Инженерный вестник Дона, №6 (2023)
ivdon.ru/ru/magazine/archive/n6y2023/8479

Метод нормализации полей внешних источников репозитория данных о кибератаках MITRE CTI

В.И. Борисов, Е.В. Федорченко

Санкт-Петербургский государственный университет телекоммуникаций им. М. А. Бонч-Бруевича

Аннотация: Растущая сложность промышленных систем существенно увеличивает поверхность возможных кибератак, и, следовательно, требует надёжных методов оценки защищенности инфраструктуры. Современные методы оценки защищенности опираются на работу с большим количеством данных, представление которых зачастую не стандартизировано. Одним из таких источников является база знаний MITRE ATT&CK, содержащая информацию об атакующих техниках в формате, позволяющим взаимодействовать с ней программно. Данная работа направлена на решение задачи нормализации полей внешних источников, описывающих атакующую технику, с целью повышения эффективности работы с вышеописанным репозиторием. Метод, предлагаемый в данной работе, основан на возможности спецификации языка STIX, используемого для описания данных, представленных в MITRE ATT&CK, к расширению и использовании открытых словарей. Разработка предлагаемого метода основывалась на данных об атакующих техниках матрицы Enterprise, как наиболее полного среди всех доменов базы знаний ATT&CK, однако предложенный метод является самостоятельным и не зависит от конкретного домена.
Ключевые слова: анализ угроз, база знаний, информационная безопасность, MITRE ATT&CK, стандартизация.

Введение

Повсеместное использование технологий в коммерческих, промышленных и государственных организациях, сопровождается усложнением общей структуры компьютерных систем: помимо традиционных компонентов в них всё чаще интегрируются мобильные устройства, применяются облачные службы [1], – всё это в свою очередь неизбежно увеличивает как поверхность атаки, так и значительно сказывается на росте их числа. Недостаточное внимание к безопасности инфраструктуры, ввиду отсутствия исчерпывающих методов её оценки, ставит под угрозу надёжное функционирование процессов в организации, а совершенные на организацию атаки негативно сказываются на ее показателях, таких, как прибыль и репутация, а для промышленных

предприятий, оснащённых киберфизическими системами, кроме всего прочего, могут приводить к нанесению физического вреда в ходе успешно реализованных злоумышленниками атак [2, 3]. Для решения данной задачи были разработаны методы оценки защищенности и анализа угроз.

Передовые решения в данной области преимущественно опираются на большое количество информации и работу с источниками данных об атаках. Одним из таких источников является репозиторий организации *MITRE*, содержащий данные, использованные при создании базы знаний *ATT&CK* [4, 5]. Каждая атакующая техника в данном наборе данных содержит внешние ссылки на связанные базы знаний, что позволяет получить исчерпывающую информацию об атакующей технике, её применимости, и возможных контрмерах. Однако данная информация представлена в ненормализованном виде, что затрудняет ее обработку. Нормализация полей внешних источников позволит исследователям и специалистам в области защиты информации более эффективно работать с вышеупомянутым набором данных, что потенциально обеспечит лучшую защищенность систем.

Описание метода

Для решения задачи, поставленной в данной работе, исследуется структура полей внешних источников, способы взаимодействия с репозиторием и предлагается метод, который позволит отобразить неупорядоченное множество внешних источников в множество, элементы которого можно использовать для формирования более детальных запросов к базе знаний *MITRE ATT&CK*.

Для хранения данных базы знаний *MITRE ATT&CK* используется язык сериализации и обмена результатами исследования угроз безопасности *STIX* [6, 7], который в свою очередь представляет собой расширения формата хранения данных *JSON*. Такой формат хранения данных базы знаний позволяет взаимодействовать с ней программно, в том числе и при помощи

специализированных библиотек, одной из которых является библиотека *stix2* языка *Python* 3. Она предоставляет возможность взаимодействовать с данными формата *STIX*, начиная от версии спецификации 2.0 и выше.

Для хранения информации об атакующих техниках используется стандартный объект *Attack Pattern*, принадлежащий семейству объектов *STIX Domain Objects (SDO)* [8], расширенный дополнительными полями, необходимыми для представления концепций, характерных для базы знаний *MITRE ATT&CK*. Такая организация данных позволяет с минимальными усилиями и максимальной степенью стандартизации представить сущности базы *ATT&CK*, в том виде, в котором с ними можно взаимодействовать уже существующими программными инструментами.

