МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра информационной безопасности

ОТЧЕТ

по научно-исследовательской работе

Тема: Методика анализа баз уязвимостей с помощью построения графов связей

Студентка гр. 8361 _____ Дымова М.Д.

Руководитель _____ Иванов Е.В.

Санкт-Петербург

2022

ЗАДАНИЕ
НА НАУЧНО-ИССЛЕДОВАТЕЛЬСКУЮ РАБОТУ

Студентка Дымова М.Д.

Группа 8361

Тема НИР: Методика анализа баз уязвимостей с помощью построения графов связей

Задание на НИР:

Разработка методики анализа существующих баз уязвимостей с помощью построения графов связей

Сроки выполнения НИР: 01.11.2022 – 01.06.2023

Дата сдачи отчета: 20.12.2022

Дата защиты отчета: 20.12.2022

Студентка _____ Дымова М.Д.

Руководитель _____ Иванов Е.В.

АННОТАЦИЯ

В данной работе рассматриваются существующие на данный момент базы уязвимостей, агрегаторы информации об уязвимостях и способы анализа с помощью графов связей.

SUMMARY

This paper discusses the currently existing databases of vulnerabilities, aggregators of information about vulnerabilities and methods of analysis using link graphs.

СОДЕРЖАНИЕ

	Введение	5
1.	Базы уязвимостей, существующие на данный момент	6
1.1.	Банк данных угроз безопасности информации (БДУ ФСТЭК России)	6
1.2.	CVE List и NVD(National Vulnerability Database)	7
1.3.	VulnDB	10
1.4.	Secunia Advisory and Vulnerability Database	10
1.5.	VND (Vulnerability Notes Database)	12
1.6.	Exploit Database	13
2.	Агрегаторы информации об уязвимостях	15
2.1.	CVEDetails	15
2.2.	Vulners	16
3.	Что такое графовые методы анализа и какие средства анализа графов существуют	18
	Заключение	21
	Список использованных источников	22

ВВЕДЕНИЕ

В 90-е годы двадцатого века вследствие резкого увеличения связности Интернета стала очевидной необходимость систематичного сбора данных об уязвимостях программного обеспечения (ПО) и их каталогизации. С ростом количества новых видов ПО, а также с ростом частоты обнаружения в них уязвимостей, потребность в базах, где собраны все известные на данный момент уязвимости, только увеличивается.

Создание реестров и баз данных с информацией об обнаруженных уязвимостях потребовало унифицированного (по крайней мере в рамках одной базы данных) способа их идентификации и классификации. Каждой обнаруженной уязвимости требовалось присвоить идентификатор, дать ей емкое и краткое описание, определить для нее список уязвимого ПО и его версий. Кроме того, требовалось оценить степень критичности данной уязвимости по ряду различных критериев: простоте обнаружения злоумышленником уязвимого ПО, легкости эксплуатации уязвимости и требуемым для этого привилегиям, а также потенциальным последствиям эксплуатации уязвимости для компьютерной системы. Эта информация могла впоследствии дополняться рекомендациями по устранению уязвимости или предотвращению ее эксплуатации и статусом уязвимости (присутствует ли она в актуальной версии ПО или была закрыта соответствующими обновлениями и патчами).

1. Базы уязвимостей, существующие на данный момент

1.1. Банк данных угроз безопасности информации (БДУ ФСТЭК России) – собственный реестр известных угроз ИБ и уязвимостей ПО Федеральной службы по техническому и экспортному контролю.

Данный реестр уязвимостей в первую очередь ориентирован на сбор и хранение информации об *угрозах и уязвимостях ПО, используемого в государственных организациях РФ*, включая информационные системы и системы управления критичными производственными процессами. Пополняется реестр ФСТЭК России путем мониторинга общедоступных источников информации – информационных бюллетеней российских и иностранных компаний, производящих ПО, а также реестров и информационных бюллетеней исследовательских организаций и компаний, предоставляющих услуги в области информационной безопасности.

