УДК 004.056

ВЛИЯНИЕ СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

Терентьев Д. Е.

магистрант,

Калужский государственный университет им. К. Э. Циолковского,

Калуга, Россия

Ткаченко А. Л.

к.т.н., доцент,

Калужский государственный университет им. К.Э. Циолковского,

Калуга, Россия

Аннотация

В статье определены предпосылки и актуальность внедрения технологий искусственного интеллекта в системы защиты информации. Приведено нормативно-правовое определение ИИ, обозначены проблемы, связанные с внедрением интеллектуальных систем. На примере внутренних и внешних (кибератаки) нарушителей, обозначены сферы, нуждающиеся в защите в виде систем искусственного интеллекта.

Ключевые слова: искусственный интеллект, киберпреступность, интеллектуальная деятельность, информационная безопасность, конфиденциальность.

INFLUENCE OF ARTIFICIAL INTELLIGENCE SYSTEMS IN INFORMATION SECURITY SYSTEMS

Terentyev D. E.

master's student,

Kaluga State University named after K. E. Tsiolkovsky,
Kaluga, Russia

Tkachenko A.L.
candidate of Technical Sciences,
Kaluga State University named after K. E. Tsiolkovsky,
Kaluga, Russia

Abstract
The article defines the prerequisites and relevance of the implementation of artificial intelligence technologies in information security systems. A regulatory and legal definition of AI is given, problems associated with the implementation of intelligent systems are identified. Using the example of internal and external (cyberattacks) violators, areas requiring protection in the form of artificial intelligence systems are identified.

Keywords: artificial intelligence, cybercrime, intellectual activity, information security, confidentiality.

Современный мир удивительно многогранен и сложен. Эволюция жизнедеятельности общества, которая проявляется в том числе посредством развития науки, привела к появлению принципиально новых механизмов, призванных обеспечить эффективность протекания процессов, возникающих в самых различных сферах и областях.

В контексте подобных изменений наиболее значительное место уделяется цифровым трансформациям, в корне изменившим привычные уклады производства, предпринимательской деятельности, здравоохранения, информационной сферы и т.д.

Тема данной статьи посвящена вопросу внедрения системы искусственного интеллекта в системы защиты информации. Актуальность темы исследования подтверждается ее практической значимостью, заключающейся в том, что защита информации на любых уровнях является залогом информационной безопасности государства, что существенно влияет на национальную безопасность страны в целом. Именно посредством использования передовых современных научных технологий, к коим можно отнести инструментарии искусственного интеллекта, возможно достижение цели по организации процесса обеспечения защиты информационных ресурсов на любых уровнях. Кроме того, защита информации является приоритетной задачей менеджмента любого коммерческого предприятия, поскольку практически всем приходится иметь дело с коммерческой тайной и персональными данными.

Для начала приведем определение искусственного интеллекта, после чего перейдем к описанию процесса по внедрению его систем в деятельность по защите информации.

ИИ с точки зрения нормативно-правовой базы трактуется как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека» [1].

Другой достаточно распространенный подход — представление об ИИ как о совокупности базовых (системообразующих) технологий ИИ: компьютерного зрения; распознавания и синтеза речи; обработки и интеллектуального анализа естественных языков; поддержки принятия решений. К базовым также можно отнести технологии: машинного и глубокого обучения; интеллектуального анализа больших данных и знаний; инженерии знаний (прежде всего на динамических графах знаний); планирования

и управления целенаправленным поведением в неструктурированных средах; мультиагентного управления [2].

Для того, чтобы искусственный интеллект смог обеспечить требуемый уровень защиты информации, он должен быть корректно имплементирован, интегрирован в существующие системы и обучен. Парадоксально, но само по себе введение в эксплуатацию интеллектуальных систем, призванных защищать данные, может привести к колоссальной по масштабу последствий «бреши» в системе защиты и существенно снизить уровень безопасности индивидуального или корпоративного пользователя.

Одной из проблем, возникающих при обучении и вводе в эксплуатацию интеллектуальной системы, может стать нарушение конфиденциальности (privacy breach). В данной связи фокус внимания смещается на новые технологии повышения конфиденциальности; к примеру, технология OPAL (open algorithms project) позволяет отказаться от пересылки данных алгоритму искусственного интеллекта за счет предоставления удаленного и контролируемого доступа к информации [3].

Мы живем в эпоху, когда центральной проблемой для пользователей, бизнес-структур и регулирующих органов стала проблема защиты персональных данных. Пользователи требуют обеспечения большей прозрачности и контроля в области сбора, хранения и использования данных, а также обмена данными. Защита информации стала одной из приоритетных задач, стоящих перед обществом. Вопрос защиты информации актуален как никогда ранее, так как масштабы киберпреступности постоянно растут. Один из возможных инструментов противодействия киберугрозам – технологии искусственного интеллекта. Большинство современных решений в сфере информационной безопасности так или иначе основаны на искусственном интеллекте. При этом внедрение искусственного интеллекта в область защиты информации сопряжено с массой рисков, в связи с чем специалисты в области обработки данных объединяют свои усилия для разработки инструментов защиты персональных

данных для систем искусственного интеллекта, которые появятся в недалеком будущем.

