СПИСОК ЛИТЕРАТУРЫ

1. *Баранов С. Н., Домарацкий А. Н., Ласточкин Н. К., Морозов В. П.* Процесс разработки программных изделий. М.: Наука — Физматлит, 2000. 176 с.
2. *Морозов В. П.* Поддержка принятия решений, ориентированная на знания эксперта // Тр. XII Санкт-Петербург. междунар. конф. „Региональная информатика (РИ-2010)“, 20—22 окт. 2010 г. СПб: СПОИСУ, 2011. С. 69—73.
3. *Журавлев Ю. И., Никифоров В. В.* Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. № 3. С. 1—11.
4. *Тележкин А. М.* Создание исторических баз данных при помощи системы САМПО+ // Тр. Юбилейной XIII Санкт-Петербург. междунар. конф. „Региональная информатика (РИ-2012)“, 24—26 окт. 2012 г. СПб: СПОИСУ, 2013. С. 84—90.

***Сведения об авторе***

***Александр Михайлович Тележкин*** — аспирант; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; E-mail: telezhkin@gmail.com

Рекомендована СПИИРАН

Поступила в редакцию
10.06.14 г.

УДК 004.056

А. В. Федорченко, А. А. Чечулин, И. В. Котенко

# ПОСТРОЕНИЕ ИНТЕГРИРОВАННОЙ БАЗЫ УЯЗВИМОСТЕЙ

Представлены результаты исследования открытых баз данных уязвимостей и описание процесса их интеграции для применения в системах оценивания защищенности компьютерных сетей. Предлагаются модель процесса формирования и структура интегрированной базы уязвимостей, а также описание и анализ разработанного прототипа.

***Ключевые слова:*** *анализ защищенности, базы данных уязвимостей, системы мониторинга безопасности.*

**Введение.** В настоящее время существует большое количество баз данных (БД) уязвимостей, как открытых для общего доступа, так и закрытых, используемых в коммерческих продуктах. Они применяются в различных системах безопасности, сканерах уязвимостей и других средствах обеспечения комплексной защиты компьютерных систем. Однако применение таких баз данных в системах оценивания защищенности компьютерных сетей в режиме реального времени недопустимо вследствие низкой скорости поиска записей уязвимостей для оперативной обработки событий, нарушающих информационную безопасность [1—4]. Также следует отметить, что формирование БД уязвимостей не стандартизовано и производится несогласованно, что влияет на точность обнаружения уязвимостей в используемом программно-аппаратном обеспечении.

Для увеличения объема уникальных записей уязвимостей и списков программно-аппаратных продуктов, соответствующих этим уязвимостям, предлагается объединение открытых баз уязвимостей. Реализация данного процесса предусматривает разработку методики интеграции баз уязвимостей и проектирование структуры интегрированной базы для адаптации ее к быстрому поиску записей уязвимостей. Конечное использование формируемой интегрированной базы уязвимостей подразумевает ее эксплуатацию в системах оценивания защищенности компьютерных сетей, анализ которых должен проводиться в режиме, близком к

реальному времени. При успешной реализации задачи интеграции баз уязвимостей и, как следствие, задачи формирования интегрированной базы предполагается повышение эффективности работы систем оценивания защищенности.

**Анализ открытых баз уязвимостей.** Для формирования интегрированной базы уязвимостей был проведен анализ ряда открытых баз уязвимостей, таких как: Общие уязвимости и воздействия (Common Vulnerabilities and Exposures — CVE) [5], Национальная база данных уязвимостей (National Vulnerabilities Database — NVD) [6], Открытая база данных уязвимостей (Open Source Vulnerabilities Data Base — OSVDB) [7].

Анализ базы данных CVE показал, что она обладает высокой степенью связности с другими источниками описания уязвимостей. Однако недостатком данной базы является отсутствие в описании уязвимостей спецификации программно-аппаратного обеспечения, для которого характерна эта уязвимость.

Более подробное описание содержащихся в базе CVE уязвимостей представлено в базе данных NVD. Эта база включает в себя описание уязвимостей с указанием подверженных им программно-аппаратных продуктов в формате CPE. Кроме того, данная база содержит показатели, характеризующие уязвимости в формате Общей системы оценивания уязвимостей (Common Vulnerability Scoring System — CVSS) [8]. Однако количество ссылок на другие базы данных, по которым производится интеграция записей уязвимостей, в базе NVD меньше, чем в базе CVE (последняя включает все ссылки базы NVD).

