Новые технологии: лучшие варианты использования искусственного интеллекта для обнаружения угроз, расследования и реагирования

27 октября 2023 г. - ID G00790125 - 22 минуты чтения

Авторы: Трэвис Ли, Мэтт Милоун, **и еще 2**

Лучшие примеры использования искусственного интеллекта в процессе TDIR подчеркивают сохраняющуюся потребность в вовлечении людей. Лидерам продуктов безопасности следует подкрепить стратегию и планы по интеграции искусственного интеллекта информацией, полученной из интервью с более чем 50 поставщиками услуг безопасности, чтобы расширить возможности аналитиков и повысить дифференциацию решений.

Обзор

Основные выводы

- Более 50% опрошенных поставщиков услуг безопасности утверждают, что уже используют искусственный интеллект на основе контролируемого машинного обучения (ML) для улучшения своих возможностей обнаружения угроз.
- Более 80% поставщиков систем безопасности активно разрабатывают или планируют интегрировать большие языковые модели (LLM) в свои платформы безопасности.
- Хотя внедрение и использование технологий искусственного интеллекта широко распространены и растут, поставщики услуг безопасности делают это с осторожностью из-за опасений по поводу раскрытия клиентских данных, корректности контента искусственного интеллекта и потенциального риска негативного воздействия на действия по исправлению.
- Использование искусственного интеллекта злоумышленниками ускоряет развитие их возможностей для атак и вынуждает разработчиков продуктов безопасности быстрее внедрять технологии искусственного интеллекта для реагирования с помощью улучшенных решений для обнаружения, расследования и реагирования на угрозы (TDIR).

Рекомендации

Безопасность продукта руководителей , ответственных за улучшение TDIR должен рассмотреть представление от опрошенных безопасности поставщиками о различных сценариев использования ИИ:

- Яне из больших наборов данных и могут быть обучены опытными экспертами по безопасности, знающими, как обнаруживать передовые методы атак.....
- Разработайте план, в котором подробно описывается, как и где вы собираетесь использовать генеративный ИИ (GenAI) для оказания помощи вашим клиентам в сортировке угроз, расследовании и реагировании.
- Планируйте использование GenAI в соответствии с проверенными политиками, чтобы убедиться, что выходные данные не раскрывают конфиденциальные данные, а действия, рекомендованные искусственным интеллектом, подтверждены экспертами-людьми.
- Определите, какие области требуют наиболее немедленной интеграции ИИ для противодействия внедрению ИИ субъектами угроз.

Анализ

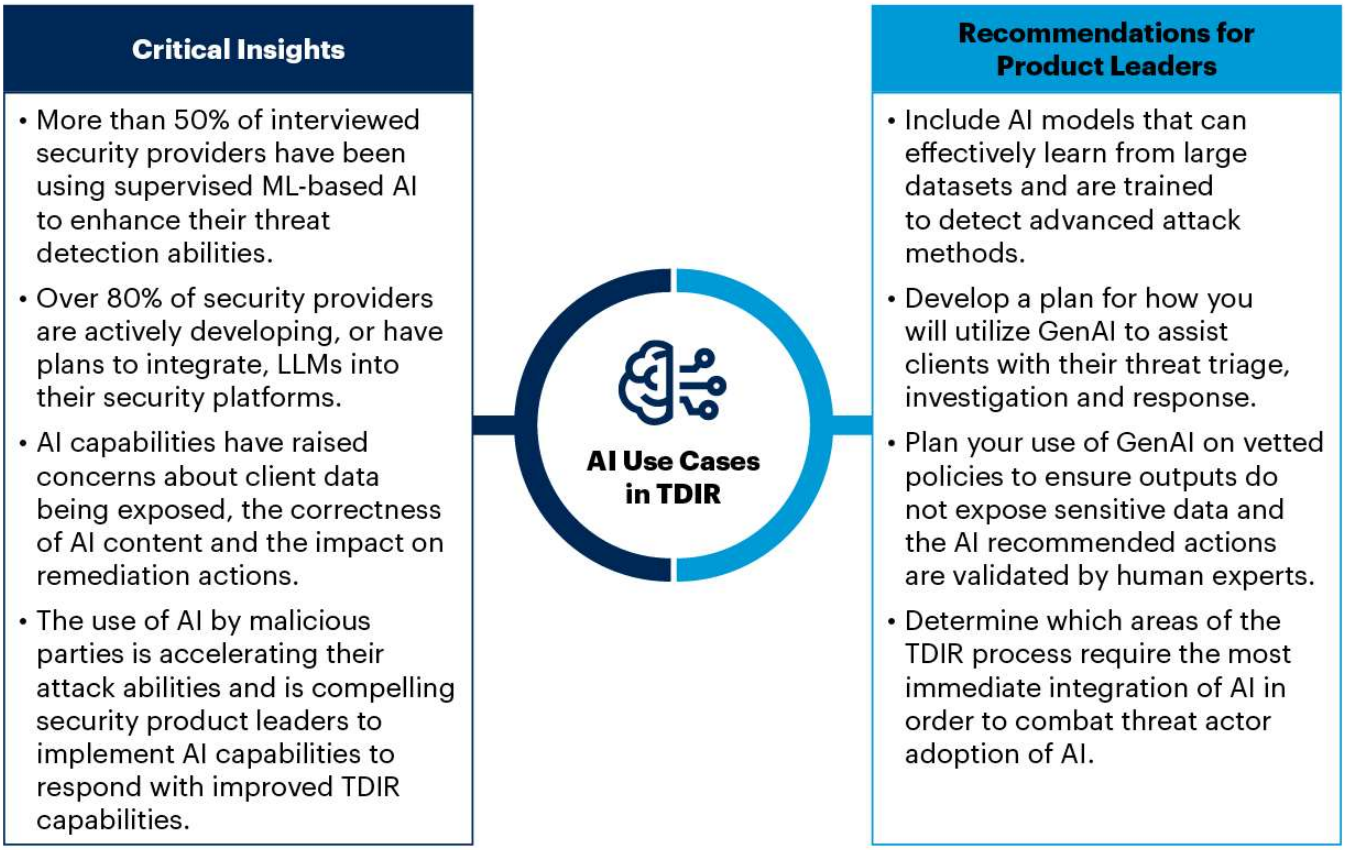
Этот документ был пересмотрен 31 октября 2023 года. Документ, который вы просматриваете, представляет собой исправленную версию. Для получения дополнительной информации смотрите на странице Исправлений на странице gartner.com.

