ОБЕСПЕЧЕНИЕ ДОВЕРИЯ И БЕЗОПАСНОСТИ ПРИ ИСПОЛЬЗОВАНИИ ИКТ

ЛЕВ ЛУНЕГОВ
ведущий технический аналитик
ФГБУ «НИИ «Интеграл»

НИКИТА НЕМЦЕВ
технический аналитик
ФГБУ «НИИ «Интеграл»

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ БОРЬБЫ С ФИШИНГОМ

ВВЕДЕНИЕ

Активная цифровизация всех сфер жизни создала широкие возможности не только для развития, но одновременно и для киберпреступности. Злоумышленники получили доступ к мощным инструментам, позволяющим проводить атаки с высокой эффективностью и в глобальном масштабе. Фишинг[1], как одна из самых распространенных угроз, значительно эволюционировал: современные атаки используют персонализацию, механизмы обхода систем обнаружения и другие методы, способствующие повышению доверия пользователя.

[1] Фишинг – это методика обмана пользователей путем имитации легитимных ресурсов с целью неправомерного получения конфиденциальных сведений для их последующего использования в совершении преступлений.

Традиционные средства защиты, основанные на статических правилах и черных списках, часто не справляются с новыми, быстро адаптирующимися тактиками злоумышленников [1]. В этих условиях искусственный интеллект (ИИ) становится критически важной технологией для противодействия фишингу, алгоритмы машинного обучения способны анализировать огромные объемы данных в реальном времени, выявлять сложные схемы мошенничества и адаптироваться к новым угрозам.

ПОЧЕМУ ЭФФЕКТИВНОСТЬ ТРАДИЦИОННЫХ МЕТОДОВ СНИЖАЕТСЯ

Несмотря на долгую историю борьбы, классические подходы к выявлению фишинговых сайтов все чаще оказываются неэффективными перед лицом современных тактик мошенников. Ключевая слабость классических подходов – статичность и реактивность в условиях динамично меняющейся угрозы [2]. Сигнатурный анализ, основанный на сравнении кода страниц или их фрагментов с базами известных вредоносных шаблонов, бесполезен против постоянно обновляемых или совершенно новых фишинговых ресурсов. Злоумышленники легко обходят его, применяя простую обфускацию кода[2] [3]: меняют структуру HTML, добавляют случайные комментарии или пробелы, используют разные скрипты для достижения одной цели, чтобы код страницы не совпал с известной сигнатурой. Даже если сигнатура создана для нового шаблона, всегда существует критическая задержка между обнаружением угрозы, созданием сигнатуры и ее обновлением на всех защищенных устройствах пользователей, которой успешно пользуются фишеры.

Черные списки доменов и URL – еще один столп традиционной защиты – имеют критический недостаток, они всегда отстают. Фишинговые кампании сегодня используют одноразовые сайты: массово создаются новые домены или используются дешевые хостинги и легитимные, но скомпрометированные веб-ресурсы для размещения фишинговых страниц. Эти страницы существуют считанные часы или дни, собирают данные жертв и исчезают задолго до того, как их URL-адрес или домен успеют внести в блоклисты [4]. Злоумышленники мастерски используют визуально похожие домены, заменяя буквы похожими символами или применяя символы идентичные латинским, которые на первый взгляд выглядят легитимно и не вызывают подозрений у систем, опирающихся только на списки запрещенных адресов [5]. Широко распространена маскировка конечного вредоносного URL-адреса через сервисы сокращения ссылок или сложные цепочки перенаправлений через легитимные, но уязвимые, или нейтральные промежуточные сайты. Пользователь видит в адресной строке или при наведении курсора безопасный короткий URL, который в итоге ведет на фишинговую страницу, и черные списки бессильны против такой маскировки.

Статические эвристические правила, пытающиеся искать на страницах подозрительные ключевые элементы, легко обходятся. Фишеры заменяют эти слова синонимами, используют изображения с текстом вместо самого текста, или просто копируют дизайн и стиль легитимных сайтов до мельчайших деталей, избегая явных ключевых фраз. Более сложные правила, применяемые при анализе структуры страницы, также ненадежны: легитимные формы ввода данных выглядят аналогично, а фишинговые страницы могут искусно маскироваться под них. Главная же проблема статических правил – их неспособность оценить контекст и правдоподобность сайта в целом. Они не могут понять, что сайт идеально копирующий дизайн банка, но расположенный на странном домене или без валидного SSL-сертификата известного центра сертификации, является подделкой [6]. Они не улавливают тонкие несоответствия в логотипах, шрифтах или общем ощущении подлинности, которые может заметить человек или более интеллектуальная система.