Следует также отметить, что база NVD единственная из исследуемых баз, содержащая логические описания конфигураций программно-аппаратного обеспечения для каждой уязвимости. Данные конфигурации определяют уязвимые системы и представляют собой списки программно-аппаратного обеспечения, объединенные логическими операторами И/ИЛИ. Считается, что конфигурация содержит зависимость в том случае, если она содержит логический оператор И (т.е. определяет уязвимую конфигурацию, содержащую одновременно не менее двух продуктов). Анализ базы NVD показал, что из всех уязвимостей зависимости содержат только 10,63 % записей. Причем каждая из этих зависимостей содержит только один логический оператор И. Такой невысокий процент зависимостей обусловливает возможность хранения конфигураций программно-аппаратного обеспечения в одной таблице, что позволяет значительно снизить вычислительные ресурсы, необходимые для поиска подходящих уязвимостей.

При анализе базы данных OSVDB было выявлено множество ссылок на сторонние источники, причем наилучшими показателями по количеству и уникальности обладают ссылки на базы CVE и NVD.

По результатам проведенного анализа именно рассмотренные базы данных стали основными источниками информации об уязвимостях для формирования интегрированной базы.

**Модель процесса интеграции данных открытых баз уязвимостей.** Данные, хранящиеся в рассмотренных открытых базах уязвимостей и необходимые для системы оценивания защищенности, условно разделяются на две группы: 1) описания уязвимостей (набор характеристик, указывающих на причины их возникновения, условия успешной реализации уязвимости и способы ее устранения); 2) описания конфигурации уязвимых продуктов, для которых данная уязвимость характерна. На этой основе в процессе интеграции открытых баз уязвимостей выделяются две составляющие: интеграция записей уязвимостей и интеграция описаний продуктов.

Интеграция записей уязвимостей производится по указанным в описании ссылкам, которые делятся на две категории: 1) прямые, т.е. непосредственно указывающие на источник информации (идентификатор записи уязвимости), используемый при интеграции; 2) косвенные, указывающие на источники информации, не используемые при интеграции. Для соответствия уязвимостей, выявленных по прямым ссылкам, необходимо и достаточно наличие одного указания на используемые источники, а для соответствия по косвенным ссылкам —

наличие двух указаний. Таким образом, формирование интегрированного списка уязвимостей можно представить в виде модели, приведенной на рис. 1.

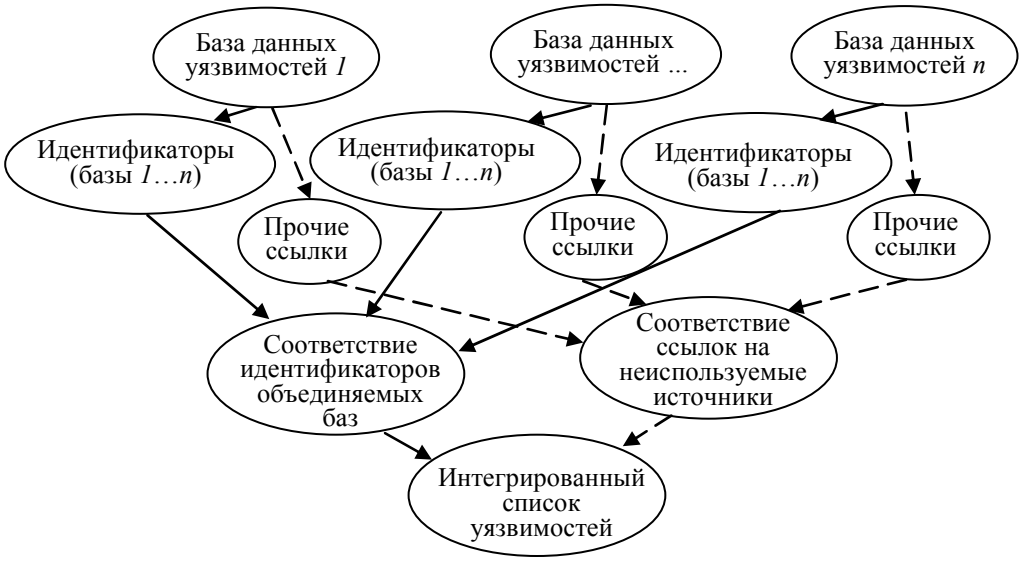


*Рис. 1*

Следует отметить, что если уязвимости в пределах одной базы содержат ссылки на записи из этой же базы, то такие уязвимости объединяются в одну запись интегрированного списка, что обеспечивает уникальность каждой записи уязвимости в формируемой базе.