Все хранящиеся в БДУ ФСТЭК России записи имеют единообразный формат и включают: текстовое описание уязвимости, дату обнаружения уязвимости, названия, версии и производителей уязвимого ПО, информацию о типе ошибки, классе уязвимости и текущем ее статусе (потенциально возможная либо подтвержденная производителями ПО или независимыми исследователями уязвимость, устранена ли уязвимость в новых версиях ПО). Также записи содержат оценку критичности уязвимости и сопутствующий вектор CVSS, пометку о наличии известных готовых сценариев эксплуатации уязвимости и возможного результата эксплуатации уязвимости, указание уязвимых аппаратных платформ или операционных систем, список возможных методов противодействия уязвимости и ссылки на источники

дополнительной информации по уязвимости (включая идентификаторы данной уязвимости в иных реестрах и базах данных).

Преимущества:

Записи в базе данных БДУ ФСТЭК России предоставляют более подробную информацию о различных аспектах, связанных с уязвимостью, чем иностранные реестры уязвимостей CVE List и NVD, предоставляемые американскими некоммерческими и государственными организациями. Кроме того, пользовательский интерфейс для выборки из базы БДУ ФСТЭК России отличается большей гибкостью настроек поиска и фильтрации результатов в сравнении с интерфейсами указанных иностранных баз данных.

Все содержимое реестра БДУ ФСТЭК России предоставляется для скачивания в форматах XLSX и XML, что обеспечивает получение информации как виде, удобном для обработки человеком (посредством популярных офисных приложений семейства MS Excel), так и в машиночитаемом варианте для различных автоматизированных средств (например, сканеров безопасности и систем обнаружения атак).

Недостатки:

К возможным недостаткам БДУ ФСТЭК России можно отнести меньшее общее количество покрытых реестром уязвимостей (в сравнении как с базами CVE List и NVD, так и с базами данных уязвимостей, созданных коммерческими компаниями), а также отсутствие какой-либо агрегации отдельных записей (которая характерна для такой базы данных, как Vulnerability Notes Database).

1.2. CVE List и NVD (National Vulnerability Database)

Стандарт Common Vulnerabilities and Exposures (CVE), разработанный американской некоммерческой исследовательской корпорацией

MITRE Corporation в 1999 году, де-факто является на сегодняшний день основным стандартом в области унифицированного именования и регистрации обнаруженных уязвимостей программного обеспечения. Данный стандарт непосредственно определяет как формат идентификаторов и содержимого записей об отдельных обнаруженных уязвимостях, так и процесс резервирования идентификаторов для новых обнаруженных уязвимостей и пополнения соответствующих баз данных.

В настоящее время поддержкой и администрированием реестра уязвимостей CVE занимается группа из 84 организаций по всему миру, в число которых входят ведущие производители программного обеспечения, телекоммуникационного оборудования и интернет-сервисов, такие как Apple, Cisco, Facebook, Google, IBM, Intel, Microsoft, Oracle и ряд компаний, специализирующихся в области информационной безопасности, например, F5 Networks, McAfee, Symantec, «Лаборатория Касперского» и др.

На текущий день базы уязвимостей MITRE CVE List и NVD содержат порядка 188 тысяч записей об отдельных уязвимостях, обнаруженных за период с 1999 года по настоящее время. При этом, хотя сами базы данных различаются на уровне функциональных возможностей, предоставляемых пользователям, сами списки записей об уязвимостях фактически идентичны друг другу. Формально CVE List выступает изначальным источником записей для базы данных NVD, а специалисты, отвечающие за поддержку базы NVD, производят уточненный анализ и сбор доступной информации по уязвимостям, зарегистрированным в CVE List (например, собирают ссылки на сторонние источники информации об уязвимости и мерах по ее устранению или предотвращению эксплуатации).

Идентификаторы CVE имеют формат CVE-YYYY-NNNN, отражая в первых четырех цифрах год регистрации уязвимости и в последующих четырех-шести цифрах – уникальный в рамках этого года номер уязвимости.

