25 июня 2025

ИИ может не только атаковать компьютеры, но и защищать их

«Безопасность операционных систем в контексте искусственного интеллекта» — так назывался круглый стол, прошедший в июне этого года в рамках XII научно-практической конференции ИСП РАН OS DAY 2025 «Изолированные среды исполнения в современных ОС». Получилось так, что день проведения этого стола совпал с известием о крупнейшей утечке паролей в истории: были скомпрометированы 16 млрд аккаунтов Apple, Google, Facebook и других компаний. Это уже не просто утечка, а заготовка для десятков атак на сервисы по всему миру, которые осуществятся в ближайшие дни и недели, говорят некоторые участники ИТ-рынка. Что лишний раз подтверждает актуальность разговоров о практических мерах решения проблем кибербезопасности.

Ведущий круглого стола Андрей Духвалов, вице-президент, директор департамента перспективных технологий «Лаборатории Касперского», поясняет: «Искусственный интеллект сейчас используется во всех сферах человеческой жизни. В том числе, это касается и деятельности злоумышленников. Например, ИИ позволяет им искать ранее неизвестные уязвимости в системном софте, в частности ОС, применять более изощрённые методы кибератак. Думаю, неправильно считать, что цель злоумышленников именно ОС. Как правило, атака на ОС — только один из

этапов проникновения. Хотя ОС есть практически на всех ИТ-устройствах, а сложность её устройства позволяет легко находить в ней уязвимости, особенно с применением ИИ. Поэтому ОС — весьма распространённый объект атаки.

Ещё один фактор заключается в том, что ИИ снижает требования к квалификации и профессионализму людей, которые его используют. Этот тренд проявляется во всех сферах, в том числе и в атаках на информационные системы. К сожалению, всё больше людей вовлекается в эту деятельность. Поэтому многие специалисты отмечают, что применение злоумышленниками ИИ влияет на ситуацию с кибербезопасностью. С 2022 г. мы наблюдаем устойчивый рост числа кибератак с применением ИИ. По оценке аналитиков Gartner, если в 2022 г. только около 10% кибератак содержали элементы использования ИИ, то к концу 2025 г. их доля может достигнуть 80%.

Однако ИИ усиливает не только атакующих, но и защитников. Например, благодаря ему появляются новые способы корреляции событий и выявления аномалий в сложных системах. Также ИИ помогает разрабатывать более устойчивый к атакам код, создавать сложные информационные системы. Всё это затрудняет жизнь злоумышленникам. С появлением новых технологий человечество всегда получало и новые проблемы. С ИИ ситуация такая же — мощность воздействия ИИ на человеческие отношения сравнима с открытием электричества, паровых двигателей, динамита, атомной энергии и интернета. Несмотря на все сложности, человечество ранее всегда обращало новые технологии в свою пользу. С ИИ тоже получится. Я уверен».

С ним согласна **Елена Скалозубова, ведущий архитектор ИБ MONS (входит в ГК «КОРУС Консалтинг»)**

Использование искусственного интеллекта в кибератаках действительно растёт. Злоумышленники активно применяют ИИ-технологии для автоматизации фишинговых атак, обхода систем безопасности и даже генерации вредоносного кода. Например, в инструментах вроде Cobalt Strike уже внедряют ИИ-функционал — на это выделяются значительные ресурсы, в том числе в даркнете. Пока доля таких атак остаётся относительно небольшой — по разным оценкам, около 5-10%. Однако тенденция очевидна, ИИ позволяет злоумышленникам проводить более изощрённые и масштабируемые атаки. К сожалению, защитные системы часто развиваются реактивно, то есть приходится „догонять“, тогда как у атакующих есть и мотивация, и финансирование для внедрения новых технологий. Скорее всего, в ближайшие годы мы увидим рост числа ИИ-атак, но пока это направление только набирает обороты.

