ПРОБЛЕМА МАСКИРОВАНИЯ И ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ В КИБЕРПРОСТРАНСТВЕ

Горбачев А.А.[1], Максимов Р.В.[2]

Цель исследования: *определение перспективных направлений научных исследований в области маскирования объектов киберпространства в контексте развития технологий машинного обучения.*

Используемые методы: *общая теория управления и моделирования, математическая статистика, общенаучные методы анализа и синтеза.*

Результат исследования: *определена научная проблема маскирования объектов киберпространства и применения технологий машинного обучения в условиях информационно-технических воздействий злоумышленников. Совершенствование методов маскирования на уровне сетевых узлов, локальных сегментов и информационных направлений с применением методов генеративного и состязательного машинного обучения позволит повысить защищенность объектов киберпространства за счет снижения эффективности сетевой разведки злоумышленников, основанной на методах и алгоритмах машинного обучения. Требуют глубокой теоретической и экспериментальной проработки вопросы: оценки формы, содержания, информативности, предварительной обработки и генерации «цифровых отпечатков» ложных и истинных информационных объектов, выбор типов и оптимальной архитектуры алгоритмов глубокого обучения, оценки качества маскирования как атак типа «уклонение» и «отравление» на алгоритмы машинного обучения потенциальных злоумышленников.*

Научная новизна: *заключается в рассмотрении концепции маскирования объектов киберпространства в условиях информационно-технического воздействия злоумышленников с позиции общей теории управления, моделирования и применения технологий машинного обучения.*

Ключевые слова: *моделирование, теория управления, проактивная парадигма защиты, сетевая разведка, машинное обучение.*

DOI:10.21681/4311-3456-2023-5-37-49

Введение

В последние десятилетия наблюдается процесс интеграции технических систем различного назначения (вычислительных и телекоммуникационных сетей, систем связи и автоматизированного управления) в единую систему, реализующую обмен информацией между гетерогенными элементами с целью решения широкого спектра задач. Это создает условия формирования единого информационного пространства или *киберпространства*, включающего в свою структуру такие понятия, как: Интернет, Интернет вещей, телекоммуникационные и вычислительные сети, процессоры, контроллеры [1]. Некоторые авторы киберпространство представляют в виде трех уровней: психо-когнитивного, программного обеспечения и приложений, а также уровня аппаратного обеспечения [2]. Киберпространство позволяет повысить доступность информационных ресурсов, оперативность информационного обмена и принятия решений, так как научно-технический прогресс приводит к росту информационных потребностей человечества и повышению требований к качеству услуг связи и автоматизированных систем управления.

В современных условиях имеет место *интеллектуализация* сфер деятельности человека, которая заключается в интенсивном развитии и внедрении систем искусственного интеллекта в связи с созданием эффективных архитектур вычислительных устройств, ростом их производительности и доступности, нако-

1 Горбачев Александр Александрович, кандидат технических наук, преподаватель Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменного училища имени генерала армии С.М. Штеменко, г. Краснодар, Россия. E-mail: infosec23.00@ mail.ru

2 Максимов Роман Викторович, доктор технических наук, профессор, профессор Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменного училища имени генерала армии С.М. Штеменко, г. Краснодар, Россия. E-mail: rvmaxim@yandex.ru

плением колоссальных объемов данных и развитием технологий их хранения.

Развитие технологий искусственного интеллекта и методов машинного обучения с одной стороны является предпосылкой, а с другой стороны – следствием изменения подходов к моделированию объектов окружающего мира вообще (рис. 1) [3, 4]. *Модель* как гомоморфное отображение существенных свойств объекта предназначена для *познания* и *управления*. Моделирование из общенаучного метода познания закономерностей окружающего мира превращается в инструмент управления для его эффективного преобразования. Рост вычислительной мощности аппаратной базы позволил развить алгоритмические и имитационные методы моделирования при исследовании объектов и процессов, имеющих принципиально стохастическую природу, так как зачастую отсутствует возможность или целесообразность проведения натурных экспериментов и построения аналитических зависимостей. С другой стороны, технологическое развитие создает потребность в оперативном и автоматизированном построении относительно точных моделей, использующих накопленные статистические наблюдения, которые имеют прикладное значение в системах поддержки принятия решений, системах управления технологическими процессами, робототехнике, системах распознавания образов, обработки естественного языка и других областях. В широком смысле технологии искусственного интеллекта предназначены для автоматизированного синтеза моделей процессов и объектов, то есть для универсальной нелинейной аппроксимации зависимостей между некоторыми входными наблюдениями и выходными количественными или категориальными переменными с целью управления этими объектами или процессами.

Рис. 1. Общие тенденции в теории и методологии математического моделирования

Отмеченные предпосылки привели к смещению методологии моделирования от создания познавательных моделей «механизмов» или «белых ящиков», имеющих детальную внутреннюю структуру, раскрывающую сущность явления к описательным моделям типа «черный ящик» или «статистическая фотография», име-

