TECHNICAL SCIENCES

CRITICAL BREAK POINTS IN THE ATTACK CYCLE: A SIX-STAGE MODEL FOR PROACTIVE DEFENSE AGAINST SOCIAL ENGINEERING

Marinov A.
PhD in Economic Sciences, associate Professor at the Cybersecurity Competence Center, Irkutsk National ResearchTechnical University, Irkutsk, Russia
https://orcid.org/0000–0003–2238–2751
Kolenchenko D.
student of group ISb-23-1, School of Information Technology and Data Science, Irkutsk National ResearchTechnical University, Irkutsk, Russia
https://orcid.org/0009-0002-7222-1462

ВЫЯВЛЕНИЕ КРИТИЧЕСКИХ ТОЧЕК РАЗРЫВА ЦИКЛА АТАКИ: ПРИМЕНЕНИЕ ШЕСТИЭТАПНОЙ МОДЕЛИ ДЛЯ ПРОАКТИВНОГО ПРОТИВОДЕЙСТВИЯ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ

Маринов А.А.
кандидат экономических наук, доцент центра компетенций по кибербезопасности, ФГБОУ ВО «Иркутский национальный исследовательский технический университет», г. Иркутск, Россия.
https://orcid.org/0000–0003–2238–2751
Коленченко Д.И.
Студент группы Ибб-23-1, Институт ИТ и АД, ФГБОУ ВО «Иркутский национальный исследовательский технический университет», г. Иркутск, Россия.
https://orcid.org/0009-0002-7222-1462
https://doi.org/10.5281/zenodo.19292675

Abstract

This paper presents an overview of modern social engineering methods enhanced by generative AI technologies. The dominant role of data in the success of cyberattacks is examined using empirical data. The inadequacy of classical linear models for describing the adaptive nature of AI-mediated threats is revealed. A six-stage conceptual model of the life cycle of a social engineering attack is proposed. The attack is presented as a cyclical, iterative process with stages of goal formulation, reconnaissance, preparation (with data verification), initiation of contact, exploitation, and completion (with reflection). The key advantages of the model are its consideration of adaptability, the allocation of data verification to AI content generation, and the ability to influence the victim's subconscious mechanisms. The practical significance of this work lies in the possibility of constructing proactive defense systems with attack interruption points at any stage of its implementation. The mathematical justification of the model using the apparatus of probability theory and Markov processes, and the corresponding flowchart of the algorithm, confirming the cyclical nature of modern threats and the possibility of their formalized description, was developed.

Аннотация

В работе представлен обзор современных методов социальной инженерии, усиливаемых технологиями генеративного ИИ. Рассмотрена доминирующая роль в успешности кибератак на основе эмпирических данных. Выявлена несостоятельность классических линейных моделей для описания адаптивной природы ИИ-опосредованных угроз. Предложена шестиэтапная концептуальная модель жизненного цикла социоинженерной атаки. Атака представлена как циклический итеративный процесс с этапами формулировки цели, рекогносцировки, подготовки (с верификацией данных), инициации контакта, эксплуатации и завершение атаки (с рефлексией). Ключевыми преимуществами модели являются учет адаптивности, выделение верификации данных для ИИ-генерации контента и возможность воздействия на подсознательные механизмы жертвы. Практическая значимость работы заключается в возможности построения проактивных систем защиты с точками прерывания атаки на любом этапе ее реализации. Математическое обоснование модели с использованием аппарата теории вероятностей и марковских процессов, а также разработана соответствующая блок-схема алгоритма, подтверждающая циклическую природу современных угроз и возможность их формализованного описания.

Keywords: cyber threat, generative artificial intelligence (generative ai), attacker, cyberattack, conceptual model of the attack lifecycle, social engineering, vulnerability, countermeasures against social engineering, defense against social engineering.

Ключевые слова: киберугроза, генеративный искусственный интеллект, злоумышленник, кибератака, концептуальная модель жизненного цикла атаки, социальная инженерия, уязвимость, противодействия социальной инженерии, критические точки разрыва цикла атаки.

Введение

Глобальная цифровизация, охватывающая все сферы человеческой деятельности с конца XX века, повышает производительность труда и качество жизни, однако одновременно актуализирует специфические проблемы информационной безопасности. Несмотря на то, что вопросы защиты на программно-техническом и правовом уровнях системно прорабатываются профильными ведомствами, а результаты НИОКР воплощаются в современных средствах защиты и адаптируемой нормативной базе, в любой защищенной системе наиболее уязвимым элементом остается пользователь – носитель индивидуальных поведенческих особенностей, объединяемых понятием «человеческий фактор». Именно это обуславливает высокую эффективность техник социальной инженерии как способа реализации кибератак: манипуляция человеческим фактором позволяет обходить классические, преимущественно технологические системы защиты и получать несанкционированный доступ к конфиденциальным данным [1, 2]. Дополнительным стимулом для злоумышленников в текущих реалиях служит обострение мировой социально-политической и экономической нестабильности, превращающее атаки на основе социальной инженерии в инструмент ведения гибридной войны [3].

Таким образом, к наиболее актуальным задачам информационной безопасности относятся как анализ методик социальной инженерии, используемых злоумышленниками, так и разработка блок-схемы, в деталях описывающая жизненный цикл современных социоинженерных атак.

Цель настоящего исследования состоит в проведении анализа современных методов и приемов социальной инженерии и разработке обоснованной концептуальной модели цикла социоинженерной атаки, отражающего специфику современных угроз.

Социальная инженерия

Термин «социальная инженерия» вошел в академический дискурс относительно недавно, однако методика получения информации путем информационно-психологического воздействия, где пользователь выступал субъектом доступа к защищаемым данным. Согласно источникам [1, 2, 3, 5], в предметной области информационной безопасности социальная инженерия определяется как комплексная методология получения несанкционированного доступа к конфиденциальной информации посредством информационно-психологического воздействия, где доступ реализуется преимущественно через эксплуатацию уязвимостей человеческого фактора, а не через прямой технический взлом. Ключевое отличие от традиционных кибератак заключается в том, что технические средства применяются лишь на отдельных этапах (доставка фишингового письма, размещение вредоносного контента), но не определяют успех операции. Основой служит психологическая манипуляция: создание правдоподобного предлога, формирование ложного контекста, использование авторитета или срочности для побуждения жертвы к целевому действию [2, 4]. Соответственно, социоинженерные атаки – это вид кибератак, направленных на получение доступа к личным данным или объектам информатизации и основанных преимущественно на эксплуатации человеческого фактора.

Высокая эффективность таких атак обусловлена тем, что любой человек, будучи субъектом информационных отношений, обладает особенностями и слабостями, воздействие на которые способно реализовать сценарий, выгодный злоумышленнику [2, 3]. Речь идет о множественных психологических техниках, образующих стратегии воздействия на сознание и поведение жертвы.

Анализ статистики актуальных киберугроз

Эмпирические данные ведущих компаний, в частности ежеквартальные отчеты Positive Technologies «Актуальные киберугрозы» [3, 4], подтверждают активный рост атак на основе социальной инженерии. Согласно отчету за I–II кварталы 2025 года, социальная инженерия доминирует в атаках на частных лиц (93%, рост на 5 п.п.) [6, 7]. В атаках на организации метод применяется в 50% успешных инцидентов, при этом в 97% таргетированных атак злоумышленники используют социальную инженерию, а в 43% случаев технические векторы не задействуются вовсе [6].