Ручной анализ и верификация сайтов специалистами безопасности не могут масштабироваться под поток новых подозрительных URL-адресов. Физически невозможно вручную в режиме реального времени проверить каждый потенциально опасный ресурс в сети Интернет или каждый сайт, на который пытается перейти пользователь. Время реакции критически велико, а нехватка экспертов делает этот подход непрактичным для массовой защиты. В итоге, все эти методы уязвимы из-за своей фундаментальной реактивности и статичности. Эти методы настроены искать некачественно сделанные мошеннические сайты или очень явные признаки подделки, но совершенно не способны проактивно выявлять новые, ранее неизвестные фишинговые ресурсы, особенно те, что искусно маскируются под легитимные. Они не могут анализировать комплексные факторы: репутацию домена в реальном времени, валидность и цепочку SSL-сертификатов, историю создания домена, визуальное сходство с брендом, поведенческие сценарии на странице – и делать выводы на основе совокупности слабых сигналов. Экономика атаки выгодна злоумышленникам: создание нового фишингового сайта или клона обходится дешево и происходит быстро, в то время как поддержание актуальности и эффективности традиционных защитных механизмов требует постоянных и значительных затрат ресурсов. Эта растущая пропасть между изощренностью и скоростью создания фишинговых сайтов и возможностями старых методов защиты делает необходимым переход к принципиально новым подходам – интеллектуальным, адаптивным, способным обучаться и анализировать

[2] Обфускация кода — это метод преобразования исходного кода приложения, который делает его трудным для понимания и анализа.

сайты комплексно, оценивая множество параметров для выявления подделок. Именно такие возможности предоставляют технологии искусственного интеллекта и машинного обучения [7].

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Осознавая ограничения традиционных подходов, ФГБУ «НИИ «Интеграл» разработало и внедрило специализированное решение на основе искусственного интеллекта, направленное на эффективное выявление фишинговых веб-ресурсов. Была проведена масштабная работа по обучению модели машинного обучения, ставшей основой программного модуля, предназначенного для глубокого анализа сайтов. Этот модуль не ограничивается поверхностной проверкой; он проводит всестороннее исследование всех доступных данных, связанных с подозрительным веб-ресурсом. Анализ начинается с тщательного изучения самого URL-адреса, выявляя паттерны, характерные для поддельных или вредоносных ссылок. Информационная система мониторинга фишинговых сайтов ИС «Антифишинг» (далее – Система) агрегирует и обрабатывает обширные данные о домене: срок его существования, информацию о регистраторе и владельце, репутационные показатели. Однако ключевая сила решения заключается в глубинном сканировании и оценке самого содержимого и поведения сайта. Модуль анализирует структуру HTML-страницы, выявляя подозрительные схемы расположения элементов, а также оценивает сценарии поведения: наличие необычно агрессивных или скрытых запросов данных пользователя, злонамеренных полей ввода, маскирующихся под легитимные, или скриптов, специально разработанных для обхода средств обнаружения, например, обфускация кода, динамическая подгрузка вредоносного контента, проверка на наличие анализатора. Для оценки самодостаточности или, наоборот, зависимости от внешних ресурсов, Система рассчитывает сложные метрики, такие как процент внутренних ссылок и процент контента, подгружаемого со сторонних, потенциально небезопасных источников. Исследуются все исполняемые скрипты на странице с целью выявления потенциально вредоносных функций или попыток эксплуатации уязвимостей браузера. Важным аспектом архитектуры Системы является то, что ИИ-модуль не действует как полностью автономный «черный ящик», выносящий окончательный и безоговорочный вердикт, вместо этого он выполняет функцию интеллектуального ассистента аналитиков. На основе комплексного анализа параметров и выявленных аномалий ИИ-модель формирует обоснованное заключение о характере ресурса и присваивает ему потенциальный уровень угрозы. Эта оценка поступает в Систему, такой подход кардинально меняет рабочий процесс операторов и аналитиков Системы. Им больше не требуется тратить значительное время на ручное, детальное изучение каждого поступающего в систему подозрительного ресурса. Система, благодаря высокой точности анализа, предварительно оценивает очевидно легитимные или, наоборот, явно вредоносные сайты, существенно сужая область для ручной проверки и фокусируя внимание специалиста на наиболее критичных моментах. Это позволило ФГБУ «НИИ «Интеграл» добиться значительного сокращения времени обработки каждого инцидента в Системе, повысить пропускную способность анализа и оперативность реагирования на новые фишинговые угрозы в веб-пространстве, переложив основную нагрузку по первичной, трудоемкой оценке на обученные алгоритмы искусственного интеллекта.

