УДК 004.056

ПРИМЕНЕНИЕ БАЗ ДАННЫХ УЯЗВИМОСТЕЙ В ЗАДАЧАХ ИССЛЕДОВАНИЯ БЕЗОПАСНОСТИ ПРОГРАММНЫХ СРЕДСТВ

Белобородов А. Ю., Горбенко А. В.

Национальный аэрокосмический университет им. Н. Е. Жуковского "Харьковский авиационный институт"

В статье рассматривается метод объединения информации об уязвимостях из различных источников и способ получения статистических данных об уязвимостях программного обеспечения. Статистические данные об уязвимостях позволяют произвести расчёт параметров системы массового обслуживания (СМО), что даёт возможность применять методы исследования систем из теории массового обслуживания для моделирования процессов появления и устранения уязвимостей программного обеспечения.

Постановка проблемы. Популяризация сервис-ориентированных систем в промышленности требует более детального изучения безопасности компонентов, из которых такие системы строятся.

Например, уязвимость операционной системы может привести к утечке коммерческой информации, что может повлечь за собой значительные финансовые потери. В таких условиях полезным было бы иметь возможность оценивать прогнозировать безопасность компьютерной системы и её компонентов. Одним из способов прогнозирования безопасности является моделирование процессов выявления и устранения уязвимостей на основе статистических данных, собранных за определённый период жизненного цикла (ЖЦ) программных средств.

В настоящее время существует целый ряд информационных ресурсов Интернет, которые предоставляют информацию об уязвимостях на страницах своих сайтов. К примеру, база данных уязвимостей NVD (www.nvd.nist.gov), позволяет точно идентифицировать уязвимый программный продукт и его версию, получить информацию о способе атаки, при которых данная уязвимость проявляет себя, разновидность угрозы и прочую полезную информацию. В то же время, база данных CVE (www.cve.mitre.org) является единым и первичным поставщиком идентификаторов уязвимостей. Эти идентификаторы используются для однозначного обозначения одной и той же уязвимости другими известными базами данных (OSVDB (www.osvdb.org), Secunia (www.secunia.com), Security Focus (www.securityfocus.com) и др.), базами эксплоитов (Exploit Database (www.exploit-db.com) и др.) и бюллетенями безопасности (Microsoft Security Bulletin [1], US-CERT Security Bulletin [2] и др.).

Однако, несмотря на общедоступность информации об уязвимостях программных продуктов, в открытом доступе, как правило, можно обнаружить лишь списки задокументированных уязвимостей и общую статистику по популярным программным продуктам. Данной информации не достаточно для моделирования процессов выявления и устранения уязвимостей.

Анализ последних исследований и публикаций. Одной из попыток статистического исследования информации об уязвимостях является предпринятая в работе [3]. Однако авторы исследования в своей работе используют лишь данные, предоставляемые базой NVD, упуская возможность получить более полноценные результаты на основе объединения информации об уязвимостях из различных источников. Во-вторых, в статье не рассмотрен вопрос моделирования состояния системы с учётом информации об уязвимостях программных компонентов.

В работах [4, 5] предложен новый подход к оценке информационной безопасности программных средств на основе описания процесса выявления и устранения уязвимостей в виде марковского процесса с помощью цепочек "гибели-размножения". В этом случае показатели оценки информационной безопасности находят своё отображения в показателях, характеризующих параметры системы массового обслуживания, которые, в свою очередь, определяются в результате статистического исследования информации об уязвимостях при комплексном использовании баз данных CVE и NVD.

Целью статьи является разработка метода комплексирования различных баз данных уязвимостей (в первую очередь CVE и NVD) и алгоритмов получения статистических данных, необходимых для дальнейшего моделирования процессов обнаружения и устранения уязвимостей программных средств.

Основные материалы исследования. Как показывает опыт работы с различными базами данных об уязвимостях, полнота представления информации об уязвимостях в этих базах существенно отличается. В данной работе анализируются две наиболее полные базы данных NVD и CVE, которые предоставляют информацию в открытом доступе в формате XML, что удобно для их дальнейшей программной обработки. Каждая уязвимость в базе данных представляет из себя некоторую структуру, состоящую из специальных полей, таких как CVE-идентификатор уязвимости, список программных продуктов, в которых данная уязвимость присутствует, описание уязвимости и др. В таблице 1 приведен перечень полей, которые используются для описания уязвимостей в исследуемых базах.

В процессе объединения данных об уязвимостях возникает задача получения даты обнаружения уязвимости и даты её устранения, т.е. даты выпуска компанией-разработчиком т.н. "заплатки". Для этого были проанализированы все упоминающиеся даты в файлах форматов CVE, CVRF и NVD.

В файлах формата CVRF и NVD встречаются только 2 типа дат: дата публикации (определяет день,

когда запись о конкретной уязвимости была размещена в файле) и дата модификации (указывает дату последнего внесения изменений в описание уязвимости). Данные типы дат для упомянутых форматов являются обязательными полями и указываются для всех записей об уязвимостях. Подробнее CVRF-формат базы данных CVE описан в статье [6].

