УДК 004.273

Искусственный интеллект в модели кибербезопасности «Нулевое доверие»

Artificial intelligence in the zero trust cybersecurity model

Комашинский / Komashinsky V.
Владимир Ильич
(kama54@rambler.ru)
доктор технических наук, доцент.
ЗАО «Институт телекоммуникаций»,
советник генерального директора.
г. Санкт-Петербург

Присяжнюк / Prisyazhnyuk S.
Сергей Прокофьевич
(office@itain.ru)
доктор технических наук, профессор,
заслуженный деятель науки РФ.
ЗАО «Институт телекоммуникаций»,
генеральный директор.
г. Санкт-Петербург

Ключевые слова: кибербезопасность – cybersecurity; архитектура нулевого доверия – zero trust architecture; искусственный интеллект – artificial intelligence; машинное обучение – machine learning; аутентификация – authentication; авторизация – authorization.

В статье рассмотрены концептуальные основы построения систем кибербезопасности, созданные на комплексном применении новой архитектуры «Нулевого доверия» и технологий искусственного интеллекта. Этот подход можно рассматривать как смену парадигмы на стратегии обеспечения кибербезопасности. Ее основная цель – предотвратить утечки данных и ограничить распространение внутренних горизонтальных угроз.

The article discusses the conceptual foundations of building cybersecurity systems based on the integrated use of the new Zero Trust architecture and artificial intelligence technologies. This approach can be considered as a paradigm shift in cybersecurity strategies. Its main goal is to prevent data leaks and limit the spread of internal horizontal threats.

Введение

Искусственный интеллект (ИИ) стремительно ускоряет развитие мира, срастаясь с остальными его кибертехническими обитателями и обеспечивая более высокие уровни автоматизации, автономности и экономической эффективности. Однако ИИ начинает использоваться и киберпреступниками, что создает новые проблемы, связанные с обеспечением безопасности современных критически важных и частных инфраструктур. Очевидно, что в этих условиях системам кибербезопасности без искусственного интеллекта никак не обойтись. Иными словами, срастание киберпреступности с ИИ уже стало реальностью, а использование новой концепции обеспечения кибербезопасности «Нулевого доверия» с применением ИИ становится новым императивом времени.

ИИ и киберпреступность

Технологии ИИ относятся к категории инструментов общего назначения, то есть их можно использовать практически для решения любой задачи. И как любой инструмент общего назначения, киберпреступники могут использовать ИИ и во вредоносных целях. Большие языковые модели (LLM), такие как OpenAI ChatGPT, Microsoft Bing Chat и Google Bard AI, продолжают удивлять нас своими уникальными возможностями обработки языка. Другие генеративные приложения ИИ, такие как Midjourney, Dall-E или Stable Diffusion, поражают нас, создавая потрясающие произведения искусства. Эти проекты становятся лучше с каждым днем, и здорово, что любой может получить к ним доступ. Однако это начинает создавать угрозы для многих людей, в частности, в последнее время ИИ широко используется для создания дипфейков, взломов паролей, проведения мошеннических транзакций, создание программ вымогателей и т. д. Например Deepfake – это сочетание «глубокого обучения» и «поддельных медиа», подразумевающее использование ИИ для создания/манипулирования аудиовизуальными медиа, чтобы они выглядели подлинными. Киберпреступники уже используют и широко применяют эту технологию для создания искусственных персонажей, похожих (визуально и аудиально) на реальных родственников, знаменитостей, политических деятелей, которые применяются для осуществления вымогательства, обмана, распространения политической дезинформации и т. д. Кроме того, киберпреступники используют машинное

обучение (МО) и ИИ для улучшения алгоритмов угадывания паролей пользователей. Хотя некоторые алгоритмы взлома паролей уже существуют, киберпреступники могут анализировать большие наборы данных паролей и генерировать реальные варианты паролей. Помимо взлома паролей, киберпреступники также используют ИИ для автоматизации и улучшения различных хакерских действий. Алгоритмы ИИ позволяют осуществлять автоматическое сканирование уязвимостей, интеллектуальное обнаружение и злонамеренное использование слабых мест систем, адаптивную разработку вредоносного ПО и т. д. Применение ИИ может способствовать расширению масштабов интенсивности DDoS-атак на бизнес-сайты и онлайн-сервисы, ботнеты, основанные на ИИ, могут координировать огромные объемы вредоносного трафика, перегружая серверы и нарушая бизнес-процессы. В ближайшей перспективе киберпреступность, взявшая себе на вооружение технологии ИИ и МО, может серьезно затормозить дальнейшие цифровизацию и интеллектуализацию транспорта, логистики и экономики в целом. В этих условиях очень важно обеспечить ускоренное развитие технологий кибербезопасности, основанной на применении ИИ, темпы их развития должны существенно опережать темпы создания интеллектуальных киберугроз. На сегодняшний день большинство архитектур сетевой безопасности используют защиту на основе периметра для изоляции внутренних сетей от внешних. Межсетевые экраны, виртуальные частные сети (VPN) и сети демилитаризованных зон предотвращают внешние атаки, создавая периметр безопасности сети. Это позволяет эффективно предотвращать внешние атаки, но предотвратить внутренние атаки сложно, потому что как только злоумышленник нарушит периметр безопасности, дальнейшие незаконные действия не будут затруднены. Кроме того, по мере ускорения развития цифровых технологий, таких как индустриальные сети, интернет вещей и облачные вычисления, количество пользователей и устройств, работающих в сети, и их проблемы безопасности растут в геометрической прогрессии, поскольку периметр сети становится все более и более размытым. Это делает защиту ресурсов более сложной, особенно по мере создания большего количества точек доступа к данным, входов и выходов информации. Таким образом, для предотвращения внутренних атак требуется иная архитектура безопасности.