Дополнительным инструментом, реализованным ФГБУ «НИИ «Интеграл» на базе обученной ИИ-модели, стал автоматизированный источник данных в Систему, работающий с ресурсами, вызывающими подозрения у пользователей сети Интернет. Механизм работает следующим образом: модуль собирает данные о веб-сайтах, на которые пользователи жалуются через различные общедоступные каналы как на потенциально мошеннические или фишинговые. Эти ресурсы, часто являющиеся еще не попавшими в фокус профессиональных аналитиков, автоматически поступают для анализа в программный модуль, основанный на искусственном интеллекте. Модуль оперативно сканирует и анализирует каждый такой сайт по всем ранее описанным параметрам. Ключевым моментом является то, что модуль не просто анализирует, а применяет свою обученную модель для оценки уровня угрозы. Если по результатам этого сайт получает оценку как опасный, он автоматически, без обязательного ручного подтверждения аналитиком на первом этапе, заносится в Систему через специальный канал. Это создает принципиально новый источник обнаружения угроз. Теперь Система получает информацию не только о ресурсах, напрямую выявленных или поступивших через официальные каналы обращений, но и о тех фишинговых сайтах, с которыми уже столкнулись и на которые пожаловались пользователи в сети Интернет, даже если эти пользователи не направляли официальное обращение непосредственно. Это позволяет значительно расширить охват Системы, оперативно включать в защиту актуальные, только что появившиеся или активно используемые мошенниками ресурсы, на которые уже обратили внимание жертвы, и блокировать их распространение на более широкую аудиторию.

ПЕРСПЕКТИВЫ СОЗДАНИЯ НОВЫХ МЕХАНИЗМОВ ПРОТИВОДЕЙСТВИЯ НА ОСНОВЕ ИИ

Опыт успешного применения ИИ-модуля для анализа сайтов открывает новые горизонты для разработки еще более совершенных инструментов, и в данный момент проводится активная разработка двух ключевых проектов. Первый проект – система предиктивного распознавания доменных имен – нацелен на выявление доменов, которые с высокой вероятностью будут использованы для фишинга, еще до начала активной фазы атаки. Используя машинное обучение для анализа паттернов, система предиктивного распознавания доменных имен сможет проактивно идентифицировать потенциально опасные ресурсы, позволяя повысить эффективность мониторинга на самом раннем этапе. Второй проект – система аналитики фишинговых сайтов на основе ИИ – предназначен для

решения задачи консолидации разрозненных инцидентов. Алгоритмы будут глубоко анализировать технические характеристики, контент, скрипты и инфраструктуру множества обнаруженных фишинговых ресурсов, выявляя скрытые связи и общие черты. Это позволит автоматически объединять отдельные сайты в целостные фишинговые кампании, выявляя преступные группы, их инструменты и тактики. Такая кластеризация критически важна для понимания полной картины угрозы, пресечения деятельности преступных групп на системном уровне и прогнозирования их следующих шагов, что выводит борьбу с фишингом на качественно новый уровень стратегического противодействия.

ЗАКЛЮЧЕНИЕ

Борьба с фишингом перешла в эпоху, где скорость, масштаб и изощренность атак сделали традиционные, реактивные методы защиты недостаточными. Как показано в статье, искусственный интеллект стал не просто дополнением, а стратегическим императивом в этой непрекращающейся битве. Его способности анализировать колоссальные объемы данных в реальном времени, выявлять сложные, скрытые паттерны и адаптироваться к новым тактикам злоумышленников предлагают принципиально иной уровень противодействия – проактивный, контекстный и масштабируемый.