Для каждой из обнаруженных уязвимостей запись в базе содержит краткое описание типа и причин уязвимости, уязвимые версии ПО, оценку

критичности уязвимости в соответствии со стандартом CVSS (Common Vulnerability Scoring System) и ссылки на внешние источники с информацией об уязвимости – чаще всего, таковыми выступают информационные бюллетени на сайтах производителей программного обеспечения или исследовательских организаций.

В плане пользовательского функционала в CVE List поддерживаются возможности простейшего поиска среди записей (по ключевым словам и CVE-идентификаторам) и скачивания архивов записей за любой выбранный год в различных форматах (HTML, XML, CVRF, CSV или Plain Text). Также возможно автоматическое получение обновлений в машиночитаемом виде через специальный data feed CVE Change Log (он позволяет как отслеживать появление новых идентификаторов CVE, так и изменения в записях для уже существующих).

Для базы NVD в свою очередь доступны продвинутые функции поиска уязвимостей по ключевым словам, временным диапазонам создания/модификации записи, компонентам CVSS-метрики и т. п. Кроме того, доступны скачивание всех записей базы данных в XML, а также получение информации об обновлениях базы в виде RSS-подписки и JSON data feed.

Преимущества:

Ежедневное обновление реестров известных уязвимостей.

Сильной стороной самого стандарта CVE является его повсеместная поддержка в современных программных продуктах и сервисах, направленных на обеспечение информационной безопасности.

Недостатки:

Отсутствие в записях об уязвимостях какой-либо информации о точном месте локализации уязвимости в коде уязвимого ПО и возможных векторах атак, посредством которых возможна эксплуатация данной уязвимости.

1.3. VulnDB

Эта база стала последовательницей OSVDB (Open Sourced Vulnerability Database), которая была создана в 2004 году, но закрыта в 2016.

База VulnDB доступна по платной подписке и содержит информацию о более 251 тыс. обнаруженных уязвимостях (включая более 80 тыс. записей об уязвимостях, отсутствующих в базах CVE List и NVD), а поддерживающие базу специалисты отслеживают появление новых уязвимостей для 19 тыс. современных программных продуктов и 2 тыс. популярных библиотечных компонентов. Одним из возможных предназначений предлагаемого сервиса может быть риск-менеджмент и оценка уровня защищенности компьютерных сетей организаций.

Преимущества:

Значительно больше записей об уязвимостях, чем в общедоступных CVE List и NVD.

Недостатки:

Доступ только по платной подписке.

1.4. Secunia Advisory and Vulnerability Database

Сходные услуги в области информационной безопасности предоставляются датской компанией Secunia Research, которая специализируется на исследованиях в области компьютерной и сетевой безопасности и в числе прочих сервисов предоставляет доступ к базе уязвимостей Secunia Advisory and Vulnerability Database.

Secunia Advisory and Vulnerability Database– это база данных с информационными бюллетенями (Secunia Advisories), содержащими сведения об обнаруженных угрозах и уязвимостях ПО. Бюллетени формируются на основе собственных исследований специалистов Secunia

Research и агрегации информации об уязвимостях, полученных из иных публичных источников.

Как и в случае с базой VulnDB, бюллетени в базе данных Secunia зачастую публикуются еще до того, как соответствующие записи появляются в базе CVE List, и уже впоследствии размечаются ссылками на соответствующие CVE-идентификаторы. При этом нередки ситуации, когда никто из CVE Numbering Authorities так и не берется за регистрацию обнаруженной уязвимости, в результате чего соответствующая запись из базы Secunia Research так и остается без выделенного CVE-идентификатора и, соответственно, не попадает в базы CVE List и NVD.