Архитектура нулевого доверия

Архитектура нулевого доверия (Zero Trust Architecture – ZTA) представляет собой смену парадигмы в стратегии кибербезопасности, основанной на принципах нулевого доверия. Ее основная цель – предотвратить утечки данных и ограничить распространение внутренних горизонтальных угроз. Нулевое доверие (Zero Trust – ZT) не является монолитной архитектурой, а воплощает в себе набор руководящих принципов, применимых к рабочим процессам, проектированию систем и эксплуатации [1, 2]. Эти принципы могут быть использованы для повышения уровня безопасности систем по всем классификациям и уровням их архитектуры. Концепция ZT в кибербезопасности эволюционировала, подтверждая фундаментальный сдвиг в подходе к безопасности. В его основе лежит концепция «никогда не доверяй, всегда проверяй» [3]. Вместо того, чтобы предполагать, что все в корпоративном брандмауэре безопасно, модель ZT работает исходя из предположения о взломе и тщательно проверяет каждый запрос, как если бы он исходил из открытой сети [4]. Модель ZT коренным образом меняет парадигму безопасности, предполагая, что ни одна граница сети не является безопасной по своей сути. Независимо от того, работают ли пользователи в сети организации или за ее пределами, они должны проходить строгую аутентификацию, авторизацию и непрерывную проверку. Модель ZT признает отсутствие традиционной сетевой периферии, предполагая, что ресурсы могут находиться где угодно – локально, в облаке или в гибридных средах. Эта структура согласуется с проблемами, связанными с удаленными сотрудниками, гибридными облачными архитектурами и постоянно меняющимся ландшафтом угроз. В модели ZT применяются политики безопасности на основе контекста, при этом применяется наименее привилегированный контроль доступа и строгая аутентификация пользователей для обеспечения надежной системы безопасности.

Ключевыми принципами работы ZT являются:

1. *Непрерывная проверка.* Это означает, что доступ всегда подтверждается, независимо от местоположения пользователя или ресурса, который он ищет, постоянная проверка гарантирует, что доступ получат только уполномоченные лица.

2. *Ограничение радиуса «взрыва».* В случае внешнего или внутреннего нарушения система нулевого доверия способствует минимизации последствий. За счет разделения доступа и сегментации ресурсов можно сдержать потенциальные последствия инцидента безопасности.

3. *Автоматизированный сбор контекста и реагирование на него.* Модель «Никому не доверяй» не полагается на догадки. Вместо этого она включает в себя поведенческие данные со всего ИТ-стека, проводит идентификацию, анализирует конечные точки, рабочие нагрузки и многое другое. Этот целостный контекст позволяет точно реагировать на события безопасности.

В модели нулевого доверия (Zero Trust Models – ZTM) принцип автоматического сбора контекста и реагирования на него играет ключевую роль. Этот принцип подчеркивает важность сбора в режиме реального времени контекста о поведении поль-

зователя или системы, такого как информация об устройстве, атрибуты пользователя и состояние сети. Контекст является важным элементом ZTM, поскольку он помогает принимать обоснованные решения [5]. Например, когда устройство пытается получить доступ к финансовым данным в сети, необходим контекст, чтобы определить, является ли это сотрудником или угрозой [5]. Местоположение устройства, личность пользователя и данные, к которым он обращается, предоставляют ценную информацию. Однако этого не хватает для того, чтобы определить, должен ли этот сотрудник иметь доступ к этим конкретным данным с этого конкретного устройства или из этого места [5]. Этот недостающий элемент и есть контекст, без него картина риска будет неполной [5].

Еще одним важным событием в области безопасности с нулевым доверием является появление модели зрелости нулевого доверия [2]. Эта модель обеспечивает структурированный подход к оценке и развитию возможностей в области нулевого доверия и состоит из пяти ключевых столпов (рис. 1), каждый из которых представляет собой важнейший аспект внедрения нулевого доверия [2]:

Идентификация. Этот столп сосредоточен на управлении идентификацией, аутентификации и авторизации. Он направлен на то, чтобы только авторизованные пользователи и устройства могли получить доступ к ресурсам. Идентификация описывает атрибуты, однозначно идентифицирующие агентство, пользователя или организацию.

Устройство. Компонент устройства подчеркивает защиту конечных точек и управление их безопасностью. Он включает в себя меры по защите от скомпрометированных устройств и несанкционированного доступа, а также активы (оборудование, программное обеспечение и прошивка) с возможностью подключения к сети.

Сеть/среда. Сеть – это открытая среда коммуникаций, включающая в себя внутренние кабельные сети, беспроводные сети и Интернет. Здесь основное внимание уделяется сегментации сети, микросегментации и видимости сети.

Приложения/рабочие нагрузки. Этот столп направлен на обеспечение безопасности приложений и защиту рабочих нагрузок. Она включает в себя защиту приложений, API и рабочих нагрузок от угроз. Приложения и рабочие нагрузки относятся к системам, программам и службам, выполняемым в различных средах.

Данные. Защита конфиденциальных данных имеет решающее значение. Этот компонент уделяет особое внимание классификации данных, шифрованию и управлению доступом. Данные включают в себя структурированные и неструктурированные файлы, находящиеся в системах, устройствах, сетях и резервных копиях.

Каждый столп также поддерживается платформой, включающей видимость и аналитику, автоматизацию и координацию, а также управление.

Видимость и аналитика относятся к наблюдаемым артефактам, возникающим в результате анализа событий в масштабах всего предприятия и данных.

Автоматизация и оркестрация – это использование автоматизированных инструментов и рабочих процессов для функций реагирования на угрозы безопасности при сохранении контроля и безопасности.

Управление обеспечивает соблюдение политик и процессов кибербезопасности по всем основным направлениям управления рисками.

Уровни зрелости ZTM подразделяются на четыре категории.

Традиционный – это ручная настройка жизненных циклов, статических политик безопасности, а также ручное реагирование и устранение рисков.

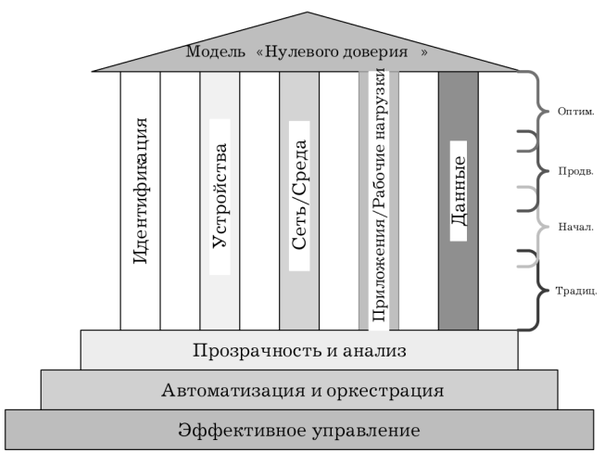

Рис. 1. Пять столпов созревания модели кибербезопасности «Нулевого доверия»

Начальный – это начало автоматизации назначения атрибутов, настройки жизненных циклов, принятия решений по политикам и применения с помощью первоначальных решений, объединяющих компоненты.