Анализ каналов доставки выявляет устойчивые различия в тактике злоумышленников: в атаках на организации доминирует электронная почта (88% инцидентов), тогда как на частных лиц – вебсайты (69%) [7]. Примечательным трендом стал рост использования мессенджеров в атаках на частных лиц на 92% [6], что обусловлено применением утекших персональных данных, взломанных аккаунтов и технологий генеративного ИИ.

Эффективность социальной инженерии подтверждается результатами атак: утечки данных зафиксированы в 52% случаев атак на организации и в 74% – на частных лиц, преимущественно похищаются учетные и персональные данные [6, 7, 8]. Рост использования мессенджеров и социальных сетей коррелирует с развитием прогрессивных методик на базе ИИ, включая дипфейки и гиперперсонализированный фишинг [7, 9]. Это свидетельствует о качественном изменении природы угроз: от массовых неперсонализированных атак к целевым операциям с автоматизированной генерацией контента. Социальная инженерия остается наиболее эффективным методом реализации кибератак, а ее эволю-

ция в сторону ИИ-опосредованных операций требует переосмысления подходов к противодействию.

Прогрессивные методы атак

Наиболее рационален подход к рассмотрению социоинженерных атак со стороны злоумышленников, использующего коммуникационные среды (Интернет, почта, соцсети) [2, 4, 5]. Поскольку методы атак на организации и физических лиц отличаются, необходима классификация с разделением цели. Аналитические данные [1, 2, 5] демонстрируют, что наиболее эффективно реализуются атаки с использованием техник фишинга, претекстинга и генеративного ИИ. Лидирующими методами по распространенности и эффективности на сегодняшний день являются Deepfake, Spear phishing, Quishing, ClickFix, Typosquatting [6, 7, 8, 9].

Технология deepfake, основанная на генеративно-состязательных сетях (GAN) и трансформерных архитектурах, позволяет создавать синтетический мультимедийный контент, визуально или акустически неотличимый от оригинала [9, 10]. В контексте социальной инженерии она реализуется в четырех модификациях: подмена лица, синтез речи (клонирование голоса), полная генерация видео «с нуля» и интерактивная видеоконференция с дипфейком. Ключевой эксплуатируемой уязвимостью является доверие к мультимодальным каналам коммуникации, что делает традиционные методы верификации неэффективными и требует внедрения дополнительных факторов аутентификации для критически важных операций [11, 12].

Спирфишинг с использованием ИИ автоматизирует подготовку атаки через большие языковые модели (LLM), заменяя ручной сбор данных и написание писем [13]. LLM генерируют гиперперсонализированные сообщения на основе открытых данных, учитывая контекст и стиль жертвы [14], что затрудняет обнаружение. По данным Avast, персонализированные фишинговые атаки с применением ИИ демонстрируют в 3-5 раз более высокий показатель конверсии по сравнению с массовыми кампаниями [15]. Традиционные модели не отражают данный уровень автоматизации и скорости создания контента.

Векторы доставки (квишинг, ClickFix) реализуют переход к пассивному «выжиданию»: создаются приманки (QR-коды, поддельные CAPTCHA), рассчитанные на автоматическое действие пользователя [7, 16]. При квишинге QR-код во вложении DOCX перенаправляет на фишинговый сайт [17], при ClickFix вредоносная команда копируется в буфер обмена через поддельную CAPTCHA для вставки пользователем. Рост методов показателен: использование мессенджеров выросло на 92%, а доля атак с ClickFix в H1 2025 увеличилась более чем на 500% относительно H2 2024 [7]. Фокус сместился на создание «ловушек», эксплуатирующих привычное поведение.

Тайпсквоттинг в экосистемах с открытым кодом (npm, PyPI) сочетает уязвимость человека и системы: разработчик загружает вредоносный пакет вследствие опечатки в названии [18]. В первом полугодии 2025 года зафиксировано более 4200 инцидентов, что на 73% превышает показатель аналогичного периода 2024 года [6, 7]. Это демонстрирует многослойность атаки, затрагивающую технические и человеческие компоненты.

Таким образом, прогрессивные методы знаменуют качественный сдвиг от массовых операций к целевым атакам с применением ИИ. Их объединяет использование современных технологических трендов и усиление психологического воздействия. Существующие модели, разработанные в эпоху ручных операций, не отражают специфику ИИ-атак (автоматизированная верификация, генерация синтетического контента, адаптивная рефлексия). Это обосновывает необходимость разработки новой концептуальной модели, адекватно описывающей жизненный цикл современных угроз.

Обзор существующих подходов к моделированию цикла атаки

Современные модели цикла атаки, включая четырехэтапную модель лаборатории Касперского [19] и подходы отечественных исследователей [4, 20, 21, 22], созданы по прицепу линейной структуры. Данный подход не отражает адаптивную природу современных ИИ-атак, требующих коррекции стратегии по обратной связи. Классическая модель (исследование, контакты, использование информации), несмотря на возможность итерации шагов, также не учитывает специфику современных угроз [23]. Ключевые недостатки традиционных моделей проявляются в трех аспектах:

1) традиционные модели игнорируют рефлексивный компонент, после завершения операции злоумышленник анализирует эффективность проведенной атаки, используя полученные сведения в дальнейшем (отсутствие этапа рефлексии делает модель неспособной описывать адаптивную природу современных угроз) [22];

2) этап подготовки в существующих моделях чрезмерно является абстрактным и игнорирует верификацию данных[1];

3) классические модели предполагают активный характер инициации контакта [2].

[1] В эпоху генеративного ИИ качество входных данных определяет убедительность контента. Современная подготовка включает сбор данных и их автоматизированную верификацию. Как отмечает А.О. Хлобыстова, этап «анализ и верификация» является ключевым [23]. Традиционные модели не отражают технологической специфики ИИ-опосредованных операций.

[2] Современные атаки часто реализуются через пассивные «ловушки», где жертва сама инициирует компрометацию. Авторами [23], [24] предлагается объединить этапы контактов в «активация общения», где происходит побуждение лица к разглашению информации.

Кроме того, существующие подходы игнорируют воздействие на подсознание жертвы: действия атакующих направлены на уровень подсознания, которое, в отличие от сознания, не сопротивляется воздействию [24]. Таким образом, модели утратили релевантность в условиях внедрения ИИ. Их неспособность отразить рефлексивность, верификацию данных и пассивные векторы обосновывает необходимость разработки новой концептуальной модели жизненного цикла ИИ-опосредованных социоинженерных атак.

Модель цикла атаки на основе социальной инженерии

Авторами представлена шестиэтапная концептуальная модель жизненного цикла социоинженерной атаки. Суть данное разработки в отличие от линейных подходов: модель представляет атаку как циклический, итеративный процесс, отражающий адаптивную природу злоумышленника. Модель включает следующие этапы: формулировка цели, рекогносцировка, подготовка атаки, инициация контакта, эксплуатация и завершение атаки (рисунок 1).

