалы, необходимые для оценки параметров фрактальной размерности.

Таблица 3

Статистические характеристики распределения *w*(*H*) для атак

Тип атаки	(N)	M_H	D_H	K_{acc}	$K_Э$
Normal	20	0.6949	0.0009	0.3137	0.4431
Analysis	9	0.6685	0.0084	0.2493	1.1772
Backdoors	8	0.6121	0.0030	0.8678	0.4829
DoS	18	0.5900	0.0051	1.0607	2.5674
Exploit	21	0.7251	0.0060	0.4528	0.3805
Fuzzers	23	0.6891	0.0045	0.1573	1.2453
Generic	15	0.6726	0.0083	0.3544	1.3438
Reconnaissance	9	0.6026	0.0013	0.1751	1.0603

В соответствии с полученными результатами в таблицу 2 признаков набора данных UNSW-NB15для обучающих и тестовых подвыборок были добавлены четыре новых признака представленных в таблице 4.

Таблица 4

Дополнительные признаки атак и нормального трафика

№	Признак	Описание
Дополнительные признаки фрактальной размерности		
50	herst_avg	Математическое ожидание ФР для распределения $w(H)$
51	herst_desp	Дисперсия D_H для распределения $w(H)$
52	herst_skew	Коэффициент асимметрии K_{acc} для распределения $w(H)$
53	herst_kurtosis	Коэффициент эксцесса $K_Э$ для распределения $w(H)$.

Если в таблице 4 запись отсутствует, то дополнительный признак принимался равным нулю. Это означает, что допол-

нительный признак отсутствует и для данной категории атаки в процессе классификации не принимается во внимание.

На рисунках 4 и 5 представлены гистограммы, позволяющие оценить значимость признаков при учете введенных дополнительных параметров ФР. Важность введенных признаков вычислялась с помощью коэффициента Джини [15,17], лежащем в основе принятия решений алгоритмов *DTC* (рис. 4) и *RF* (рис. 5).

Сравнение гистограмм 4а и 4б показывает, что для алгоритма *DTC* учет только одного дополнительного атрибута в виде среднего значения параметра Херста *herst_avg* ставит его на второе место по значимости при классификации атак. Однако если появляется возможность оценить дополнительные параметры ФР, то наибольшей значимостью будут обладать атрибуты *herst_desp* и *herst_kurtosis*.

В соответствии с таблицей 4 атрибут *herst_desp* характеризует разброс параметра Херста относительно среднего значения. Параметр *herst_kurtosis*, характеризующий форму распределения параметра Херста $w(H)$ имеет хотя и важное, но существенно меньшее значение.

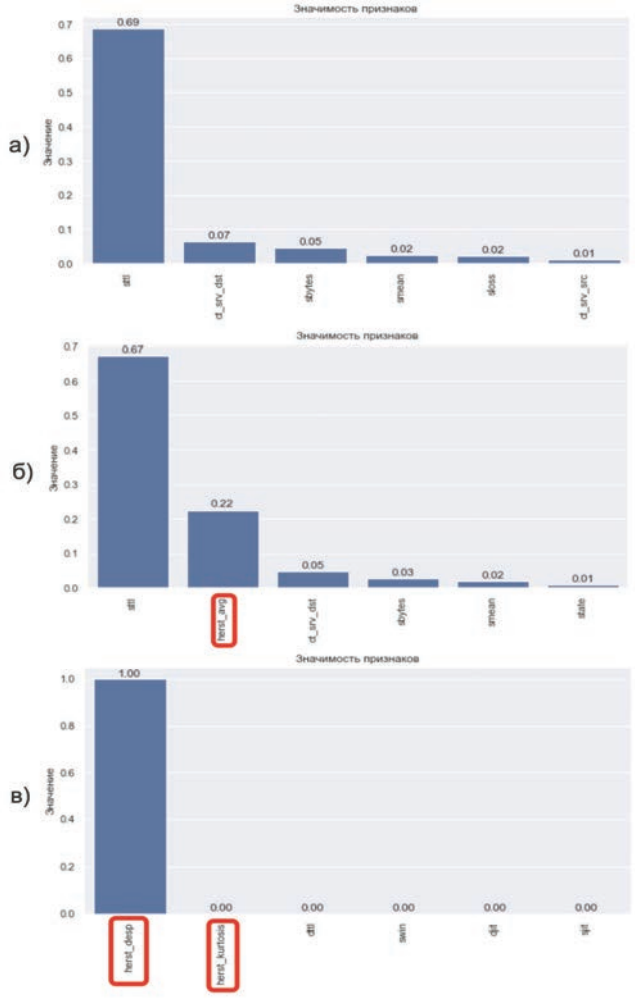


Рис. 4. Значимость первых 6 признаков для алгоритма *DTC* в задаче классификации а) без учета ФР; б) с учетом параметра *herst_avg*(50); в) с учетом всех статистических параметров ФР из таблицы 5