Практический пример – разработки ФГБУ «НИИ «Интеграл». Внедренный ИИ-модуль для комплексного анализа сайтов значительно ускорил обработку инцидентов в информационной системе мониторинга фишинговых сайтов ИС «Антифишинг». Модуль выносит предварительную оценку угрозы, фокусируя внимание аналитиков на наиболее критичных случаях. Автоматический источник данных на основе ИИ-анализа сайтов, отмеченных в пользовательских жалобах, оперативно включает в защиту актуальные угрозы, с которыми уже столкнулись пользователи.

Таким образом, технологии ИИ показывают свою ключевую роль в борьбе с фишингом, предлагая не просто обнаружение, а проактивную, масштабируемую защиту. Развитие интеллектуальных систем в настоящее время – необходимое направление для обеспечения безопасности в цифровой среде.

СПИСОК ЛИТЕРАТУРЫ

1. Гордиенко В. В., Жданов Д. М. МЕТОДЫ ЗАЩИТЫ ОТ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ И ФИШИНГА. ИХ ДОСТОИНСТВА И НЕДОСТАТКИ // Auditorium. 2024. №2 (42). URL: *https://cyberleninka.ru/article/n/metody-zaschity-ot-sotsialnoy-inzhenerii-i-fishinga-ih-dostoinstva-i-nedostatki* (дата обращения: 09.06.2025).

2. Архипова А. Б., Нечаев Д. А. ТЕХНОЛОГИЯ ФОРМИРОВАНИЯ ИНТЕГРИРОВАННОЙ АНТИФИШИНГОВОЙ СИСТЕМЫ В ЦИФРОВОМ ОБЩЕСТВЕ // Вестник СибГУТИ. 2023. №2. URL: *https://cyberleninka.ru/article/n/tehnologiya-formirovaniya-integrirovannoy-antifishingovoy-sistemy-v-tsifrovom-obschestve* (дата обращения: 09.06.2025).

3. Войнов А. С. ИССЛЕДОВАНИЕ УЯЗВИМОСТЕЙ И УГРОЗ В МОБИЛЬНЫХ ПРИЛОЖЕНИЯХ: СТРАТЕГИИ ПРОТИВОДЕЙСТВИЯ // Вестник науки. 2023. №9 (66). URL: *https://cyberleninka.ru/article/n/issledovanie-uyazvimostey-i-ugroz-v-mobilnyh-prilozheniyah-strategii-protivodeystviya* (дата обращения: 09.06.2025).

4. Корнюхина С. П., Лапонина О. Р. ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ЗАЩИТЫ ОТ ФИШИНГОВЫХ АТАК // International Journal of Open Information Technologies. 2023. №6. URL: *https://cyberleninka.ru/article/n/issledovanie-vozmozhnostey-algoritmov-glubokogo-obucheniya-dlya-zaschity-ot-fishingovyh-atak* (дата обращения: 10.06.2025).

5. Данько О. С., Медведева Т. А. ИССЛЕДОВАНИЕ ТЕХНИК ФИШИНГА И МЕТОДОВ ЗАЩИТЫ ОТ НЕГО // Молодой исследователь Дона. 2021. №3 (30). URL: *https://cyberleninka.ru/article/n/issledovanie-tehnik-fishinga-i-metodov-zaschity-ot-nego* (дата обращения: 10.06.2025).

6. Афанасьева Н. С., Елизаров Д. А., Мызникова Т. А. КЛАССИФИКАЦИЯ ФИШИНГОВЫХ АТАК И МЕРЫ ПРОТИВОДЕЙСТВИЯ ИМ // ИВД. 2022. №5 (89). URL: *https://cyberleninka.ru/article/n/klassifikatsiya-fishingovyh-atak-i-mery-protivodeystviya-im* (дата обращения: 10.06.2025).

7. Баратбаев И. З. ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ В ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ // In The World Of Science and Education. 2025. №15 март ТН. URL: *https://cyberleninka.ru/article/n/primenenie-mashinnogo-obucheniya-v-informatsionnoy-bezopasnosti* (дата обращения: 10.06.2025).