Таблица 1 – Сравнение баз CVE и NVD

Информация об уязвимости	База данных уязвимостей	
	CVE	NVD
CVE ид. уязвимости	Есть	Есть
Описание	Есть	Есть
Критичность	Нет	Есть
Тип угрозы	Нет	Есть
Даты записей	Есть	Есть
Список уязвимого ПО	Нет	Есть (список CPE)
Влияние на готовность, целостность и конфиденциальность	Нет	Есть
Способ эксплуатации уязвимости	Нет	Есть
Ссылки на дополнительные источники	Есть	Есть

С файлом формата CVE дело обстоит иначе. Во-первых, этот формат содержит поле для одной единственной даты последней модификации информации об уязвимости с указанием стадии уязвимости. Это означает, если записи об уязвимости была присвоена одна дата (к примеру, дата публикации), а с течением времени были внесены изменения в запись, и, соответственно, дата изменилась, то предыдущая дата будет замещена новой (например, датой модификации записи). В таком случае, если не сохранить первоначальную дату, то её значение будет безвозвратно потеряно. Данный формат предусматривает следующие стадии записи об уязвимости: *assignment phase* (начальная стадия записи об уязвимости), *proposal phase* (уязвимость предложена на рассмотрение экспертам), *voting phase* (стадия голосования экспертами по данной уязвимости), *interim decision* (рассмотрение уязвимости экспертами завершено, и необходимо принять окончательное решение по данной уязвимости: каждый эксперт должен выбрать одно из предлагаемых решений: "принять" факт наличия уязвимости или "отклонить"), *final decision* (фаза окончательного принятия решения по уязвимости, после неё уязвимость считается "принятой" или "отклонённой" с указанием причины).

Для объединения информации об уязвимостях была разработана специальная утилита VulnerabilityCollector, целью которой является, используя идентификатор CVE, объединить информацию о каждой уязвимости и получить даты обнаружения и устранения уязвимостей на основе информации, полученной в результате обработки файлов CVE, CVRF и NVD. Полученная информация сохраняется в реляционной базе данных для дальнейшей статистической обработки. Диаграммы сущностей обобщенной базы данных представлена на рисунке 1. В результате объединения XML-файлов баз данных уязвимостей получаем нормализованную информацию об уязвимостях, которую можно в дальнейшем обрабатывать для получения различного рода статистики.

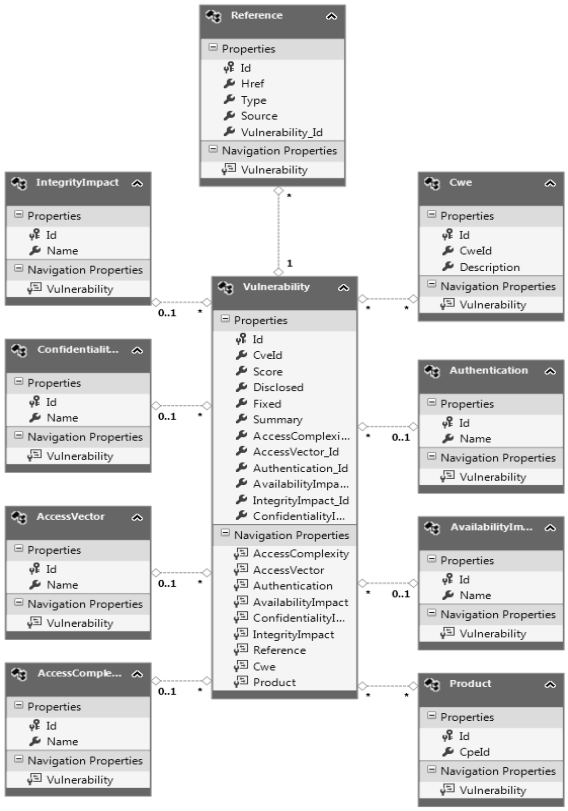

Рисунок 1 – Диаграмма сущностей

Исследования дат, содержащихся в файлах форматов CVE, CVRF и NVD, показывают, что публикация одной и той же уязвимости в NVD происходит значительно позже (как правило, спустя несколько недель, а иногда и месяцев), чем в CVE или CVRF. Данный факт может говорить о том, что специалистам, которые сопровождают базу NVD требуется больше времени перед публикацией информации об уязвимости, чтобы собрать больше информации о ней (метрики критичности уязвимости, список уязвимых продуктов, тип угрозы, представляемой уязвимостью и прочую полезную информацию). Сравнение дат публикации информации об уязвимостях в базе данных NVD с информационными бюллетенями разработчиков уязвимых программных продуктов, в которых анонсируется выпуск исправлений для устранения уязвимостей показывает, что эти даты в большинстве случаев совпадают. Таким образом было выдвинуто допущение о том, что дату публикации информации об уязвимости в базе NVD можно считать датой выпуска "заплатки", устраняющей уязвимость. В то же время, дату публикации информации об уязвимости в базе данных CVE (в формате CVRF) можно считать датой официального выявления (раскрытия) уязвимости.