Продвинутый – это применимые автоматизированные средства управления для назначения жизненного цикла и конфигурации с координацией между столпами.

Термин «оптимальный» относится к полностью автоматизированным жизненным циклам «точно в срок» и назначению атрибутов на основе автоматизированных/наблюдаемых триггеров.

Логическая архитектура модели «Нулевого доверия»

Концептуальное представление логической архитектуры нулевого доверия [2] отражает основные отношения между компонентами и их взаимодействием (рис. 2). При этом важно отметить, что точка принятия решения о политике разбита на два логических компонента: механизм генерации политик и администратор политик. Кроме того, логические компоненты используют отдельную плоскость управления для взаимодействия, в то же время данные приложений передаются через плоскость данных.

К основным компонентам модели архитектуры «Нулевого доверия» (рис. 2) относятся механизмы формирования политик, включающие механизм формирования политик и администрирования политик и точку применения политик (ТПП).

Механизм формирования политик (МФП) отвечает за окончательное решение о предоставлении доступа к ресурсу для данного субъекта. Плоскость управления использует корпоративные политики, а также информацию из внешних источников в качестве входных данных для работы алгоритма доверия для предоставления, отказа или отзыва доступа к ресурсу. Механизм формирования политик работает в паре с компонентом администратора политики. Механизм формирования политик принимает и регистрирует решение, а администратор политики выполняет решение.

Администратор политик (АП) отвечает за установление и/или отключение канала связи между субъектом и ресурсом (на основе команд от точки принятия решений). Он тесно связан с МФП и зависит от его решения о том, разрешить или отклонить сессию. Если сеанс авторизован и запрос аутентифицирован, АП настраивает точку применения политик так, чтобы разрешить запуск сеанса. Если сеанс отклоняется (или предыдущее одобрение отменяется), АП подает сигнал на ТПП о необходимости завершения соединения. В некоторых реализациях МП и АП могут рассматриваться как единая служба, в нашем случае она делится на две логические составляющие.

Точка применения политик (ТПП) отвечает за активизацию, мониторинг и, в конечном итоге, прекращение связей между субъектом и ресурсом. Точка ТПП связывается с АП для пересылки запросов и/или получения обновлений политики от АП. Этот логический компонент может быть разбит на две части: клиентскую (например, агент на ноутбуке) и ресурсную часть (например, компонент шлюза перед ресурсом, который контролирует доступ) или один компонент портала, который действует как привратник для путей связи. За ТПП находится зона доверия, в которой размещен корпоративный ресурс. Функ-

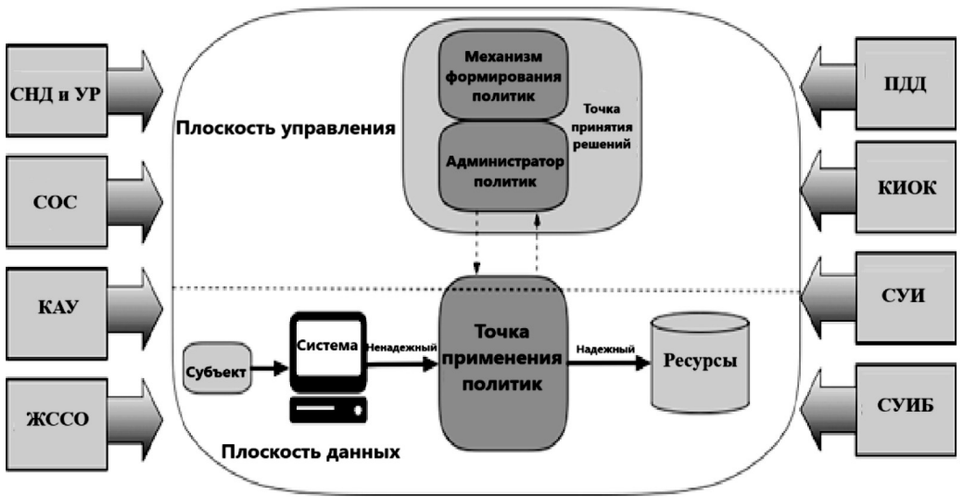

Рис. 2. Основные логические компоненты модели архитектуры «Нулевого доверия»

ционирование основных компонентов поддерживается дополнительными компонентами архитектуры, которые являются источниками входных данных, используемых механизмом политик при принятии решений о доступе.

Дополнительные компоненты модели архитектуры «Нулевого доверия»

К дополнительным компонентам архитектуры (рис. 2) относятся как локальные, так и внешние (т. е. неконтролируемые или специально созданные) источники данных, такие как [2]:

Система непрерывной диагностики и устранения рисков (СНД и УР). Эта система собирает информацию о текущем состоянии корпоративного актива и применяет обновления к конфигурации и программным компонентам. Система СНД и УР предоставляет обработчику политик информацию об активе, делающем запрос на доступ, например о том, работает ли он под управлением соответствующей исправленной операционной системы (ОС), о целостности утвержденных программных компонентов или о наличии неутвержденных компонентов, а также о наличии у актива каких-либо известных уязвимостей. Системы СНД и УР также отвечают за идентификацию и потенциальное применение подмножества политик на некорпоративных устройствах, применяемых в корпоративной инфраструктуре.

Система отраслевого соответствия (СОС) гарантирует, что предприятие остается в соответствии с любым нормативным режимом, под который оно может подпадать (например, требования к информационной системе государственного управления, безопасности транспорта, здравоохранения или банковской сферы). Она включает в себя все правила политики, которые предприятие разрабатывает для обеспечения соответствия.

Каналы аналитики угроз (КАУ) предоставляют информацию из внутренних или внешних источников, которая помогает подсистеме политик принимать решения о доступе. Это могут быть несколько сервисов, которые получают данные из внутренних и/или нескольких внешних источников и предоставляют информацию о недавно обнаруженных атаках или уязвимостях, а также недавно обнаруженные недостатки в программном обеспечении, недавно обнаруженные вредоносные программы и сообщения об атаках на другие активы, доступ к которым обработчик политик захочет запретить из корпоративных активов.