Одним из стандартных полей атакующей техники является поле *external_references*, которое представляет собой список внешних источников, содержащих ту или иную информацию об атакующей технике.

Формат элементов данного списка представлен в таблице 1 [8].

Обязательным для элемента такого списка является только поле *source_name*. Однако спецификация уточняет, что хотя остальные поля и помечены как опциональные, но помимо *source_name*, у элемента должно существовать хотя бы одно из следующих полей: *description*, *url*, *source_id*.

Количество внешних ссылок у действующих атакующих техник по доменам представлено в таблице 2. Под действующими объектами следует понимать те атакующие техники, которые не были помечены, как устаревшие или отозванные. За эти свойства отвечают логические поля *x_mitre_deprecated* и *x_mitre_is_revoked* [9].

Структура элемента внешней ссылки позволяет определять подлинность объектов благодаря полю *hashes*, находить ссылку по внешнему идентификатору благодаря связке полей *external_id* и *source_name*, или же ссылаться на внешний сетевой ресурс, предоставляющий информацию об

атакующей техники. Такая структура не предоставляет никаких средств по категоризации внешних ссылок или содержимого, на которые эти внешние ссылки указывают.

Таблица № 1

Формат элементов списка внешних источников атакующей техники

Наименование поля элемента	Тип поля элемента	Описание
source_name (обязательное)	строка	Наименование источника, для которого определён элемент, представляющий собой внешнюю ссылку
description (опциональное)	строка	Описание внешнего источника
url (опциональное)	строка	Ссылка на ресурс, содержащий внешнюю ссылку в формате URL
hashes (опциональное)	ассоциативный массив	Данное поле указывает на ассоциативный массив хэш-строк для содержимого, расположенного по адресу, который содержится в поле url. Таким образом, данное поле может существовать для элемента, у которого существует поле url. Ключами такого массива обязаны быть элементы открытого словаря hash-algorithmov.
external_id (опциональное)	строка	Идентификатор для содержимого внешней ссылки

Таблица № 2

Количество внешних ссылок у объектов Attack Pattern

Матрица атак	Число внешних ссылок
Enterprise	2836
Mobile	344
ATT&CK для ICS	212

Для работы с данными репозитория базы знаний была использована библиотека для языка Python 3 – stix2. Для загрузки объектов из файла, расположенного локально в файловой системе, используется объект *MemorySource*. Основным методом, который используется для получения данных из этого объекта выступает метод *MemorySource.query()*. Данный метод на вход принимает список объектов *Filter*, содержащий условия, которым должны отвечать запрашиваемые объекты, и возвращает список из объектов, соответствующих заданному условию [10].

Спецификация STIX определяет категорию типа данных, представляющую собой заранее объявленный набор строк – «STIX Vocabularies». Целью типов данной категории является повышение взаимной совместимости результатов анализа угроз, благодаря использованию одних и тех же строк для представления общих концепций. Иерархия данной категории типов представлена на рис.1.

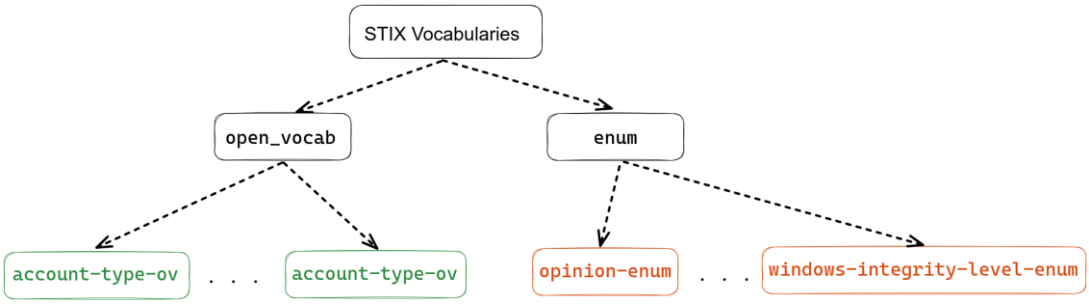

Рис. 1 – Иерархия категории типов STIX Vocabularies

STIX Vocabularies содержит в себе следующие абстрактные типы:

— Тип *open_vocab*. Конкретные подтипы данного типа определяют набор строк, который рекомендован к использованию в качестве значений, но может быть расширен другими строковыми значениями при необходимости. Наименование подтипов данного типа сопровождается суффиксом «*-ov*»;

— Тип *enum*. Конкретные подтипы данного типа определяют набор строк, который обязателен к использованию, и не может быть расширен

другими значениями, с целью гарантии наличия в свойствах некоторых объектов заранее известного множества возможных строк. Наименование подтипов данного типа сопровождается суффиксом «*-enum*».