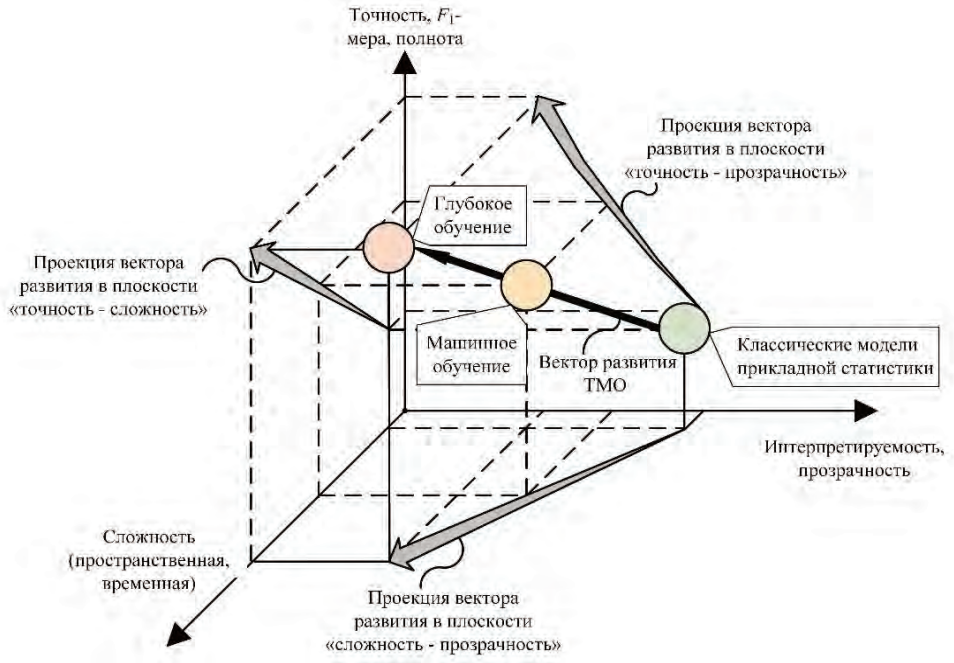

Рис. 2. Тенденции развития технологий машинного обучения и моделирования

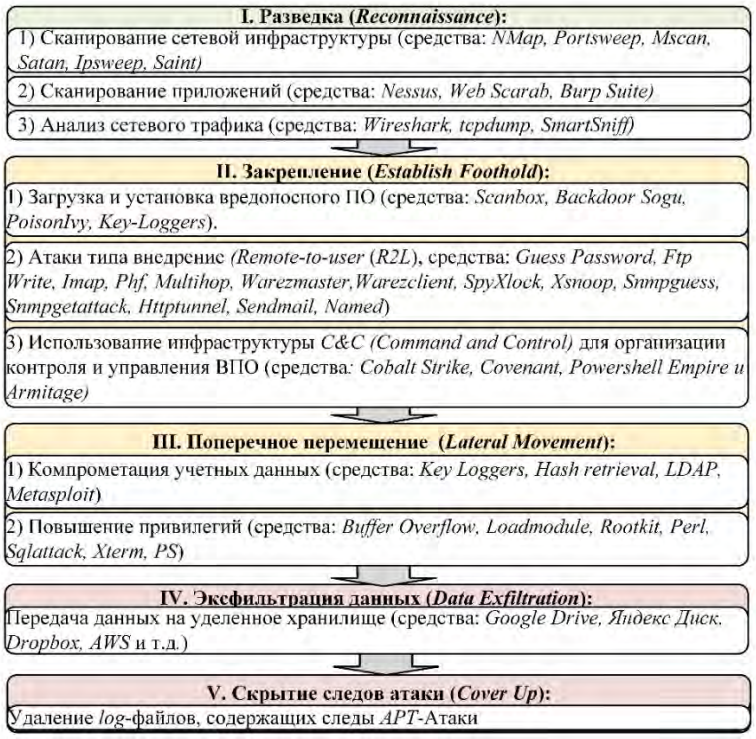

Рис. 3. Вариант реализации APT-атаки на объекты киберпространства

ющим высокую точность, обобщающую способность относительно обучающей выборки, но не раскрывающим сущности описываемых процессов (рис. 2).

Повышение качества статистической аппроксимации моделями глубокого машинного обучения осуществляется за счет их способности обрабатывать большие объемы неструктурированных данных, отсутствия теоретической интерпретации архитектуры модели, то есть ее инвариантности относительно природы объекта. Иными словами, качество аппроксимации наибольшим образом зависит не от архитектуры алгоритма (модели), а от качества обучающих и тестовых наборов данных. Так смещенность, недостаточная представительность обучающих и тестовых выборок могут привести к тому, что доверие к системе искусственного интеллекта не будет обеспечено[3]. В свою очередь наличие подобных уязвимостей процесса идентификации параметров (обучения) моделей инициирует исследование злоумышленниками методов атак на алгоритмы машинного обучения (отравление, уклонение, оракул), которые используются в системах безопасности (обнаружение вредоносного программного обеспечения, системы аутентификации) [5, 6].

3 ГОСТ Р 59276-2020. Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения. Москва: Федеральное агентство по техническому регулированию и метрологии, 2020. 25 с.

Возможности злоумышленников и основные парадигмы защиты

Оборотной стороной указанных условий и факторов является использование технологий машинного обучения для реализации методов информационно-технического воздействия на элементы информационного пространства в форме компьютерных атак и морально-психологического воздействия на определенные группы лиц с целью достижения экономических, политических и военных целей. В статье рассматривается составляющая киберпространства, представленная информационно-телекоммуникационными и вычислительными сетями, а также информационно-технические воздействия в форме компьютерных атак. В мировой практике принято детализировать процесс реализации целевых компьютерных атак (*APT*-атак) на последовательность этапов (рис. 3) [7].

Ключевым этапом, на котором сфокусировано внимание в данной работе, является проведение злоумышленниками *разведки* (сетевой или компьютерной), обеспечивающей достижение целей информационно-технического воздействия на узлы (сетевые информационные объекты) вычислительных сетей за счет определения свойств программного и аппаратного обеспечения, то есть посредством *моделирования* объектов.

	Базовые парадигмы защиты		
	Пассивная	*Реактивная*	*Проактивная*
Характеристика	Физическое или логическое *дистанцирование* со злоумышленником. *Методы*: 1. Создание физически и логически обособленной инфраструктуры. 2.Создание и организационно-техническое обеспечение контролируемой зоны. 3. Логическое разграничение доступа к ресурсам (фильтрация).	Физическое или логическое *дистанцирование* со злоумышленником *при обнаружении* признаков компьютерных атак. *Методы*: 1. Обнаружение вредоносного программного обеспечения. 2. Обнаружение аномалий сетевого трафика, поведения пользователей, содержания файлов и т.д.	*Управление* структурно-функциональными характеристиками объектов вычислительной сети. *Методы*: 1. Защита с использованием подвижных целей (динамичность, многообразие, избыточность). 2. Маскирование информационного обмена и структуры (имитация, мимикрия); 3. Стеганография и шифрование.
Недостатки	1. Сложность обеспечения растущих потребностей и требований к качеству услуг связи и АСУ *между обособленными системами*; 2. Высокие капиталовложения на создание и поддержание инфраструктуры.	1. Реактивный (запаздывающий) характер по отношению к воздействию. 2. Вычислительная ресурсоемкость технических решений. 3. Низкая эффективность относительно угроз 0-дня.	1. Вычислительная ресурсоемкость технических решений. 2. Отрицательное влияние на качество услуг связи. 3. Отсутствие абсолютной защиты без использования пассивных или реактивных средств. 4. Стойкость к вскрытию, затраты на распределение ключей.

Рис. 4. Характеристика базовых парадигм защиты

Если рассматривать меры защиты объектов киберпространства по признаку *активности* по отношению к потенциальным воздействиям злоумышленника, то они реализуются с позиции трех *базовых парадигм защиты*: *пассивной*, *реактивной* и *проактивной* (рис. 4).