При формировании интегрированного словаря продуктов за основу берется схема „Общее перечисление платформ“ (Common Platform Enumeration — CPE) [9]. CPE — это структурированная схема именования компьютерных систем и платформ, основанная на синтаксисе Универсальных идентификаторов ресурсов (Uniform Resource Identifiers — URI) [10]. Выбор словаря CPE обусловлен тем, что он содержит большее количество описаний продуктов по сравнению с аналогичными словарями, а формат представления продуктов является лучшим на данный момент. На первом этапе интеграции описаний продуктов производится сравнение записей, содержащихся в словаре CPE, с записями, имеющимися в пространствах имен используемых баз уязвимостей. Затем оба вида записей объединяются в список и в результате формируют интегрированный словарь продуктов. Модель процесса формирования интегрированного словаря продуктов представлена на рис. 2.

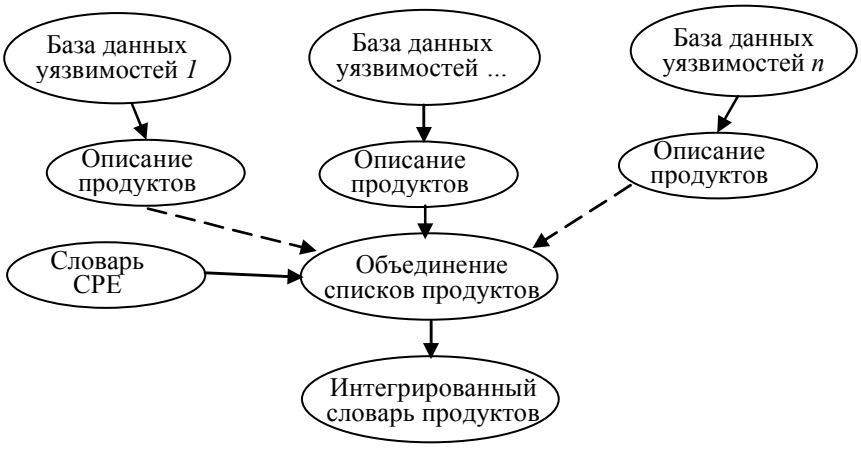


*Рис. 2*

Несмотря на то, что в качестве основы для формирования интегрированного словаря продуктов был использован словарь CPE, формат самих записей был изменен и приобрел следующий вид:

{тип}:{производитель}:{продукт}:{версия}:{модификация}:{ редакция}.

**Структура интегрированной базы уязвимостей.** На основе предложенной модели интеграции записей уязвимостей и списков продуктов была разработана структура интегрированной базы уязвимостей. Под структурой в данном случае понимается реляционная модель базы данных [11]. В результате выполнения процесса интеграции вся информация об уязвимостях и программно-аппаратном обеспечении записывается в реляционные таблицы 1) производителей, 2) продуктов, 3) конечных описаний продуктов (включает поля „тип“, „версия“, „модификация“, „редакция“) и 4) уязвимостей, а также таблицу связности (5) записей уязвимостей и описаний продуктов (включает поле „параметр зависимости“): см. рис 3. При этом связь содержащихся в таблицах данных осуществляется за счет использования первичных ключей PK (однозначно определяющих каждую запись в таблице) и внешних ключей K (хранящих значение первичного ключа из другой таблицы).