По результатам анализа внешних ссылок репозитория данных MITRE ATT&CK предлагается использовать открытый словарь формата STIX совместно с пользовательским полем, расширяющим спецификацию STIX 2.1.

Данное поле расширяет элемент списка *external_references*, имеет тип *open_vocab*, и в качестве значений может принимать одно из строковых значений из числа элементов, объявленных в предлагаемом открытом словаре.

Предлагаемый словарь имеет идентификатор *extref-type-ov* и объявление, представленное в таблице 3.

Таблица № 3

Объявление открытого словаря для категоризации внешних ссылок

Значение словаря	Описание
document	Указывает на то, что внешняя ссылка представляет собой определённый документ, такой, как статья, технический отчёт, или книга.
developer-network	Данное значение указывает на то, что внешняя ссылка является ссылкой на документацию разработчика.
document	Данное значение указывает на то, что внешняя ссылка является некоторым документом. Оно может использоваться для обозначения технической литературы, спецификаций и других технических документов, а также академических работ.
knowledge-base	Данное значение указывает на то, что внешняя ссылка указывает на запись в базе знаний. Для уточнения конкретной записи используется поле external_id.

Значение словаря	Описание
mailing-list	Данное значение указывает на то, что внешняя ссылка является цепочкой сообщений в почтовой рассылке группы разработки того или иного решения. Зачастую почтовые рассылки разработчиков open-source решений являются открытыми и указать ссылку до сетевого ресурса, на котором они размещены, можно в поле url.
proof-of-concept	Данное значение указывает на то, что внешняя ссылка представляет собой некоторую утилиту или демонстрационную среду, с помощью которой можно воспроизвести работу атакующей техники, к которой принадлежит данная внешняя ссылка.
security-blog	Данное значение указывает на то, что внешняя ссылка является частным или корпоративным блогом, содержащим информацию об атакующей технике.
src-code	Данное значение указывает на то, что внешняя ссылка представляет собой ссылку на некоторый фрагмент исходного кода, зачастую open-source решения. Этот фрагмент может как содержать исправления уязвимости, используемой атакующей техникой, так и сам код, содержащий уязвимость.
social-media	Данное значение указывает на то, что внешняя ссылка представляет собой некую запись в социальных сетях. Нередко зарубежные специалисты предпочитают делиться обнаруженной информацией об атакующих техниках и/или уязвимостях в социальных сетях, не прибегая к публикации на специализированных ресурсах.

В ходе разработки метода было рассмотрено альтернативное решение, поставленной задачи. Использование отдельного массива под каждую категорию элементов внешних ссылок способом, представленным на рис.2, позволило бы упростить процесс поиска по категориям с точки зрения количества итераций, совершаемых, чтобы получить все объекты одной категории.

Хотя альтернативный метод избавляет от необходимости производить итерацию по всем внешним ссылкам для получения всех элементов целевой категории, он обладает существенным архитектурным недостатком.

```
"external_references" : [
                            "developers_networks" :          [ ... ],
                            "documents" :                    [ ... ],
                            "knowledge_bases" :              [ ... ],
                            "mailing_lists" :                [ ... ],
                            "proof_of_concept" :             [ ... ],
                            "security_blogs" :               [ ... ],
                            "social_media" :                 [ ... ],
                            "src_code" :                     [ ... ]
                        ]
```

Рис. 2. – Альтернативный метод категоризации ссылок

Такое решение требует поддерживать отдельный массив в объекте для каждой категории, что ухудшает читаемость данных, нарушая основную концепцию языка STIX.

Заключение

Таким образом, в данной работе было предложено решение, которое позволяет с минимальной модификацией стандартных типов нормализовать элементы внешних ссылок. А использование подтипа open_vocab, в противовес типу enum, определенному в спецификации STIX, позволит расширять словарь в случае появления новых форм источников.

Литература

1. Ястремская Н.Ю., Фролова Л.А. Развитие информационного рынка как условие и результат становления информационной экономики // Инженерный вестник Дона, 2013, №4. URL: ivdon.ru/ru/magazine/archive/n4y2013/2141.