Для качественной оценки парадигм защиты рассмотрим процесс информационно-технического воздействия с точки зрения общей теории управления[4] как взаимодействие объекта воздействия (атаки), окружающей среды и системы управления информационно-технического воздействия злоумышленника или субъекта атаки (рис. 5). Окружающая среда (информационно-телекоммуникационная сеть, сопряженная с объектом) воздействует на объект посредством компоненты X (сетевой трафик, генерируемый средой). Объект в свою очередь осуществляет воздействие на среду посредством компоненты Y (сетевой трафик, генерируемый объектом). При этом объект может представлять собой как отдельный элемент или узел вычислительной сети, так и целую подсеть (сегмент, систему) узлов, на которые возложены взаимосвязанные цели функционирования Y_o и защиты Y_d, которые достигаются посредством выполнения алгоритмов функционирования и защиты объекта из множеств алгоритмов Ω_o и Ω_d соответственно.

Цель злоумышленника на этапе разведки состоит в построении *адекватной модели* F_a объекта атаки с использованием средств и алгоритмов разведки Ω_r. Иными словами, целью сетевой разведки является *идентификация* (структурная и параметрическая) модели объекта в широком смысле:

$$\left|Y' - F_a(X', U_r, U_a)\right| \to \min_{F_a \in Q_a, U_r \in \Omega_r, U_a \in \Omega_a}, \qquad (1)$$

где, Q_a – множество модельных операторов и их параметров, находящихся в распоряжении у субъекта атаки; F_a – искомый модельный оператор, связывающий свойства среды и объекта (X' и Y') с воздействиями средства разведки U_r и средствами информационно-технического воздействия U_a.

Построенная модель объекта в форме модельного оператора $F_a(X', U_r, U_a)$ используется в блоке управления для определения оптимального управляющего воздействия субъекта U_a^* из множества алгоритмов (способов) информационно-технического воздействия Ω_a на объект с целью достижения требуемого состояния системы с выходом Y_a.

Моделирование объекта в ходе разведки осуществляется на основании измерений X' и Y', производимых с использованием средств, выполняющих функцию датчиков D_1 и D_2. Датчик D_1 (канальная среда) предназначен для пассивного анализа среды (*пассивная разведка*: анализ сетевого трафика), взаимодействующей с объектом исследования (подконтрольный злоумышленнику элемент, выполняю-

4 Растригин Л.А., Марков В.А. Кибернетические модели познания. Вопросы методологии / Издательство «Зинатне». Рига. 1976 г. 264 с.

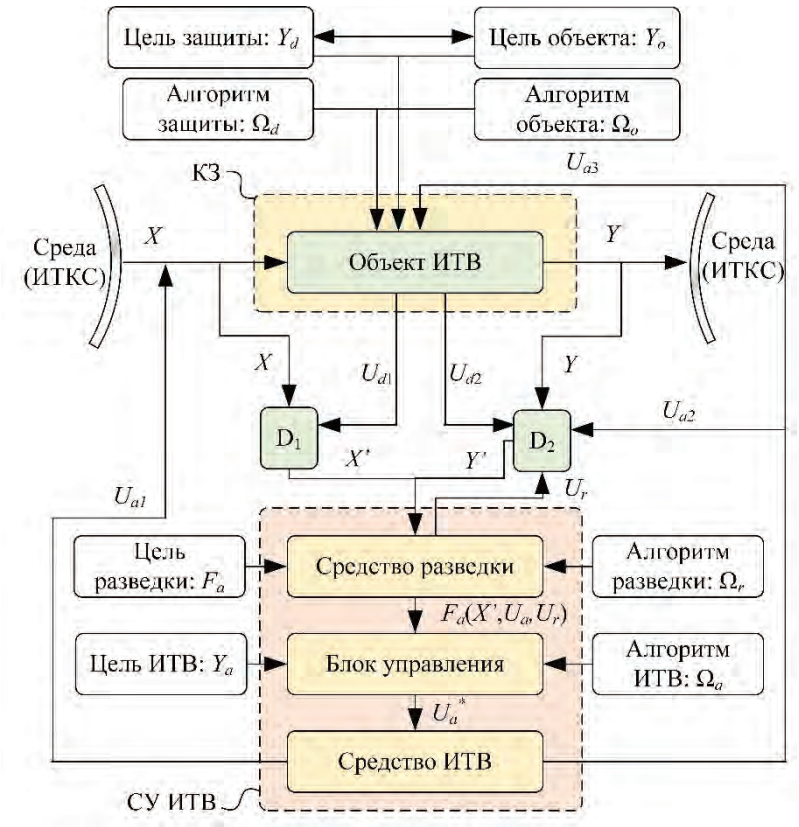

Рис. 5. Общая схема реализации процесса сетевой разведки, информационно-технического воздействия (ИТВ) и защиты в вычислительных сетях (КЗ – контролируемая зона, ИТКС – информационно-телекоммуникационная сеть)

щий функцию захвата сетевого трафика). Датчик D_2 (физический или логический интерфейс взаимодействия объекта с окружающей средой) позволяет осуществить измерение показателей свойств объекта при непосредственном взаимодействии средства разведки и целевого узла системы (*активная разведка*: сканирование).

На основании построенной модели, субъект воздействия обладает возможностью применения способов информационно-технического воздействия опосредованно U_{a1} (атаки типа «отказ в обслуживании», распределенные атаки), непосредственно взаимодействуя с объектом через сетевые интерфейсы U_{a2} (атаки типа: подбор пароля, эксплуатация уязвимостей), а также с использованием способов воздействия U_{a3} в пределах контролируемой зоны (внедрение вредоносного кода с использованием методов социальной инженерии, использование недекларированных возможностей программного и аппаратного обеспечения).

Изменение компонент X' и Y' позволяет системе защиты объекта противодействовать процессу информационного-технического воздействия.

Для построения адекватной модели объекта необходимо соблюдение общих принципов (постулатов) моделирования:

- *наблюдаемость*, которая заключается в возможности измерения интересующих показателей свойств (инвариантов) объекта исследования (вне контролируемой зоны наблюдаемость обеспечивается датчиками D_1 и D_2);
- *стабильность*, которая заключается в стационарности (статистической устойчивости) в узком смысле интересующих показателей свойств (инвариантов) объекта исследования во времени;
- *экстраполируемость*, которая заключается в возможности использования синтезированной модели объекта исследования для последующего управления в иных условиях (зависит от обобщающей способности модельного оператора F_a и стационарности инвариантов объекта исследования);
- *конечность*, которая заключается в конечномерности параметров, подлежащих оценке, в соответствии с требованиями аппаратного и программного обеспечения к сложности алго-

ритма моделирования (конечность определяется видом модельного оператора F_a, многообразием измеряемых характеристик объекта исследования через датчики D_1, D_2, временными и вычислительными ресурсами субъекта моделирования);

— *согласованность* объекта и субъекта, характеризующая наличие понятийного аппарата для качественной интерпретации результатов моделирования;

— *измеримость* указывает на существование системы мер, посредством которой производится измерение показателей свойств объекта исследования.