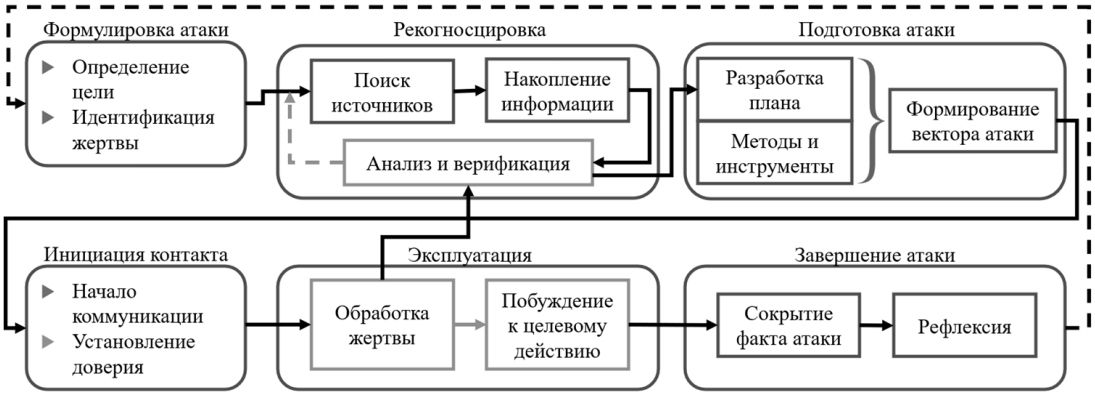

Рис. 1. Концептуальная модель жизненного цикла социоинженерной атаки

В рамках предлагаемой модели атака представляется как циклический процесс, основные этапы представлены в табл. 1.

Таблица 1.

Описание этапов социоинженерной атаки

Этап	Описание	Формулы для описания
1 – формулировка цели	На данном этапе злоумышленник определяет общую цель атаки (например, кража данных, финансовая выгода, репутационный ущерб) и идентифицирует потенциальную жертву или группу жертв. Здесь же формируется первичный «портрет цели» на основе доступных данных, в дальнейшем данные «жертвы» ранжируются по убыванию полезности для выбора оптимальной цели.	Процесс формализуется через расчет функции полезности атакующего – **формула 9**. Формирование первичного портрета – **формула 10**.
2 – рекогносцировка	Классический этап сбора данных, который, однако, в современных условиях часто автоматизирован и/или ИИ-адаптирован. Он включает поиск всевозможных источников информации (социальные сети, корпоративные сайты, банки утечек данных) и агрегацию информации о жертве, такой как ее интересы, связи и поведенческие паттерны. При необходимости злоумышленник может дуплицировать все значимые действия этого этапа (поиск, накопление, анализ данных), поскольку именно он является основой для последующей персонализации атаки. Если качество данных ниже порогового значения, счетчик увеличивается и сбор повторяется. После достижения порога качества проводится верификация атрибутов через байесовскую оценку по формуле (14). Если достоверность ниже порога, алгоритм возвращается к повторному сбору данных, обеспечивая высокую надежность входной информации.	Подбор данных рассмотрен как итерационный цикл: инициализируется счетчик, выполняется сбор данных, **формула – 11**. Расчет метрики качества **формула – 12**.

3 – подготовка атаки	Этап является ключевым для современных ИИ-опосредованных атак и детализирует классический этап подготовки. Он включает два подпроцесса: - верификация данных: автоматизированная проверка достоверности и согласованности собранных на этапе рекогносцировки данных. Это критически важно для генеративных моделей, которые требуют качественной входной информации для создания убедительного контента; - разработка плана: на основе верифицированных данных формируется вектор атаки (выбор канала: почта, мессенджер), разрабатывается детальный план (сценарий коммуникации, психологические триггеры) и подбираются инструменты (фишинговые шаблоны, deepfake-генераторы, вредоносные скрипты).	Выбор оптимального вектора атаки осуществляется путем многокритериальной оптимизации **формула – 16**. Расчет вероятности успеха – **формула 17**. Генерация сценария коммуникации опирается на модель социального влияния – **формула 19**.
4 – инициация контакта	На данном этапе злоумышленник создает условия для первого взаимодействия с жертвой. Цель – установить доверие за счет использования персональных типированных сведений, собранных на предыдущих этапах. Этот этап предполагает реализацию злоумышленником конкретного метода – отправку фишингового письма, инициацию видеозвонка с поддельным «руководителем» или размещение QR-кода на физическом носителе.	Уровень доверия формализуется через показатель, рассчитываемый – формула 23. Данные моделируются под марковский процесс принятия решений (MDP) – формула 20; 21. *Если данные ниже порогового значения, осуществляется коррекция сценария и попытка повторяется, что позволяет адаптивно подстраиваться под реакцию жертвы.*
5 – эксплуатация	Это этап получения желаемого результата от жертвы. Здесь могут применяться уже тонкие психологические приемы и воздействие на стандартные медиаторы (страх, срочность, любопытство, жадность) для побуждения жертвы к целевому действию: вводу данных, переводу средств, установке ПО или выполнению вредоносного кода. Именно здесь проявляется уязвимость человеческого фактора как основного источника рисков.	Вероятность целевого действия вычисляется по формуле 24. Успешность эксплуатации оценивается бинарной функцией – формуле 25. В случае неудачи применяется адаптивная коррекция по принципу ε-greedy (предполагается смена тактики или исследование альтернативных стратегий) –формула 26.
6 – завершение атаки	Этап включает сокрытие следов атаки (удаление логов, смена инфраструктуры) и, что наиболее важно, рефлексию. Злоумышленник анализирует эффективность использованных тактик, документирует успешные приемы и собирает обратную связь для работы над ошибками. Полученные сведения дорабатываются и используется в дальнейших схемах. Алгоритм проверяет условие сходимости: если оно не выполнено и доступны новые цели, осуществляется возврат к ***Этапу 1*** с обновленными параметрами, замыкая цикл атаки.	Реализуется через обновление параметров модели злоумышленника – формула 27. Расчет функции потерь – формула 28. Вероятности успешного завершения атаки – формула 29. *Финальный вывод алгоритма включает вектор успешности, обновленные параметры модели, стационарные вероятности и общий риск.*

На рисунке 2 представлена блок-схема модели цикла социоинженерной атаки. В отличие от традиционных линейных схем, модель учитывает динамические аспекты взаимодействия злоумышленника и жертвы.