За основу при объединении информации об уязвимостях из разных источников (CVE, CVRF и NVD),

были приняты следующие допущения: 1) датой раскрытия уязвимости будем считать минимальную из дат в рассмотренных источниках данных; 2) датой устранения уязвимости будем считать максимальную из дат в рассмотренных источниках данных. При этом стоит отметить, что даты модификации записи об уязвимости не участвуют в поиске дат раскрытия и устранения, поскольку они не отражают стадию уязвимости, а только лишь тот факт, что данная запись была изменена.

Получение дат выявления и устранения уязвимостей позволяет применить аппарат марковского моделирования и систем массового обслуживания для оценки уровня уязвимости компьютерной системы в целом. Одними из ключевых характеристик СМО являются интенсивность поступления заявок (интенсивность обнаружения уязвимостей) и интенсивность обслуживания заявок (интенсивность выпуска заплаток для уязвимостей) [5]. В целом алгоритм получения этих параметров на основе статистического анализа обобщенной базы данных об уязвимостях состоит из трёх основных этапов: 1) фильтрация уязвимостей по исследуемому программному продукту и интересующему интервалу времени; 2) группировка уязвимостей по одной из дат, в зависимости типа от параметра, который нужно получить в итоге, и вычисление количества уязвимостей в каждой группе (поле "Count"); 3) группировка результатов предыдущего шага по полю "Count".

Для получения статистических данных об интенсивности обнаружения уязвимостей достаточно (в терминах языка SQL) сгруппировать список уязвимостей исследуемого программного продукта по полю "Disclosed date" (дата раскрытия уязвимости). Тогда в результате получится таблица с полем даты и полем с количеством обнаруженных в этот день уязвимостей ("Count"). Эти данные можно использовать для проверки статистических гипотез о законе распределения случайной величины. Для получения значений интенсивностей появления уязвимостей в случае группового обнаружения уязвимостей, необходимо сгруппировать результаты, полученные на предыдущем шаге по полю "Count" и разделить каждое значение на длительность исследуемого периода.

Для получения параметра интенсивности исправления уязвимостей нужно проделать те же самые действия, как и для первого рассмотренного параметра СМО, за исключением того, что данные второго шага должны группироваться по полю "Fixed date" (дата выпуска заплатки, устраняющей уязвимость).

Выводы. В статье изложен подход к объединению информации об уязвимостях из двух наиболее популярных открытых источников информации об уязвимостях (NVD, CVE), предоставляющих дампы своих баз в виде XML-файлов; исследованы особенности извлечения дат выявления и устранения уязвимостей из файлов форматов NVD, CVE и CVRF рассмотренных баз данных уязвимостей. Результаты исследования могут быть использованы для параметризации системы массового обслуживания и исследования информационной безопасности (защищённости от атак на уязвимости) компьютерных систем.

Список использованных источников

1. Microsoft Security Bulletin [Электронный ресурс] – Режим доступа: https://technet.microsoft.com/en-us/security/bulletin/dn602597.aspx
2. 2015 Bulletins | US-CERT [Электронный ресурс] – Режим доступа: https://www.us-cert.gov/ncas/bulletins
3. OS Diversity for Intrusion Tolerance: Myth or Reality? [Электронный ресурс] – Режим доступа: http://www.di.fc.ul.pt/~bessani/publications/dsn11-osdiversity.pdf
4. Щеглов А. Ю. Безопасность современных ОС "в цифрах" [Электронный ресурс] / А. Ю. Щеглов – Режим доступа: http://www.itsec.ru/articles2/Inf_security/bezopasnost-OS.
5. Белобородов, А. Ю. Применение апарата теории массового обслуживания для исследования процес сов выявления и устранения уязвимостепй программных средств / А. Ю. Белобородов, А. В. Горбенко, В. С. Харченко // Радиоэлектронные и компьютерные системы. – 2014. – №5(69). – с 65-69
6. CVE Usage of CVRF [Электронный ресурс] – Режим доступа: https://cve.mitre.org/cve/cvrf.html

Анотація

ЗАСТОСУВАННЯ БАЗ ДАНИХ ВРАЗЛИВОСТЕЙ У ЗАДАЧАХ ДОСЛІДЖЕННЯ БЕЗПЕКИ ПРОГРАМНИХ ЗАСОБІВ

Білобородов О. Ю., Горбенко А. В.

У статті розглянутий метод об'єднання інформації про вразливості з різних джерел та спосіб отримання статистичних даних про вразливості програмного забезпечення. Статистичні дані про вразливості дозволяє провести розрахунок параметрів системи масового обслуговування (СМО), що дає можливість застосовувати методи дослідження систем із теорії масового обслуговування для моделювання процесів виявлення та усунення вразливостей програмного забезпечення.

Abstract

USAGE OF VULNERABILITY DATABASES IN CASE OF STUDY SOFTWARE SECURITY

O. Biloborodov, A. Gorbenko

A method of combining information about vulnerabilities from different sources and a method of gathering software vulnerabilities statistics are proposed in the article. Statistics allows calculating parameters of queuing system which brings opportunity to apply the queuing theory methods in order to study and model vulnerabilities disclosure and elimination processes.