Елена Скалозубова,

Ведущий архитектор по информационной безопасности MONS (ГК «КОРУС Консалтинг»)

Впрочем, не все столь категоричны. Вот как комментирует ситуацию руководитель группы защиты инфраструктурных ИТ-решений компании «Газинформсервис» Сергей Полунин: «Я бы не сказал, что число атак с использованием технологий ИИ стремительно увеличивается. Это слово сложно кодифицировать. В то же время с уверенностью можно говорить, что с помощью ИИ автоматизируются отдельные элементы атак. Например, на этапе поиска уязвимостей ИИ хорошо справляется с автоматическим анализом кода или конфигурационных файлов, поиском открытых портов, где есть какая-то перспектива для атаки. С другой стороны, ИИ неплохо себя показывает уже сегодня для организации атак с помощью социальной инженерии — все эти дипфейки, фишинг и персонализированные атаки выходят совсем на другой уровень. Ну и отчасти ИИ помогает обходить традиционные средства защиты, такие как EDR/XDR, хотя это пока экзотика. В отчетах Microsoft, IBM X-Force и Mandiant за 2024-2025 гг. отмечается рост использования ИИ APT-группами по всему миру, поэтому справедливо говорить, что ИИ занимает свое достойное место в арсенале киберпреступников».

Похожего мнения придерживается архитектор девопс-компании Git In Sky Дмитрий Лерх: «ИИ применяется для автоматизации и масштабирования рутинных задач: генерации фишинговых писем с более убедительным текстом и без ошибок, создания вредоносного кода, анализа больших объемов данных для поиска уязвимостей или адаптации атак под конкретную цель. Утверждать, что „ИИ атакует инфраструктуру“ — все равно, что говорить „Google помог атаковать“, потому что злоумышленник искал в нем эксплойты.

Технологии атаки фундаментально не изменились. ИИ делает злоумышленников более продуктивными и усложняет обнаружение некоторых векторов атак, но не создает принципиально новых угроз из ничего. Таким образом, рост инцидентов с упоминанием ИИ — это скорее следствие его доступности и эффективности как инструмента для автоматизации старых методов, а не появления неких неуязвимых „искусственных“ хакеров. Защитникам необходимо фокусироваться на базовых принципах кибергигиены, своевременном обновлении систем и обучении пользователей, оставаясь в курсе новых тактик, которые ИИ позволяет реализовать быстрее».

Константин Мельников, руководитель Департамента специальных сервисов компании Infosecurity (с 2024 г. входит в ГК Softline), отметил: «Искусственный интеллект сам по себе не является методом кибератаки, но он служит мощным инструментом, позволяющим злоумышленникам повышать качество атак и масштабировать их. С демократизацией ИИ-технологий эти решения стали доступны широкому кругу пользователей, в том числе киберпреступникам, что привело к заметной эскалации угроз в цифровом пространстве.

Одним из наиболее заметных последствий стал рост фишинговой активности. Во-первых, автоматизация рутинных процессов через ИИ привела к увеличению количества мошеннических ресурсов. Во-вторых, злоумышленники начали активно применять модели для создания фото- и видеоконтента. Особую популярность приобрели дипфейки, в частности, так называемые „фейкбоссы“ — цифровые двойники руководителей. Бизнес уже

активно использует цифровые аватары для публичных выступлений и записей, что снижает необходимость личного присутствия. Однако эта же технология стала инструментом мошенников, создающих убедительные подделки для социальной инженерии.

Кроме того, ИИ упрощает разработку фишинговых интерфейсов. Ранее злоумышленникам приходилось вручную копировать элементы сайтов, но теперь генеративные модели автоматически воспроизводят не только дизайн сайтов, но и такие детали, как favicon и формы ввода данных. Например, в прошлом году мы обнаружили профессионально выполненную копию банковского сайта. Более того, некоторые группировки уже внедрили автоматизированные сервисы, генерирующие фишинговые страницы за считанные минуты — достаточно указать целевой сайт, и система формирует готовый шаблон с необходимыми элементами.

Хотя подобные случаи пока единичны, а качество кибератак остается неоднородным, злоумышленники активно совершенствуют свои методы. Особую опасность представляют дипфейки, достигшие высокого уровня реалистичности. Преступники постоянно экспериментируют с новыми технологиями, снижая затраты и повышая эффективность своих операций, что неизбежно ведет к увеличению количества инцидентов».