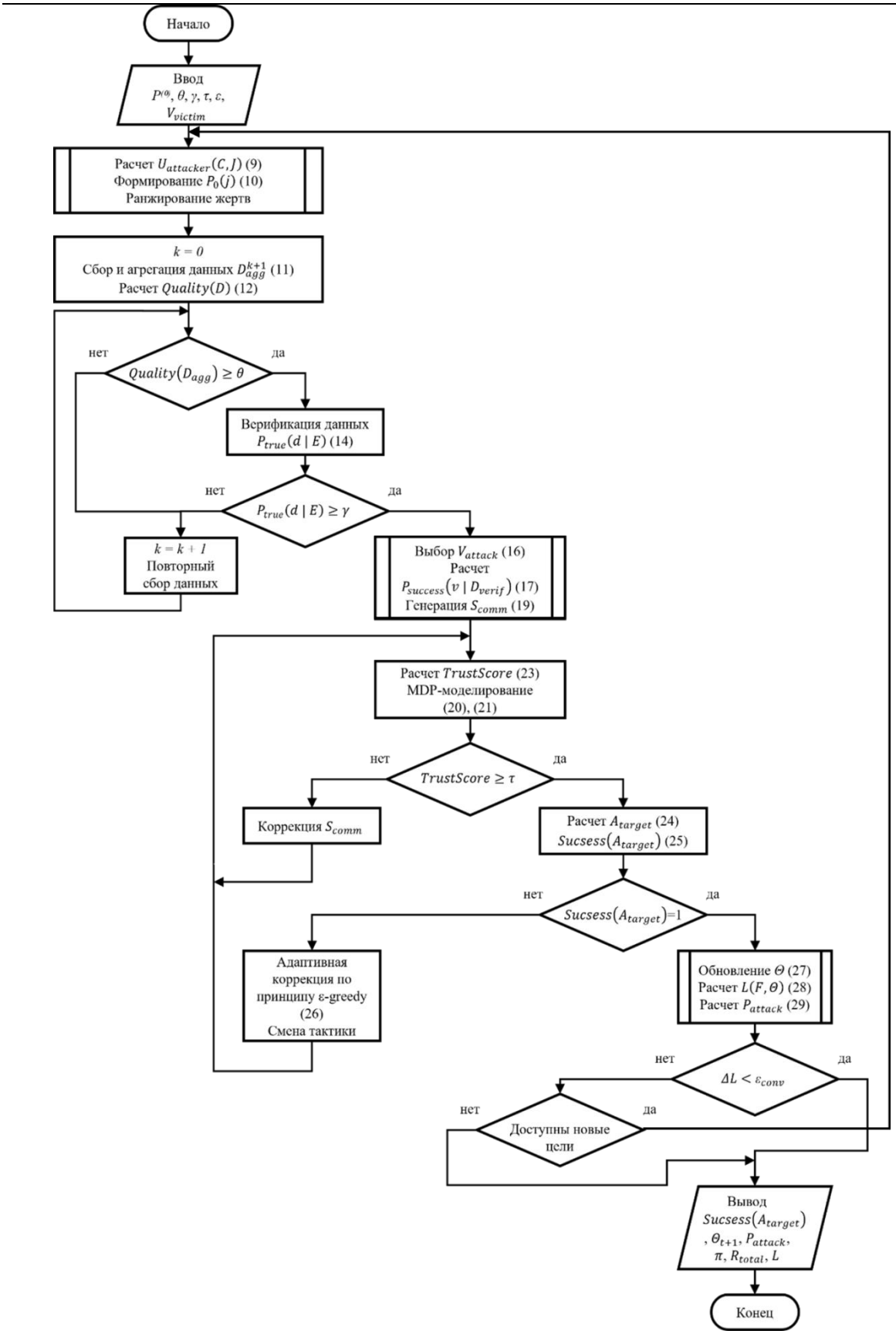

Рис. 2. Блок-схема модели цикла социоинженерной атаки

Ключевые преимущества модели заключаются во включении рефлексивного компонента (система обучения злоумышленника), выделении верификации данных (критично для ИИ-генерации) и признании итеративной природы атаки как цикла коррекции действий. В отличие от классических линейных построений, предложенная модель формализует процесс атаки как ориентированный граф с вероятностными переходами, что позволяет количественно оценивать риски на каждом этапе и прогнозировать развитие инцидента. Модель также

учитывает подсознательные аспекты манипуляции, что делает социальную инженерию особенно опасной [21] и позволяет описывать техники, эксплуатирующие бессознательные реакции. Интеграция виктимной модели жертвы, основанной на поколенческих уязвимостях и психологических профи-

Математическое обоснование разрабатываемой модели

В отличие от линейных описаний, предлагаемая модель трактует атаку как циклический итеративный процесс, что требует применения методов теории вероятностей, теории марковских процессов и моделей оценки рисков. Обоснование структуры модели базируется на представлении атаки как ориентированного графа, где вершинами выступают этапы взаимодействия злоумышленника с жертвой, а дугами – вероятностные переходы между ними. Такой подход позволяет учесть адаптивную природу современных угроз, когда злоумышленник корректирует свои действия на основе обратной связи, полученной в ходе реализации атаки [25].

В основе формализации жизненного цикла атаки лежит базовое представление риска как функции от вероятности реализации угрозы и величины потенциального ущерба:

$$R = P \cdot L, \quad (1)$$

где R – базовый риск социоинженерной атаки;

P – вероятность реализации угрозы;

L – величина потенциального ущерба (в условных единицах).

Данная формула является фундаментальной для всех последующих вычислений и согласуется с общепринятыми подходами к оценке рисков информационной безопасности [26].

Для обоснования уязвимости человеческого фактора на этапах рекогносцировки и эксплуатации необходимо построить формальную модель жертвы. В основе виктимной модели лежит интегральная схема социального влияния, предложенная Т.В. Тулупьевой в [27], где злоумышленник выступает агентом влияния, а пользователь – реципиентом.

Виктимная модель жертвы V_{victim} представляет собой входные параметры для всего алгоритма, включающий вектор психологических параметров жертвы (доверчивость, склонность к риску, когнитивная нагрузка, импульсивность), вектор поведенческих характеристик (цифровая грамотность, осведомленность о безопасности, история соблюдения правил), вектор технического контекста (защищенность устройства, методы аутентификации, сетевая экспозиция) и вектор ситуационных факторов (временное давление, социальное доказательство, сигналы авторитета) [27]. При этом функция интегральной уязвимости вычисляется как взвешенная линейная комбинация нормализованных параметров:

$$U(V_{victim}) = \sum_{i=1}^{K} \lambda_i \cdot \widetilde{x_i}, \quad (2)$$

где K – общее количество параметров виктимной модели;

λ_i – весовой коэффициент i-го параметра ($\sum \lambda_i = 1$);

$\widetilde{x_i}$ – нормализованное значение i-го параметра.

Нормализация параметров осуществляется по формуле:

$$\widetilde{x_i} = \frac{x_i - x_{min}}{x_{max} - x_{min}}, \quad (3)$$

где x_i – актуальное значение i-го параметра;

x_{min} – минимальное возможное значение параметра;

x_{max} – максимальное возможное значение параметра.

Для учета нелинейных взаимодействий параметров применяется аппарат нечеткой логики [28]. Функция принадлежности определяет степень соответствия конкретного значения лингвистическому термину (например, «высокий уровень угрозы»). Выходное значение модели нечеткой логики вычисляется по формуле:

$$U_{fuzzy} = \frac{\sum_{i=1}^{n} w_i \cdot c_i}{\sum_{i=1}^{n} w_i}, \quad (4)$$

где U_{fuzzy} – итоговое значение уязвимости после дефаззификации (метод центра тяжести)[3];

w_i – весовой коэффициент i-го правила нечеткого вывода (степень активации правила);

c_i – центр функции принадлежности i-го правила;

n – общее количество правил нечеткой базы.

Данная модель согласуется с исследованиями Г.А. Остапенко и других авторов [29], которые выделяют особенности уязвимостей. Вероятность успешной атаки с учетом поколенческой специфики оценивается по формуле:

$$P_v^{gen} = 1 - \left(1 - \frac{1}{n}\right)^m, \quad (5)$$

где P_v^{gen} – вероятность реализации атаки на поколение gen с учетом уязвимостей социального характера;

m – количество уязвимостей для текущей возрастной категории;

n – общее количество уязвимостей.

выражение $\left(1 - \frac{1}{n}\right)$ представляет вероятность того, что одна конкретная уязвимость не приведет к успешной атаке. Возведение в степень m дает вероятность того, что ни одна из m уязвимостей не сработает. Вычитание из единицы дает искомую вероятность успеха атаки [29].

[3] формула (4) реализует метод центра тяжести для дефаззификации, где числитель представляет собой взвешенную сумму центров функций принадлежности, а знаменатель – сумму весов всех активированных правил. Данный подход позволяет преобразовать нечеткие выводы в конкретное числовое значение уязвимости.

