3. Одноядерный ПК взломал «квантовое» шифрование всего за 60 минут [Электронный ресурс]. URL: https://4pda.to/2022/08/04/402330/odnoyadernyj_pk_vzlomal_kvantovoe_shifrovanie_vsego_za_60_minut/ (Дата обращения: 15.03.2024)

Кабанов Дмитрий Дмитриевич, студент магистратуры каф. информационной безопасности. E-mail: dima_kabanov_2001@mail.ru

УДК 004.056.5:378.3

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ МАРКОВСКИХ ПРОЦЕССОВ ПРИНЯТИЯ РЕШЕНИЙ ПРИ МОДЕЛИРОВАНИИ АТАК НА СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

В. В. Подтопельный
Калининградский государственный технический университет, г. Калининград

Одним из наиболее распространенных методов, используемых злоумышленником для осуществления атак на системы искусственного интеллекта (СИИ), является внедрение ложных данных в подсистему классификации путем передачи модифицированных легитимных (правильных) данных.

Системы, при получении данных, определяют их принадлежность к определённой группе информации (нейтральные, заслуживающего доверия, ложные) с учетом принятой классификации, и реагируют соответствующим образом. Цель подобных атак заключается в изменении параметров (диапазонов значений) механизма классификации в соответствии с требованиями злоумышленника без потери доверия к нему, как к источнику данных. Для выявления и блокирования атак подобного типа важно выстраивать оптимальную стратегию защиты с учетом знаний о возможных действиях (оптимальной политики вектора атаки) со стороны нападающего на СИИ [1].

Определить стратегию поведения злоумышленника возможно, используя моделирование на основе марковских процессов принятия решений (МППР). При этом, политики являются стационарными (не зависящими от времени). Процесс моделирования предполагает использование параметров, описывающих затраты на реализацию стратегии нападения, которое подразумевает ввод в заблуждение механизмов классификации СИИ на уровне весов нейросети (модифицирование классифицирования) и на уровне дифференциации источников данных на доверенные и недоверенные (навязывание ложных обучающих данных). Соответственно, требуется произвести анализ параметров, влияющих на выработку оптимальных решений

злоумышленником [2]. Оптимальная стратегия нападения строится с использованием уравнения Беллмана.

Модель на основе МППР для общей атаки на системы с элементами искусственного интеллекта предполагает четыре состояния:

– контролируемое взаимодействие (*C* - передача легитимных обучающих данных с фильтрацией);
– легальное взаимодействие (*L* -передача легитимных обучающих данных);
– доверенное взаимодействие (*T* - передача легитимных обучающих данных без фильтрации);
– блокировка (*B*).

При первом получении данных механизмы фильтрации СИИ обычно находятся в нейтральном состоянии по отношению к принимаемой информации. Некоторые системы, основанные на эксплуатации нейронных сетей, контролируют источники сообщений при обучении в заданный для этого период времени [3,4]. Если системами определяется, что полученные ими данные являются подлинными, то есть соответствуют результатам контрольной выборки, то источники информации считаются «доверенными» и в дальнейшем данные от них будут фильтроваться с меньшей интенсивностью.

При атаке злоумышленник может выбрать следующие действия с учетом выбранной им стратегии:

– обман;
– легальное взаимодействие (передача легитимных обучающих данных);
– сброс.

При выборе стратегии атаки в рамках моделирования с использованием МППР также учитываются параметры «вознаграждения» и «затрат».

Нейтральное состояние является начальным в развертываемом векторе атаки. Действия, связанные с осуществлением атаки методом навязывания ложных обучающих данных, будут зафиксированы СИИ, что вызовет ответную реакцию.