Как видно из рисунка 5а для алгоритма *RF* на качество классификации влияет большее число признаков, по сравнению с алгоритмом *DTC*. Однако и в этом случае учет только одного дополнительного атрибута *herst_avg* ставит его на первое место по значимости.

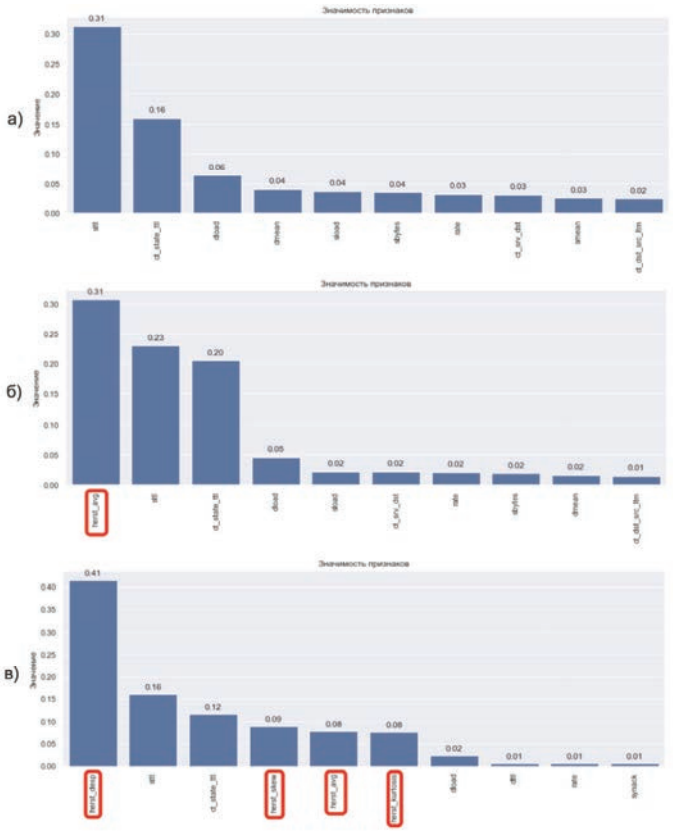


Рис. 5. Значимость первых 10 признаков для алгоритма *RF* в задаче классификации: а) без учета ФР; б) с учетом параметра *herst_avg*(50); в) с учетом всех статистических параметров ФР из таблицы 4

Однако, если появляется возможность оценить дополнительные параметры ФР, характеризующие форму и параметры распределения Херста $w(H)$, наибольшей значимостью будут обладать атрибут *herst_desp*.

В соответствии с таблицей 4 атрибут *herst_desp* характеризует разброс параметра Херста относительно среднего значения. Параметры *herst_skew*, *herst_kurtosis*, характеризующие форму распределения $w(H)$ имеют несколько меньшее значение, занимают по степени важности 4-е и 6-е место.

Из представленных гистограмм можно видеть, что дополнительные статистические атрибуты, представленные в таблице 4, оказывают существенное влияние на алгоритм принятия решения.

4. Результаты бинарной классификации

Рассмотрим результаты сравнительного анализа влияния статистических характеристик $w(H)$ на качество бинарной

классификации атак. Для бинарной классификации все категории атак были приведены к одной категории “Attack”.

В результате классификация сводится к задаче идентификации двух классов: Attack и Normal. Анализировались три режима работы.

1. Классификация только при использовании исходных признаков 1…49 приведенных в таблице 1. Результаты классификации представлены на *рисунках 6-7 (а)*;
2. Классификация при добавлении к набору признаков 1…49 одного дополнительного признака №50 - *herst_avg (50)* - среднего значения показателя Херста. Результаты классификации, соответствующие этому случаю, приведены на рисунках 6-7 (б);
3. Классификация при добавлении к набору признаков 1…49 всех четырех статистических признаков № 50…53 приведенных в таблице 4*: herst_stat* **(***herst_avg*; *herst_desp*; *herst_skew*; *herst_kurtosis***)**. Результаты классификации, соответствующие этому случаю, приведены на рисунках 6-7 (в).