Соблюдение последних трех принципов обусловлено научно-техническим обеспечением злоумышленника, его вычислительными, временными ресурсами, ассортиментом моделей и методов, аппроксимирующих соответствующие отображения.

Меры защиты объекта, принятые в рамках базовых парадигм защиты, влияют на первые четыре принципа моделирования свойств объекта средством сетевой разведки. Для *парадигмы пассивной защиты* (физического или логического дистанцирования с потенциальным злоумышленником) свойственно редукция компонент X, Y, X' и Y' в связи с полной изоляцией объекта от внешней среды. Пассивная защита не позволяет средству разведки обеспечить принцип наблюдаемости при идентификации модели объекта, так как вне контролируемой зоны отсутствует возможность снятия показаний с датчиков D_1 и D_2. При этом, с одной стороны, не обеспечивается цель объекта по поддержанию заданного уровня показателей качества функционирования Y_0 (в смысле обеспечения обмена информацией с узлами распределенной информационной системы за пределами контролируемой зоны с требуемым качеством), с другой стороны, единственным способом реализации потенциала средств информационно-технического воздействия остается применение способов воздействия $U_{аз}$ в пределах контролируемой зоны, то есть поиск субъектом возможностей для реализации внутренних угроз.

Парадигма реактивной защиты, основанная на принципах реагирования на инциденты безопасности, не накладывает ограничений на компоненты X, Y, X' и Y' в связи с тем, что существует возможность информационного обмена со средой (удаленными узлами), соответственно все основные принципы моделирования (наблюдаемости, стабильности и экстраполируемости) соблюдаются. И лишь при обнаружении признаков атак (вторжений) осуществляется физическое (логическое) отключение защищаемого объекта от среды или блокирование действий субъекта. Принципиальной уязвимостью данного подхода является запаздывающий характер реакции на попытки воздействия злоумышленником, что связано как с принципами настройки средств защиты, так и их функционирования. Основная задача злоумышленника при реализации компьютерных атак сводится к таким воздействиям, которые классифицируются средствами защиты как легитимные (ошибка II рода). В случае эксфильтрации данных злоумышленником может пройти длительный промежуток времени (от нескольких часов до нескольких месяцев) прежде, чем будет обнаружен факт вторжения. Также существует принципиальная возможность блокирования легитимных взаимодействий с объектом (ошибка I рода). Техническая реализация реактивного подхода представлена широким классом средств защиты: системами обнаружения вторжений или атак (СОВ, СОА), средствами антивирусной защиты, фильтрации и анализа данных (трафика), обнаружения аномалий поведения. Несмотря на принципиальные недостатки, разработка и применение средств и способов защиты в рассмотренных подходах являются необходимыми и перспективными направлениями в области информационной безопасности.

Парадигма проактивной защиты предполагает постоянное влияние на компоненты X, Y, X' и Y' без необходимости физического и логического дистанцирования с системами потенциального злоумышленника. В пределах данной парадигмы различают основные подходы – это *защита с использованием подвижной цели*, *маскирование*, *стеганография* и *шифрование*.

Защита с использованием подвижной цели (*moving target defense*) это подход, который использует динамическое изменение структурно-функциональных характеристик защищаемых узлов вычислительной сети. Принципиально данный подход основан на *динамичности* (*shuffling* – изменение параметров объекта во времени), *многообразии* (*diversity* – генерация многообразия возможных значений параметров объекта) и *избыточности* (*redundancy* – синтез дополнительных объектов). Динамичность свойств объекта позволяет нарушить соблюдение принципа стабильности и экстраполируемости при определении модельного оператора объекта атаки F_a злоумышленником, а повышение многообразия и избыточности влияют на конечность измеряемых характеристик посредством изменения структуры и мощности мно-

Рис. 6. Характеристика способов реализации проактивной парадигмы защиты в рамках концепций защиты с использованием подвижной цели и маскирования

жества способов воздействия Ω_a. Данный подход предполагает принципиальную недостижимость абсолютной защищенности, то есть такого состояния объекта защиты, в котором реализация угрозы безопасности информации является невозможным событием [8]. Недостижимость абсолютной защиты связана с тем, что воздействия злоумышленников носят принципиально неопределенный, случайный характер и невозможно достоверно предсказать характеристики этих воздействий (за исключением вырожденного случая, в котором объект перестает выполнять свои функции по предназначению).

Маскирование (*маскировка*), аналогом которого в зарубежной литературе является *киберобман* (*Cyber Deception*) [9], это подход, направленный на создание ложного представления о свойствах объекта атаки методами *мимикрии* и *имитации*. Маскирование и защита с использованием подвижной цели используют общий фундаментальный принцип: динамический характер свойств объекта защиты. Особенностью маскирования является целевой характер изменения характеристик объекта, направленный на создание у злоумышленника *заданного ложного представления* об объекте атаки. В терминах общей теории управления оно направлено на идентификацию злоумышленником *заданного* ложного модельного оператора F'_a и соответственно синтез и применение неоптимального способа информационно-технического воздействия из *заданного* подмножества Ω'_a множества алгоритмов Ω_a.

Одной из проблем при реализации маскирования является неопределенность, связанная с выбором того свойства, характеристики или набора свойств объекта, по которому потенциальный злоумышленник классифицирует реальную или ложную цель, поэтому важным различием подходов является то, что маскирование основано на управлении *произвольным цифровым отпечатком* объекта. Этот отпечаток может содержать любое подмножество или его отображение из множества возможных параметров управления истинным или ложным объектом.

Технические реализации данного подхода предполагают управление параметрами аппаратного и программного обеспечения объектов вычислительных сетей (рис. 6).

Маскирование свойств объекта защиты в форме *мимикрии* позволяет настроить функционирование существующего объекта защиты (узла или системы узлов) в некотором смысле (с точки зрения некоторого функционала степени близости) похожего на какой-либо другой целевой объект (*редукция многообразия*). Например, *сервер-приманка* имеет близкие статические и динамические признаки функционирования с признаками сервера критической информационной инфраструктуры, что приводит к повышению неопределенности для злоумышленника относительно идентификации наиболее важной цели для атаки. Мимикрия может быть использована без создания ложных информационных объектов с целью снижения информативности демаскирующих признаков, свойственных критически важным узлам.

Второй формой маскирования является *имитация*, которая представляет собой создание ложных объектов с заданными характеристиками (*порождение многооб-*

разия). Предназначением ложных объектов является их использование в качестве мишеней для снижения возможностей противника по реализации атак на реальные объекты, а также для проведения *сетевой контрразведки*, то есть для получения информации относительно возможных алгоритмов (способов) Ω_a информационно-технического воздействия злоумышленника.