Описание тенденции

Одним из четких выводов, полученных в результате исследования Gartner на основе конкретных случаев (CBR), проведенного в период с мая по август 2023 года, было сосредоточение внимания на том, как опрошенные поставщики планируют использовать искусственный интеллект для улучшения процесса TDIR и оказания помощи аналитикам (см. Примечание 1 к нашей методологии CBR). Проект CBR включал интервью с более чем 50 поставщиками услуг безопасности и технологий. Это исследование дает представление о том, сколько из этих поставщиков продвигают свои услуги или продукты TDIR с помощью искусственного интеллекта. На рисункее 1 представлены важнейшие выводы и рекомендации в данной области.

Рисунок 1: Критическая информация об использовании искусственного интеллекта для обнаружения, расследования угроз и реагирования на них

Critical Insights for AI Use in Threat Detection, Investigation and Response



Source: Gartner
790125_C

Gartner

Ключом к такому акценту на ИИ является то, что службы безопасности и поставщики технологий продолжают использовать новые методы ИИ и подходят к ним, руководствуясь принципом “действовать быстро” и в то же время “быть в безопасности”. (Смотрите 3 императива генеративного ИИ для ИТ-директоров в 2023 году.)

Одна из причин, по которой рбродяги “продвигаются быстро” , заключается в том, что их подстегивают злоумышленники, которые активно используют искусственный интеллект для разработки новых методов атаки. Если они не будут действовать быстро, то не смогут угнаться за расширяющимися возможностями злоумышленников, которые используют искусственный интеллект для поспешной разработки.

Скорее всего, вы уже видите влияние:

- Труднообнаруживаемые варианты вредоносных программ
- Усовершенствованный контент для фишинга и компрометации деловой электронной почты
- Поиск уязвимостей, которые можно использовать, во всех областях цифровой экосистемы

В то же время в ответах опрошенных поставщиков услуг рассказывалось о том, как они внедряют или планируют внедрять искусственный интеллект, вдумчиво рассматривая “быть в безопасности” этику защиты конфиденциальных данных клиентов и компании. Если они не будут действовать с осторожностью, потенциальное раскрытие клиентских данных и действия, предпринятые в связи с некорректной галлюцинацией искусственного интеллекта, могут повлиять на них самих и их клиентов. Это требует тщательной разработки политик, гарантирующих, что модели искусственного интеллекта не передадут данные непреднамеренным или злонамеренным участникам и что подготовленные наборы данных защищены от несанкционированного доступа.

Учитывая императивы “действовать быстро” и “быть в безопасности”, провайдеры дополняют свои текущие модели, разрабатывая новые способы использования искусственного интеллекта для повышения эффективности TDIR.

В интервью было подчеркнуто, что поставщики систем безопасности позиционируют свои текущие и новые решения с поддержкой искусственного интеллекта, чтобы дифференцировать свои возможности по улучшению результатов обнаружения за счет повышения точности обнаружения истинно положительных инцидентов и значительного снижения частоты ложноположительных и ложноотрицательных срабатываний. Эти улучшения не только повышают эффективность их работы, но и, что более важно, повышают ценность для их клиентов, поскольку обеспечивают надлежащее обнаружение угроз и их быстрое устранение.

Они также ориентированы на варианты использования, которые:

- Повышениеточностиобнаружения
- Обеспечивает более быстрое время отклика
- Яповышаю производительность труда сотрудников
- Повышение удовлетворенности работой аналитика

Критический анализ: Поставщики услуг электронной безопасности все больше полагаются на контролируемую ML для расширения своих возможностей обнаружения угроз.

В течение многих лет поставщики систем безопасности интегрировали искусственный интеллект в свои продукты и процессы обнаружения угроз, что позволяет им выявлять аномалии и закономерности, масштаб которых превосходит человеческие возможности. Фактически, более 50% опрошенных поставщиков услуг безопасности и технологий утверждают, что уже используют контролируемый искусственный интеллект на основе ML в сочетании с контролируемым обучением для улучшения своих возможностей обнаружения угроз.

Благодаря интеграции телеметрии угроз с конечных точек, сети, электронной почты, облака, идентификации, Интернета вещей, операционных технологий (OT) и других областей безопасности в платформах расширенного обнаружения и реагирования (XDR) и решениях для управления информацией о безопасности и событиями (SIEM) становится возможным объединять эти сигналы для эффективного обнаружения угроз. Однако многие точечные решения, такие как endpoint detection and response (EDR), network detection and response (NDR), cloud detection and response (CDR), а также предложения по управляемой безопасности, выдают оповещения только на основе контекста из их соответствующего domain insight, а не с помощью интегрированной модели сигналов.

