МЕТОДИКА ОЦЕНКИ ЗАЩИЩЕННОСТИ ВЕБ-РЕСУРСОВ НА БАЗЕ МЕТРИКИ CVSS

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Давлатов Ш.Р.

Кучинский П.В. – доктор физико-математических наук

В рамках данной работы была создана система для сбора технической информации об активных веб-ресурсах в сети интернет из общедоступных каталогов и реестров. На основе анализа данных об уязвимостях веб-ресурсов и метрики CVSS определяется распределение усредненной величины оценки уязвимости для каждого ресурса. Разработаны шаблоны поиска с помощью RegExp выражений языка JavaScript для точного определения версий технологий, которые были использованы для создания веб-сайтов. На базе полученных данных установлены процентные соотношения используемых технологий, доменов верхнего уровня и географическое расположение серверов, которые обслуживают веб-ресурсы. Данная разработка была апробирована на примере 19 тысяч наиболее популярных веб-ресурсов Беларуси.

С развитием веб-технологий растет и число потенциальных уязвимостей в онлайн-ресурсах. Широкое использование инструментов для реализации угроз информационной безопасности в Интернете определяет актуальность использования систем для анализа безопасности веб-ресурсов. Специалисты по защите информации широко используют объективные количественные показатели защищенности, которые вычисляются на основе метрик открытой системы оценки CVSS [1] (Common Vulnerability Scoring System). Существуют открытые базы данных, где хранится информация об уязвимостях определенных версий технологий в формате <название технологии, версия, оценка CVSS>. Метрика CVSS предлагает простой инструментарий для расчета числового показателя уязвимости по десятибалльной шкале [2]. Чем выше значение метрики, тем более оперативная реакция требуется для исправления проблемы безопасности системы.

В рамках данной работы были разработаны NodeJS скрипты для автоматического сбора информации о веб-ресурсах из открытых источников shodan.io и censys.io. В результате процесса сканирования удалось собрать данные более 19 тысяч веб-ресурсов Беларуси, которые были разделены на 5 категорий. Для каждого отдельного домена была получена техническая информация в формате: IP-адреса, открытые порты, географическое расположение веб-серверов и заголовки ответов HTTP. Далее, с помощью RegExp выражений языка JavaScript определяются версии технологий, на базе которых были созданы веб-ресурсы. Для решения поставленной задачи достаточно получить исходный код страницы веб-сайта в формате HTML и заголовки ответов сервера [3]. Результаты запросов и заголовки ответов сервера были сохранены в локальной базе данных для последующих процессов обработки и анализа данных. Данный скрипт также может обнаруживать типы систем управления контентом (CMS), платформы электронной коммерции, версии веб-фреймворков, серверное программное обеспечение и аналитические инструменты. Последним этапом является проверка безопасности каждого веб-ресурса по отдельности на базе публичных API.

На основе полученных данных о версиях технологий была создана диаграмма, показывающая процентное соотношение используемых технологий для разработки веб-ресурсов в шести категориях (рис. 1).

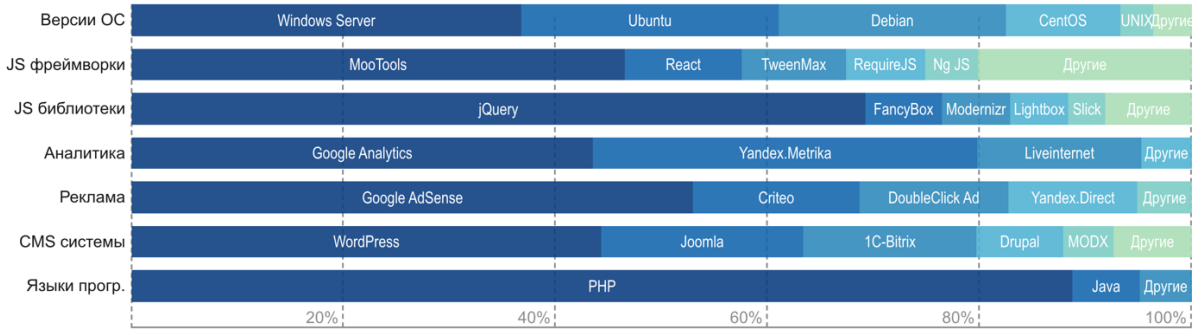

Рисунок 1 – Процентное соотношение используемых технологий в веб-ресурсах

Данные о географических расположениях серверов, обслуживающие веб-ресурсы из нашей исходной выборки, представлены в виде наглядной столбчатой диаграммы (рис. 2). На оси X расположены названия стран, где сосредоточено наибольшее количество веб-серверов: Беларусь, Россия, США, Германия и другие страны (Нидерланды, Польша, Украина и Великобритания).

Каждый столбец показывает количество серверов, которые обслуживают веб-ресурсы определенной категории в той или иной стране.