По своей структуре записи в базе Secunia сходны с содержимым баз CVE List и NVD и содержат дату регистрации уязвимости, тип и краткую классификацию уязвимости, критичность уязвимости (описывается с помощью перечислимого типа Secunia Research Criticality Rating вместо скалярной оценки по стандарту CVSS), списки уязвимого ПО и его версий, ссылки на внешние источники информации и рекомендации по устранению угрозы (как правило, установку патчей от производителя ПО или апгрейд уязвимого ПО до безопасной версии – в этом случае бюллетень содержит упоминание минимального номера безопасной версии).

Характерно, что в бюллетенях Secunia Advisories принято агрегировать в одной записи информацию о множестве отдельных уязвимостей, одновременно обнаруженных в одном и том же программном обеспечении. Это означает, что одной записи из базы данных Secunia может соответствовать множество различных CVE-идентификаторов. В настоящее время база данных Secunia Research содержит приблизительно 100 тысяч записей об уязвимостях, обнаруженных начиная с 2003 года.

Бесплатный доступ к Secunia Advisories производится только на условиях некоммерческого использования предоставленной информации и только в формате html (по всей видимости, это делается для того, чтобы затруднить автоматизированное извлечение информации из базы).

Для коммерческого использования доступ к базе данных от Secunia Research предоставляется посредством специализированного средства Software Vulnerability Manager и соответствующей подписки на сервис компании Flexera, которой и принадлежит с 2015 года Secunia Research.

1.5. VND (Vulnerability Notes Database)

Примером иного подхода к организации базы уязвимостей является база VND (Vulnerability Notes Database), поддерживаемая центром реагирования CERT/CC при Институте программной инженерии в университете Карнеги-Меллона, США.

Преимущества:

Каждая запись в базе VND агрегирует информацию о множестве сходных уязвимостей для какого-либо конкретного ПО, ссылаясь на множество соответствующих CVE идентификаторов. Данная агрегация является характерным отличием базы VND от баз CVE List и NVD, позволяя проверить целое множество уязвимостей сходной природы в конкретном уязвимом ПО или его компоненте. Кроме этого, записи в базе VND зачастую содержат подробное и детальное руководство по устранению уязвимостей и/или предотвращению их эксплуатации злоумышленником, а также обзор текущей ситуации с наличием или успешным закрытием обнаруженной уязвимости различными вендорами (в случае, когда уязвимым оказывается библиотечный компонент, используемый несколькими производителями ПО). Ещё один приятный момент – это разрешение на бесплатное использование всех материалов базы для любых целей и возможностью полного скачивания всех записей базы в формате JSON с помощью специального бесплатно предоставляемого программного обеспечения.

Недостатки:

Редкие обновления (единицы раз в месяц) и слабый охват всех существующих уязвимостей (в том числе даже зарегистрированных в CVE List), что существенно ограничивают полезность данного каталога уязвимостей для оперативного реагирования на новые уязвимости ПО. В частности, в настоящий момент в базе зарегистрировано лишь около 3,5 тысяч записей.

1.6. Exploit Database

Альтернативным подходом к каталогизации информации об обнаруженных уязвимостях ПО является регистрация не самих уязвимостей, а сценариев их эксплуатации (эксплойтов, exploits) или примеров эксплуатации уязвимости (Proof of Concept). Примером реализации такого подхода является база Exploit Database, сформированная организацией Offensive Security Team, специализирующейся на проведении тренингов в области информационной безопасности и тестирования компьютерных систем на проникновение.

База Exploit Database на настоящий момент содержит порядка 39 тысяч записей, разбитых на различные категории (эксплойты для веб-приложений, удаленной и локальной эксплуатации уязвимостей, примеры атак Denial of Service и исполнимые фрагменты кода shellcode для различных уязвимостей переполнения стека или доступа к памяти). Данные записи покрывают множество уязвимостей, обнаруженных с 2000 года по настоящее время.