В результате первоначального действия при развертывании атаки на СИИ, состояние рассматриваемой системы (следовательно, и вектора атаки) будет изменено: при вводе ложных данных злоумышленником, находящимся в нейтральном состоянии по отношении к атакуемой системе, ее состояние изменится на доверенное, либо на контролируемое. Таким образом, любое действие приведет систему в новые состояния. Атакующий узнает о своем переходном состоянии на основе отклика СИИ. Если системы уточняют корректность поставляемых данных, атакующий (вносящий некорректные изменения в классификатор метом навязывания ложных данных в общей череде правильных) определяет его состояние как «контролируемое взаимодействие» (*C)*. Злоумышленник может вернуться

в нейтральное состояние, если решит завершить ложное обучение (например, из-за нехватки ресурсов, то есть превышения доступных для совершения атаки ресурсов, или, наоборот, получения достаточного количества информации от СИИ). Атакуемая система может классифицировать полученные сообщения как угрожающее (обман), и, соответственно, перевести поставщика данных (злоумышленника) в заблокированное состояние B. Таким образом, злоумышленник может узнать, что он находится в заблокированном состоянии.

При выявлении актуальной политики для злоумышленника может использоваться метод итерации, применимый в МППР [5]. При определении функций награды используются три основных состояния (кроме состояния блокировки): контролируемое взаимодействие, легальное взаимодействие, доверенное взаимодействие. Начальные значения награды (R) переходов по вершинам графа атаки для всех состояний устанавливаются равными нулю. Далее для каждого состояния вычисляются новые значения. Общая функция ценности выглядит следующим образом:

$$V_{i+1}^{*}(s) = \max_{a \in A} \sum_{s' \in S} P(s,a,s') \left[R(s,a,s') + \gamma V_i^{*}(s')\right], \ \forall s' \in S, \quad (1)$$

где:

– $V_{i+1}^{*}(s)$ - функция ценности в состоянии s из множества возможных состояний S, если взять оптимальное действие.

– $P(s,a,s')$ - вероятность перехода, начиная от состояния s и заканчивая состоянием s' после выполнения действия a.

– $R(s,a,s')$ — ожидаемые награды, полученные после состояния переход от s к s' после выполнения действия aс учетом дисконтирования $\gamma V_i^{*}(s')$.

Этот процесс повторяется до тех пор, пока значения награды не достигнут равновесного состояния, перестав изменяться. Кроме того, учитывается максимальное количество повторений, чтобы избежать попадания в бесконечный цикл, когда значения меняются очень незначительно. Эти значения награды (коэффициенты стимуляции оптимальной политики) должны быть масштабированы в заранее определенном диапазоне, чтобы избежать поверхностного эффекта больших вознаграждений, и, одновременно, они должны быть значимыми, представляя ценность, полученную при переходе из одного состояния в другое. Когда атакующему отправляется отказ в приеме данных, вознаграждение сводится к минимальным значениям при расчете вероятностей наступления тех состояний, которые необходимы злоумышленнику (в соответствии с графом атаки), поскольку наблюдается отсутствие успешного влияния на систему.

Учитывая логику развития атаки, которая соответствует описанию ФСТЭК, можно предположить то, что злоумышленник будет сначала, передавать легитимные данные с целью завоевать доверие СИИ [6].

Например, атакующий выбирает записи из обучающей выборки, которые модель СИИ правильно классифицирует, и далее добавляет возмущения (отклонения в точности) к ним, приводящие в итоге к неправильному классифицированию данных в последующем. Также злоумышленник должен получить достаточное вознаграждение R за то, что смог завоевать доверие атакуемых систем. При этом начальные значения вознаграждения отражают выгоды и потери, свойственные состоянию, при котором нападающий не достиг легального взаимодействия с атакуемым объектом. Оптимальная политика в этом случае описывается следующим образом:

$$\pi^*(s) = \underset{a \in A}{\operatorname{argmax}} \sum_{s' \in S} P(s, a, s') \left[R(s, a, s') + \gamma V_i^*(s') \right], \qquad (2)$$

Таким образом, в процессе моделирования, при определении значений ожидаемых вознаграждений *R,* нужно учитывать следующие особенности:

1. Стоимость введения в заблуждение СИИ может оказаться намного выше, чем стоимость легального взаимодействия.
2. Выгоды, получаемые при переходе между состояниями, зависят от исходных и целевых вершин графа атаки.
3. Вознаграждений *R* не будет в случае, когда злоумышленник сначала отправляют ложные данные, а затем его сообщения блокируют независимо от каких-либо действий.