Стеганография и *шифрование* (криптографическое преобразование информации) являются обособленными методами, рассмотрение которых выходит за рамки данной работы, которые позволяют либо скрыть сам факт передачи информации, либо скрыть содержание передаваемой информации. При этом стойкость стеганографии определяется секретностью алгоритма (способа) скрытой передачи данных, а во втором случае секретностью, характеристиками ключа и криптоалгоритма.

Текущее состояние и перспективы исследований

С учетом типовой архитектуры клиент-серверных вычислительных сетей процесс маскирования структуры и процесса функционирования объектов защиты целесообразно рассмотреть на уровнях: *сетевых узлов*, *локального сегмента* (в пределах контролируемой зоны), *информационных направлений* (между обособленными локальными сегментами за пределами контролируемой зоны).

Анализ предметной области показал, что на уровне маскирования свойств *сетевых узлов* особое внимание уделялось вопросам: имитации ложными и мимикрии реальными сетевыми объектами свойств канальной среды низкого качества с целью исчерпания временных и вычислительных ресурсов средств сетевой разведки при взаимодействии с ними (управление потоком данных и параметрами фрагментации сообщений на сетевом, транспортном и прикладном уровнях стека протоколов *TCP/IP*), а также динамическому изменению структурно-функциональных характеристик узлов (управление *IP*, *MAC*-адресами, *TCP/UDP* портами, частотой их смены) [10-12].

На уровне маскирования свойств *локального сегмента* основными задачами явились: поиск оптимальных режимов многоадресного сетевого соединения и динамической адресации с учетом ресурсных ограничений (управление маской подсети; временем аренды *IP*-адресов, таблицами маршрутизации сети) [13-16].

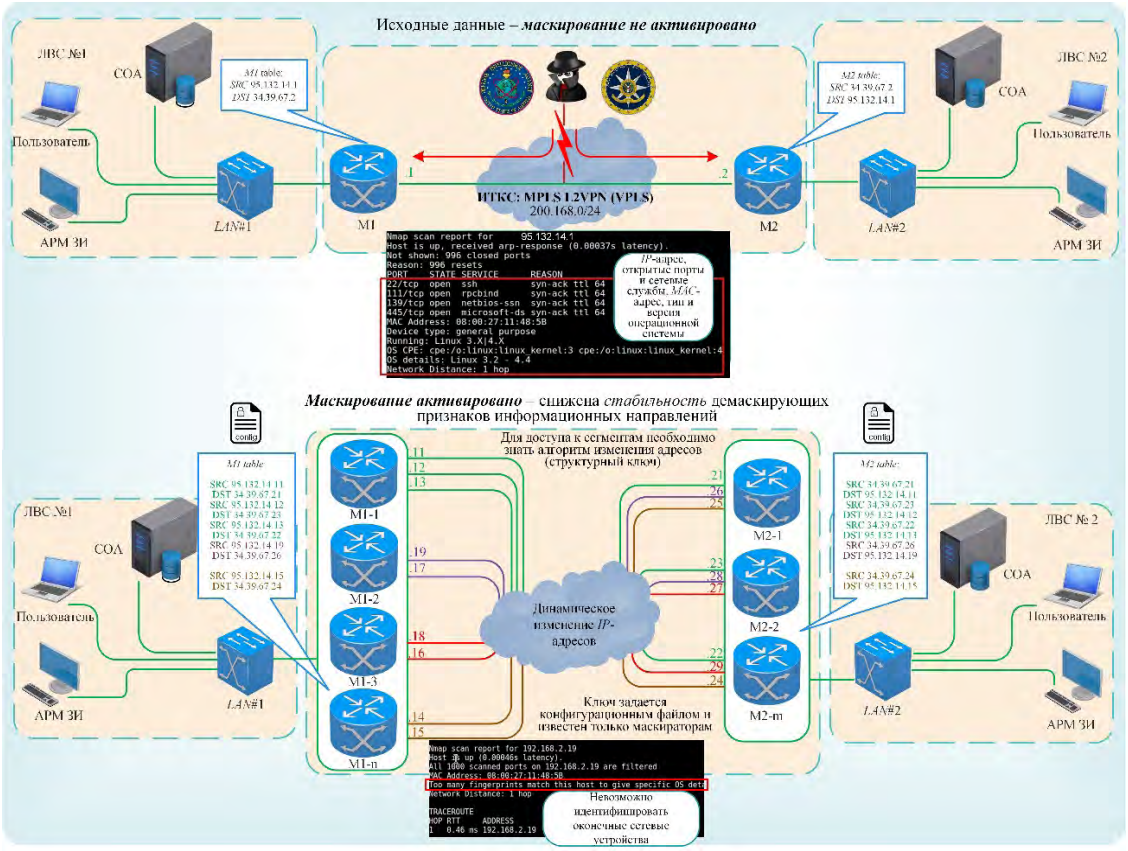

Рис. 7. Принципиальная схема маскирования информационных направлений

Исследования по маскированию свойств *информационных направлений* (рис. 7) были сосредоточены: на мимикрии реальной или имитации ложной иерархической структуры взаимодействующих локальных сегментов и узлов под заданную структуру (управление маршрутизацией между сегментами); моделировании динамических свойств для синтеза ложных (имитация), а также конфигурировании реальных сетевых информационных объектов (мимикрия) посредством генерации ложного (маскирующего) трафика) [17-19].

Научное обоснование вышеописанных разработок осуществлялось на основе: методов математической статистики (первичная обработка экспериментальных данных, проверка статистических гипотез, оценка неизвестных параметров статистических моделей), теории случайных процессов (марковские, полумарковские случайные процессы с дискретным пространством состояний для оценки вероятностно-временных характеристик неблагоприятных состояний информационного объекта, а также для использования полученных моделей в качестве уравнений связи или функционала качества в оптимизационных задачах), теории графов (алгоритмы построения минимальных остовных деревьев, синтез моделей информационных процессов в форме орграфов, решение транспортной задачи для маршрутизации трафика), методов прогнозирования и моделирования временных рядов (имитация ложного трафика с заданными динамическими характеристиками с использованием линейных регрессионных моделей, критериев самоподобия), теории алгоритмов (построение алгоритмов и оценка их свойств), теории оптимизации (использование точных и приближенных методов поиска экстремума функционала качества, заданного аналитически или алгоритмически), а также теории игр. Техническая реализация идей маскирования представлена в программно-аппаратных комплексах, генераторах ложного трафика, ложных сетевых информационных объектах («песочницы», *honeypot*, *honeynet*), специализированном программном обеспечении для реализации мимикрии в существующих системах.