Журналы сетевых и системных операций (ЖССО). Эта корпоративная система объединяет журналы активов, сетевой трафик, действия по доступу к ресурсам и другие события, которые обеспечивают обратную связь в режиме реального времени (или почти в реальном времени) о состоянии безопасности корпоративных информационных систем.

Политики доступа к данным (ПДД) – это атрибуты, правила и политики доступа к корпоративным ресурсам. Этот набор правил может быть закодирован (через интерфейс управления) или динамически сгенерирован механизмом политик. Эти политики являются отправной точкой для авторизации доступа к ресурсу, поскольку они предоставляют основные привилегии доступа для учетных записей и приложений/служб на предприятии. Эти политики должны основываться на определенных миссиях, ролях и потребностях организации.

Корпоративная инфраструктура открытых ключей (public key infrastructure – PKI). Эта система отвечает за создание и регистрацию сертификатов, выданных предприятием ресурсам, субъектам, службам и прило-

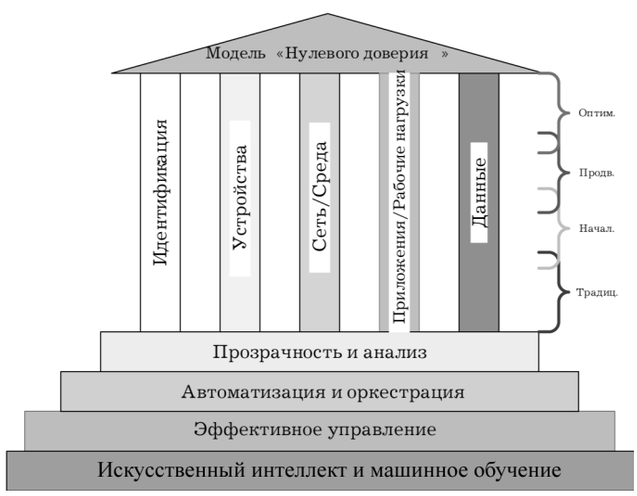

Рис. 3. Искусственный интеллект в модели кибернетической безопасности «Нулевого доверия»

жениям. Также включает в себя глобальную экосистему центров сертификации и федеральную PKI, которая может быть интегрирована или не быть интегрирована с корпоративной PKI, также это может быть PKI, которая не основана на сертификатах X.509.

Система управления идентификаторами (СУИ) отвечает за создание, хранение и управление учетными записями пользователей предприятия и записями удостоверений. Эта система содержит необходимую информацию о субъекте (например, имя, адрес электронной почты, сертификаты) и другие характеристики, такие как роль, атрибуты доступа и назначенные активы. Эта система часто использует другие системы (например, PKI) для артефактов, связанных с учетными записями пользователей. Эта система может быть частью более крупного федеративного сообщества и может включать в себя не сотрудников или ссылки на некорпоративные активы для совместной работы.

Система управления информационной безопасностью и событиями безопасности (СУИБ). Эта система собирает информацию, связанную с безопасностью, для последующего анализа. Затем эти данные используются для уточнения политик и предупреждения о возможных атаках на активы предприятия.

Особенности применения искусственного интеллекта в архитектуре кибербезопасности с нулевым доверием

Развитие технологии обеспечения безопасности на основе реализации концепции «Нулевого доверия» рассматривается как новый эффективный путь повышения защищенности информационных и телекоммуникационных систем и не только их. Конвергенция архитектур безопасности с нулевым доверием и искусственного интеллекта имеет двухсторонний эффект, с одной стороны это способствует дальнейшему усилению кибербезопасности, а с другой стороны формирует новое направление в обеспечении защищенности самого искусственного интеллекта. При этом в основании пяти столпов созревания модели нулевого доверия (рис. 1) формируется новый мощный слой искусственного интеллекта и машинного обучения (рис. 3), который оказывает влияние на все остальные составные части конструкции.

Одним из важных направлений применения ИИ в модели кибербезопасности с нулевым доверием является подсистема управления доступом к идентификационным данным (Identity Access Management – IAM).

ИИ в управлении доступом к идентификационным данным

Фундаментальные принципы реализации архитектуры «Нулевого доверия» включают в себя как аутентификацию, так и контроль доступа. Эти элементы служат механизмами, с помощью которых проверяется личность пользователя и определяются его разрешения на выполнение различных операций с защищенными ресурсами. Цель этой проверки – облегчить пользователю переход из неопознанного состояния в идентифицированное. В быстро разви-

