УДК 004.056.5

К ВОПРОСАМ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В БЕЗОПАСТНОСТИ

Васильев Н.А.[1], заместитель начальника отдела научно-исследовательского центра, к.т.н.,
Ситдиков Д.С.[1], младший научный сотрудник, Лещинский Б.С.[1], адъюнкт
[1]Военная академия связи
г. Санкт-Петербург

Современные информационные технологии прочно вошли в повседневную жизнь общества, охватив практически все его сферы. В результате данной цифровой трансформации резко возросли как объемы данных, так и угрозы, связанные с их защитой. Кибербезопасность стала одной из приоритетных задач для компаний и государственных организаций. При этом традиционные методы защиты данных, такие как антивирусные программы и системы предотвращения вторжений, стали менее эффективными в условиях современных атак.

В связи с этим искусственный интеллект (далее ИИ), благодаря своим уникальным способностям к анализу больших данных, обнаружению аномалий и автоматизации процессов, становится одним из самых перспективных инструментов в области обеспечения безопасности. ИИ позволяет обнаруживать сложные угрозы, которые невозможно распознать с помощью традиционных методов. Он способен обучаться на исторических данных, предсказывать возможные атаки и адаптироваться к новым угрозам. Настоящая статья посвящена роли искусственного интеллекта в обеспечении безопасности: анализу его применений, оценке эффективности и рисков, а также обсуждению будущих перспектив развития. В статье приводятся примеры реальных приложений ИИ в кибербезопасности, рассматриваются релевантные научные работы.

Рассмотрим ключевые направления, в которых использование ИИ дает существенные преимущества.

1. Обнаружение угроз

Одной из важнейших задач в области безопасности является своевременное обнаружение угроз, будь то атаки на сеть, вредоносное ПО или утечки данных. Искусственный интеллект может эффективно решать эту задачу за счет использования методов машинного обучения и анализа данных. Применение алгоритмов ИИ позволяет:

1.1 Анализировать сетевой трафик на предмет аномалий [1]. Например, ИИ может выявлять отклонения в поведении пользователей или систем, указывающие на возможные атаки.

1.2 Распознавать новые типы угроз на основе анализа паттернов и корреляций между событиями.

1.3 Автоматизировать процесс обнаружения за счет сокращения числа ложных срабатываний и повышения точности анализа данных.

Примером может служить использование ИИ для обнаружения атак типа «отказ в обслуживании» (DDoS) [2]. Традиционные системы могут не справиться с высокими объемами трафика, однако ИИ позволяет быстрее выявить и блокировать вредоносные запросы, основываясь на их отклонении от нормального поведения.

2. Прогнозирование атак

ИИ может не только обнаруживать текущие угрозы, но и предсказывать возможные будущие атаки. Это достигается за счет анализа исторических данных и выявления закономерностей, которые указывают на высокую вероятность появления новых угроз. Прогнозирование атак основано на использовании методов предсказательной аналитики, которая позволяет выявлять уязвимости в системе безопасности и принимать меры по их устранению до того, как будет совершена атака.

Алгоритмы машинного обучения могут анализировать данные о прошлых атаках, чтобы находить скрытые зависимости и уязвимости в сетях и системах [3,4,7]. Эти данные можно использовать для создания моделей, которые будут предупреждать о потенциальных атаках на основе схожих условий.

3. Автоматизация мониторинга и реагирования

Одним из самых больших вызовов в кибербезопасности является необходимость оперативно реагировать на возникающие инциденты. Чем быстрее будет выявлена и устранена угроза, тем меньше ущерб. ИИ может автоматизировать процесс мониторинга и реагирования на инциденты [5-6]. В отличие от традиционных систем, которые требуют вмешательства специалистов, системы на базе ИИ могут самостоятельно обнаруживать инциденты, оценивать их серьезность, принимать меры для минимизации ущерба.

Примером является автоматизация обработки инцидентов: ИИ может сразу заблокировать учетную запись при подозрительной активности или ограничить доступ к определенным данным в случае их компрометации. Это значительно снижает время реакции на угрозы и уменьшает нагрузку на службы безопасности.

4. Борьба с фишингом

Фишинг остается одним из самых распространенных методов атак, направленных на кражу данных и компрометацию учетных записей пользователей. Атаки на основе социальной инженерии, такие как фишинг, используют психологические уловки для того, чтобы заставить пользователей раскрыть конфиденциальную информацию или скачать вредоносные файлы.

Искусственный интеллект помогает эффективно бороться с фишингом за счет анализа огромных объемов данных, включая электронные письма, текстовые сообщения и веб-сайты [8,9]. Основные механизмы борьбы с фишингом с использованием ИИ:

4.1 Анализ текста сообщений**.** ИИ обучен распознавать подозрительные фразы, заголовки и контент, который характерен для фишинговых писем. Такие системы могут анализировать миллионы писем в реальном времени и блокировать потенциально опасные сообщения до их получения пользователем.