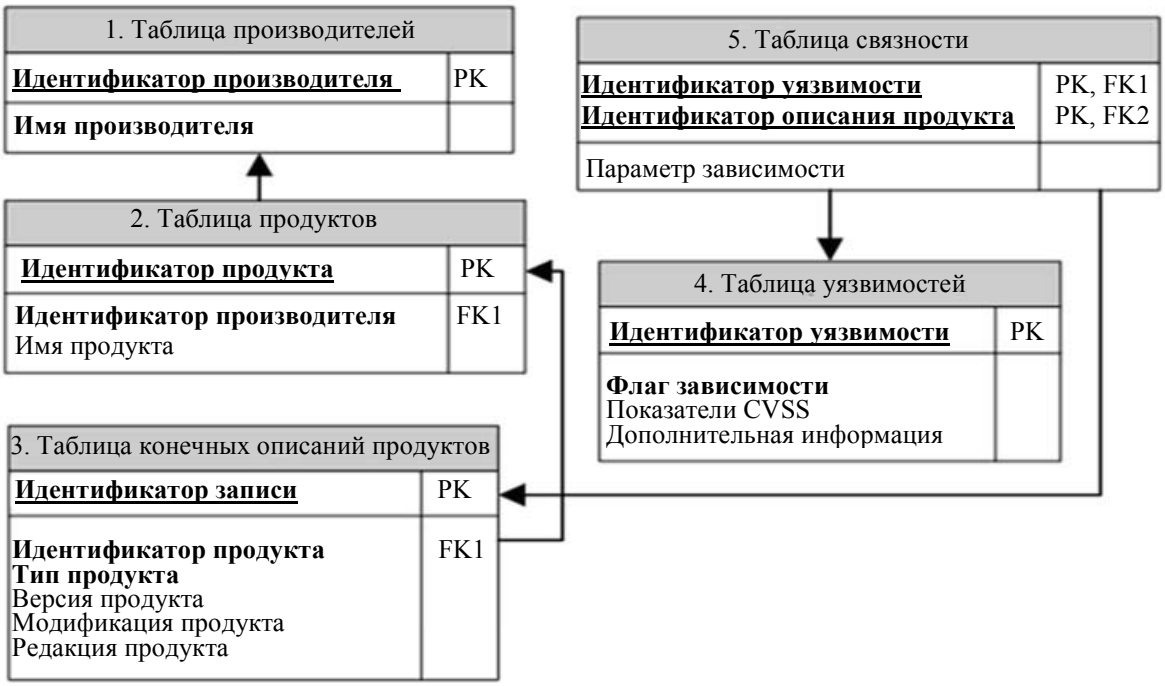


*Рис. 3*

Параметр зависимости в 5-й реляционной таблице (см. рис. 3) является единственным признаком, определяющим связь между уязвимым программно-аппаратным обеспечением и продуктами, при которой уязвимость может быть реализована. Таким образом, если указанный в качестве ключа продукт имеет уязвимость без такой зависимости, то поле „параметр зависимости“ этой записи будет иметь значение „0“. Если же уязвимость, характерная для некоторого продукта, реализуется только в случае наличия другого продукта, то такая зависимость описывается следующим образом:

— группа продуктов, для которых могут быть реализованы уязвимости в зависимости от наличия в системе любого из продуктов другой группы, является группой *A*;

— группа продуктов, влияющих на возможность реализации уязвимости продуктов группы *A*, является группой *B*;

— значение поля „параметр зависимости“ (в 5-й реляционной таблице) для продуктов группы *A* будет нечетным, начиная со значения „1“;

— значение поля „параметр зависимости“ (в 5-й таблице) для продуктов группы *B* будет четным, начиная со значения „2“.

Рассмотренная структура интегрированной базы данных уязвимостей позволяет с помощью единственного SQL-запроса к ней осуществлять поиск уязвимостей, присущих программно-аппаратным платформам, задаваемым в качестве списка идентификаторов описаний продуктов при реализации запроса.

**Прототип интегрированной базы уязвимостей.** В качестве практической реализации интегрированной БД уязвимостей был разработан прототип, для формирования которого использовались следующие источники описания уязвимостей и продуктов: CVE, CPE, NVD, OSVDB. Для оценки эффективности построения интегрированной БД уязвимостей следует воспользоваться количественными показателями, приведенным в табл. 1.

*Таблица 1*

| Источник данных | Общее количество записей уязвимостей | Общее количество описаний продуктов | Количество уникальных записей уязвимостей |
|---|---|---|---|
| CVE | 67 966 | — | 48 621 |
| CPE | — | 82 987 | — |
| NVD | 59 779 | 147 505 | 48 621 |
| OSVDB | 98 625 | 127 564 | 59 827 |
| Интегрированная база уязвимостей | **73 908** | **183 321** | **73 908** |

Опираясь на представленные результаты, можно вывести показатель ($P$), характеризующий преимущество интегрированной базы данных относительно использованных для ее формирования источников информации. Результаты сравнительного анализа представлены в табл. 2.

*Таблица 2*

| Источник данных | $P$, % | |
|---|---|---|
| | Интегрированный список уязвимостей | Интегрированный словарь продуктов |
| CVE | 34 | — |
| CPE | — | 55 |
| NVD | 34 | 20 |
| OSVDB | 19 | 30 |

Представленный прототип интегрированной базы данных был использован как базовый компонент для построения автоматизированной системы моделирования атак и анализа защищенности компьютерных сетей [4, 12]. Для проверки эффективности функционирования интегрированной БД была проведена серия экспериментов по моделированию различных сценариев атак.