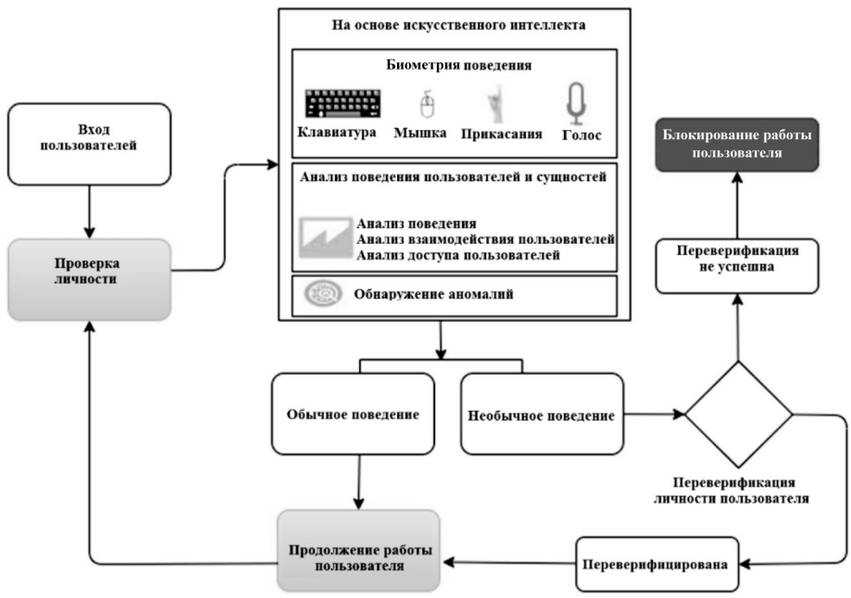

Рис. 4. Интеллектуальная, адаптивная и непрерывная аутентификация пользователя

вающейся области облачных вычислений важность надежных систем IAM значительно возросла из-за роста угроз кибербезопасности и сложных характеристик цифровых удостоверений личности [6]. По мере того, как утечки данных и сложные киберугрозы становятся все более распространенными, организации ищут инновационные решения для укрепления своих стратегий IAM. Одной из таких новых технологий, которая набирает обороты в этой области, является искусственный интеллект, который позволяет машинам учиться, адаптироваться и генерировать новые данные. Искусственный интеллект приобретает все большее значение в IAM, поскольку он все шире применяется в системах защиты облачных вычислений. В настоящее время все шире проводятся исследования по применению методов машинного обучения и развертывание генеративных моделей ИИ (GenAI), что приводит к быстрой трансформации ландшафта IAM, включая внедрение автоматизированной генерации политик защиты [7]. Функционирование систем IAM базируется на четырех ключевых элементах: аутентификации, авторизации, администрировании и аудите. Каждый элемент играет жизненно важную роль в создании безопасной и эффективной системы управления доступом.

Аутентификация

Аутентификация направлена на проверку личности пользователей, гарантируя, что доступ к системе получают только действительные доверенные лица. *Интеллектуальная, адаптивная и непрерывная аутентификация пользователя.* Этот метод предусматривает непрерывное отслеживание поведения пользователей на протяжении всего сеанса и запрашивает повторную аутентификацию в случае обнаружения аномалии. Жизненный цикл систем непрерывной аутентификации начинается с моделирования поведения пользователей во время их взаимодействия со своими устройствами в течение определенного периода времени. После сбора данных они проходят предварительную обработку и сохраняются в наборе данных, содержащем актуальную информацию о моделях поведения пользователей. Для создания точного набора данных для профиля пользователя решающее значение имеет тщательный отбор характеристик или функций из различных измерений устройств, таких как датчики, приложения, коммуникации, экранные жесты и т. д. Наконец, последний шаг включает в себя сравнение текущего использования с устоявшимся поведением пользователя, хранящимся в наборе данных [8]. Искусственный интеллект позволяет обеспечить адаптивную и непрерывную аутентификацию пользователей в сочетании с поведенческой биометрией, обнаружением аномалий и аналитикой поведения пользователей и прикладных сущностей. На рис. 4 показаны этапы интеллектуальной адаптивной и непрерывной аутентификации пользователей.

Когда пользователь входит в систему, его личность проверяется, и система на основе искусственного интеллекта и машинного обучения постоянно отслеживает поведение пользователя. Поведенческая биометрия обнаруживает тонкие различия в том, как человек печатает, пользуется мышью, держит телефон или взаимодействует с сенсорными экранами. Аналитика поведения пользователей и прикладных сущностей анализирует данные из всех возможных корпоративных источников, таких как межсетевые экраны, маршрутизаторы, виртуальные частные сети, решения для управления доступом к удостоверениям, антивирусы, программное обеспечение для защиты от вредоносных программ, обнаружение и реагирование на конечные точки, управление информацией о безопасности и событиями безопасности и каналы аналитики угроз. Система обнаружения аномалий в режиме реального времени выявляет подозрительные отклонения от базового уровня и запрашивает повторную верификацию личности пользователя, если пользователю не удалось пройти повторную проверку, то система блокирует сеанс пользователя. Интеллектуальная система IAM может постоянно оценивать поведение пользователей и динамически изменять требования к аутентификации в зависимости от уровня риска, обеспечивая баланс между безопасностью и удобством пользователя. Кроме того, для получения точного решения используется поведенческий анализ и установление связей между, казалось бы, не связанными между собой действиями, тем самым заблаговременно предотвращая атаки еще до того, как будет нанесен какой-либо вред или боковое перемещение [9].

Распознавание голоса и речи. Голосовая аутентификация – это биометрическая технология, которая проверяет пользователей на основе их отличительных характеристик голоса. Технологии ИИ могут изучать и различать голоса отдельных пользователей, повышая точность голосовой аутентификации и делая ее более устойчивой к попыткам спуфинга. Кроме того, ИИ и МО обладают способностью обрабатывать обширные наборы данных и повышать эффективность за счет автономного обучения и адаптации к изменениям окружающей среды [10]. Например, машины можно обучить различать различные акценты, диалекты, контексты и эмоциональные сигналы [11].

Распознавание лиц – это метод идентификации человеческих лиц с использованием технологий и биометрии, часто путем сопоставления черт лица с фотографий или видео. Система сравнивает текущую информацию с базой данных известных лиц, чтобы определить совпадение. Распознавание лиц уже широко применяется, начиная от аэропортов и мобильных телефонов до учебных аудиторий, социальных сетей и предприятий. Примечательно, что некоторые организации заменили традиционные бейджи безопасности

системами распознавания лиц [12]. Искусственный интеллект может создавать сложные системы распознавания лиц, а также генерировать и анализировать данные о лицах, обеспечивая идентификацию и аутентификацию пользователей на основе их уникальных черт лица.

Обнаружение аномалий. Обнаружение аномалий на основе искусственного интеллекта выявляет отклонения от нормы в данных, тем самым повышая безопасность. Обнаружение аномалий на основе машинного обучения использует непомеченные данные для изучения закономерностей. Такие методы, как кластеризация *k*-средних, решающий лес и одноклассовые методы опорных векторов (SVM) выявляют аномалии и отмечают отклонения от нормального поведения как потенциальные угрозы. В сфере онлайн-безопасности модели и алгоритмы на основе искусственного интеллекта уже сегодня играют решающую роль в упреждающем обнаружении угроз и реагировании на них [13]. Система искусственного интеллекта может непрерывно отслеживать поведение пользователей и выявлять необычные закономерности, указывающие на потенциальные угрозы безопасности или скомпрометированные учетные записи

Авторизация. Авторизация определяет объем доступа, разрешенный каждому аутентифицированному пользователю, гарантируя, что он может получить доступ только к ресурсам, относящимся к его ролям и обязанностям. Можно эффективно использовать ИИ для авторизации пользователей и управления доступом на основе ролей в развертываниях организациями IAM. Искусственный интеллект в этих областях улучшает управление доступом, оптимизирует распределение ролей и повышает общую безопасность, в частности, на основе интеллектуального назначения ролей.