Типичная запись в базе Exploit Database содержит краткое описание уязвимости, указание уязвимых версий приложений или их компонентов, уязвимую программную платформу (операционную систему или фреймворк веб-приложения), CVE-идентификатор, присвоенный данной уязвимости (при его наличии), и ссылки на сторонние источники информации об уязвимости. Однако самая важная и содержательная часть записи – это детальное описание самих причин возникновения уязвимости, места

локализации уязвимости в коде (с непосредственной демонстрацией уязвимого фрагмента кода, если код приложения публично доступен) и описание работоспособных сценариев эксплуатации уязвимостей или сценариев Proof of Concept (PoC).

Кроме этого, поддерживается архив уязвимых версий приложений для того, чтобы исследователи, использующие базу Exploit Database, имели возможность воспроизвести наличие уязвимости и проверить работоспособность нацеленного на нее эксплойта.

Преимущества:

Наибольшую пользу подобные базы с эксплойтами и PoC-сценариями могут принести специалистам, занятым тестированием компьютерных сетей на проникновение, в составе инструментальных средств проверки наличия уязвимостей в исследуемых сетях.

Также доступные в базе эксплойты могут быть использованы в качестве дидактического материала для начинающих исследователей и специалистов в области информационной безопасности в рамках образовательного процесса или повышения квалификации.

Наконец, подобная база с набором работоспособных сценариев эксплуатации уязвимостей для веб-приложений и удаленной эскалации привилегий могла бы быть полезна и в качестве источника информации для компаний, занятых разработкой сигнатурных систем обнаружения атак и подобных средств мониторинга трафика.

Недостатки:

Использование информации может быть затруднено в силу отсутствия в Exploit Database интерфейса для получения обновлений базы, возможностей для скачивания архива всех записей и, в некоторых случаях, соглашениями об использовании предоставляемых материалов.

2. Агрегаторы информации об уязвимостях

Разнообразие различных реестров и баз данных уязвимостей вызывает у специалистов в области информационной безопасности естественное желание использовать различного рода агрегаторы информации, которые бы обеспечивали автоматизированный сбор доступной информации об уязвимостях и дополнительные функции поиска и фильтрации интересующей информации.

Подобного рода агрегаторы информации об уязвимостях существуют и представлены различного рода сервисами, начиная от специализированного агрегатора CVE-релевантной информации и до агрегатора с интерфейсом полноценной поисковой машины, адаптированной под предметную область.

2.1. CVEDetails

Примером сервисов первого типа является CVEDetails. Это простой специализированный агрегатор информации об уязвимостях, который собирает всю доступную из публичных реестров и баз данных уязвимостей информацию по конкретному CVE-идентификатору и объединяет ее в единую запись.

Фактическим функционалом данного сервиса является автоматизация поиска всей доступной информации по CVE-идентификатору с дополнительными функциями поиска по вендорам, типам уязвимостей, оценке критичности по метрикам CVSS и т. п. Также реализованы сбор и хранение различного рода статистики по уязвимостям, например, распределение уязвимостей по степени критичности (согласно метрике CVSS), распределение уязвимостей по вендорам ПО и др. – с удобным переходам к списку уязвимостей, удовлетворяющих выбранному критерию.

Что касается интерфейса, то CVEDetails в целом ориентирован на компактное и удобное для восприятия человеком табличное представление данных, а для автоматизированных систем поддерживает формирование

RSS-подписки (в формате JSON) для получения обновленных данных об уязвимостях выбранных категорий, например, для всех новых уязвимостей класса SQL-инъекций или XSS.

2.2. Vulners

Интересным примером другого подхода является Vulners. Это разработанный российскими специалистами и весьма популярный среди экспертов в области информационной безопасности сервис с собственной базой данных, предназначенный для поиска информации по самым разным материалам в области информационной безопасности.

Фактически Vulners представляет собой поисковый движок с собственной базой данных, адаптированный под предметную область. Таким образом он покрывает гораздо более широкое множество сущностей, чем простые агрегаторы уязвимостей.

В настоящее время база данных Vulners агрегировала в себя порядка 2 млн записей об уязвимостях и примерно 250 тысяч записей об известных эксплойтах. По данному массиву информации возможны поиск по ключевым словам и фильтрация результатов как по источнику информации (организации, опубликовавшей запись об уязвимости), так и по дате публикации записи, CVSS-оценке критичности уязвимости и другим подобным параметрам.