2. Горбачева А.А., Купченко А.Ю. Противоречия управления в обеспечении продовольственной безопасности России // Инженерный вестник Дона, 2014, №2. URL: ivdon.ru/ru/magazine/archive/n2y2014/2334.

3. Kumar P., Gupta G. P., Tripathi R. Toward design of an intelligent cyber attack detection system using hybrid feature reduced approach for iot networks // Arabian Journal for Science and Engineering. 2021. T. 46. pp. 3749-3778.

4. Rajesh P., Alam M., Tahernezhadi M., Monika A., Chanakya G. Analysis Of Cyber Threat Detection And Emulation Using MITRE Attack Framework // 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA). IEEE, 2022. pp. 4-12.

5. Möller D.P.F. NIST Cybersecurity Framework and MITRE Cybersecurity Criteria // Guide to Cybersecurity in Digital Transformation: Trends, Methods, Technologies, Applications and Best Practices. Cham: Springer Nature Switzerland, 2023. pp. 231-271.

6. Briliyant O.C., Tirsa N.P., Hasditama M.A. Towards an automated dissemination process of cyber threat intelligence data using STIX // 2021 6th International Workshop on Big Data and Information Security (IWBIS). IEEE, 2021. pp. 109-114.

7. Aviad A., Węcel K. Cyber treat intelligence modeling // Business Information Systems: 22nd International Conference, BIS 2019, Seville, Spain, June 26–28, 2019, Proceedings, Part I 22. Springer International Publishing, 2019. pp. 361-370.

8. STIX Version 2.1. Edited by Bret Jordan, Rich Piazza, and Trey Darley. 10 June 2021. OASIS Standard. URL: docs.oasisopen.org/cti/stix/v2.1/stix-v2.1.html (date assessed: 12 мая 2023 г.)

9. Руководство по использованию репозитория данных MITRE ATT&CK. URL: github.com/mitre-attack/attack-stix-data/blob/master/USAGE.md (дата обращения: 13 мая 2023 г.)

10. Справочник прикладного программного интерфейса библиотеки stix2. URL: stix2.readthedocs.io/en/latest/api_ref.html (дата обращения: 13 мая 2023 г.)

References

1. Jastremskaja N.Ju., Frolova L.A. Inzhenernyj vestnik Dona, 2013, №4. URL: ivdon.ru/ru/magazine/archive/n4y2013/2141.

2. Gorbacheva A.A., Kupchenko A.Ju. Inzhenernyj vestnik Dona, 2014, №2. URL: ivdon.ru/ru/magazine/archive/n2y2014/2334.

3. Kumar P., Gupta G. P., Tripathi R. Arabian Journal for Science and Engineering. 2021. V. 46. pp. 3749-3778.

4. Rajesh P., Alam M., Tahernezhadi M., Monika A., Chanakya G. 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA). IEEE, 2022. pp. 4-12.

5. Möller D.P.F. Guide to Cybersecurity in Digital Transformation: Trends, Methods, Technologies, Applications and Best Practices. Cham: Springer Nature Switzerland, 2023. pp. 231-271.

6. Briliyant O.C., Tirsa N.P., Hasditama M.A. 2021 6th International Workshop on Big Data and Information Security (IWBIS). IEEE, 2021. pp. 109-114.

7. Aviad A., Węcel K. Business Information Systems: 22nd International Conference, BIS 2019, Seville, Spain, June 26–28, 2019, Proceedings, Part I 22. Springer International Publishing, 2019. pp. 361-370.

8. STIX Version 2.1. Edited by Bret Jordan, Rich Piazza, and Trey Darley. 10 June 2021. OASIS Standard. URL: docs.oasisopen.org/cti/stix/v2.1/stix-v2.1.html (accessed: May, 12, 2023).

9. Rukovodstvo po ispol'zovaniju repozitorija dannyh MITRE ATT&CK [Guide to using the MITRE ATT&CK Data Repository]. URL: github.com/mitre-attack/attack-stix-data/blob/master/USAGE.md (accessed: May 12, 2023).

10. Spravochnik prikladnogo programmnogo interfejsa biblioteki stix2 [Manual of the application programming interface of the stix2 library]. URL: stix2.readthedocs.io/en/latest/api_ref.html (accessed: May, 12, 2023).