Интеллектуальное назначение ролей включает в себя динамическое назначение ролей пользователям на основе контекстуальных факторов, улучшая управление доступом и безопасность. Искусственный интеллект в интересах назначения ролей может анализировать исторические данные о доступе и поведение пользователей. Система искусственного интеллекта может выявлять закономерности и сходства между пользователями с аналогичными должностными функциями, способствуя более точному и эффективному распределению ролей.

Автоматизированное управление доступом на основе ролей (role-based access control – RBAC) – это метод управления доступом к системам, сетям или ресурсам путем назначения разрешений на основе роли отдельного пользователя в организации. В системах RBAC сотрудникам предоставляется доступ только к информации, относящейся к их должностным обязанностям. Появление новых типов удостоверений, таких как машины, устройства, API, приложения и микросервисы потребовало разработки сложных методов контроля доступа. Традиционные решения RBAC в значительной степени полагаются на ручное обнаружение и моделирование ролей. При этом ручной подход сталкивается со значительными проблемами, пытаясь идти в ногу с постоянно меняющимся ландшафтом идентификаций в современной динамичной среде. Частая смена ролей сотрудников и организационные сдвиги усугубляют эту сложность. Последствия использования исключительно ручных процессов RBAC включают в себя избыточный доступ, «бесхозные» учетные записи и коварное явление, известное как расползание прав, все это усугубляет как внутренние, так и внешние угрозы безопасности. Применение RBAC в архитектуре нулевого доверия с использованием ручных процессов и статических данных исключительно сложно из-за присущих им ловушек. Применение интеллектуального RBAC включает в себя использование ИИ для обнаружения и анализа шаблонов доступа к ролям в масштабах всего предприятия. Таким образом, он выявляет роли с высоким риском и комбинации ролей, определяет высококачественные роли на основе надежных шаблонов доступа, настраивает критерии риска, а также предоставляет рекомендации по ролям и результаты анализа влияний. По сути, RBAC на основе искусственного интеллекта прокладывает путь к созданию более безопасной и адаптивной структуры управления доступом в контексте ZTA. Системы искусственного интеллекта предоставляют возможность постоянно отслеживать поведение пользователей при навигации по сети с соблюдением их прав авторизованного доступа [14]. Кроме того, эти системы обладают способностью обнаруживать аномальное, нелогичное или непредсказуемое поведение [14]. Например, они могут идентифицировать случаи, когда пользователи заходят в системные разделы, которые они обычно не посещают, или извлекать необычно большое количество файлов по сравнению с их обычными шаблонами. Система искусственного интеллекта может автоматически инициировать проверку ролей при обнаружении аномалий или изменений в поведении пользователей, гарантируя, что разрешения на доступ остаются актуальными и соответствуют обязанностям пользователей.

Ролевой майнинг играет ключевую роль в контексте RBAC. Его основная цель заключается в эффективном определении ролей в организации за счет использования разрешений, уже назначенных пользователям. Таким образом, он помогает ИТ-администраторам повысить кибербезопасность, решая проблему пользователей, имеющих привилегии доступа, выходящих за рамки их должностных требований. Процесс интеллектуального анализа ролей с использованием искусственного интеллекта позволяет выявить потенциальные иерархии ролей и зависимости между ними. С помощью этих методов организации могут иден-

тифицировать и классифицировать различные роли, а также связанные с ними привилегии доступа и права на ресурсы. Система искусственного интеллекта может анализировать шаблоны доступа и атрибуты пользователей, чтобы предложить слияние или разделение ролей, уменьшить беспорядок ролей и повысить безопасность.

Динамическое управление доступом. Динамическое управление доступом на основе применения искусственного интеллекта позволяет учитывать контекстуальные факторы и принимать решения по контролю доступа на основе нескольких критических параметров. К ним относятся географическое местоположение пользователя, время суток и конкретное устройство, которое он использует. Искусственный интеллект может динамически корректировать права доступа на основе действий пользователя и контекстуальных факторов, гарантируя, что пользователи имеют соответствующий уровень доступа в любой момент времени.

Администрирование включает в себя эффективное управление учетными записями пользователей, ролями и привилегиями доступа, что снижает сложность IAM для ИТ-администраторов. Искусственный интеллект можно использовать для автоматизации администрирования удостоверений в IAM. В частности, отмена инициализации – это один из важнейших процессов в IAM, который включает в себя отзыв привилегий или доступа из учетных записей пользователей. К этому действию подталкивают различные факторы, в том числе внутренние переводы сотрудников, увольнения или угрозы безопасности. Во время отмены инициализации учетные записи могут быть отключены или полностью уничтожены. При этом пользователи удаляются из всех групп или ролей, с которыми они были связаны.

Автоматизируя регулирование доступа, искусственный интеллект сводит к минимуму ручные операции и снижает риск человеческой ошибки. Искусственный интеллект может поддерживать порталы самообслуживания для управления пользователями, позволяя сотрудникам запрашивать доступ, обновлять свои профили или сбрасывать пароли без вмешательства человека. Искусственный интеллект может обрабатывать эти запросы, проверять их легитимность и автоматически выполнять их, когда это необходимо.

Системы автоматизации запросов на основе ИИ могут динамически применять политики IAM с учетом требований и ограничений, специфичных для пользователя. Таким образом они снижают нагрузку на ИТ-специалистов, которые в противном случае вручную определяли бы «наименьшие привилегии» для каждого варианта использования [15, 16]. Принцип наименьших привилегий в системах безопасности нулевого доверия, когда пользователям предоставляются только минимальные разрешения, необходимые для выполнения их задач, имеет решающее значение для безопасности. Политики IAM на основе искусственного интеллекта обеспечивают точную настройку доступа, повышая эффективность и безопасность. Технологии ИИ позволяют автоматизировать аутентификацию доступа с низким уровнем риска, внимательно отслеживая шаблоны действий пользователей. В результате некоторые задачи управления IAM облегчаются, в то же время защищая пользователей от «усталости от безопасности» [17].