Таким образом, в основу проведенных исследований положено выявление закономерностей с использованием классических математических моделей, имеющих теоретическую интерпретируемость результатов с учетом относительно грубых допущений (ограниченное последействие, нормальное распределение ошибки в значениях временных рядов), но основным недостатком традиционных методов с точки зрения синтеза моделей управления является необходимость структуризации данных, определения существенных свойств объектов. Структура и функционирование информационных объектов характеризуются значительными объемами данных, имеющими в общем случае нелинейные зависимости, оценка информативности которых крайне затруднительна.

Основной проблемой маскирования объектов вычислительных сетей является поиск и устранение *демаскирующих признаков,* по которым осуществляется классификация ложных узлов от истинных, узлов одной категории уязвимости от другой. Несмотря на глубокую теоретическую и прикладную проработку остаются практически не затронутыми методы, модели и алгоритмы машинного обучения при исследовании вопросов маскирования статических и динамических свойств (признаков) объектов, в частности методы *генеративного* и *вредоносного машинного обучения* (*состязательного*), подходящих под специфику научных задач.

Особенности алгоритмов машинного обучения, а именно зависимость их качества от свойств обучающих и тестовых наборов данных позволяет сделать предположение, что маскирование имеет значительное влияние на качество моделей, используемых в процессе компьютерной разведки. Маскирование инвариантов объектов вычислительных сетей по отношению к системам машинного обучения злоумышленника может быть интерпретировано как вредоносное машинное обучение, включающее в себя класс атак типа *уклонение* и *отравление данных* [20]. В терминах вредоносного машинного обучения уклонение использует *вредоносные образы*, которые позволяют избежать правильной классификации объекта алгоритмом машинного обучения. Сущность маскирования заключается в создании подобных вредоносных образов: изменении свойств или цифровых отпечатков сетевых информационных объектов, по которым осуществляется их классификация.

Отравление данных актуально при их использовании в процессе обучения классификаторов, то есть с целью изменения значений параметров моделей, оптимизируемых при обучении с маркированными данными. Так как неизвестно, когда и какие данные могут быть использованы злоумышленниками для обучения алгоритмов, то маскирование структуры и процессов функционирования информационных объектов приводит к уклонению в случае, если классификаторы злоумышленников уже обучены и к отравлению в противном случае.

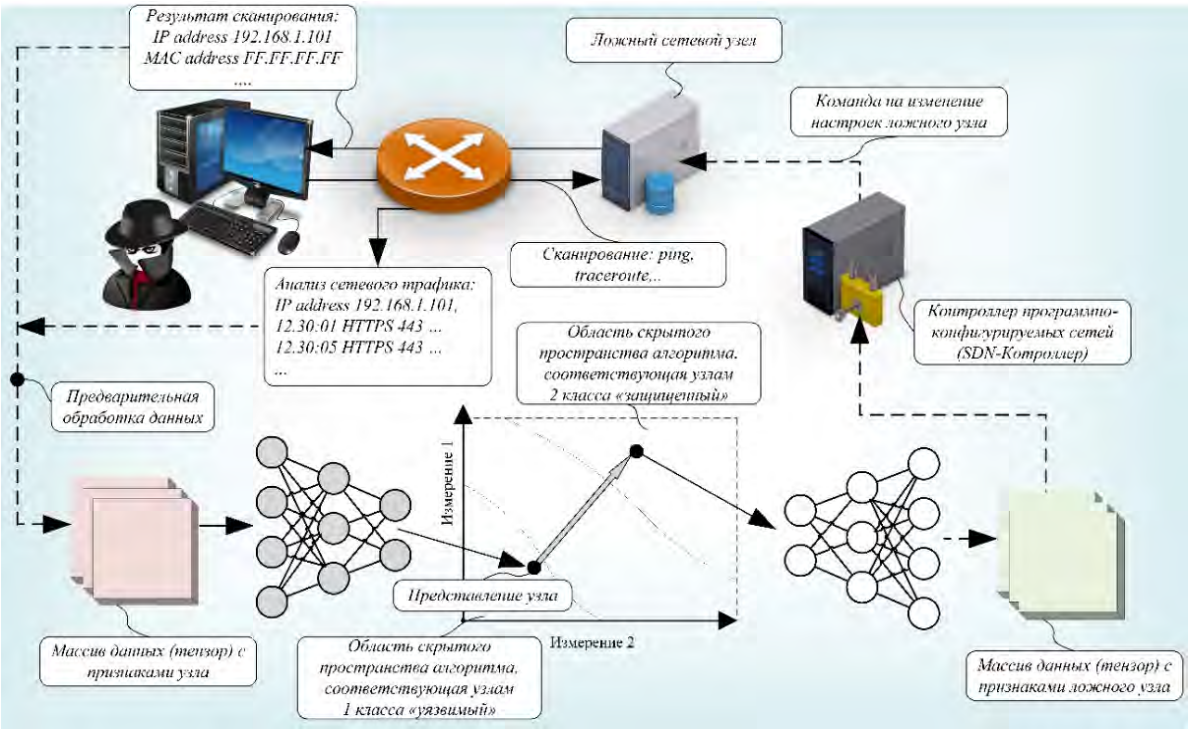

Рис. 8. Принципиальная схема использования алгоритмов машинного обучения для классификации и генерации узлов вычислительных сетей с заданными свойствами

Компьютерная разведка на уровне *сетевого узла* с использованием сканирования и анализа сетевого трафика позволяет получить множество статических (*IP*, *MAC*-адреса, *TCP/UPD*-порты, версии операционных систем, типы и версии сетевых протоколов) и динамических структурно-функциональных характеристик (распределение пакетов от одного источника по протоколам, объемы трафика за фиксированные интервалы времени), образующих *цифровой отпечаток узла* в необработанном виде. Методами глубокого обучения существует принципиальная возможность обработки и обобщения подобных неструктурированных массивов данных для получения сведений о распределении различных объектов вычислительной сети в метрике *скрытого* (*кодового*) *представления* алгоритма (рис. 8).

С точки зрения развития методов маскирования данное обстоятельство может быть использовано: для *генерации ложных узлов* с заданной степенью близости к реальным узлам вычислительной сети; для оценки *значимости* (информативности) первичных демаскирующих признаков посредством построения *карт значимости* признаков. Конкретная форма предварительной обработки (кодирования, масштабирования) и алгоритмов глубокого обучения зависят от специфики первичных данных о структурно-функциональных характеристиках информационных объектов. Сгенерированная конфигурация ложных или истинных сетевых узлов и сегментов сети может быть реализована технологией программно-конфигурируемых сетей.

На уровне *локальных сегментов* и *информационных направлений* использование технологий машинного обучения позволяет дополнительно к рассмотренным тензорам о свойствах узлов учесть сведения, относительно *топологии* и особенностях информационного обмена между узлами и подсетями (интенсивности и распределения трафика по узлам и сегментам, типы и версии сетевых протоколов, матрицы смежности пограничных маршрутизаторов).

