УДК 004.8

СОЦИАЛЬНАЯ ИНЖЕНЕРИЯ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Пантелеев И.Ю.
Московский авиационный институт
(национальный исследовательский университет)
г. Москва

Социальная инженерия (СИ) представляет собой метод манипуляции человеческим поведением с целью получения конфиденциальной информации, обхода систем безопасности или выполнения действий, выгодных атакующему. В отличие от классических кибератак, эксплуатирующих уязвимости ПО, СИ фокусируется на человеческих слабостях: доверии, неосведомлённости и когнитивных искажениях. Её главные цели включают получение данных (пароли, банковские реквизиты, персональная информация), обход защиты (проникновение в закрытые помещения или компьютерные системы) и мошенничество (несанкционированные переводы денег, установка вредоносного ПО). Одним из самых распространённых методов социальной инженерии является фишинг — рассылка поддельных писем или сообщений, имитирующих легитимные источники, такие как банки, социальные сети или коллеги. Например, жертва может получить письмо с текстом: «Ваш аккаунт взломан! Перейдите по ссылке для восстановления», которое ведёт на фальшивый сайт для кражи данных. Другой метод — претекстинг, где злоумышленник создаёт ложный предлог для установления доверия. Это может быть звонок «от технической поддержки» с просьбой сообщить пароль для «проверки системы безопасности» или сотрудник службы доставки, требующий доступа в офис под предлогом срочной посылки.

С развитием искусственного интеллекта (ИИ) методы социальной инженерии трансформировались, став масштабнее, изощрённее и опаснее. Технологии ИИ коренным образом изменили ландшафт киберугроз. Если традиционные методы СИ, такие как фишинг и претекстинг, требовали прямого взаимодействия с жертвой, то современные инструменты на базе ИИ позволяют автоматизировать и персонализировать атаки. Например, генеративные модели вроде ChatGPT создают убедительные фишинговые письма с персонализированными деталями (имена, должности, ссылки на реальные события), а алгоритмы анализа Big Data собирают для них информацию из социальных сетей для точечного воздействия. Глубокое обучение (Deep Learning) лежит в основе deepfake-технологий, генерирующих поддельные видео и аудиозаписи, которые практически неотличимы от реальных. Обычно их используют для шатажа или дезинформации. Также одним из ключевых примеров эволюции СИ стала автоматизация атак. Боты, имитирующие живое общение в мессенджерах или социальных сетях, способны массово вовлекать пользователей в мошеннические схемы. Это демонстрирует, как ИИ преодо-

левает «естественные барьеры» доверия, делая атаки не только эффективными, но и психологически разрушительными.

Особую опасность представляют гибридные атаки. Так, например, злоумышленники могут использовать ChatGPT для генерации фишинговых писем, которые затем будут рассылать через Telegram ботов, имитирующих техподдержку банков. Традиционные методы защиты, основанные на проверке URL, в таких случаях бессильны — требуется анализ семантики и контекста. Противодействие социальной инженерии, усиленной искусственным интелектом, требует симбиоза технологических и образовательных мер. Со стороны технологий актуальны NLP-модели для анализа текста на фишинговые паттерны, GAN-сети для детекции deepfake и ИИ-тренажёры для симуляции атак, обучающие сотрудников распознавать угрозы в безопасной среде. Не менее важны и регуляторные инициативы. Европейский AI Act и GDPR уже устанавливают рамки для этичного использования ИИ, запрещая скрытую манипуляцию поведением. Однако глобальная координация остаётся проблемой: в странах с мягким законодательством продолжают действовать теневые платформы вроде WormGPT, предлагающие ИИ-инструменты для киберпреступников.

ИИ в социальной инженерии поднимает острые этические вопросы. С одной стороны, технологии защиты на базе ИИ спасают компании от многомиллионных убытков. С другой — их двойное применение угрожает приватности и демократическим институтам. Например, deepfake-видео политиков может спровоцировать социальные волнения, а сбор данных для персонализированных атак превращается в инструмент слежки. Ключевой вызов на сегодняшний момент — это размывание доверия (erosion of trust). Когда любое видео, голосовое сообщение или email могут оказаться фальшивыми, исчезает основа цифровой коммуникации. Уже сегодня большинство пользователей сомневаются в подлинности получаемого контента. Не исключено, что в дальнейшем это может привести к паранойе и социальной изоляции людей.

Искусственный интеллект стал катализатором как рисков, так и возможностей в области социальной инженерии. Современные киберугрозы используют ИИ для анализа поведения пользователей, автоматизированного фишинга и персонализированных атак, что делает традиционные методы защиты недостаточными. Если ранее защита сводилась к обучению сотрудников и установке антивирусов, то сегодня требуется комплексный подход, включающий интеграцию ИИ-систем, способных предугадывать сценарии атак, анализировать поведенческие аномалии и адаптироваться к новым угрозам в реальном времени.

Системы машинного обучения позволяют выявлять подозрительные паттерны взаимодействий, анализировать большие объемы данных о кибе-

ратаках и автоматически принимать меры по их предотвращению. Например, продвинутые системы безопасности способны распознавать фишинговые письма на основе лексического анализа, выявлять поддельные сайты по незначительным отличиям в коде и блокировать потенциальные угрозы до того, как пользователь с ними столкнется.

С другой стороны, злоумышленники также активно используют ИИ для генерации реалистичных мошеннических сообщений, автоматизации атак и обхода традиционных систем защиты. В связи с этим кибербезопасность становится гонкой вооружений, где ключевую роль играет скорость адаптации и способность прогнозировать новые методы социальной инженерии. Современные системы защиты включают поведенческую аналитику, многослойную аутентификацию, а также автоматизированные ИИ-алгоритмы для мониторинга и предотвращения угроз.

Таким образом, развитие искусственного интеллекта в сфере социальной инженерии требует не только совершенствования методов защиты, но и активного внедрения передовых технологий для предотвращения атак на самых ранних этапах.

Список литературы:

1. Salahdine F., Kaabouch N. Social engineering attacks: A survey //Future internet. – 2019. – Т. 11. – №. 4. – С. 89.

2. Другач Ю.С. Контролируемый взлом. Библия социальной инженерии.. - 2 изд. - Санкт-Петербург: БХВ-Петербург, 2024. - 208 с.

3. Hadnagy C. Social engineering: The art of human hacking. – John Wiley & Sons, 2010.