Интеллектуальное управление профилем пользователя (профилирование пользователей) включает в себя извлечение скрытой информации о пользователях на основе наблюдаемых данных, связанных с их действиями или вербальными выражениями [18]. Технологии ИИ позволяют создавать интеллектуальные профили пользователей на основе исторического поведения и шаблонов доступа, и эти профили помогают принимать решения на основе данных о правах доступа и политиках безопасности.

Аудит, управление и соответствие нормативным требованиям. Аудит включает в себя непрерывный мониторинг и запись событий доступа для выявления потенциальных рисков безопасности, обеспечения соблюдения нормативных требований и ведения всестороннего учета действий пользователей. Искусственный интеллект для управления политиками доступа может анализировать политики доступа и исторические данные доступа, чтобы выявлять закономерности и аномалии. На основе этого анализа ИИ может предложить обновления или оптимизацию политик доступа, гарантируя, что разрешения соответствуют принципу минимальных привилегий и требованиям соответствия. Проверки доступа на основе ИИ упрощают соблюдение нормативных требований, устраняют постоянные привилегии и повышают безопасность без необходимости ручного вмешательства. Непрерывный мониторинг предоставляет ценную информацию, включая модели использования, оценки рисков и индикаторы, связанные с должностями и отделами, что помогает принимать обоснованные решения по анализу. Кроме того, система автоматически утверждает доступ с низким уровнем риска и может выполнять отмену доступа. Комплексные отчеты, доступ к результатам сертификации и исправления действия тщательно отслеживаются и легко доступны аудиторам благодаря оптимизированной отчетности.

Управление привилегированным доступом (Privileged Access Management – PAM), усиленное ИИ и машинным обучением, позволяет анализировать шаблоны входа в систему привилегированных пользователей и обучаться на их основе. Устанавливая базовый уровень нормального поведения, эти технологии могут выявлять аномалии, которые могут сигнализировать о рисках безопасности. Одно из наиболее эффективных применений искусственного интеллекта и машинного обучения в PAM заключается в их

предсказательных возможностях. Тщательно изучая исторические данные и выявляя закономерности, они могут прогнозировать потенциальные угрозы безопасности до того, как они материализуются, что позволяет организациям заблаговременно устранять их. В рамках платформы PAM ИИ может оптимизировать и обезопасить процессы повышения привилегий и делегирования. Кроме того, эффективное решение PAM должно включать оценку рисков на основе индивидуального поведения пользователя и исторического контекста. Анализ запросов на доступ в режиме реального времени позволяет гибко принимать решения, выходящие за рамки жестких правил. Кроме того, ИИ позволяет легко интегрироваться с каналами аналитики угроз, расширяя возможности решений PAM по выявлению новых угроз и реагированию на них.

Непрерывный мониторинг соответствия является систематическим процессом, включающим в себя непрерывное сканирование, мониторинг и оценку стандартов безопасности и соответствия требованиям в ИТ-инфраструктуре организации. Такая практика гарантирует, что организация последовательно придерживается нормативных требований и лучших отраслевых практик. Инструменты на основе искусственного интеллекта могут обрабатывать огромные наборы данных в соответствии с постоянно меняющимися нормативными требованиями. Эти усовершенствованные алгоритмы повышают точность, эффективность и адаптируемость, обеспечивая соответствие системным требованиям. Инструменты проверки личности повышают точность и эффективность проверки и мониторинга действий клиентов.

Адаптивные реакции на соответствие требованиям. Системы, управляемые искусственным интеллектом, не являются статичными, они адаптируются и обучаются [19]. Анализируя данные и тенденции соответствия, эти системы постоянно совершенствуются и могут рекомендовать изменения в политиках доступа или требованиях к аутентификации на основе развивающихся нормативных требований или передовых практик безопасности.

Автоматизированное создание контрольного журнала. Искусственный интеллект позволяет создавать подробные контрольные журналы для действий IAM, помогая организациям вести полный учет событий и изменений доступа. Такое автоматизированное создание контрольного журнала упрощает составление отчетов о соответствии нормативным требованиям и снижает административную нагрузку на ИТ-отделы.

Прогнозный анализ соответствия, основанный на искусственном интеллекте, использует исторические тенденции соблюдения нормативных требований для предоставления прогнозной аналитики в отношении потенциальных будущих рисков соответствия. Предвидя критически важные проблемные области, ИИ заблаговременно формирует рекомендации для обеспечения непрерывного соответствия.

Ограничения по применению ИИ и МО в архитектурах безопасности нулевого доверия

Несмотря на то, что исследования, посвященные изучению влияния ИИ на технологии нулевого доверия, дают оптимистичные результаты, важно отметить некоторые ограничения:

Быстро меняющийся ландшафт киберугроз постоянно трансформируется, при этом регулярно появляются новые угрозы и уязвимости. Этот динамизм создает проблему для того, чтобы исследования оставались актуальными. Чтобы быть в курсе меняющихся угроз, требуется непрерывная совместная научная работа ученых и ИТ администраторов, вооруженных технологиями ИИ и МО.

Вариативность в реализации. Эффективность моделей нулевого доверия значительно варьируется в зависимости от их реализации в той или иной организации. На результаты влияют такие факторы, как организационный контекст, инфраструктура и методы работы. Иными словами, сделать окончательные выводы о всеобщей значимости модели нулевого доверия непросто.

Ограниченный объем. Некоторые исследования могут быть узко сосредоточены на конкретных аспектах модели нулевого доверия и, хотя глубина исследований имеет ценность, она может непреднамеренно упустить из виду иные важные аспекты.

Необходимость постоянного сотрудничества. Устранение рассмотренных ограничений требует постоянного сотрудничества между научными кругами, промышленностью и регулирующими органами. Коллективные усилия улучшают общее понимание модели нулевого доверия и ключевой роли в ней технологий ИИ и МО. Несмотря на имеющие место ограничения, продолжение исследований позволит сформировать ориентиры на пути к более надежным и эффективным стратегиям безопасности.