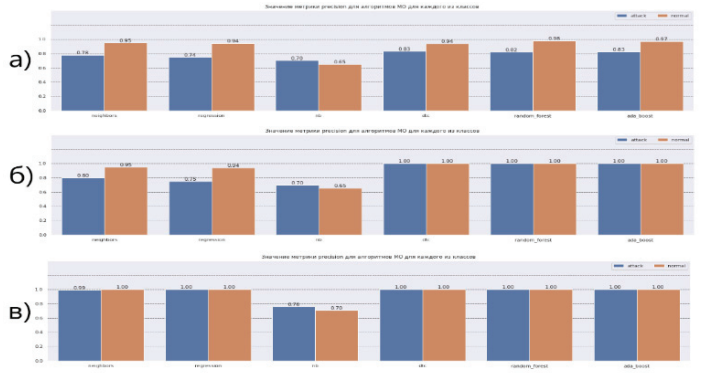


Рис. 6. Значения метрики precision для классификаций а) без учета ФР, б) с учетом параметра *herst_avg*(50) в) с учетом всех статистических параметров из *herst_stat* таблицы 4

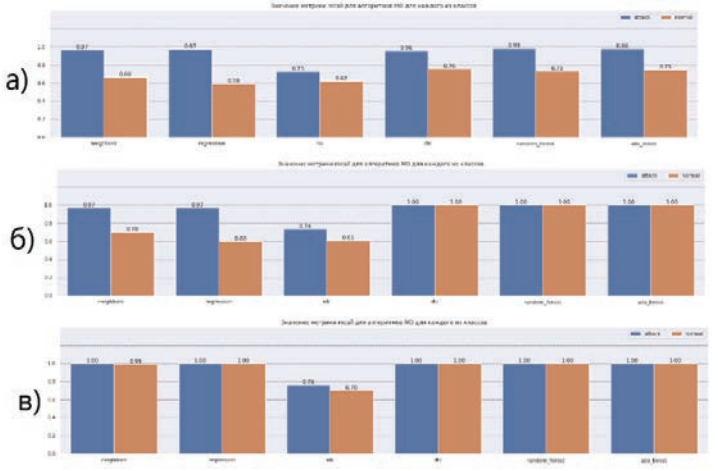


Рис. 7. Значения метрики *recall* для классификаций а) без учета ФР, б) с учетом параметра herst_avg(50) в) с учетом всех статистических параметров *herst_stat* из таблицы 4

При классификации не использовались признаки 1-4, 29, 30 из таблицы 2, поскольку они отсутствовали в исходных обучающих и тестовых выборках.

Из представленных результатов видно, что эффективность использования дополнительных атрибутов в виде статистических параметров ФР атак *herst_stat* и нормального трафика наиболее заметна для алгоритмов классификации *k-NN* и *LR*. Для этих алгоритмов выигрыш от использования дополнительных атрибутов достигает 21% для метрики *precision* при наличии атак и 41% при их отсутствии.

Выигрыш в метрике *f1-score* скромнее и составляет около 7%. Для метрики *AUC-PR* выигрыш составляет 7-8%.

Наибольший эффект достигается от использования в качестве дополнительного признака среднего значения фрактальной размерности – M_H . При использовании алгоритмов классификации *DTC* и *RF* выигрыш от использования дополнительного атрибута M_H оставляет около 15-20% практически для всех рассмотренных метрик.

Более существенным выигрыш от использования дополнительных атрибутов оказывается в сокращении времени обучения и тестирования. Эти результаты приведены в таблице 5.

Таблица 5

Быстродействие алгоритмов классификаций

Алгоритмы\ доп признаки	нет			herst_avg			herst_stat		
	обуч.	предск.	всего	обуч.	предск.	всего	обуч.	предск.	всего
k-NN	76,01	34,84	110,85	94,03	45,78	139,81	**67,57**	**27,44**	**95,01**
LR	6,84	0,01	6,85	7,17	0,01	7,18	**4,65**	**0,005**	**4,65**
NB	0,56	0,01	0,57	0,53	0,02	0,55	**0,48**	**0,01**	**0,49**
DTC	2,32	0,09	2,41	1,47	0,11	1,58	**0,59**	**0,09**	**0,68**
RF	16,49	0,35	16,84	8,21	0,27	8,48	**4,59**	**0,24**	**4,83**
Ada Boost	547,08	14,87	561,95	596,65	15,22	611,87	**469,23**	**13,49**	**482,72**

Наиболее эффективными здесь оказываются также алгоритмы DTC и RF. В случае алгоритма DTC использование одного дополнительного параметра в виде среднего значения ФР привело к снижению времени на обучение и тестирование более чем в 1,5 раза, а для алгоритма «случайный лес» в 1,98 раза.