Жизненный цикл социоинженерной атаки представляется как ориентированный граф:

$$G = (V, E, P, W), \quad (6)$$

где $V = (v_1, v_2, \dots, v_6)$ – множество вершин графа, соответствующих шести этапам атаки;

$E \subseteq \mathrm{V} \times \mathrm{V}$ – множество дуг, представляющих возможные переходы между этапами;

$P = (p_{ij})$ – матрица вероятностей переходов размером 6×6, где $p_{ij} \in [0, 1]$ – вероятность перехода с этапа i на этап j;

$W = (w_1, w_2, \dots, w_6)$ – вектор весов вершин, характеризующих временные и ресурсные затраты на выполнение этапа i.

Для учета динамической важности отдельных этапов атаки и их взаимосвязей вводится коэффициент важности этапа, аналогичный подходу, предложенному Коцыняком и коллегами для таргетированных кибератак [30]:

$$A_i(t) = A_i^0 \cdot exp\left(\frac{t}{\tau_{\text{возд}}}\right), \quad (7)$$

где $A_i(t)$ – динамический коэффициент важности i-го этапа атаки;

A_i^0 – базовая важность этапа (начальное значение);

t – текущее время реализации атаки;

$\tau_{\text{возд}}$ – среднее значение цикла воздействия (временной горизонт планирования атаки).

Данный коэффициент позволяет учесть, что важность отдельных этапов может изменяться в процессе реализации атаки, что особенно критично для адаптивных социоинженерных операций.

Также вводится коэффициент связности между этапами, характеризующий функциональную взаимосвязь последовательных стадий атаки:

$$\alpha_{ij} = \frac{\lambda_{ij}}{\sum_k \lambda_{ikr}}, \quad (8)$$

где α_{ij} – коэффициент связности между этапами i и j $(0 < \alpha_{ij} \leq 1)$;

λ_{ij} – вес (интенсивность) перехода от этапа i к этапу j в общем потоке атаки;

$\sum_k \lambda_{ikr}$ – суммарная интенсивность всех переходов из этапа i (сумма по всем возможным целевым этапам k).

Использование данных коэффициентов позволяет более точно моделировать распределение ресурсов злоумышленника по этапам атаки и прогнозировать наиболее критические точки для вмешательства систем защиты [30].

Такая формализация позволяет трактовать атаку как стохастический процесс с памятью, где каждое состояние содержит достаточную информацию для определения вероятностей будущих переходов [28]. На стратегическом уровне злоумышленник определяет цель атаки и идентифицирует жертву, что формализуется через функцию полезности атакующего:

$$U_{attacker}(C, J) = \sum_{i=1}^{n} w_i \cdot E[R_i(C_i, J_i)] - \lambda \cdot Cost(C, J), \quad (9)$$

где $U_{attacker}(C, J)$ – функция полезности атакующего;

$C \in \{C_{data}, C_{fin}, C_{rep}\}$ – множество стратегических целей (кража данных, финансовая выгода, репутационный ущерб);

$J \in \{j_1, j_2, \dots, j_n\}$ – множество потенциальных жертв;

R_i – случайная величина выгоды от атаки на жертву j_i с целью C;

w_i – весовой коэффициент значимости жертвы;

λ – коэффициент чувствительности к затратам;

$Cost(C, J)$ – функция ресурсных затрат на реализацию атаки.

Первичный портрет цели формируется как отображение доступных данных в пространство признаков:

$$P_0(j) = \Phi(D_{avail}) = (\vec{x}_{dem}, \vec{x}_{behav}, \vec{x}_{tech}, \vec{x}_{psych}), \quad (10)$$

где $P_0(j)$ – первичный портрет цели j;

D_{avail} – множество доступных данных о жертве;

Φ – функция извлечения признаков;

$\vec{x}_{dem}, \vec{x}_{behav}, \vec{x}_{tech}, \vec{x}_{psych}$ – векторы демографических, поведенческих, технических и психологических характеристик соответственно.

Процесс сбора и агрегации данных на этапе рекогносцировки описывается стохастической моделью с итеративной коррекцией качества:

$$D_{agg}^{k+1} = \alpha \cdot D_{agg}^{k} + (1 - \alpha) \cdot F_{new}(SRC, J), \quad (11)$$

где D_{agg}^{k} – агрегированный набор данных на итерации k;

$\alpha \in [0, 1]$ – коэффициент инерции накопления знаний;

F_{new} – функция извлечения новых данных из источников SRC о жертве J.

Критерий качества данных определяется через метрику полноты и достоверности:

$$Quality(D) = \beta_1 \cdot Completeness(D) + \beta_2 \cdot Consistency(D) + \beta_3 \cdot Freshness(D), \quad (12)$$

где $Quality(D)$ – интегральный показатель качества данных (от 0 до 1);

$Completeness(D)$ – доля заполненных атрибутов в наборе данных D;

$Consistency(D)$ – доля непротиворечивых записей;

$Freshness(D)$ – актуальность данных (обратная пропорциональность времени сбора);

$\beta_1, \beta_2, \beta_3$ – весовые коэффициенты (их сумма равна 1).

Итерационный цикл этапа 2 продолжается до достижения порогового значения θ:

$$Quality(D_{agg}) \geq \theta, \quad (13)$$

где θ – минимально допустимый уровень качества данных для перехода к следующему этапу.

Для верификации данных применяется байесовская оценка достоверности каждого атрибута:

$$P_{true}(d \mid E) = \frac{P(E \mid d) \cdot P_{prior}(d)}{\sum P_{prior}(d')}, \quad (14)$$

где $P_{true}(d \mid E)$ – апостериорная вероятность истинности атрибута d при наличии свидетельств E;

$P(E \mid d)$ – вероятность наблюдения свидетельств E при истинности d;

$P_{prior}(d)$ – априорная вероятность истинности атрибута d.

Атрибут включается в верифицированный набор при условии:

$$P_{true}(d \mid E) \geq \gamma, \quad (15)$$

где γ – порог доверия, устанавливаемый в зависимости от критичности атаки.

Выбор вектора атаки на этапе подготовки формализуется как задача многокритериальной оптимизации:

$$V_{attack} = max_{v \in V}\left[P_{success}\left(v \mid D_{verif}\right) \cdot impact(v) - Cost(v)\right], \quad (16)$$

где V_{attack} – оптимальный вектор атаки;

V – множество допустимых векторов (email, messenger, voice, physical);

$P_{success}\left(v \mid D_{verif}\right)$ – вероятность успеха вектора v при верифицированных данных D_{verif};

$impact(v)$ – ожидаемый ущерб от применения вектора v;

$Cost(v)$ – ресурсные затраты на реализацию.

Вероятность успеха моделируется на основе уязвимости жертвы и эффективности техники [26]:

$$P_{success}\left(v \mid D_{verif}\right) = max_{j \in J}\left[V_j \cdot T_v \cdot \left(1 - D_{org}\right)\right], \quad (17)$$

где V_j – интегральная уязвимость жертвы j, вычисляемая по формуле (3);

T_v – эмпирическая эффективность техники v (отношение успешных атак к общему числу попыток);

D_{org} – интегральный уровень организационной защиты.

Уровень организационной защиты рассчитывается как:

$$D_{org} = \sum \theta_I \cdot M_I, \quad (18)$$

где M_I – эффективность l-й меры защиты (0–1);

θ_I – весовой коэффициент l-й меры (сумма всех весов равна 1).