Выводы

Всемирная история убедительно показывает, что избыточное доверие часто создавало, создает и, видимо, будет продолжать создавать угрозы отдельным людям, организациям, странам и миру в целом. Окружающий мир становится с одной стороны все боле совершенным и комфортным, а с другой стороны все более разнообразным, сложным, хрупким и опасным. Появление технологий искусственного интеллекта и машинного обучения рассматривается в качестве технологии, которая позволит эффективно управлять разнообразием, эффективно использовать сложность, обеспечивать надежность и повышать безопас-

ность окружающего мира [20]. Все это может стать возможным, если доверие станет интеллектуально обоснованным. Искусственный интеллект и машинное обучение могут выявлять аномальные закономерности, обнаруживать тонкие отклонения и прогнозировать потенциальные нарушения безопасности, будь то выявление попыток несанкционированного доступа или пометка подозрительного поведения. Модель нулевого доверия совместно с технологиями искусственного интеллекта улучшает прогнозирование инцидентов и эффективность реагирования на них, сокращает время ликвидации последствий и понесенные убытки. При этом политики, связанные с доверием пользователям, атрибутам устройств и контекстуальным факторами становятся более динамичными, адаптивными и эффективными.

Несмотря на огромные перспективы ИИ, этические дилеммы сохраняются, модели ИИ могут выдавать ложные срабатывания (неправильное пометка безобидных действий как угрозы) или ложноотрицательные (отсутствие реальных угроз). Соблюдение правильного баланса имеет решающее значение, чтобы избежать ненужных оповещений и при этом не пропустить критические инциденты. Алгоритмы ИИ могут наследовать предубеждения от обучающих данных. Обеспечение справедливости и прозрачности имеет важное значение для предотвращения дискриминационных результатов. Организации должны находить баланс между использованием возможностей ИИ и защитой индивидуальных прав. В целом нужно подчеркнуть важность и преобразующее влияние ИИ и парадигм нулевого доверия в построении перспективных систем общей и кибернетической безопасности, необходимость проведения непрерывных исследований, этических соображений и стратегий реализации.

Литература

1. Automation and orchestration of zero trust architecture: potential solutions and challenges / Y. Cao, S.R. Pokhrel, Y. Zhu [et al.] // Machine Intelligence Research. – 2024. – Vol. 21 (2). – P. 294–317.

2. Zero trust maturity model / CISA. – 2023. – URL : https://www.cisa.gov/sites/default/files/2023-04/CISA_Zero_Trust_Maturity_Model_Version_2_508c.pdf (дата обращения: 09.07.2023).

3. Federated zero trust architecture using artificial intelligence / M. Hussain, S. Pal, Z. Jadidi [et al.] // IEEE Wireless Communications. – 2024. – Vol. 1 (2). – P. 30–35.

4. What is zero trust architecture? / Sans. – 2023. – URL : https://www.sans.org/blog/what-is-zero-trust-architecture/ (дата обращения: 10.09.2023).

5. When implementing zero trust, context is everything / Security Intelligence. – 2020. – URL : https://securityintelligence.com/posts/when-implementing-zero-trust-context-is-everything/ (дата обращения: 30.07.20230).

6. The impact of generative artificial intelligence on socioeconomic inequalities and policy making / V. Capraro, A. Lentsch, D. Acemoglu [et al.]. – 2023. – ArXiv abs/2401.05377.

7. Futuristic person re-identification over internet of biometrics things (IoBT): Technical potential versus practical reality / N.K.S. Behera, T.K. Behera, M. Nappi [et al.] // Pattern Recognition Letters. – 2021. – Vol. 151, No. C. – P. 163–171.

8. Improving the security and QoE in mobile devices through an intelligent and adaptive continuous authentication system / J.M.J. Válero, P.M.S. Sánchez, L.F. Maimy [et al.] // Sensors. – 2018. – Vol. 18 (11). – P. 3769.

9. An introduction to context-aware security and user entity behavior analytics / Y. Arohan, A. Yadav, A. Pandey [et al.] // International Journal of Advance Research, Ideas and Innovations in Technology. – 2020. – Vol. 6 (5). – P. 27–33.

10. Investigating the prospects of generative artificial intelligence / M. Mandapuram, S.S. Gutlapalli, A. Bodepudi, M. Reddy // Asian Journal of Humanity, Art and Literature. – 2018. – Vol. 5 (2). – P. 167–174.

11. Voice recognition systems in the cloud networks: has it reached its full potential / A. Bodepudi, M. Reddy, S.S. Gutlapalli, M. Mandapuram // Asian Journal of Humanity, Art and Literature. – 2019. – Vol. 8 (1). – P. 51–60.

12. What is facial recognition and how does it work? // Norton [сайт]. – 2023. – URL : https://us.norton.com/blog/iot/how-facial-recognition-software-works (дата обращения: 12.08.2023).

13. Ali, T. HuntGPT: Integrating Machine Learning-Based Anomaly Detection and Explainable AI with Large Language Models (LLMs) / T. Ali, P. Kostakos // arXiv abs/2309.16021. – 2023.

14. Automation and orchestration of zero trust architecture: potential solutions and challenges / Y. Cao, S.R. Pokhrel, Y. Zhu [et al.] // Machine Intelligence Research. – 2024. – Vol. 21 (2). – P. 294–317.

15. Frankish, K. The Cambridge handbook of artificial intelligence / K. Frankish, W.M. Ramsey. – Cambridge : Cambridge University Press, 2014. – 365 p.

16. Mohammed, I. A. The interaction between artificial intelligence and identity and access management: an empirical study / I.A. Mohammed // International Journal of Creative Research Thoughts. – 2021. – Vol. 3, Iss. 1. – P. 668–671.

17. Artificial intelligence and machine learning are transforming Iam // Identity Management Institute [сайт]. – 2021. – URL : https:// ident ityma nagem entin stitu te.org/ artificial intelligence and machine learning are trans forming iam/ (дата обращения: 12.08.2023).

18. Zukerman, I. Predictive statistical models for user modeling / I. Zukerman, D.W. Albrecht // User Modeling and User-Adapted Interaction. – 2001. – Vol. 11 (1). – P. 5–18.

19. Rangaraju, S. Secure by intelligence: enhancing products with AI-driven security measures / S. Rangaraju // EPH – International Journal of Science And Engineering. – 2023. – Vol. 9 (3). – P. 36–41.

20. Комашинский, В. И. Нейронные сети и их применение в системах управления и связи / В.И. Комашинский, Д.А. Смирнов. – Москва : Горячая линия – Телеком, 2003. – 94 с.