Относительный размер вознаграждений *R,* полученных злоумышленником за каждое изменение (переход в иное состояние), независимо от возможных действий отвечает следующим закономерностям: *R(C−T) > R(L−T)> R(L−C)> R(C−C)≥ R(C-L) ≥ R(T−T)> R(T−C) ≥ R(T−L)> R(C−B).*

Правила выбора значений вероятности перехода предполагают ограничения:

1. Суммирование вероятностей перехода из одного состояния в другое для каждого действия должно быть равно 1. В случае получения суммирования, не равного единице, требуется стандартизировать значения, чтобы гарантировать, что они попадают в надлежащий интервал [0,1].
2. При выполнении совместного действия вероятность перехода в доверенное состояние должна быть больше, чем вероятность перехода в оспариваемое состояние. Кроме того, вероятность перехода в оспариваемое состояние должна быть больше, чем вероятность перехода в заблокированное состояние.
3. Вероятность перехода в заблокированное состояние *B* должна быть больше, чем вероятность перехода в состояние *C* (контролируемое взаимодействие). При выполнении обманного действия вероятность перехода в оспариваемое состояние должна быть больше, чем вероятность перехода в доверенное состояние.

Приведенные особенности моделирования атак на СИИ с использованием МППР, учитывая требования оптимальной политики

нападения, позволяют более точно построить вектор атак, связанных с навязыванием ложных обучающих данных, а также модифицированием классифицирования. При этом выявляется тенденция: максимально снижать стоимость «отравления данными» и поддерживать относительно низкие затраты на легальное взаимодействие.

Список использованных источников

1. Горбачев И.Е., Глухов А.П. Моделирование процессов нарушения информационной безопасности критической инфраструктуры// Труды СПИИРАН. – Москва, 2015. – Вып. 1(38). – С. 112 – 135.

2. Панченко А.А. Анализ подходов к построению системы защиты информации на базе модели процесса обработки данных / А.А, Панченко, М.В. Аникиенко, В.Н. Пржегорлинский // Вестник Рязанского гос. радиотехнического ун-та. – 2005. – № 16. – С. 120–123.

3. Горюнов М. Н. Синтез модели машинного обучения для обнаружения компьютерных атак на основе набора данных CICIDS2017 / М.Н. Горюнов, А.Г. Мацкевич, Д.А. Рыболовлев // Труды ИСП РАН. – 2020. – Т. 32, вып. 5. – С. 81–94.

4. Горюнов М.Н. Оценка применимости методов машинного обучения для обнаружения компьютерных атак / М.Н. Горюнов, А.А. Рыболовлев, Д.А. Рыболовлев // Информационные системы и технологии (Орел). – 2020. – № 6. – С. 103–111.

5. Кохендерфер М., Уилер Т., Рэй К. Алгоритмы принятия решений / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2023. – 684 с.

6. Методика оценки угроз безопасности информации Методический документ ФСТЭК России: утв. ФСТЭК России России 5 февраля 2021 г.

Подтопельный Владислав Владимирович, Ст. преп. Института цифровых технологий государственного технического университета (КГТУ), ionpvv@mail.ru.

УДК 004.056.52; 004.891.3

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПОВЕДЕНЧЕСКИХ ПАТТЕРНОВ В DLP-СИСТЕМАХ

Е.А. Марченко, С.В. Жуков
«Самарский национальный исследовательский университет имени академика С.П. Королёва», г. Самара

Ключевые слова: DLP-система, система контроля, рекуррентная нейронная сеть, нейросетевые технологии, поведенческий анализ.

DLP-системы (Data Loss Prevention) - это специализированные программные решения, разработанные для предотвращения утечек конфиденциальных данных, обеспечения соблюдения правил безопасности и защиты информации в организациях. Они играют ключевую роль в предотвращении утечек конфиденциальной информации, соблюдении нормативных требований по защите данных и минимизации рисков