Генерация сценария коммуникации опирается на модель социального влияния [27]:

$$S_{comm} = max_{s \in S} \sum \omega_m \cdot E[Response(s, m, P_{victim})], \quad (19)$$

где S_{comm} – оптимальный сценарий коммуникации;

S – множество возможных сценариев;

$M = \{M_{fear}, M_{urgency}, M_{curiosity}, M_{greed}, \dots\}$ – множество психологических медиаторов;

ω_m – весовой коэффициент медиатора m;

$Response(s, m, P_{victim})$ – функция отклика жертвы на сценарий s с медиатором m при профиле P_{victim}.

Процесс установления доверия на этапе инициации контакта моделируется как марковский процесс принятия решений (MDP) [25]:

$$M = (S, A, R, P), \quad (20)$$

где $S = \{Neutral, Trusted, Challenged, Blocked\}$ – пространство состояний взаимодействия;

$A = \{Cooperate, Deceive, Reset\}$ – множество действий злоумышленника;

$P(s' \mid s, a)$ – вероятность перехода из состояния s в s' при действии a;

$R(s, a, s')$ – функция немедленного вознаграждения.

Оптимальная стратегия определяется через уравнение Беллмана:

$$V(s) = max_{a \in A}[\sum R(s, a, s') + \gamma \cdot V(s')], \quad (21)$$

где$V(s)$ – оптимальная функция ценности состояния s;

$\gamma \in [0, 1]$ – коэффициент дисконтирования будущих вознаграждений.

Переход к этапу эксплуатации осуществляется при достижении состояния $Trusted$ с вероятностью, превышающей порог τ:

$$P(TrustScore \geq \tau \mid P_0, S_{comm}, Context) \geq \tau, \quad (22)$$

где $TrustScore$ вычисляется как:

$$TrustScore = \delta \cdot Authority(M) + \varepsilon \cdot Urgency(M) + \zeta \cdot Personalization(V), \quad (23)$$

где $Authority(M)$ – сила авторитета сообщения $M \epsilon (0; 1)$;

$Urgency(M)$ – уровень срочности сообщения $M \epsilon (0; 1)$;

$Personalization(V)$ – степень персонализации по данным жертвы $V \epsilon (0; 1)$;

$\delta, \varepsilon, \zeta$ – весовые коэффициенты (сумма равна 1).

Действие жертвы на этапе эксплуатации моделируется как функция от психологического состояния, внешнего воздействия и контекстных факторов:

$$A_{target} = \sigma(\sum \alpha_m \cdot M_m \cdot S(P_{victim}, m) + \varepsilon), \quad (24)$$

где A_{target} – целевое действие жертвы;

$\sigma(\cdot)$ – сигмоидальная функция активации (вероятность выполнения действия);

α_m – интенсивность применения медиатора m;

$S(P_{victim}, m)$ – функция восприимчивости жертвы к медиатору m;

$\varepsilon \sim N(0, \sigma^2)$ – стохастическая компонента, отражающая неопределенность человеческого поведения.

Успешность эксплуатации оценивается через бинарную функцию:

$$Sucsess\left(A_{target}\right) = I\left[U_{attacker}\left(A_{target}\right) > T\right], \quad (25)$$

где $Sucsess\left(A_{target}\right)$ – результат эксплуатации (1 – успех, 0 – неудача);

$U_{attacker}\left(A_{target}\right)$ – полезность действия для атакующего;

T – пороговое значение полезности для считания атаки успешной;

$I(\cdot)$ – индикаторная функция (равна 1 при истинности условия, иначе 0).

В случае неудачи система осуществляет адаптивную коррекцию по принципу ε-greedy:

$$P(explore) = \varepsilon, P(exploit) = 1 - \varepsilon, \quad (26)$$

где $P(explore)$ – вероятность исследования альтернативной стратегии;

$P(exploit)$ – вероятность эксплуатации текущей оптимальной стратегии;

ε – параметр исследования $(0 < \varepsilon < 1)$.

Рефлексивный компонент модели на этапе завершения атаки реализуется через механизм обновления базы знаний злоумышленника:

$$\Theta_{t+1} = \Theta_t + \eta \cdot grad\big(L(F_t, \Theta_t)\big), \quad (27)$$

где Θ_{t+1} – обновленные параметры модели злоумышленника на шаге t+1;

Θ_t – параметры модели злоумышленника на шаге t (стратегии, шаблоны, оценки уязвимостей);

η – скорость обучения$(0 < \eta < 1)$;

L – функция потерь на шаге t;

F_t – функция потерь на шаге t.

Функция потерь определяется как:

$$L(F, \Theta) = (E_{Outcome} - A_{Outcome})^2, \quad (28)$$

где $L(F, \Theta)$ – значение функции потерь;

$E_{Outcome}$ – прогнозируемый результат атаки;

$A_{Outcome}$ – фактический результат атаки.

Вероятность успешного завершения атаки с учетом временных характеристик вычисляется по полумарковской модели [28]:

$$P_{attack} = 1 - \exp(-\Lambda \cdot T), \quad (29)$$

где P_{attack} – вероятность успешной атаки за время T;

Λ – матрица интенсивностей переходов между этапами;

T – общее время реализации атаки.

Матрица интенсивностей переходов имеет вид:

$$\Lambda = \begin{vmatrix} -\lambda_1 & 0 & \lambda_3 Q_{13} \\ \lambda_1 Q_{21} & -\lambda_2 & 0 \\ \lambda_1 Q_{31} & \lambda_2 Q_{32} & -\lambda_3 \end{vmatrix}, \quad (30)$$

где λ_i – интенсивность завершения этапа i;

Q_{ij} – вероятность выбора перехода из этапа i в этап j.

В отличие от линейных моделей, где процесс завершается на этапе получения результата, предложенная модель предполагает, что злоумышленник анализирует итоги успешной атаки (рефлексия) и использует полученную информацию для формулировки целей следующей атаки. Согласно теории вероятностей, вероятность совместного осуществления цепи зависимых событий равна произведению вероятностей перехода между каждым соседним состоянием [25, 28]. Соответственно, математически это представляется как цепь последовательных событий, где переход к следующему этапу возможен только при успешном завершении предыдущего:

$$P_{cycle} = p_{61} \cdot p_{12} \cdot p_{23} \cdot p_{34} \cdot p_{45} \cdot p_{56}, \quad (31)$$

где P_{cycle} – вероятность полного цикла атаки с последующей итерацией.

Формула (31) представляет полный цикл социоинженерной атаки. Индексы в обозначении соответствуют номерам этапов модели (1-6). Множитель обозначает вероятность успешного перехода с этапа на этап.

Ключевым параметром, который обеспечивает цикличность модели, выступает коэффициент p_{61}, обозначающий вероятность возврата от этапа завершения атаки к этапу формулировки цели новой итерации; он отражает решение злоумышленника не завершать деятельность окончательно, а на основе рефлексии и анализа успеха или неудачи сформулировать новую цель для повторной атаки, что принципиально отличает предложенную модель от классических линейных схем.

Стационарное распределение вероятностей нахождения системы в каждом состоянии вычисляется как решение системы уравнений:

$$\rho = \pi \cdot P, \quad (32)$$

где $\pi = (\pi_1, \pi_2, \dots, \pi_6)$ – вектор стационарных вероятностей;

P – матрица переходных вероятностей размером 6×6.

При условии нормировки:

$$\sum_{i=1}^{6} \pi_i = 1, \quad (33)$$

где π_i – стационарная вероятность нахождения на этапе i.

