«Сейфуллин оқулары-18(2): «XXI ғасыр ғылымы – трансформация дәуірі» Халықаралық ғылыми-практикалық конференция материалдары = Материалы международной научно-практической конференции «Сейфуллинские чтения – 18(2): «Наука XXI века - эпоха трансформации» - 2022.- Т.I, Ч.III. - С.92-94.

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ В СФЕРЕ ОБЕСПЕЧЕНИЯ КИБЕРБЕЗОПАСНОСТИ И ВЫЯВЛЕНИИ АНОМАЛИЙ

Рзаев Б.Т., докторант 3 курса
Бельдеубаева Ж.Т., доктор PhD

Казахский агротехнический университет им. С.Сейфуллина, г. Нур-Султан

Увалиева И.М., доктор PhD

Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск

Развитие систем машинного обучения (далее - МО) растет с каждым днем. Существующие методы совершенствуются, и их способность понимать и отвечать на реальные вопросы находится на очень высоком уровне. Эти достижения привели к внедрению машинного обучения в области компьютерного зрения, медицинского анализа, компьютерных игр, социальных сетей и маркетинга [1]. Машинное обучение доказало свою способность решать наиболее распространенные проблемы [2]. В некоторых случаях, техники машинного обучения являются более предпочтительными чем традиционные алгоритмы на основе стандартных правил, и могут справиться с задачами, которые не под силу сетевому администратору [1]. Достижения в области машинного обучения положили начало совершенно новой эре. Эра, в которой почти любая часть собранных данных обрабатывается и анализируется с помощью алгоритмов, зависящих от машинного обучения [3].

Кибербезопасность. С момента изобретения интернет-технологии, киберпространство стало центром создания кибератак. Развитие технологий еще больше способствует тому, что хакеры могут обнаруживать уязвимости и создавать вирусы и вредоносное программное обеспечение (далее – ПО), что постоянно бросает вызов индустрии кибербезопасности. Киберпреступность распространяется повсеместно, эксплуатируя все виды уязвимостей вычислительной среды. Этические хакеры уделяют больше внимания в направлении оценки уязвимостей и выработки рекомендаций по методологиям смягчения последствий. Разработка эффективных методов является насущной потребностью в деле обеспечения кибербезопасности. В последнее

время в силу своей эффективности обучение с помощью машин для обеспечения кибербезопасности приобрело большое значение.

По определению, кибербезопасность - это совокупность технологий, процессов и методов, предназначенных для защиты сетей, компьютеров, программ и данных от атак, повреждения или неавторизованного доступа. Один из самых сложных элементов кибербезопасности - это быстрый и постоянно меняющийся характер рисков безопасности [4]. Угрозы нарастают быстрее, чем мы успеваем за ней. Угроза меняется быстрее, чем наша идея оценки риска. Больше невозможно написать большую белую книгу о риске для определенной системы. Необходимо будет постоянно переписывать белую бумагу [4]. Несмотря на это, машинные методы обучения успешно применяются в области кибербезопасности, для разработки высокоэффективных систем [2].

Из всего множества, можно выделить несколько хорошо апробированных функций МО для кибербезопасности:

- интерпретирование действий и событий для категоризации атак;
- увеличение уровня обнаружения кибератак;
- уменьшение количества ложных срабатываний;
- сортировка сообщений для качественного уменьшения количества сетевых атак;
- анализ продолжительных повторяющихся событий;
- выявление мошеннических атак;
- выявление неочевидных для администратора сети закономерностей;
- обнаружение спама и фишинга.

Хотя машинное обучение не может полностью автоматизировать систему кибербезопасности, она помогает определять методологии, ориентированные на программное обеспечение, и, таким образом, уменьшает бремя для аналитиков по безопасности. Постоянно меняющаяся природа киберугроз ставит перед исследователями задачи по исследованию идеальных сочетаний глубоких знаний в области кибербезопасности и информатики, чтобы можно было спрогнозировать действия злоумышленников [2]. Говоря о прогнозировании, машинное обучение, как технология, охватило все киберпространство. Следуя простому принципу предсказания, алгоритмы машинного обучения решают задачи для любого типа проблемы, возникающей в пределах всего технологического пространства. Глядя на глубокие возможности машинного обучения, она начала успешную адаптацию в сферу кибербезопасности [5].

Выявление аномалий. Подавляющее большинство данных, связанных с безопасностью, могут быть обработаны с помощью простого обнаружения на основе известных правил. Эти правила работают быстро, дешево и точно, и эксперты хорошо понимают их и поддерживают. Однако, правила не достаточно сильны, чтобы справиться с неизвестными и новыми образцами входных данных. Хотя они представляют собой лишь небольшую часть из общих входных образцов, такие данные могут нанести существенный ущерб, если их вовремя не остановить. Машинные алгоритмы обучения могут играть

здесь ключевую роль, потенциально обнаруживая совершенно новые, ранее невидимые образцы атаки и неизвестные нам аномалии [5].