Цель использования искусственного интеллекта - ускорить процесс разработки и уточнения моделей обнаружения, чтобы основывать обнаружение на интеграции обширных наборов данных и поведенческих паттернов. (См. Новые технологии: безопасность — улучшите обнаружение угроз и реагирование на них с помощью поведенческих индикаторов на основе искусственного интеллекта.) Хотя модели обнаружения ИИ могут анализировать большие наборы данных лучше, чем люди, эти модели должны находиться под наблюдением человека, чтобы убедиться, что они обучены надлежащим образом и выходные данные проверены. Существует ошибочное представление о том, что ИИ может жить и развиваться сам по себе — текущие модели не демонстрируют этой способности. Продуктовые лидеры должны инвестировать в квалифицированный персонал (специалистов по обработке данных и опытных аналитиков), который может постоянно направлять развитие моделей искусственного интеллекта и соответствовать методам злоумышленников.

Искусственный интеллект может позволить поставщикам технологий интегрировать телеметрию безопасности в:

- Расширьте возможности обнаружения новых методов атак
- Разработка моделей обнаружения, которые идентифицируют подозрительные устройства или действия пользователей
- Определение опасного поведения до нарушения
- Разработка новых средств защиты

И искусственный интеллект может позволить поставщикам услуг:

- Поддержите большее количество клиентов текущим уровнем аналитики
- Позволяет аналитикам сосредоточить свое внимание на том, что важнее всего, а не на рутинной сортировке и обобщении данных
- Экономьте время аналитиков, избегая огромного количества ложных срабатываний
- Быстрое расследование инцидентов для принятия более быстрых и эффективных ответных решений
- Рекомендует принимать ответные решения на основе опыта принятия ответных мер в соответствии с передовым опытом

Краткосрочные последствия для продуктовых лидеров

Руководители продуктов служб безопасности должны действовать оперативно, чтобы использовать весь потенциал данных из всей цифровой экосистемы для максимального повышения ценности решений на основе искусственного интеллекта для своих клиентов. Интеграция телеметрии из нескольких источников данных предоставляет решениям искусственного интеллекта новые возможности для контекстуализации данных более разнообразными способами и корреляции с многодоменными угрозами. Это, в свою очередь, повышает точность и расширяет спектр предоставляемых клиентам услуг.

Механизмы обнаружения угроз на основе искусственного интеллекта должны постоянно контролироваться и обучаться надежными экспертами, чтобы гарантировать правильное обнаружение инцидентов и совершенствовать модели по мере развития методов злоумышленников. Сегодняшние модели будут менее эффективны в течение нескольких месяцев, поскольку киберпреступники используют новые методы и продолжают использовать искусственный интеллект для собственного продвижения. Чтобы оставаться актуальной и эффективной, требуется участие персонала для обеспечения постоянного сбора информации об угрозах и проверки результатов.

Рекомендуемые действия на следующие шесть-18 месяцев

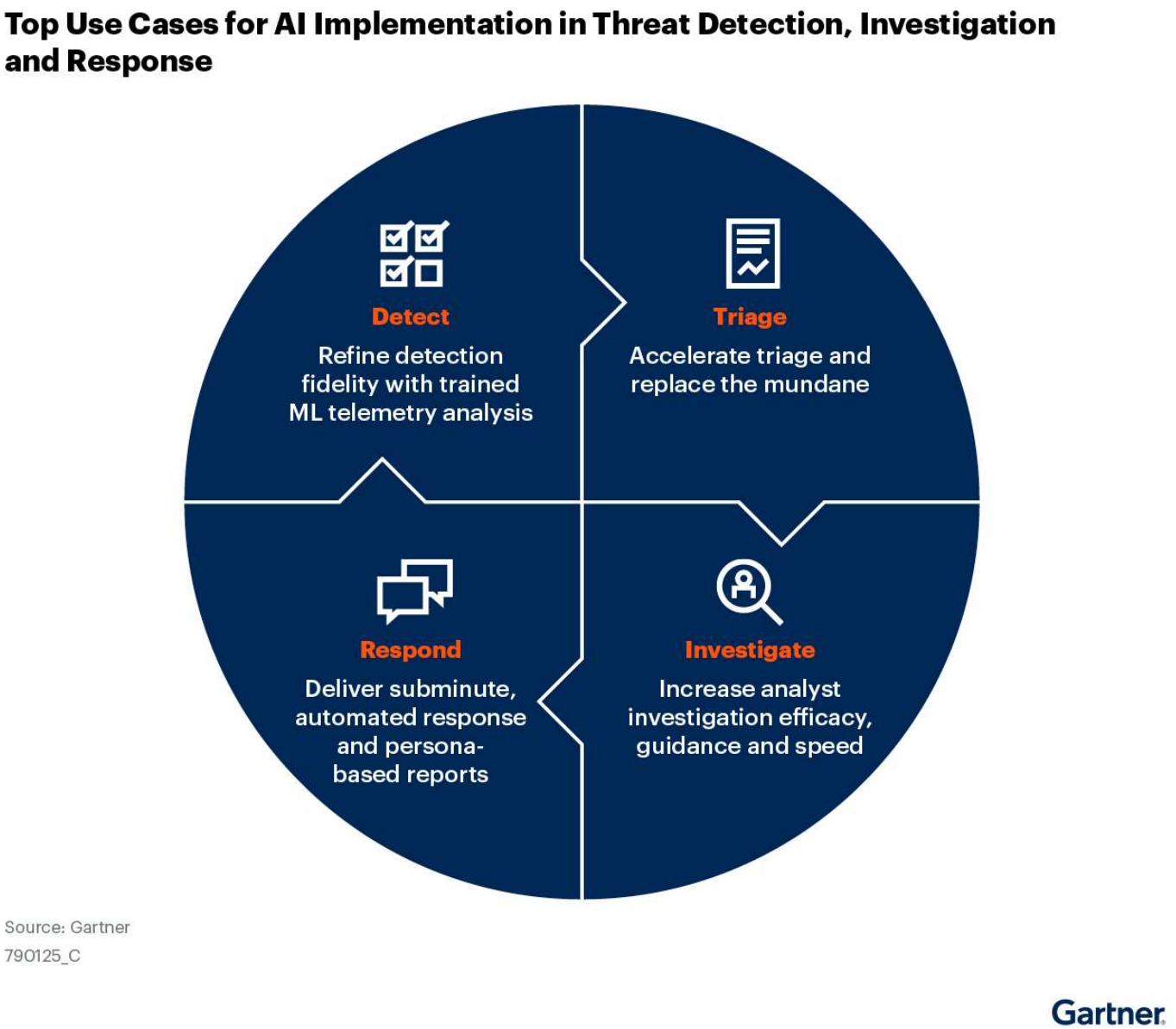
- В первую очередь сосредоточьтесь на разработке или интеграции существующих алгоритмов обнаружения, которые анализируют телеметрию из как можно большего числа источников данных, имеющих отношение к безопасности, чтобы предоставлять клиентам более точные данные о безопасности.
- Выстраивайте процессы, нанимайте или перераспределяйте сотрудников и приобретайте навыки для непрерывного контроля и обучения результатов работы механизма обнаружения искусственного интеллекта для обнаружения новых методов атак и обеспечения минимизации ложных срабатываний / негативов.
- Постоянно инвестируйте в совершенствование модели обнаружения с помощью искусственного интеллекта благодаря расширяющемуся анализу угроз и доказательствам новых методов атак.