Данная формализация подтверждает итеративную природу модели, где злоумышленник может многократно проходить цикл с корректировкой стратегии на основе накопленного опыта. Для оценки рисков на каждом этапе применяется базовая формула анализа рисков информационной безопасности:

$$R = P_{threat} \cdot P_{vulnerability} \cdot Impact, \quad (34)$$

где R – риск (в условных единицах ущерба);

P_{threat} – вероятность возникновения угрозы;

$P_{vulnerability}$ – вероятность использования уязвимости;

$Impact$ – величина потенциального ущерба.

Формула может быть детализирована с учетом эффективности защитных мер:

$$R = P_{threat} \cdot P_{vulnerability} - \big(1 - E_{prevent}\big) \cdot Impact \cdot \big(1 - E_{mitigate}\big), \quad (35)$$

где $E_{prevent}$ – эффективность мер, предотвращающих уязвимость;

$E_{mitigate}$ – эффективность мер, снижающих последствия.

Общий риск для всей атаки вычисляется как сумма рисков по всем этапам:

$$R_{total} = \sum_{i=1}^{6} R_i, \quad (36)$$

где R_i – риск на каждом i-ом этапе атаки.

На основании приведенных математических формулировок формируется пошаговый алгоритм. На этапе инициализации параметров модели задаются начальные значения матрицы переходных вероятностей P, устанавливаются пороги θ, γ, τ для критериев качества, доверия и перехода, инициализируется виктимная модель жертвы V_{victim}. При формулировке цели вычисляется функция полезности $U_{attacker}(C, J)$ по формуле (9), формируется первичный портрет цели $P_0(j)$ по формуле (10), жертвы ранжируются по убыванию полезности. На этапе рекогносцировки инициализируется счетчик

итераций k = 0, выполняется сбор данных по формуле (11), вычисляется качество данных $Quality(D)$ по формуле (12). Если $Quality \geq \theta$, k увеличивается на 1 и повторяется сбор, иначе переходится к верификации.

При подготовке атаки верифицируются данные по условию (14). Если доля верифицированных данных меньше ρ, осуществляется возврат к этапу рекогносцировки, иначе вычисляется оптимальный вектор атаки по формуле (16) и генерируется сценарий коммуникации по формуле (19). На этапе инициации контакта вычисляется $TrustScore$ по формуле (23). Если $TrustScore < \tau$, адаптируется сценарий и повторяется попытка, иначе переходится к этапу эксплуатации. При эксплуатации вычисляется вероятность целевого действия по формуле (24), оценивается успешность по формуле (25). При неудаче применяется адаптивная коррекция по формуле (26).

На этапе завершения атаки обновляются параметры модели злоумышленника по формулам (27) и (28), вычисляется вероятность успешного завершения по формуле (29). Если доступны новые цели, осуществляется возврат к этапу формулировки цели с обновленными параметрами Θ_{t+1}. Проверка условия сходимости осуществляется вычислением изменения функции потерь $\Delta L = |L_t - L_{t-1}|$. Если $\Delta L < \varepsilon$, алгоритм завершен, иначе увеличивается счетчик итераций и осуществляется переход к этапу формулировки цели. Условие сходимости алгоритма:

$$\lim_{t\to\infty} \Delta L = 0, \qquad (37)$$

что гарантирует стабилизацию модели злоумышленника и достижение оптимальной стратегии атаки.

Таким образом, нами представлено математическое обоснование шестиэтапной концептуальной модели жизненного цикла социоинженерной атаки. Атака представлена как ориентированный граф с вероятностными переходами, что позволяет описать адаптивную природу современных угроз. Циклическая природа модели доказана через стационарное распределение вероятностей марковской цепи, где $p_{61} > 0$ обеспечивает возможность многократных итераций. Разработан пошаговый алгоритм с условием сходимости, гарантирующим стабилизацию модели злоумышленника и достижение оптимальной стратегии атаки.

Заключение

Проведенное исследование подтвердило, что в условиях глобальной цифровизации и геополитической нестабильности социальная инженерия остается наиболее критическим вектором киберугроз. Анализ эмпирических данных за первое полугодие 2025 года свидетельствует о доминировании методов манипуляции человеческим фактором: 97% таргетированных атак на организации и 93% атак на частных лиц реализуются посредством методов социальной инженерии. Ключевым трендом эволюции угроз стал переход к технологиям искусственного интеллекта (Deepfake, LLM-фишинг, автоматизированная верификация данных и другие).

Установлено, что существующие модели проведения атак утратили релевантность, поскольку не учитывают «адаптивную природу» современных злоумышленников, игнорируют этап автоматизированной верификации данных и не отражают комплексной взаимосвязи (обучение атакующего на основе обратной связи). В ответ на выявленные ограничения в работе разработана и обоснована шестиэтапная концептуальная модель цикла социоинженерной атаки. Модель прошла процедуру математической формализации с использованием аппарата теории вероятностей, марковских процессов и нечеткой логики. На основе математического обоснования построена блок-схема алгоритма реализации модели, отражающая все логические переходы, условия сходимости и итерационные циклы.

Практическая значимость работы заключается в возможности использования предложенной модели в качестве новых теоретических решений для разработки систем защиты информации. Ключевое преимущество таких систем – способность идентифицировать критические точки разрыва в цикле атаки, в том числе на ранних стадиях. Реализация предложенной методики точечного противодействия на уровне отдельных субъектов или организаций позволяет трансформировать техническую защиту (киберзащиту): устранение последствий инцидентов к упреждающему прерыванию атак на каждой из фаз жизненного цикла. Кроме того, это позволит актуализировать обучающие программы для пользователей с учётом психологических аспектов принятия решений.

Список литературы:

1. Мельников А. И. Социальная инженерия в цифровой эпохе: анализ методов манипуляции человеческим фактором в целях кибератак // Социально-гуманитарные знания. – 2024. – № 1. – С. 50–54.

2. Наумова К. Д. Исследование основных методов противодействия атакам, основанным на методах социальной инженерии, на предмет их эффективности и применимости к современной ситуации в РФ / К. Д. Наумова, В. Ю. Радыгин // Инновационные механизмы управления цифровой и региональной экономикой: Материалы V Международной студенческой научной конференции, Москва, 15–16 июня 2023 года. – Москва: Национальный исследовательский ядерный университет «МИФИ», 2023. – С. 145–158.

3. Турков Б. А. Анализ методов воздействия в социальной инженерии и способы защиты информации / Б. А. Турков // Морская стратегия и политика России в контексте обеспечения национальной безопасности и устойчивого развития в XXI веке: сборник научных трудов. Вып. 5. – Севастополь: Черноморское высшее военно-морское училище имени П. С. Нахимова, 2022. – С. 224–229.

4. Маринов, А. А. Противодействие преступлениям, связанным с нарушением тайны телефонных переговоров / А. А. Маринов, Е. Г. Усов, В. В. Загайнов // Вестник СибГУТИ. – 2021. – № 4(56). – С. 69-75.

5. Маринов, А. А. Социоинженерные атаки в корпоративных информационных системах: перспективы противодействия с использованием средств искусственного интеллекта / А. А. Маринов, В. С. Федоров // Экономический альманах : Материалы IX Всероссийской научно-практической конференции «Экономика инфраструктурных преобразований: проблемы и перспективы развития», Иркутск, 30 ноября 2022 года. – Иркутск: Иркутский национальный исследовательский технический университет, 2023. – С. 643-652.