Системы обнаружения аномалий имеют два основных преимущества по сравнению с системами обнаружения вторжений на основе сигнатур. Первым преимуществом является их способность обнаруживать неизвестные атаки, поскольку они могут моделировать нормальную работу системы и обнаруживать отклонения от этой модели. Вторым преимуществом является возможность настройки профилей нормальной деятельности для каждой системы, приложения и сети. Это увеличивает сложность для злоумышленника в понимании того, какие действия можно выполнять без его обнаружения [6]. Обнаружение аномалий, в данном контексте подразумевает документирование предметов, событий или наблюдений, которые не соответствуют ожидаемому образцу или другим элементам в наборе данных.

Повышение точности при применении МО происходит за счет предоставления большего объема данных; иными словами, предоставление ему знаний, необходимых для улучшения его эффективности. После, методы машинного обучения, основываясь на статистических предположениях о распределении входных данных и полагаясь на учебные данные, полученные на основе исходных данных, строят модель в целях дальнейшего анализа. Предоставляя доверительные интервалы для их предсказаний, методы обучения могут определять приоритетность данных, которые должны проверяться экспертами вручную, и, таким образом, в значительной степени повысить производительность аналитиков данных [5].

Устранение аномальных данных в контролируемом обучении приводит к статистически значимому повышению точности. Система обнаружения и прогностического анализа на уровне приложений обнаруживает определенных пользователей по их поведению, чтобы предсказать, являются ли их действия нормальными или нет [4].

Алгоритмы МО дают почти точный результат при подаче и обучении огромного количества данных, чтобы помочь обнаружить вредоносные закономерности. Данные должны быть консистентными для алгоритма МО чтобы работать по полную мощь. Объединение вывода МО с другими устройствами инфраструктуры, такими как IPS и брандмауэр, усилит корреляцию и будет содействовать в валидации атаки на приложения. После выявления и анализа закономерностей, оно может быть интегрировано в системы сбора и обработки данных для полного централизованного управления [4].

Обнаружение аномалий само по себе или в сочетании с функцией прогнозирования может быть эффективным средством для выявления мошенничества и обнаружения странной активности в больших и сложных наборах данных. Это может иметь решающее значение для банковской безопасности, медицины, маркетинга, естественных наук и обрабатывающей промышленности, которые зависят от бесперебойной и безопасной работы. С помощью искусственного интеллекта предприятия могут повысить эффективность и безопасность своих цифровых операций [7].

Машинные алгоритмы обработки данных позволяют компьютерным системам выполнять выбранные задачи путем выявления закономерностей и аномалий в огромных объемах данных, преобразуя сложные данные в компактное представление, известную как модель. Без создания настоящего искусственного интеллекта (ИИ), являющегося его конечной целью, машинное обучение рассматривается как одна из технологий, которая может быть ключом к ее достижению [3].

В области безопасности хорошо зарекомендовало себя именно машинное обучение, анализируя данные для поиска закономерностей, чтобы мы могли лучше обнаруживать вредоносные программы в зашифрованном трафике, находить инсайдерские угрозы, предсказывать неблагоприятные события, чтобы обеспечить безопасность людей и активов компании, или защищать данные в "облаке", обнаруживая подозрительное поведение пользователей.

Ландшафт киберугроз заставляет организации постоянно отслеживать и соотносить миллионы внешних и внутренних точек данных по всей своей инфраструктуре и пользователям. Просто невозможно управлять таким объемом информации, имея в своем распоряжении только команду аналитиков. Становится необходимым автоматизировать их рутинный труд, применением машины, способной постоянно обучаться, и выполнять работу аналитика в десятки, а то и сотни раз быстрее, и затратив на это меньше времени. Автоматизируя анализ, кибер-машины могут быстро обнаруживать угрозы и изолировать ситуации, требующие более глубокого человеческого анализа [6].

Но все же более безопасный и сбалансированный подход к корпоративной кибербезопасности заключается в развертывании многоуровневого решения, которое может использовать силу и потенциал машинного обучения - но поддерживает его совместно другими технологиями обнаружения и предотвращения, а также использует человеческий опыт [3].

Список использованной литературы

1 G. Apruzzese, L. Ferretti, M. Marchetti, M. Colajanni, A. Guido. On the Effectiveness of Machine and Deep Learning for Cyber Security [Текст] / Материал из конференции: «2018 10th International Conference on Cyber Conflict», 2018.

2 A. Lakshmanarao, M. Shashi. A Survey On Machine Learning For Cyber Security [Text] / Статья из международного журнала: «International Journal Of Scientific & Technology Research, -Vol.9. 2020 .

3 Machine-learning era in cybersecurity: a step towards a safer world or the brink of chaos? [Текст] / Материал из журнала « Machine-learning era in cybersecurity», февраль, -2019.

4 Nilaykumar K., Haroot Z. Machine Learning in Application Security [Text] / Материал из книги «Advances in Security in Computing and Communications», 2017.

5 Anthony D. Machine Learning Methods for Computer Security [Text] / Joseph, P. Laskov, F. Roli, J. D. Tygar, B. Nelson. // Материал из семинара «Dagstuhl Perspectives Workshop», 9-14 сентября, 2012.

6 What Is Machine Learning in Security? / www.cisco.com. 2020.

7 Anomaly Detection, A Key Task for AI and Machine Learning, Explained [Text] / www.kdnuggets.com. 2019.