Критическая внутренняя информацияборьба: поставщики услуг безопасности используют (или планируют использовать) GenAI и LLMs для расширения возможностей аналитики.

Более 80% поставщиков систем безопасности, опрошенных Gartner, заявили, что они либо в настоящее время разрабатывают, либо планируют использовать GenAI и LLMs для расширения возможностей аналитики.

После усовершенствования механизмов обнаружения он „Большинство компаний, с которыми Gartner провела интервью, заявили, что они либо планируют, либо уже начали проекты, ориентированные на эти области. Смотрите Рисунок 2 для обзора основных вариантов использования.

Рисунок 2: Основные варианты использования искусственного интеллекта для обнаружения, расследования и реагирования на угрозы



Сортировка

Несколько опрошенных компаний отметили, что при работе с новыми клиентами они обнаружили, что клиенты часто тратят больше времени на сортировку и изучение предупреждений, а не на реагирование на угрозы или поиск угроз и / или улучшение процессов. И во многих

случаях они тратят время на проверку предупреждений, которые позже оказываются ложноположительными.

Лучший пример использования сортировки оповещений, которым поделились опрошенные поставщики услуг безопасности, заключается в использовании искусственного интеллекта (ML и GenAI), чтобы стать более эффективными при сортировке оповещений и отфильтровывании ложноположительных оповещений или с низким приоритетом. Основываясь на дорожных картах и планах многих поставщиков систем безопасности, Gartner прогнозирует, что к 2026 году процесс сортировки оповещений будет почти полностью осуществляться машиной под наблюдением человека.

Это не только позволит поставщикам услуг безопасности и оперативным группам внутренней безопасности сосредоточить свой ограниченный персонал на расследованиях с участием людей, но и сократит количество оповещений, требующих расследования.

Расследование

Искусственный интеллект может поддержать и без того напряженных и перегруженных аналитиков по безопасности, предоставляя контекст и предлагая варианты повышения эффективности процессов расследования.

Лучшие варианты использования для расследования инцидентов, предложенные опрошенными компаниями, направлены на то, чтобы дать возможность команде безопасности:

- Уделите время более высокоуровневому анализу расследования, а не сбору подробностей об инциденте
- Ускорьте процесс расследования, чтобы получить рекомендацию по ответу в течение 20-30 минут, а не часов

Первый вариант использования для расследования заключается в том, что GenAI curate преобразует сведения о сортированных оповещениях в сведения об инцидентах, “готовых к расследованию”, и экономит часы ручной исследовательской работы. Искусственный интеллект может собирать данные из больших наборов данных и по нескольким каналам в готовые к расследованию кейсы, чтобы аналитики могли сосредоточить свое время на более высоком уровне мышления. Опрошенные поставщики услуг планируют автоматизировать ручные и повторяющиеся исследовательские задачи, которые в противном случае были бы возложены на аналитиков. Такая информация, как задействованные идентификационные данные, исходящие IP-адреса и устройства, внешние или внутренние каналы атаки, затронутые организации, вовлеченные пользователи и данные, к которым был получен доступ, может быть быстро обработана GenAI в случае инцидента. Это не только ускоряет время проведения расследования, но и повышает удовлетворенность работой аналитиков и снижает их эмоциональное выгорание за счет устранения рутинных задач.