Следующим немаловажным направлением применения искусственного интеллекта в вопросе информационной безопасности является работа с внутренними нарушителями. Зная типичное поведение пользователя, система может отправить предупреждение аналитикам ИБ в случае существенного изменения модели работы сотрудника (посещение подозрительных сайтов, длительное отсутствие за рабочим ПК, изменение круга общения при переписке в корпоративном мессенджере и т.д.). Системы защиты, оснащенные компьютерным зрением и обработкой речи, смогут оперативно оповещать охрану о попытках прохода через проходную посторонних или сотрудников по чужим пропускам, анализировать рабочую активность сотрудников с помощью веб-камер, оценивать корректность общения менеджеров с клиентами по телефону [4].

В целом, ИИ работает правильно только в том случае, если обучающие данные и данные пользователя, используемые в работе, качественно совпадают, т. е. с точки зрения их распределения. То есть, это правило одинаково справедливо по отношению к обоим вариантам, описанным выше.

Кроме того, обучающие данные и пользовательские данные должны обладать признаком, который необходимо отфильтровать. В результате возникают предпосылки для успешного использования, описанные ниже, которые в то же время указывают на проблемы или возможности для атаки:

1. Стабильность: это означает, что ситуация не должна быстро меняться. В этом случае систему необходимо заново обучить, используя новые данные.

2. Целостность обучающих данных. Обучающие данные также должны обладать целостностью в том смысле, что ими не манипулировал злоумышленник. Манипуляция может происходить, например, путем влияния на пользователей, а также за счет того, что злоумышленник уже активен на этапе обучения, так что впоследствии он не будет распознан.

3. Целостность процесса обучения. Также должно быть невозможно, чтобы злоумышленник намеренно манипулировал процессом обучения, вводя неблагоприятные примеры. Следует помнить, что «восприятие» систем ИИ сильно отличается от человеческого. Изменение нескольких пикселей на изображении может превратить лицо в автомобиль для ИИ, в то время как человек сразу же распознает разницу.

4. Маркировка. Часто не хватает информации, необходимой для обучения. В сложных ИТ-системах это трудно определить и параметризировать. Есть много вопросов, на которые необходимо ответить: как определяются примеры «хорошего случая»? Какие возможности существуют для создания этих «хороших примеров» в смысле максимизации результатов с течением времени и предоставления их для изучения? И какие рамочные условия, жесткие и мягкие, должны быть предоставлены такой системе ИИ?

5. Для успешного обучения необходим определенный объем данных, который включает «хорошие» и «плохие» случаи. Система не сможет научится распознавать ошибки, если их никогда не встречала [5].

Резюмируя вышеизложенное, можно сделать вывод относительно того, что несмотря на колоссальный объем усилий, который вкладывается в развитие искусственного интеллекта в контексте обеспечения потребностей информационной безопасности, по сегодняшний день существует множество научных и методических пробелов, из раза в раз вскрываемых посредством изощренных преступных схем злоумышленников. Единственным вариантом в данной ситуации является адекватная и своевременная реакция на потенциальные угрозы.

Абзац представленный выше свидетельствует о том, что, несмотря на всю свою развитость, искусственный интеллект не может быть совершены, поскольку алгоритмы, заложенные в основе его работы, создаются человеком. Следовательно, если возникнет прецедент, с которым не сталкивались специалисты в сфере IT, инструментарии искусственного интеллекта будут

бесполезны. Этот факт подтверждается тем, что даже крупные государственные ведомства мировых стран-лидеров, достаточно часто сталкиваются с успешными хакерскими атаками, в результате которых происходило хищение информационных ресурсов [6].

Библиографический список:

1. Национальная стратегия развития искусственного интеллекта на период до 2030 года. Указ Президента РФ №490 от 10.10.2019.

2. Забежайло М.И., Борисов В.В. Об интерпретациях понятия «искусственный интеллект» // Речевые технологии/Speech Technologies. 2022. №1. [Электронный ресурс]. URL: https://cyberleninka.ru/article/n/ob-interpretatsiyah-ponyatiya-iskusstvennyy-intellekt (дата обращения: 15.01.2024).

3. Шананин В.А. Применение систем искусственного интеллекта в защите информации // Инновации и инвестиции. 2022. №11. [Электронный ресурс]. URL: https://cyberleninka.ru/article/n/primenenie-sistem-iskusstvennogo-intellekta-v-zaschite-informatsii (дата обращения: 15.01.2024).

4. Рахметов Р. Искусственный интеллект в информационной безопасности. [Электронный ресурс]. URL: https://www.securityvision.ru/blog/iskusstvennyy-intellekt-v-informatsionnoy-bezopasnosti (дата обращения: 15.01.2024).

5. Гумерова Л.Д., Ефимова Ю.А., Файзуллин Р.В. Информационная безопасность в системах с искусственным интеллектом. [Электронный ресурс]. URL: clck.ru/3FotQL (дата обращения: 15.01.2024).

6. Хакеры взломали систему при помощи уязвимости четырехлетней давности. [Электронный ресурс]. URL: https://www.gazeta.ru/tech/news/2023/03/17/19990483.shtml (дата обращения: 15.01.2024).

Оригинальность 76%