**Заключение.** Результаты анализа открытых баз уязвимостей продемонстрировали имеющиеся в них проблемы, которые, в свою очередь, негативно сказываются на процессе мониторинга компьютерных систем на предмет наличия уязвимого программного и аппаратного обеспечения. Для преодоления выявленных проблем в данной статье предложена модель интеграции данных открытых баз уязвимостей, определены форматы данных и осуществлено проектирование структуры интегрированной базы данных уязвимостей. В качестве практической реализации разработан ее прототип.

Эксперименты, проведенные с использованием прототипа, доказали преимущество и конкурентоспособность сформированной интегрированной базы по отношению к лидерам в области баз уязвимостей, что делает ее подходящей в качестве компонента для использования в системах оценивания защищенности.

Направлением дальнейших исследований является разработка и модификация интегрированной базы данных уязвимостей за счет увеличения числа используемых открытых баз, присвоения показателей доверия источникам ссылок и расширенного анализа уникальности уязвимостей для обеспечения наиболее качественной интеграции.

Статья подготовлена по результатам работы, выполняемой при финансовой поддержке Российского фонда фундаментальных исследований (гранты 13-01-00843, 13-07-13159, 14-07-00697, 14-07-00417, 14-37-50735), программы фундаментальных исследований ОНИТ РАН (контракт № 2.2), проекта ENGENSEC программы Европейского сообщества TEMPUS и государственных контрактов № 14.604.21.0033, 14.604.21.0137, 14.604.21.0147 и 14.616.21.0028.

СПИСОК ЛИТЕРАТУРЫ

1. *Kotenko I., Stepashkin M.* Network security evaluation based on simulation of malefactor's behavior // Proc. of the Intern. Conf. on Security and Cryptography (SECRYPT'06). Setubal, Portugal, 2006. P. 339—344.
2. *Котенко И. В., Степашкин М. В., Богданов В. С.* Архитектуры и модели компонентов активного анализа защищенности на основе имитации действий злоумышленников // Проблемы информационной безопасности. Компьютерные системы. 2006. № 2. С. 7—24.
3. *Ruiz J. F., Harjani R., Maña A., Desnitsky V., Kotenko I., Chechulin A.* A methodology for the analysis and modeling of security threats and attacks for systems of embedded components // Proc. of the 20th Euromicro Intern. Conf. on Parallel, Distributed and Network-Based Processing (PDP 2012). Garching, Germany. 2012. C. 261—268.
4. *Kotenko I. V., Chechulin A. A.* Common framework for attack modeling and security evaluation in SIEM systems // IEEE Intern. Conf. on Green Computing and Communications, Conf. on Internet of Things, and Conf. on Cyber, Physical and Social Computing; Besançon, France, 11—14 Sept., 2012; Los Alamitos, CA: IEEE Computer Society, 2012. P. 94—101.
5. Common Vulnerabilities and Exposures (CVE) [Электронный ресурс]: <http://cve.mitre.org/>, 05.05.2014.
6. National Vulnerabilities Database (NVD) [Электронный ресурс]: <http://nvd.nist.gov/>, 05.05.2014.
7. Open Source Vulnerabilities Data Base (OSVDB) [Электронный ресурс]: <http://osvdb.org/>, 05.05.2014.
8. Common Vulnerability Scoring System (CVSS) [Электронный ресурс]: <http://www.first.org/cvss>, 05.05.2014.
9. Common Platform Enumeration (CPE) [Электронный ресурс]: <http://cpe.mitre.org>, 05.05.2014.
10. *Котенко И. В., Дойникова Е. В.* Анализ протокола автоматизации управления данными безопасности SCAP // Защита информации. Инсайд. 2012. № 2. С. 56—63.
11. *Фиайли К.* SQL “Quick Start”. М.: ДМК Пресс, 2003. 456 с.
12. *Kotenko I., Chechulin A.* Attack modeling and security evaluation in SIEM systems // Intern. Transact. on Systems Science and Applications. 2012. Vol. 8, Dec. P. 129—147.

***Сведения об авторах***

***Андрей Владимирович Федорченко*** — СПИИРАН, лаборатория проблем компьютерной безопасности; мл. научный сотрудник; E-mail: fedorchenko@comsec.spb.ru

***Андрей Алексеевич Чечулин*** — канд. техн. наук; СПИИРАН, лаборатория проблем компьютерной безопасности; E-mail: chechulin@comsec.spb.ru

***Игорь Витальевич Котенко*** — д-р техн. наук, профессор; СПИИРАН, лаборатория проблем компьютерной безопасности; E-mail: ivkote@comsec.spb.ru

Рекомендована СПИИРАН

Поступила в редакцию
10.06.14 г.