Второй популярный вариант использования заключается в оказании помощи экспертам по безопасности в процессе анализа путем предоставления рекомендаций по сложным вопросам в доступной для аналитиков форме. Многие опрошенные компании анонсировали помощников по кибербезопасности на базе GenAI, которые преобразуют поисковые запросы аналитиков в формальный синтаксис поиска, что позволяет им быстро собирать релевантную информацию из больших наборов данных для ускорения процесса расследования. Помощь в расследовании позволит аналитикам выполнять поиск по многоуровневым оповещениям о событиях, командным строкам, сценариям злоумышленников и возвращать рекомендации по устранению угроз, полученные из GenAI.

Один из опрошенных поставщиков объединил эксперта по реагированию на инциденты с специалистом по обработке данных, чтобы определить, каким должен быть наилучший подход к реагированию, основанный на конкретном случае инцидента. Такое сочетание обеспечит эффективную разработку модели, а постоянный мониторинг модели искусственного интеллекта экспертами по безопасности обеспечит необходимое обучение. Такой подход к оказанию помощи аналитикам повысит эффективность их расследований и обеспечит лучшие результаты реагирования для их клиентов.

Подход помощника аналитика на базе GenAI поможет начинающим аналитикам учиться и проводить исследования без технических знаний, которые в настоящее время требуются для выполнения анализа более высокого уровня, такого как выполнение SQL-запросов, скриптов, команд powershell и многого другого. Помощь аналитиков позволит аналитикам задавать вопросы по модели или использовать существующие подсказки, такие как:

- Существуют ли какие-либо другие подобные оповещения?
- Кто мои самые рискованные пользователи?
- Можете ли вы рассказать мне больше об этом IP и хэше?
- Какие политики следует включить для защиты от этой новой угрозы?
- Встречается ли подобное где-нибудь еще в окружающей среде?

Эти основанные на исследованиях примеры использования GenAI позволят аналитикам:

- Получите четкое представление о подготовленной информации об инцидентах, которую им не нужно собирать, и позвольте им немедленно приступить к анализу
- Задавайте интуитивно понятные вопросы, которые не требуют обширной подготовки и подкрепляются опытом аналитиков
- Определите наилучший ответ на основе анализа, проведенного человеком, и / или с помощью рекомендаций GenAI model как можно быстрее

Ответить

После завершения расследования инцидента под руководством человека и определения ответных действий или одобренного руководства GenAI, рекомендованного для реагирования, целью ИИ является ускорение выполнения ответных действий. Предварительно разработанные рабочие процессы реагирования могут выполняться механизмом принятия решений с помощью искусственного интеллекта, способным выбирать наименее вредный или наиболее эффективный курс действий в зависимости от склонности к риску и других внешних факторов.

Целью усовершенствований в процессе TDIR является принятие ответных мер до того, как злоумышленник сможет совершить нарушение и вырваться наружу или нанести вред окружающей среде. Среднее время, затрачиваемое злоумышленником на прорыв, сократилось с 84 минут в 2022 году до 79 минут в 2023 году. [1] Это среднее значение требует обнаружения угроз и реагирования на них, чтобы обеспечить автоматическое реагирование на атаки продолжительностью менее 60 минут.

Опрошенные поставщики услуг работают над возможностью предоставлять аналитикам рекомендации о том, что делать в конкретном случае, над которым они работают, на основе того, что эксперты-аналитики использовали в прошлом. Это экспертное руководство используется для обучения модели искусственного интеллекта тому, что рекомендовать при реагировании на инциденты того же типа.

Помощникам по искусственному интеллекту будет предоставлена возможность предоставлять рекомендуемые действия, которые позволят изолировать и сдержать возникающую атаку.

Из предоставленных ответов на интервью следует, что обычные действия se заключаются в том, чтобы:

1. Определите разработанные сборники автоматизированных ответов для “одобренных клиентом” вариантов использования, когда определенные поставщиком действия разрешены без одобрения или участия клиента.
2. Координируйте рекомендации по реагированию для “ориентированных на клиента” вариантов использования, которые требуют, чтобы клиент рассмотрел и либо одобрил, либо отклонил рекомендуемое ответное действие. В случае одобрения сценарий может быть выполнен. В случае отклонения клиенту предоставляется рекомендация по ответу, но он берет на себя ответственность за выполнение самостоятельно.
3. Определите варианты использования “ответа только для клиента”, когда клиент не разрешает никаких действий в ответ поставщику, но будет владеть действием в ответ. Эти случаи должны включать рекомендацию поставщика услуг, предложенную искусственным интеллектом и проверенную аналитиками по инциденту, и объяснение прогнозируемого воздействия подтвержденной угрозы.

ИИ должен превратить реагирование на инциденты из механического сценария “если это произойдет, предпримите это действие” в сценарий динамической координации действий, при котором обогащенные данные и человеческий интеллект в сочетании с анализом ИИ проверенных ответных действий предоставляют аналитикам одобренные рекомендации по реагированию.

Еще одним планируемым вариантом использования реагирования на инциденты с помощью искусственного интеллекта является создание отчетов о результатах после инцидента на основе персоны получателя. GenAI может создавать полностью объяснимые отчеты об инцидентах, которые предоставляют получателям легко читаемую визуальную информацию, рекомендации и подробные сведения об инцидентах в зависимости от личности предполагаемого получателя. Эти отчеты необходимо будет просмотреть аналитикам или экспертам перед распространением предполагаемому получателю, но даже с учетом времени проверки это сэкономит аналитикам часы на одно расследование.