4.2 Проверка подлинности отправителя. Системы на основе ИИ могут сверять данные о доменах отправителей с известными базами данных, проверять признаки подделки адресов электронной почты и обнаруживать фальсификацию.

4.3 Поведенческий анализ**.** ИИ может отслеживать поведение пользователей, определяя, когда они взаимодействуют с подозрительными веб-сайтами или ссылками. Это помогает выявлять попытки обмана и предотвращать переходы на фальшивые страницы.

Например, многие современные почтовые системы, такие как Google Mail и Microsoft Outlook, используют ИИ для фильтрации фишинговых писем, помогая предотвратить атаки на этапе получения сообщений.

5. Управление доступом и аутентификацией

Управление доступом и защита аутентификационных данных — важный аспект в области безопасности. Кража учетных записей и их использование злоумышленниками представляют собой одну из наиболее частых угроз. Системы на базе ИИ способны повысить уровень безопасности за счет использования адаптивных методов аутентификации:

5.1 Биометрическая аутентификация**.** ИИ используется для анализа отпечатков пальцев, изображения лица, голоса и других биометрических данных для подтверждения личности пользователя. Эти данные трудно подделать, что делает такие методы более надежными по сравнению с традиционными паролями [10].

5.2 Многофакторная аутентификация с ИИ**.** ИИ может улучшить многофакторную аутентификацию, анализируя дополнительные параметры, такие как поведение пользователя или местоположение [11]. Например, если система обнаружит попытку входа с необычного устройства или из необычного места, она может запросить дополнительное подтверждение.

Современные системы управления доступом могут также использовать машинное обучение для анализа активности пользователей и адаптироваться к изменениям, что позволяет эффективно предотвращать атаки с использованием скомпрометированных учетных данных.

Научные исследования и разработки в области применения ИИ в безопасности играют важную роль в развитии технологий киберзащиты. Рассмотрим несколько ключевых исследований и практических примеров использования ИИ в защите информационных систем.

1. Обнаружение атак «нулевого дня»

Атаки «нулевого дня» представляют собой особую угрозу, поскольку они эксплуатируют уязвимости, которые еще не были публично раскрыты или исправлены. В исследовании Arun A. и коллег [12] было показано, что использование методов глубокого обучения позволяет значительно улучшить обнаружение таких атак. В рамках исследования был разработан алгоритм на основе LSTM, способный распознавать новые схемы атак и выявлять малейшие отклонения от обычного поведения, что подтверждается тщательным тестированием. В результате получается мощная система обнаружения атак «нулевого

дня», которая повышает точность и скорость реагирования. Это снижает потенциальный ущерб, который могут нанести неизвестные уязвимости.

2. ИИ против фишинга

В работе Kalla D. и соавторов [13] изучались различные методы ИИ для борьбы с фишингом. Исследование представляет алгоритм Phishcatch, который показал значительный успех в выявлении фишинговых писем и оповещении потребителей о мошеннических попытках. Phishcatch изучает поведение пользователей на веб-сайтах и ограничивает доступ в случае обнаружения поведения. Точность и уровень обнаружения алгоритма составили 90%.

3. Анализа больших данных в ИИ-системах обнаружения вторжений для выявления кибератак в ICS-сетях

Исследование Ali B. S. и коллег [14] посвящено применению искусственного интеллекта (ИИ) для обнаружения кибератак в промышленных системах управления (ICS), таких как SCADA-сети, где основной проблемой являются несбалансированные наборы данных, затрудняющие классификацию атак, особенно в многоклассовых сценариях. Исследователи разработали модель ICS-IDS — систему обнаружения вторжений, адаптированную для ICS-сетей, применив методы машинного обучения, включая нормализацию и метод Фишера для уменьшения размерности данных, а также k-NN (метод k-ближайших соседей) для перебалансировки датасета. Использование instance-based learning (обучения на примерах) позволило выявлять скрытые паттерны атак, что привело к достижению высокой точности: 99% accuracy и 99% detection rate (DR) на реальных данных ICS-сетей, демонстрируя эффективность предложенного подхода для защиты критической инфраструктуры.

Несмотря на все преимущества ИИ, его использование в кибербезопасности связано с рядом проблем. Рассмотрим основные из них.

1. Ложные срабатывания

Одной из наиболее частых проблем ИИ в безопасности является высокая вероятность ложных срабатываний. Несмотря на то, что ИИ способен обнаруживать аномалии, некоторые из них могут не быть реальными угрозами. Это может привести к неправильной блокировке пользователей или систем, что создает дополнительные сложности для администраторов.

2. Атаки на ИИ

ИИ сам может стать объектом атак. Злоумышленники могут попытаться подделать данные, на которых обучается система ИИ [15], чтобы изменить ее поведение и заставить ошибаться в процессе обнаружения угроз. Это требует разработки новых механизмов защиты ИИ от целенаправленных атак.