Выводы

Научно-технический прогресс вызвал смещение принципов моделирования объектов окружающего мира от синтеза познавательных моделей типа «белый ящик» к автоматическому синтезу описательных моделей типа «черный ящик», предназначенных для целей управления. Моделирование или идентификация адекватной модели объекта вычислительной сети в ходе сетевой разведки является ключевым этапом, от которого зависит успех реализации информацион-

но-технических воздействий злоумышленников.

Проактивная защита как одна из базовых парадигм защиты позволяет обеспечить растущие информационные потребности с учетом качества услуг связи и автоматизации, а также защищенности объектов киберпространства за счет нарушения принципов моделирования их свойств. Маскирование и защита с использованием подвижной цели основаны на общих принципах управления структурно-функциональными характеристиками объектов.

Автоматизированный синтез цифровых отпечатков или инвариантов сетевых узлов, локальных сегментов и информационных направлений с применением технологий генеративного и состязательного машинного обучения, открывает новые возможности для развития средств и методов генерации ложных информационных объектов с заданными динамическими и статическими свойствами, соответствующими замыслу системы защиты.

Дальнейшие исследования в данной области будут направлены на разработку научно-методического аппарата и экспериментальную оценку эффективности маскирования объектов вычислительной сети с использованием различных архитектур алгоритмов глубокого обучения и технологий программно-конфигурируемых сетей.

Литература

1. Стародубцев Ю.И., Закалкин П.В., Иванов С.А. Техносферная война как основной способ разрешения конфликтов в условиях глобализации // Военная мысль. 2020. № 10. С. 16-21.
2. Daniel Ventre. Artificial Intelligence, Cybersecurity and Cyber Defense. Wiley. 2020. 237 p. ISBN 978-1-78630-467-4.
3. Сейновски, Т. Антология машинного обучения: «Издательство «Эксмо», 2022. 509 с.
4. Дауни, Аллен Б. Изучение сложных систем с помощью Python / пер. с анг. Д.А. Беликова. – М.: ДМК Пресс, 2019. 160 с.
5. Грибунин В.Г., Кондаков С.Е. К вопросу о защите информации в интеллектуализированных образцах вооружения // Вопросы кибербезопасности. 2021. № 5(45). С. 5–11. DOI: 10.21681/2311-3456-2021-5-5-11.
6. Aneesh Sreevallabh Chivukula, Xinghao Yang, Bo Liu, Wei Liu, Wanlei Zhou. Adversarial Machine Learning. Attack Surfaces, Defense Mechanisms, Learning Theories in Artificial Intelligence. Springer. 2023. 302 p. ISBN 978-3-030-99771-7.
7. Wang G., Ciptadi A., Ahmadzadeh A. Deployable Machine Learning for Security Defense. Communications in Computer and Information Science. 2020. Vol. 1271. 163 p.
8. Jin-Hee Cho, Dilli P. Sharma, Hooman Alavizadeh, Seunghyun Yoon, Noam Ben-Asher, Terrence J. Moore, Dong Seong Kim, Hyuk Lim, and Frederica F. Nelson. Toward Proactive, Adaptive Defense: A Survey on Moving Target Defense. arXiv:1909.08092v1 [cs.NI] 12 Sep 2019.
9. Kristin E. Heckman, Frank J. Stech. Roshan K. Thomas, Ben Schmoker, Alexander W. Tsow. Cyber Denial, Deception and Counter Deception. A Framework for Supporting Active Cyber Defense. Springer. 2023. 302 p. ISBN 978-3-030-99771-7.
10. Максимов Р.В., Орехов Д.Н., Соколовский С.П. Модель и алгоритм функционирования клиент-серверной информационной системы в условиях сетевой разведки // Системы управления, связи и безопасности. 2019. № 4. С. 50-99.
11. Лебедкина Т.В., Хорев Г.А. Модель функционирования и алгоритм конфигурирования адресации ложных сетевых информационных объектов в условиях сетевой разведки // Системы управления, связи и безопасности. 2023. № 2. С. 23-62.
12. Горбачев А.А. Модель и параметрическая оптимизация проактивной защиты сервиса электронной почты от сетевой разведки // Вопросы кибербезопасности. 2023. № 3(49). С. 69–81. DOI: 10.21681/2311-3456-2023-3-69-81.
13. Максимов Р.В., Соколовский С.П., Ворончихин И.С. Алгоритм и технические решения динамического конфигурирования клиент-серверных вычислительных сетей // Информатика и информатизация. 2020. № 5. С. 1018-1049.
14. Fraunholz D., Anton S.D., Lipps C., Reti D., Krohmer D., Pohl F., Tammen M., Schotten H.D. Demystifying Deception Technology: A Survey. pp. 1-25. arXiv:1804.06196v1 [cs.CR] 17 Apr 2018.
15. Ворончихин И.С., Иванов И.И., Максимов Р.В., Соколовский С.П. Маскирование структуры распределенных информационных систем в киберпространстве // Вопросы кибербезопасности. 2019. № 6(34). С. 92-99. DOI: 10.21681/2311-3456-2019-6-92-99.
16. Москвин А.А., Максимов Р.В., Горбачев А.А. Модель, оптимизация и оценка эффективности применения многоадресных сетевых соединений в условиях сетевой разведки // Вопросы кибербезопасности. 2023. № 3(55). С. 13-22.
17. Кучуров В.В., Максимов Р.В., Шерстобитов Р.С. Модель и методика маскирования адресации корреспондентов в киберпространстве // Вопросы кибербезопасности. 2020. № 6(40). С. 2-13. DOI: 10.21681/2311-3456-2020-6-2-13.
18. Соколовский С.П., Теленьга А.П. Методика формирования ложного сетевого трафика информационных систем для защиты от сетевой разведки // Вестник компьютерных и информационных технологий. 2022. № 2(212). С. 40-47.
19. Шерстобитов Р.С., Шарифуллин С.Р., Максимов Р.В. Маскирование интегрированных сетей связи ведомственного назначения // Системы управления, связи и безопасности. 2018. № 4. С. 136-175.
20. Уорр Кэти. Надежность нейронных сетей: укрепляем устойчивость ИИ к обману / СПб.: Питер, 2021. — 272 с.: ил. ISBN 978-5-4461-1676-8.