Использование всех четырех дополнительных атрибутов *herst_stat* **(**herst_avg; herst_desp; herst_skew; herst_kurtosis), представленных в таблице 4 повысило их значимость и привело к снижению времени на обучение и тестирование для алгоритма «дерево решения» в 3,54 раза, а для алгоритма «случайный лес» в 3,48 раза. Абсолютные цифры оказались меньше у алгоритма «дерево решений» и составили 0,68 сек, в то время как для «случайный лес» - 4,83 сек.

Выводы

Введение дополнительных статистических параметров фрактальной размерности, характеризуемых средним значением параметра Херста M_H , дисперсией D_H , K_{acc} , и $K_Э$ характеризующими форму распределения $w(H)$ положительно влияет на качество и скорость бинарной классификации атак.

Для оценки этих параметров могут быть применены традиционные статистические методы. Размер выборок позволяющий провести оценку указанных параметров n должен позволять оценить указанные параметры с заданной погрешностью как на этапе обучения, так и на этапе тестирования.

Сравнительный анализ дополнительных атрибутов показал, что наибольшей значимостью при использовании алгоритма *DTC* являются атрибуты M_H и D_H. При использовании алгоритма *RF* наибольшей значимостью обладает атрибут D_H. Однако велико значение и атрибутов K_{acc}, и $K_Э$ характеризующих форму распределения $w(H)$.

Использование в качестве дополнительных информационных признаков среднего значения, дисперсии, коэффициентов асимметрии и эксцесса, характеризующих форму и параметры распределения статистических характеристик распределения ФР позволяет повысить эффективность бинарной классификации в среднем на 10%.

Наибольший эффект от учета дополнительных статистических параметров ФР заметен для алгоритмов классификации *k- NN* и *LR*.

Для алгоритмов DTC и RF наибольший эффект от использования дополнительных атрибутов (M_H, D_H, K_{acc} и $K_Э$) оказывается в сокращении времени обучения и тестирования и составляет около 3,5 раз для каждого из алгоритмов.