Следует отметить, что Vulners не предоставляет некой единой сводки информации по конкретной уязвимости с заданным CVE-идентификатором (или иным внутренним идентификатором одного из альтернативных реестров), а возвращает множество записей, релевантных поисковому запросу в стиле классического поискового движка. При этом наличие фильтрации результатов по организации-источнику информации (например, type:cvelist) позволяет производить выборку записей только из указанной базы данных.

Все результаты поисковой выдачи из базы данных Vulners могут быть получены не только в удобном для человека, но и в машиночитаемом виде (в формате JSON) через соответствующий API поискового запроса.

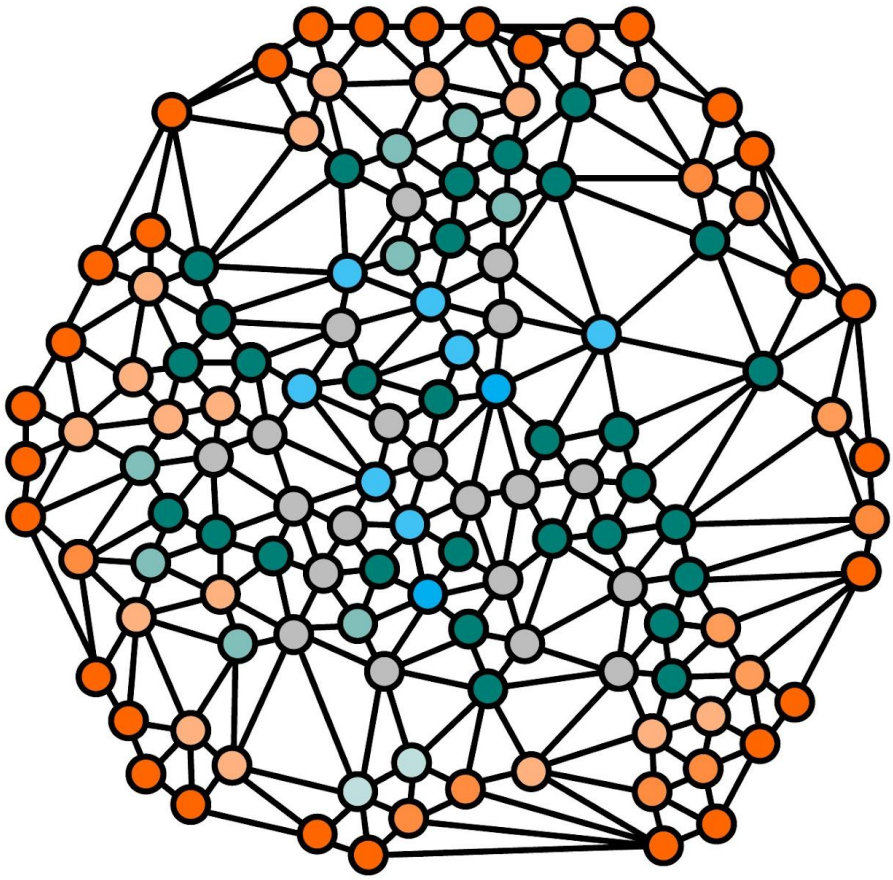
3. Что такое графовые методы анализа и какие средства анализа графов существуют

Графовый (также сетевой) анализ - это набор методов, направленных на изучение связей между сущностями. При помощи этих методов исследуется структура графа и выявляются неочевидные зависимости.

Графовый анализ эффективен, когда мы рассматриваем объекты в контексте связей с другими объектами.

Что такое граф?

Граф - математический объект, который изображает отношения между сущностями. Граф состоит из **вершин** (объектов) и **рёбер** (связей). С помощью графов можно представить разные ситуации: например, пользователей соцсети, которые находятся друг у друга в друзьях; клиентов банка, которые переводят друг другу денежные средства; географические объекты и пути между ними.