6. Громов Ю. Ю. Информационная безопасность и защита информации: учебное пособие / Ю. Ю. Громов, В. О. Драчев, О. Г. Иванова. – Старый Оскол: ТНТ, 2018. – 384 с.

7. Компания Positive Technologies. Актуальные киберугрозы: I–II кварталы 2025 года / А. Вяткина, А. Осипова // [Электронный ресурс]; URL: https://ptsecurity.com/research/analytics/aktual-nye-kiberugrozy-i-ii-kvartaly-2025-goda/#id1 (Дата обращения 10.02.2026).

8. Компания Positive Technologies. Актуальные киберугрозы: IV квартал 2024 года – I квартал 2025 года / А. Голушко // [Электронный ресурс]; URL: https://ptsecurity.com/research/analytics/aktualnye-kiberugrozy-iv-kvartal-2024-goda-i-kvartal-2025-goda/#id1 (Дата обращения 10.02.2026).

9. Отчет об угрозах Avast за первый квартал 2024 г. / Компания Avast // [Электронный ресурс]; URL: https://decoded.avast.io/threatresearch/avast-q1-2024-threat-report/ (Дата обращения 10.02.2026).

10. Тренды фишинговых атак на организации в 2022–2023 годах / Компания Positive Technologies // [Электронный ресурс]; URL: https://www.ptsecurity.com/ru-ru/research/analytics/phishing-attacks-on-organizations-in-2022-2023/ (Дата обращения 10.02.2026).

11. Raza A. A Comprehensive Review of Deepfake Detection Techniques: From Traditional Machine Learning to Advanced Deep Learning Architectures / A. Raza, A. Basit, A. Amin, Z. A. Arfeen, M. I. Masud, U. Fayyaz, T. A. Jumani // AI. – 2026. – Т. 7. – С. 68.

12. Singh S. Unmasking digital deceptions: An integrative review of deepfake detection, multimedia forensics, and cybersecurity challenges / S. Singh, A. Dhumane // MethodsX. – 2025. – Т. 15. – С. 103632.

13. Pedersen K. T. Deepfake-Driven Social Engineering: Threats, Detection Techniques, and Defensive Strategies in Corporate Environments / K. T. Pedersen, L. Pepke, T. Stærmose, M. Papaioannou, G. Choudhary, N. Dragoni // Journal of Cybersecurity and Privacy. – 2025. – Т. 5. – С. 18.

14. With generative AI, social engineering gets more dangerous – and harder to spot / IBM // [Электронный ресурс]; URL: https://www.ibm.com/think/insights/generative-ai-social-engineering (Дата обращения 13.02.2026).

15. Jabir R. Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors / R. Jabir, J. Le, C. Nguyen // AI. – 2025. – Т. 6. – С. 174.

16. Toapanta F. AI-Driven Vishing Attacks: A Practical Approach / F. Toapanta, B. Rivadeneira, C. Tipantuña, D. Guamán // Engineering Proceedings. – 2024. – Т. 77. – С. 16. Think before you Click(Fix): Analyzing the ClickFix social engineering technique / Компания Microsoft // [Электронный ресурс]; URL: https://www.microsoft.com/en-us/security/blog/2025/08/21/think-before-you-clickfix-analyzing-the-clickfix-social-engineering-technique/ (Дата обращения 13.02.2026).

17. Top Cyber Industry Predictions and Cybersecurity Trend Reports for 2025 - Part 1 / D. Lohrmann // [Электронный ресурс]; URL: https://www.linkedin.com/pulse/top-cyber-industry-predictions-cybersecurity-trend-reports-lohrmann-70r2e (Дата обращения 13.02.2026).

18. Kumarage T. Personalized Attacks of Social Engineering in Multi-turn Conversations: LLM Agents for Simulation and Detection / T. Kumarage, C. Johnson, J. Adams, L. Ai, M. Kirchner, A. Hoogs, J. Garland, J. Hirschberg, A. Basharat, H. Liu // [Электронный ресурс]; URL: https://arxiv.org/abs/2503.15552 (Дата обращения 13.02.2026).

19. Что такое социальная инженерия в 2021 году / Компания «Лаборатория Касперского» // [Электронный ресурс]; URL: https://www.kaspersky.ru/resource-center/definitions/what-is-social-engineering (Дата обращения 13.02.2026).

20. Степанов М. О. Социальная инженерия: методы атак и способы предотвращения / М. О. Степанов, Н. Х. Нагаев // Актуальные проблемы науки и образования в условиях современных вызовов: сборник материалов XXVIII Международной научно-практической конференции, Москва, 1 марта 2024 г. – Санкт-Петербург: Печатный цех, 2024. – С. 130–136.

21. Марзоев А. А. Социальная инженерия / А. А. Марзоев, В. А. Софронов, Р. С. Щербаков // Повышение обороноспособности государства 2022: сборник научных трудов по материалам заочной научной конференции. – Санкт-Петербург: ООО «Полторак», 2022. – С. 25–31.

22. Исследование механизмов социальной инженерии и анализ методов противодействия / В. Ю. Евглевский, М. М. Путято, А. С. Макарян, И. В. Володин // Научные труды КубГТУ. – 2021. – № 2. – С. 57–68.

23. Хлобыстова А. О. Концептуальная модель цикла социоинженерной атаки: современные подходы и архитектура прототипа программного комплекса // Компьютерные инструменты в образовании. – 2021. – № 3. – С. 17–28.

24. Компания Positive Technologies. Киберугроза №1, или Как бороться с социальной инженерией / А. Мальнев // [Электронный ресурс]; URL: https://www.securitylab.ru/analytics/500877.php (Дата обращения 13.02.2026).

25. Abri F. Markov Decision Process for Modeling Social Engineering Attacks and Finding Optimal Attack Strategies / F. Abri, J. Zheng, A.S. Namin, K.S. Jones // San Jose State University Faculty Research. – 2022. – doi:10.1109/ACCESS.2022.3213711

26. Șandor A. A Mathematical Model for Risk Assessment of Social Engineering Attacks / A. Șandor, G. Tont, E. Simion // TEM Journal. – 2022. – Vol. 11, № 1. – P. 334–338. – doi:10.18421/TEM111-47

27. Тулупьева Т.В. Модель социального влияния в анализе социоинженерных атак / Т.В. Тулупьева, М.В. Абрамов, А.Л. Тулупьев // Вопросы кибербезопасности. – 2024. – № 3. – С. 45–56.

28. Гавдан Г.П. Полумарковская модель кибератак, воздействующих на информационную систему / Г.П. Гавдан, А.А. Голубев, А.Г. Голубев, И.Ю. Жуков, Э.П. Рыбалко // Безопасность информационных технологий. – 2025. – Т. 32, № 3. – С. 26–43. – doi:10.26583/bit.2025.3.03

29. Остапенко Г.А. Риски информационной безопасности при атаках социальной инженерии: способы обнаружения и регистрации / Г.А. Остапенко, А.А. Остапенко, М.А. Грамыкин, М.В. Кондратьев, Н.С. Харламов // Информация и безопасность. – 2025. – Т. 28, № 2. – С. 255–268. – doi:10.36622/1682-7813.2025.28.2.011

30. Коцыняк М.А. Математическая модель таргетированной компьютерной атаки / М.А. Коцыняк, О.С. Лаута, Д.А. Иванов // Вопросы кибербезопасности. – 2025. – № 1. – С. 12–24. – URL: https://cyberrus.info/wp-content/uploads/2025/03/vokib-2025-1-cc.pdf