Литература

1. *Sheluhin O., Smolskiy S., Osin A.* Self-Similar Processes in Telecommunications, John Wiley & Sons, 2007.
2. *Atayero A.A., Sheluhin O.I.* Integrated Model for Information Communication Systems and Networks. Desigtn and Development. IGI Global. USA, 2013. P. 462.
3. *Park K., Willinger W.* (Eds.), Self-similar Network Traffic and Performance Evaluation, John Wiley & Sons. 2000.
4. *Monowar H. Bhuyan, Bhattacharyya D. K. Kalita J. K.* Network Anomaly Detection: Methods, Systems and Tools // IEEE Communications surveys & tutorials. 2013. Vol. 60(1). Pp. 303–336.
5. *Wang, X. and B.X. Fang.* An exploratory development on the Hurst parameter variety of network traffic abnormity signal. J. Harbin Inst. Technol., 2005, no. 37, pp.1046-1049.
6. *Mohiuddin A., Abdun Naser M., Jiankun H.* A survey of network anomaly detection techniques // J. Network and Comp. App. 2015. No .60. P. 21.
7. *Z. Sheng, Z. Qifei, P. Xuezeng and Z. Xuhui*, Detection of Low-rate DDoS Attack Based on Self Similarity, in 2010 Second International Workshop on Education Technology and Computer Science, vol. 1 , pp. 333-336, 2010.
8. *Gagandeep Kaur, Vikas Saxena and Jay Prakash Gupta*. Study of Self-Similarity for Detection of Rate-based Network Anomalies. International Journal of Security and Its Applications Vol. 11, No. 8 (2017), pp. 27-44.
9. *Sheluhin O.I., Lukin I.Yu.* Network traffic anomalies detection using fixing method of jumps of multifractal dimension in the real-time mode. Automatic Control and Computer Sciences, September 2018. Vol. 52, Issue 5, pp. 421-430, DOI 10.3103/S0146411618050115
10. *Sheluhin O.I., Kazhemskiy M.A.* Influence of Fractal Dimension on Network Anomalies Binary Classification Quality using Machine Learning Methods, Automatic Control and Computer Sciences, 2020. Vol. 54, No. 3, pp. 216-228, DOI: 10.3103/S0146411620030074
11. KDD Cup 1999 Data https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.
12. NSL-KDD Dataset https://www.unb.ca/cic/datasets/nsl.html
13. Australian Center for Cyber Security (ACCS). (2014). Retrieved from http://www.accs.unsw.adfa.edu.au/
14. *Moustafa N., Slay J*. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set), Military Communications and Information Systems Conference (MilCIS), 2015, At Canberra, Australia, DOI: 10.1109/MilCIS.2015.7348942
15. *Шелухин О.И., Ерохин С.Д., Ванюшина А.В*. Классификация IP-трафика методами машинного обучения / Под ред. профессора О. И. Шелухина. М.: Горячая линия – Телеком, 2018. 282 с. ISBN 978-5-9912-0719-5
16. *Шелухин О.И., Рыбаков С.Ю., Ванюшина А.В*. Модификация алгоритма обнаружения сетевых атак методом фиксации скачков фрактальной размерности в режиме online. Труды учебных заведений связи. 2022. Т. 8. № 3. С. 117-126. DOI:10.31854/1813-324X-2022-8-3-117-126
17. *Шелухин О.И., Раковский Д.И.* Прогнозирование профиля функционирования компьютерной системы на основе многозначных закономерностей // Вопросы кибербезопасности. 2022. № 6(52). С. 53-70. DOI 10.21681/2311-3456-2022-6-53-70.

INFLUENCE OF FRACTAL DIMENSION ON QUALITY CLASSIFICATION OF COMPUTER ATTACKS BY MACHINE LEARNING METHODS

OLEG I. SHELUHIN
Moscow, Russia

SERGEY YU. RYBAKOV
Moscow, Russia

ANNA V. VANYUSHINA
Moscow, Russia

KEYWORDS: *a system of simultaneously and independently operating generators, synergy, emergence, harmonic signal, signal phase, frequency stability, frequency estimation, non-bias and efficiency of estimates, QAM signals.*

ABSTRACT

Introduction. For building an effective network protection system in computer network against attacks, a promising direction is joint use of fractal analysis and data mining. It is proposed to increase the efficiency of network attacks classification by introducing additional fractal dimension (FD) statistics of attacks along with other attributes. In contrast to the well-known works, it is proposed to further improve the efficiency of classifying network attacks by using not only the average value, but also other statistical characteristics of the DF of attacks and normal traffic as information features. These can be variance, skewness and kurtosis coefficients that characterize the shape and parameters of the distribution of the RF. The effectiveness of the proposed

method is evaluated using machine learning algorithms by assessing the quality of the binary classification of network attacks and normal traffic using the UNSW-NB15 database as an example. The following classification algorithms were used to classify the dataset: k-nearest neighbors (k-NN), multiple logistic regression (LR), decision tree (DTC), random forest (RF), ada boost. The following metrics were used to evaluate the effectiveness of the constructed models: accuracy (precision), recall (recall), F-score (F-score), ROC-curves, AUC-ROC. It is shown that the use of mean value, variance, skewness and kurtosis coefficients, which characterize the shape and distribution parameters of the statistical characteristics of the FD distribution as additional information features, makes it possible to increase the efficiency of attack classification by an average of 10%. k-NN and LR classification algorithms. For the DTC and RF algorithms, the greatest effect from the use of additional attributes is in reducing the training and testing time and is about 3.5 times for each of the algorithms.

REFERENCES

1. O. Sheluhin, S. Smolskiy, A. Osin. Self-Similar Processes in Telecommunications, John Wiley & Sons, 2007.

2. A.A. Atayero, O.I. Sheluhin. Integrated Model for Information Communication Systems and Networks. *Desigtn and Development*. IGI Global. USA, 2013. P. 462.

3. K. Park, W. Willinger (Eds.), Self-similar Network Traffic and Performance Evaluation, John Wiley & Sons. 2000.

4. Monowar H. Bhuyan, Bhattacharyya D. K. Kalita J. K. Network Anomaly Detection: Methods, Systems and Tools. *IEEE Communications surveys & tutorials*. 2013. Vol. 60(1). Pp. 303-336.