THE PROBLEM OF MASKING AND APPLYING OF MACHINE LEARNING TECHNOLOGIES IN CYBERSPACE

Gorbachev A.A.[5], Maximov R.V.[6]

The purpose of the study: *is to identify promising areas of scientific research in the field of masking cyberspace objects in the context of machine learning technologies.*

Methods used: *methods of general control theory and modeling, mathematical statistics, general scientific methods of analysis and synthesis.*

The result of the study: *the scientific problem of masking cyberspace objects and applying of machine learning technologies in the conditions of information and technical influences of intruders is determined. Improving masking methods at the level of network nodes, local segments and information directions using generative and adversarial machine learning methods will increase the security of cyberspace objects by reducing the effectiveness of network intelligence of intruders based on machine learning methods and algorithms. The following issues require deep theoretical and experimental study: evaluation of the form, content, informativity, preprocessing and generation of «digital fingerprints» of fake and true information objects, selection of types and optimal architecture of deep learning algorithms, evaluation of the quality of masking methods as «evasion» and «poisoning» attacks on machine learning algorithms of potential attackers.*

Scientific novelty: *consists in considering the concept of masking cyberspace objects in the conditions of information and technical impact of intruders from the standpoint of the general theory of control, modeling and application of machine learning technologies.*

Keywords: *modeling, control theory, proactive protection paradigm, network intelligence, machine learning.*

References

1. Starodubcev YU.I., Zakalkin P.V., Ivanov S.A. Tekhnosfernaya vojna kak osnovnoj sposob razresheniya konfliktov v usloviyah globalizacii // Voennaya mysl'. 2020. № 10. S. 16-21.
2. Daniel Ventre. Artificial Intelligence, Cybersecurity and Cyber Defense. Wiley. 2020. 237 p. ISBN 978-1-78630-467-4.
3. Sejnovski, T. Antologiya mashinnogo obucheniya: «Izdatel'stvo «Eksmo», 2022. 509 s.
4. Dauni, Allen B. Izuchenie slozhnyh sistem s pomoshch'yu Python / per. s ang. D.A. Belikova. – M.: DMK Press, 2019. 160 s.
5. Gribunin V.G., Kondakov S.E. K voprosu o zashchite informacii v intellektualizirovannyh obrazcah vooruzheniya // Voprosy kiberbezopasnosti. 2021. № 5(45). S. 5–11. DOI: 10.21681/2311-3456-2021-5-5-11.
6. Aneesh Sreevallabh Chivukula, Xinghao Yang, Bo Liu, Wei Liu, Wanlei Zhou. Adversarial Machine Learning. Attack Surfaces, Defense Mechanisms, Learning Theories in Artificial Intelligence. Springer. 2023. 302 p. ISBN 978-3-030-99771-7.
7. Wang G., Ciptadi A., Ahmadzadeh A. Deployable Machine Learning for Security Defense. Communications in Computer and Information Science. 2020. Vol. 1271. 163 p.
8. Jin-Hee Cho, Dilli P. Sharma, Hooman Alavizadeh, Seunghyun Yoon, Noam Ben-Asher, Terrence J. Moore, Dong Seong Kim, Hyuk Lim, and Frederica F. Nelson. Toward Proactive, Adaptive Defense: A Survey on Moving Target Defense. arXiv:1909.08092v1 [cs.NI] 12 Sep 2019.
9. Kristin E. Heckman, Frank J. Stech. Roshan K. Thomas, Ben Schmoker, Alexander W. Tsow. Cyber Denial, Deception and Counter Deception. A Framework for Supporting Active Cyber Defense. Springer. 2023. 302 p. ISBN 978-3-030-99771-7.
10. Maksimov R.V., Orekhov D.N., Sokolovskij S.P. Model' i algoritm funkcionirovaniya klient-servernoj informacionnoj sistemy v usloviyah setevoj razvedki // Sistemy upravleniya, svyazi i bezopasnosti. 2019. № 4. S. 50-99.
11. Lebedkina T.V., Horev G.A. Model' funkcionirovaniya i algoritm konfigurirovaniya adresacii lozhnyh setevyh informacionnyh ob"ektov v usloviyah setevoj razvedki // Sistemy upravleniya, svyazi i bezopasnosti. 2023. № 2. S. 23-62.
12. Gorbachev A.A. Model' i parametricheskaya optimizaciya proaktivnoj zashchity servisa elektronnoj pochty ot setevoj razvedki // Voprosy kiberbezopasnosti. 2023. № 3(49). S. 69–81. DOI: 10.21681/2311-3456-2023-3-69-81.
13. Maksimov R.V., Sokolovskij S.P., Voronchihin I.S. Algoritm i tekhnicheskie resheniya dinamicheskogo konfigurirovaniya klient-servernyh vychislitel'nyh setej // Informatika i informatizaciya. 2020. № 5. S. 1018-1049.

5 Alexander A. Gorbachev, Ph.D., Assistant Professor, Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M. Shtemenko, Krasnodar, Russia. E-mail: infosec23.00@mail.ru

6 Roman V. Maximov, Dr.Sc., Professor, Professor, Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M. Shtemenko, Krasnodar, Russia. E-mail: rvmaxim@yandex.ru

14. Fraunholz D., Anton S.D., Lipps C., Reti D., Krohmer D., Pohl F., Tammen M., Schotten H.D. Demystifying Deception Technology: A Survey. pp. 1-25. arXiv:1804.06196v1 [cs.CR] 17 Apr 2018.
15. Voronchihin I.S., Ivanov I.I., Maksimov R.V., Sokolovskij S.P. Maskirovanie struktury raspredelennyh informacionnyh sistem v kiberprostranstve // Voprosy kiberbezopasnosti. 2019. № 6(34). S. 92-99. DOI: 10.21681/2311-3456-2019-6-92-99.
16. Moskvin A.A., Maksimov R.V., Gorbachev A.A. Model', optimizaciya i ocenka effektivnosti primeneniya mnogoadresnyh setevyh soedinenij v usloviyah setevoj razvedki // Voprosy kiberbezopasnosti. 2023. № 3(55). S. 13-22.
17. Kuchurov V.V., Maksimov R.V., SHerstobitov R.S. Model' i metodika maskirovaniya adresacii korrespondentov v kiberprostranstve // Voprosy kiberbezopasnosti. 2020. № 6(40). S. 2-13. DOI: 10.21681/2311-3456-2020-6-2-13.
18. Sokolovskij S.P., Telen'ga A.P. Metodika formirovaniya lozhnogo setevogo trafika informacionnyh sistem dlya zashchity ot setevoj razvedki // Vestnik komp'yuternyh i informacionnyh tekhnologij. 2022. № 2(212). S. 40-47.
19. SHerstobitov R.S., SHarifullin S.R., Maksimov R.V. Maskirovanie integrirovannyh setej svyazi vedomstvennogo naznacheniya // Sistemy upravleniya, svyazi i bezopasnosti. 2018. № 4. S. 136-175.
20. Uorr Keti. Nadezhnost' nejronnyh setej: ukreplyaem ustojchivost' II k obmanu / SPb.: Piter, 2021. — 272 s.: il. ISBN 978-5-4461-1676-8.