Краткосрочные последствия для продуктовых лидеров

Благодаря разработке моделей искусственного интеллекта, использующих GenAI / LLMs, платформы TDIR смогут использовать контекст для расширенных наборов данных по всем доступным каналам безопасности. ИТ-отдел сможет помогать аналитикам в процессе расследования и поиска информации, а также надлежащим образом реагировать на инциденты как можно быстрее. Это позволит производственным лидерам дифференцировать свои продукты на конкурентном рынке TDIR.

Продуктовые лидеры, стремящиеся внедрить искусственный интеллект в процесс обнаружения угроз, должны понимать, что конечная цель заключается не в замене участия человека, а в повышении эффективности работы экспертов по безопасности. Также следует сосредоточиться на повышении удержания аналитиков, сделав их работу более увлекательной и менее рутинной.

Рекомендуемые действия на следующие шесть-18 месяцев

- Помните, что аналитик останется жизненно важным элементом процесса обеспечения безопасности, и его работа будет оптимизирована с помощью искусственного интеллекта.
- Убедитесь, что для обучения моделей искусственного интеллекта точности и эволюции с течением времени используютсяалгоритмов, а также правильно используйте помощь искусственного интеллекта для творческого реагирования на возникающие угрозы.
- Разработайте систему поддержки аналитиков с помощью ограждений, обеспечивающих правильный ответ в зависимости от цели запроса. Например, запрос “Является ли этот IP-адрес вредоносным?” должен направляться не в LLM, а в телеметрию, хранящуюся в озере данных. Тем временем вопрос “Как мне ответить на эту атаку Cobalt strike?” следует направить в магистратуру.

Критическое значениensight: возможности искусственного интеллекта вызывают опасения по поводу раскрытия клиентских данных, корректности контента искусственного интеллекта и потенциального риска негативного воздействия на действия по исправлению положения.

Поскольку вендоры систем безопасности и поставщики услуг стремятся интегрировать возможности искусственного интеллекта в свои возможности TDIR, они придерживаются подхода "будь в безопасности" к своим дорожным картам и действиям.

Из интервью со многими компаниями на эту тему следует, что основные области внедряемых подходов "будь в безопасности" заключаются в следующем.

Убедитесь, что у вас действует определенная политика.

Как отмечалось в первых двух важных выводах выше, есть действие "действуй быстро", которое должно быть частью вашего плана интеграции искусственного интеллекта. Безусловно, есть области, над внедрением которых вам следует работать уже сейчас и как можно быстрее. Но при этом все равно следует руководствоваться хорошо продуманными, прошедшими внутреннюю проверку и одобренными политиками в отношении того, как вы будете использовать данные в модели искусственного интеллекта. Эта политика должна контролироваться советом по управлению данными, представляющим различные бизнес-подразделения вашей компании. Руководящий совет должен возглавляться лицом, несущим в вашей компании окончательную ответственность за кибербезопасность, и в него должны входить руководители служб безопасности, ИТ, маркетинга, продаж и отдела кадров.

Эта политика должна быть разработана до вашего участия в интеграции искусственного интеллекта и должна пересматриваться ежеквартально в ближайшей перспективе, поскольку вы хотите "быстро продвигаться" в определенных областях вашей интеграции искусственного интеллекта. Со временем к нему можно будет возвращаться раз в полгода.

Убедитесь, что данные, доступные движку GenAI, не имеют доступа к конфиденциальным данным.

Критически важным для вашей политики в отношении данных искусственного интеллекта является защита и доступность конфиденциальной информации. Необходимо в обязательном порядке не допускать публичного доступа к клиентским, корпоративным или другим конфиденциальным данным для любого решения GenAI / LLM, которое вы решите использовать. Даже если вы решите использовать решение с частным доступом, вы должны убедиться, что конфиденциальные данные и корпоративный IP недоступны. Используя этот подход с нулевым доверием, вы должны использовать только те данные, которые необходимы для эффективной реализации модели GenAI. Это усиливает необходимость в обученных экспертах, которые постоянно оценивают, какие данные доступны, и дают рекомендации относительно того, что следует или не следует использовать.

Существует множество моделей GenAI/ LLM, которые могут быть реализованы в частной архитектуре, не позволяющей другим пользователям получать доступ к данным, которые вы сделали доступными для модели. Этот частный подход рекомендуется в целях обеспечения безопасности.

Запретить автоматическую публикацию из GenAI systems.

Большинство опрошенных компаний, критически относящихся к подходу "быть в безопасности" при использовании искусственного интеллекта, определили, что им требуется проверка экспертом по предметной области любого контента, который будет предоставлен клиенту или кому-либо за пределами их компании. Использование такого контента GenAI без экспертной оценки является безответственным, особенно в случаях использования, когда контент потенциально может повлиять на реакцию клиента или аналитика в среде клиента. В ходе проверки также будет проверено, не были ли повреждены обучающие данные или алгоритм искусственного интеллекта.