5. X. Wang, B.X. Fang. An exploratory development on the Hurst parameter variety of network traffic abnormity signal. J. Harbin Inst. Technol., 2005, no. 37, pp. 1046-1049.

6. Mohiuddin A., Abdun Naser M., Jiankun H. A survey of network anomaly detection techniques. *J. Network and Comp. App*. 2015. No 60. P. 21.

7. Z. Sheng, Z. Qifei, P. Xuezeng and Z. Xuhui, Detection of Low-rate DDoS Attack Based on Self Similarity. *2010 Second International Workshop on Education Technology and Computer Science*, vol. 1 , pp. 333-336, 2010.

8. Gagandeep Kaur, Vikas Saxena and Jay Prakash Gupta. Study of Self-Similarity for Detection of Rate-based Network Anomalies. International Journal of Security and Its Applications Vol. 11, No. 8 (2017), pp.27-44.

9. O.I. Sheluhin, I.Yu. Lukin. Network traffic anomalies detection using fixing method of jumps of multifractal dimension in the real-time mode. *Automatic Control and Computer Sciences*, September 2018, Vol. 52, Issue 5, pp 421-430, DOI 10.3103/S0146411618050115

10. O.I. Sheluhin., M.A. Kazhemskiy Influence of Fractal Dimension on Network Anomalies Binary Classification Quality using Machine Learning Methods. *Automatic Control and Computer Sciences*, 2020, Vol. 54, No. 3, pp. 216-228, DOI: 10.3103/S0146411620030074

11. KDD Cup 1999 Data https://kdd.ics.uci.edu/databases/kddcup99/kddcup99

12. NSL-KDD Dataset https://www.unb.ca/cic/datasets/nsl.html.

13. Australian Center for Cyber Security (ACCS). (2014). Retrieved from http://www.accs.unsw.adfa.edu.au/

14. N. Moustafa, J. Slay, UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set), *Military Communications and Information Systems Conference (MilCIS)*, 2015, At Canberra, Australia, DOI: 10.1109/MilCIS.2015.7348942

15. O.I. Sheluhin, S.D. Erokhin, A.V. Vanyushina. IP traffic classification by methods of machine learning. Moscow: Hotline – Telecom, 2018. ISBN 978-5-9912-0719-8

16. O. Sheluhin, S. Rybakov, A. Vanyushina. Modified Algorithm for Detecting Network Attacks Using the Fractal Dimension Jump Estimation Method in Online Mode. *Proc. of Telecom. Universities.* 2022;8(3), pp. 117-126. (in Russ.) DOI:10.31854/1813-324X-2022-8-3-117-126

17. O.I. Sheluhin, D.I. Rakovskiy. Prediction of the profile functioning of a computer system based on multivalued patterns. *Voprosy kiberbezopasnosti*. 2022. No. 6(52), pp. 53-70. DOI 10.21681/2311-3456-2022-6-53-70.

INFORMATION ABOUT AUTHORS:

Oleg I. Sheluhin, Moscow Technical University of Communications and Informatics, Moscow, Russia
Sergey Y. Rybakov, Moscow Technical University of Communications and Informatics, Moscow, Russia
Anna V. Vanyushina, Moscow Technical University of Communications and Informatics, Moscow, Russia

For citation: Sheluhin O.I., Rybakov S.Yu., Vanyushina A.V. Influence of fractal dimension on quality classification of computer attacks by machine learning methods. H&ES Reserch. 2023. Vol. 15. No. 1. P. 57-64. doi: 10.36724/2409-5419-2023-15-1-57-64 (In Rus)

XVI

МЕЖДУНАРОДНЫЙ
НАВИГАЦИОННЫЙ ФОРУМ
И КОНГРЕСС «СФЕРА»

РОСКОСМОС

1
ДЕНЬ

1500
ДЕЛЕГАТОВ

75
ДОКЛАДЧИКОВ-ЭКСПЕРТОВ

400
КОМПАНИЙ

+7 (495) 641 57 17

glonass-forum.ru

13
апреля
2023
ЦВК «Экспоцентр»

Автонет
Национальная технологическая инициатива

ЭКСПОЦЕНТР

Российская неделя высоких технологий
РНВТ

НАВИТЕХ · NAVITECH