Извлекать из графов полезную информацию позволяют графовые алгоритмы, которые можно условно поделить на несколько семейств. Рассмотрим эти семейства на примере социальной сети:

- **Обнаружение сообществ (community detection)** - это выделение тесно связанных между собой групп людей (к примеру, через большое количество общих связей). При этом совсем не обязательно, чтобы все участники взаимодействовали друг с другом напрямую. Это семейство алгоритмов является своего рода аналогом кластеризации.
- **Алгоритмы центральности (centrality algorithms)** поможет выявить лидеров мнений и влиятельных людей в сообществах. Под центральностью мы подразумеваем некоторую меру значимости вершины или ребра.Алгоритмы центральности и сообществ можно применять для создания новых предикторов в ML-pipeline.
- **Предсказание связей (link prediction)** оценивает вероятность наличия связи между двумя отдельными людьми в том случае, если её не существует на графе. Связи, подобранные таким образом, могут помочь в рекомендации друзей.
- **Алгоритмы сходства (similarity algorithms)** пригодятся, чтобы найти похожие группы людей. Это может быть полезно, чтобы собрать аудиторию для рекламы по принципу lookalike или выявить поддельные учетные записи, основываясь на свойствах их окружения.
- **Поиск путей (path detection)** окажется полезен для того, чтобы найти кратчайшую цепочку знакомств между людьми. В качестве меры расстояния (весов ребер) можно использовать характеристики взаимодействия пользователей - например, частоту их общения.

Работа с графами

Взаимодействие с графами отличается от взаимодействия с привычными таблицами. Для этого существуют специальные программные решения, которые перечислены ниже.

- Графовые базы данных

  По сравнению с реляционными и документарными БД, графовые базы позволяют создавать гибкую структуру, в которую можно вносить любые изменения, не ломая её общую архитектуру. Для общения с графовыми СУБД существуют отдельные языки запросов, например, Cypher (Neo4j) и SPARQL. Графовые СУБД выигрывают у

реляционных в скорости в тех случаях, когда мы работаем со связями и перемещаемся по графу. Это обусловлено тем, что каждая вершина графа вместе со своими связями хранится в оперативной памяти, и не используется JOIN.

- Другие инструменты работы с графами

  Помимо графовых СУБД, для работы с графами существуют специальные программные библиотеки (например, для Python написаны популярные библиотеки NetworkX и igraph). Также при необходимости логику графовых вычислений можно частично реализовать и в реляционных базах данных.

ЗАКЛЮЧЕНИЕ

Мы рассмотрели существующие на данный момент базы уязвимости, проанализировали их преимущества и недостатки. Также рассмотрели существующие на данный момент агрегаторы информации об уязвимостях. А также рассмотрели графовые методы анализа, которые мы будем применять при анализе баз уязвимостей.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Андрей Сапожников. Общий обзор реестров и классификаций уязвимостей (CVE, OSVDB, NVD, Secunia). – 2019. – URL: https://safe-surf.ru/specialists/article/5228/607311/

2. Топ-7 крупных баз уязвимостей для поиска и отслеживания новых. – 2019. - URL: https://itsecforu.ru/2019/09/25/%F0%9F%A6%9F-%D1%82%D0%BE%D0%BF-7-%D0%BA%D1%80%D1%83%D0%BF%D0%BD%D1%8B%D1%85-%D0%B1%D0%B0%D0%B7-%D1%83%D1%8F%D0%B7%D0%B2%D0%B8%D0%BC%D0%BE%D1%81%D1%82%D0%B5%D0%B9-%D0%B4%D0%BB%D1%8F-%D0%BF%D0%BE%D0%B8/

3. Что такое CVE и какие угрозы там хранятся? – 2022. - URL: https://habr.com/ru/company/pvs-studio/blog/678410/

4. Графовый анализ — обзор и области применения. – 2021. - URL: https://habr.com/ru/company/glowbyte/blog/594221/