Ближайшиепоследствия для продуктовых лидеров

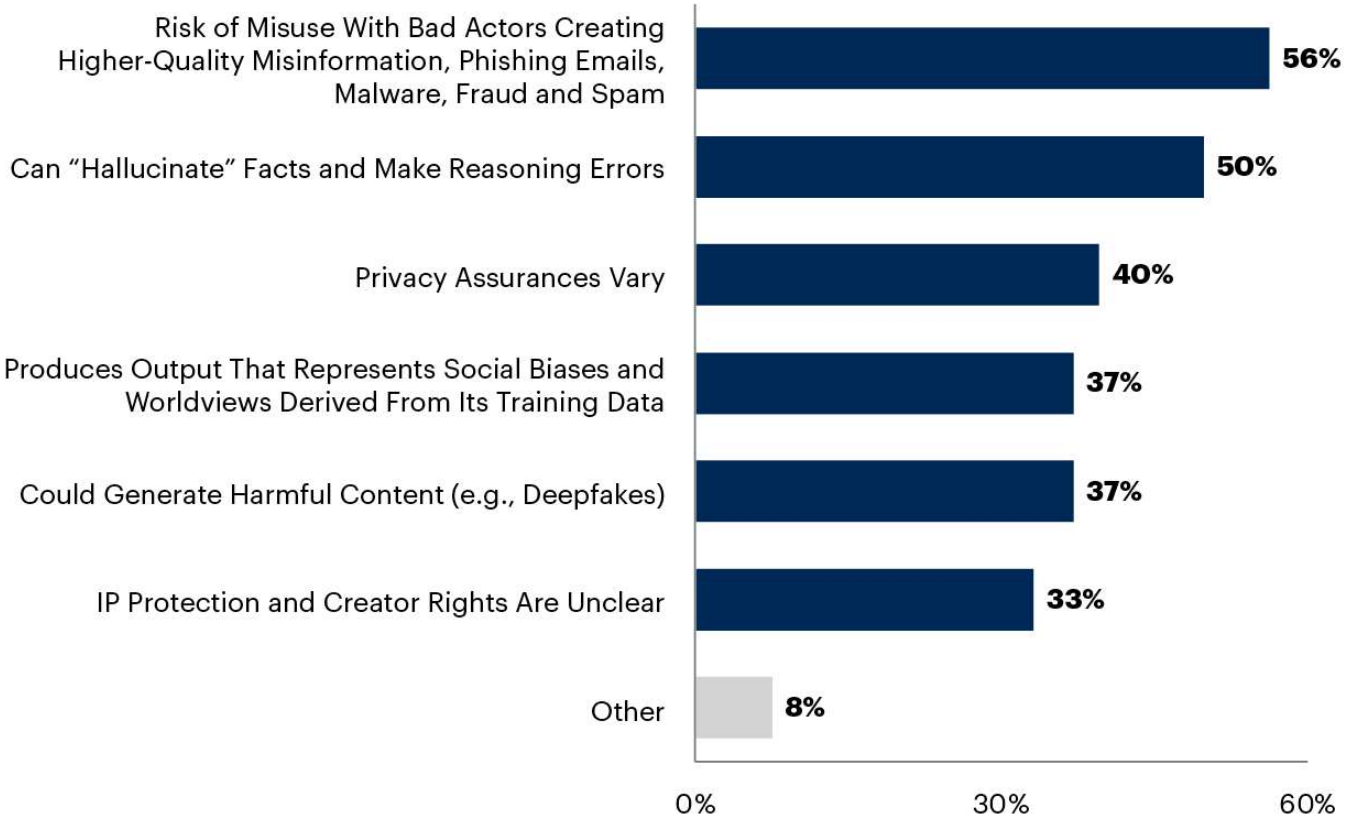
Продуктовым лидерам необходимо проявлять бдительность при разработке, утверждении и соблюдении политики в области искусственного интеллекта, которая обеспечивает защиту клиентских и корпоративных данных и гарантирует действия, проверенные человеком.

Короче говоря, конфиденциальность, точность и безопасность должны быть в центре внимания любой интеграции искусственного интеллекта.

В обзоре по генеративному ИИ Gartner за 2023 год перечислены ответы ИТ-директоров, которые их больше всего беспокоят в связи с использованием GenAI в их компании. Обеспечение конфиденциальности указано в списке как третья по значимости задача. Смотрите Рисунок 3, где ИТ-директор оценивает потенциальные риски, связанные с ИИ.

Рисунок 3: Потенциальные риски, связанные с генеративным ИИ

Potential Risks of Generative AI
Multiple Responses Allowed



n = 78; CIOs, excluding "not sure"

Q. Finally, which of these potential risks of generative AI are you most concerned about in the context of your enterprise?
Source: 2023 Gartner CIO Generative AI Survey
Note: Respondents are Gartner's Research Circle members and external participants.
790125_C

Gartner

Рекомендуемые действия на следующие шесть-18 месяцев

- Обеспечьте прозрачность внедрения вашего искусственного интеллекта, чтобы добиться уверенности клиентов в том, что искусственный интеллект заслуживает доверия, беспристрастен и поддается аудиту.
- Убедитесь, что конфиденциальные данные не включены в доступные данные, чтобы они не стали общественным достоянием.
- Убедитесь в точности ответов, чтобы клиенты и аналитики были уверены в результатах GenAI.

Критический вывод: использование искусственного интеллекта субъектами угроз вынуждает лидеров в области продуктов безопасности внедрять возможности искусственного интеллекта для улучшения своих возможностей TDIR.

Технологические компании - не единственные, кто "быстро" использует потенциальные преимущества новых достижений в моделях искусственного интеллекта. Злоумышленники активно ищут способы его использования и воплощения в жизнь.

Мы уже видели случаи, когда программист разрабатывал WormGPT, бот, похожий на ChatGPT, который может помогать покупателям создавать фишинговые сообщения и вредоносное ПО. Также существует группа, которая создала вредоносного чат-бота под названием FraudGPT. Этот чат-бот с искусственным интеллектом использует возможности генеративных моделей для создания вводящего в заблуждение и реалистичного контента, который приводит к эксплуатации ничего не подозревающих пользователей.

некоторые ключевые области, в которых злоумышленники в настоящее время используют GenAI, следующие:

- Доступность — Есть преступники, у которых нет технических навыков для написания сложного кода, но которые с помощью доступных в настоящее время и планируемых к разработке инструментов GenAI могут повысить свои способности и усовершенствовать свою тактику.