3. Этические проблемы

Применение ИИ в безопасности вызывает вопросы этического характера [16]. Например, системы, которые анализируют поведение пользователей, могут нарушать их приватность, что требует более строгого регулирования и обеспечения прозрачности использования ИИ в кибербезопасности.

Искусственный интеллект значительно изменил подходы к кибербезопасности, предлагая новые возможности для обнаружения и предотвращения

угроз, автоматизации процессов и повышения эффективности защиты. Однако вместе с этим возникают и новые вызовы, такие как этические вопросы, атаки на ИИ и необходимость улучшения точности моделей. Тем не менее, будущее ИИ в безопасности представляется перспективным: его интеграция с другими технологиями, такими как Интернет вещей и большие данные, будет продолжать развиваться, предоставляя новые инструменты для защиты информации и систем от современных угроз.

Список литературы:

1. Зуев В. Н. Обнаружение аномалий сетевого трафика методом глубокого обучения //Программные продукты и системы. – 2021. – Т. 34. – №. 1. – С. 91-97.

2. Пирлиев К. и др. Эффективные подходы к обнаружению и предсказанию DDOS-атак с использованием машинного обучения //Всемирный ученый. – 2024. – Т. 1. – №. 24. – С. 160-166.

3. Лебедев Д. В. Прогнозирование сценариев кибератак с использованием методов машинного обучения без учителя //Весенние дни науки ИнЭУ: сборник докладов Международной конференции студентов и молодых ученых.—Екатеринбург, 2024. – Издательство Издательский Дом «Ажур», 2024. – С. 87-91.

4. Пирлиев К. и др. Эффективные подходы к обнаружению и предсказанию DDOS-атак с использованием машинного обучения //Всемирный ученый. – 2024. – Т. 1. – №. 24. – С. 160-166.

5. Андрияхов Я. В. Автоматизация процессов классификации и реагирования на инциденты информационной безопасности на основе нейросетевых технологий //Высшее образование в России. – 2021. – №. 6. – С. 121-131.

6. Частикова В. А., Козачёк К. В. Применение нейронных сетей в платформе реагирования на инциденты как эффективное средство управления кибербезопасностью //Вестник УрФО. Безопасность в информационной сфере. – 2023. – Т. 4. – №. 50. – С. 70-76.

7. Peter H. Automation in Enterprise Security: Leveraging AI for Threat Prediction and Resolution. – 2024.

8. Чичерина К. С. Методы машинного обучения и анализа данных в системе обнаружения фишинга и спама //Научный руководитель. – 2023. – С. 264.

9. Киргизбаев С. П., Киргизбаев В. П. Практическое применение искусственного интеллекта для распознавания и предотвращения фишинговых атак //Современная Наука 3. – 2024. – С. 24.

10. Сулавко А. Е. Высоконадежная биометрическая аутентификация на основе защищенного исполнения нейросетевых моделей и алгоритмов искусственного интеллекта //Системная инженерия и информационные технологии. – 2024. – Т. 6. – №. 2 (17). – С. 11-32.

11. Chennuri K. M. R. Adaptive multi-factor authentication systems: a comprehensive analysis of modern security approaches //International Journal Of Computer Engineering And Technology (IJCET). – 2024. – Т. 15. – №. 6. – С. 787-795.

12. Arun A., Nair A. S., Sreedevi A. G. Zero day attack detection and simulation through deep learning techniques //2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence). – IEEE, 2024. – С. 852-857.

13. Kalla D. et al. Phishing detection implementation using databricks and artificial Intelligence //International Journal of Computer Applications. – 2023. – Т. 185. – №. 11. – С. 1-11.

14. Ali B. S. et al. ICS-IDS: application of big data analysis in AI-based intrusion detection systems to identify cyberattacks in ICS networks //The Journal of Supercomputing. – 2024. – Т. 80. – №. 6. – С. 7876-7905.

15. Киберполигон: симуляция реальных угроз в виртуальном мире / Б. С. Лещинский, Н. А. Васильев, О. С. Лаута, Д. С. Ситдиков // Молодежь. Техника. Космос : Труды XVI общероссийской молодёжной научно-технической конференции, приуроченной к 90-летию Юрия Алексеевича Гагарина, летчика-космонавта СССР, Героя Советского Союза, первого человека, отправившегося в космическое пространство. В 4-х томах, Санкт-Петербур, 25–29 марта 2024 года. – Санкт-Петербург: Балтийский государственный технический университет "ВОЕНМЕХ" им. Д.Ф. Устинова, 2024. – С. 241-243.

16. Вопросы применения искусственного интеллекта в безопасности / Б. С. Лещинский, Н. А. Васильев, Р. Р. Хабушев, А. Р. Назаров // Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века : Сборник статей по материалам Девятой всероссийской научно-практической конференции с международным участием, Пермь, 17–18 октября 2024 года. – Пермь: Пермский государственный национальный исследовательский университет, 2024. – С. 77-81.