- например, фишинговые атаки будет сложнее распознать из-за усовершенствования, которое инструменты GenAI предоставляют ранее менее способным злоумышленникам.
- Скорость разработки — Атаки будут увеличиваться по мере того, как киберпреступники увеличивают скорость своих атак и воздействие с помощью инструментов GenAI, которые не имеют ограждений, препятствующих их разработке со злым умыслом.

Краткосрочные последствия для продуктовых лидеров

Производственным лидерам необходимо действовать быстро, чтобы противодействовать атакам киберпреступников с помощью искусственного интеллекта. По мере того, как киберпреступники будут разрабатывать все более изощренные векторы атак, организации будут становиться уязвимыми для неизвестных угроз, которые останутся незамеченными современными системами обнаружения. Искусственный интеллект позволит поставщикам технологий совершенствовать свои существующие решения, чтобы иметь возможность обнаруживать и предотвращать неизвестные угрозы, включая уязвимости, которые еще предстоит выявить или исправить поставщикам программного обеспечения.

Системы искусственного интеллекта могут обрабатывать и понимать огромные объемы данных, недоступные специалистам по безопасности. Таким образом, организации могут автоматически обнаруживать новые угрозы среди огромных объемов данных и сетевого трафика, которые могли бы остаться незамеченными традиционными системами.

Рекомендуемые действия на следующие шесть-18 месяцев

Следуйте приведенным выше рекомендациям, и как только внутренние политики в области искусственного интеллекта будут определены и одобрены, быстро переходите к интеграции моделей GenAI / LLM, которые позволят:

- Обнаружение контента GenAI и новых методов атаки.
- Упростите и ускорьте работу аналитиков по расследованию атак и реагированию на них.
- Укрепите доверие клиентов к тому, что вы используете их данные.

Доказательства

[1] Некуда спрятаться: отчет об отслеживании угроз за 2023 год, Crowdstrike.

Опрос ИТ-директора Gartner по генеративному ИИ в 2023 году. Этот опрос проводился онлайн с 16 мая по 15 июня 2023 года, чтобы получить базовое представление о том, что ИТ-директора думают о GenAI и в чем, по их мнению, будет заключаться их роль. Всего в нем приняли участие 80 ИТ-директоров. Семьдесят шесть были членами Исследовательского круга ИТ-директоров Gartner, группы, управляемой Gartner, и четверо были из внешнего опроса, ссылка на который была предоставлена через социальные каналы и контакты аналитиков. Участники Исследовательского кружка были из Северной Америки (n = 41), региона EMEA (n = 21), Азиатско-Тихоокеанского региона (n = 8) и Латинской Америки (n = 4). Отказ от ответственности: Результаты этого опроса не отражают глобальные выводы или рынок в целом, но отражают настроения респондентов и опрошенных компаний.

Примечание 1: Методология исследования, основанная на конкретных примерах Gartner

В период с мая по август 2023 года Gartner провела исследование на основе обоснованной теории (CBR), посвященное управляемым службам безопасности. В исследовании приняли участие более 40 случайно выбранных поставщиков, которые придерживались обоснованного теоретического подхода. Gartner задала каждому участнику стандартный набор открытых вопросов как с точки зрения поставщика, так и с точки зрения внедрения. Участники предоставили письменный ответ перед брифингом.

Для каждого поставщика было проведено два брифинга с поставщиками:

- Брифинг поставщика — освещает предстоящие инновации (на горизонте от одного до трех с лишним лет), возможности, уникальные функции и проблемы.
- Брифинг для разработчиков — обзор трех или более проверяемых реальных вариантов использования, демонстрирующих внедрение результатов инноваций поставщика. (Варианты использования были проверены с помощью одной или нескольких из следующих мер: опросов конечных пользователей, интервью, отчетов сторонних производителей и данных, предоставленных поставщиками.)

Gartner собрала письменные ответы, информационные заметки и вспомогательные документы, предоставленные участниками, в специально созданное приложение, предназначенное для сбора качественных данных с помощью компьютерных и смешанных методов. Затем Gartner применила подход, основанный на обоснованной теории, к качественному анализу, идентифицируя и маркируя важные концепции внутри категорий. Эти концепции и коды выявляют закономерности в различных случаях, развивают выводы и помогают обосновать и сформулировать исследовательские позиции. Данные, выводы и позиции в этом исследовании основаны на анализе конкретных случаев.

Примечание: Участие какой-либо одной организации не означает, что тема этого исследования связана с ней.

Среди поставщиков, опрошенных для этого новогоисследования, следующие:

- Акцент
- Adlumin
- Кибербезопасность Aiuken

- Арктический Волк
- AT&T
- Avertium
- Бинарная защита
- Bitdefender
- Критический старт
- Краудстрит
- Cytellix
- Deloitte
- Суть
- ESET
- Доказательства
- Исключить
- Fortinet
- LMNTRIX
- Режим смешивания
- Mphasis
- ДАННЫЕ NTT
- Obrela
- OccamSec
- Открытые системы (постоянно)
- Orange Cyberdefense
- Сети Пало-Альто
- Точка восприятия
- Rapid7
- Красная Канарейка
- Надежность
- Scythe
- Secureworks
- Sophos
- Звездный Кибер
- TCS
- Tech Mahindra
- Trellix
- Trustwave
- Ультрафиолетовый Кибер
- НеобычныЙ вариант
- Vectra AI
- Vodafone