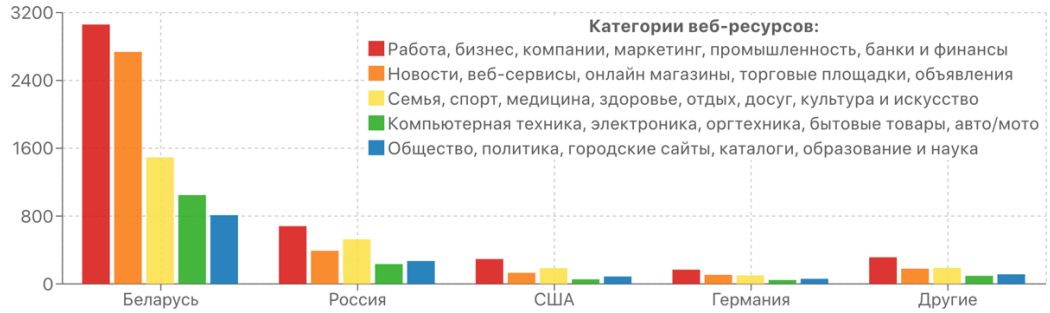

Рисунок 2 – Распределение веб-ресурсов по категориям и странам расположения серверов

Процентное соотношение доменов верхнего уровня показаны на рисунке 3 в виде круговой диаграммы. Данные были получены из исходной выборки веб-ресурсов путем приведения доменных имен к каноническому виду. Из рисунка видно, что домены BY составляют примерно 75% из всех имеющихся записей в базе данных. А в свою очередь домены COM и RU составляют 8,8% и 8,2% соответственно. Следует отметить, что все остальные домены (NET, БЕЛ, ORG и другие) были объединены в одну категорию с процентной долей 7,5%.

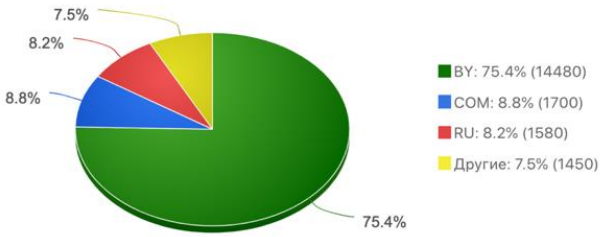

Рисунок 3 – Процентное соотношение доменов верхнего уровня

В рамках данного исследования также был проведен анализ оценки уязвимостей всех веб-ресурсов из нашей исходной выборки. Для каждого веб-ресурса была использована усредненная оценка CVSS в соответствии с выражением $Si = (A1+A2+\ldots+Am)/m$, $1 \le i \le n$, где n – количество элементов в нашей выборке; m – количество распознанных версий технологий и ЯП для заданного веб-ресурса; Aj – оценка уязвимости определенной версии технологии. Для вычисления значения Aj была выбрана функция максимума по всем известным оценкам CVSS: $Aj = \max(Ck)$, $1 \le k \le r$, где r – количество найденных уязвимостей для определенной версии технологии в общедоступной базе vulners.com. Для исследования распределения была создана случайная выборка из исходной базы данных веб-ресурсов, состоящая из N = 2000 элементов (около 10% записей). На основе этих данных была построена диаграмма эмпирического распределения усредненной оценки уязвимости Si веб-ресурсов (рис. 4).

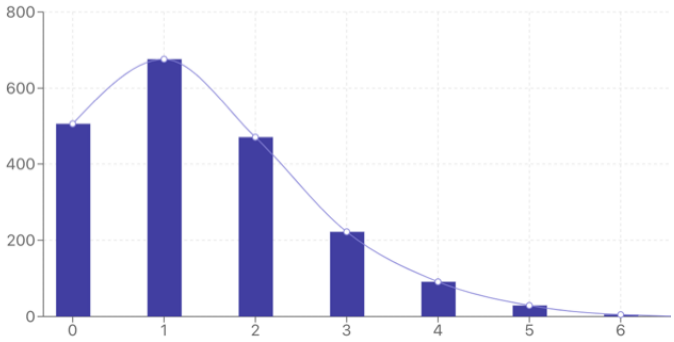

Рисунок 4 – Диаграмма эмпирического распределения оценок уязвимостей веб-ресурсов

Следует отметить, что в исходной генеральной совокупности, порядка 90% веб-ресурсов имеют значение усредненной оценки CVSS в интервале [0, 5]. Было выявлено, что остальные 10% ресурсов имеют высокую критическую оценку уязвимости из-за использования устаревших версий таких технологий как: CMS WordPress, ЯП PHP, веб-сервера nginx и JavaScript библиотеки jQuery.

Список использованных источников:

1. Дойникова, Е. В., Чечулин, А. А., Котенко, И. В. Оценка защищенности компьютерных сетей на основе метрик CVSS // Информационно-управляющие системы, 76-87. DOI: 10.15217/issn1684-8853.2017.6.76.

2. Li, H., Zhao, L. Study on the distribution of CVSS environmental score // 5th International Conference on Electronics Information and Emergency Communication. May 2015. DOI: 10.1109/ICEIEC.2015.7284502.

3. Bostic, T., Stanley J., Higgins, J., Chudnov, D., Montgomery, B., Brunell, J. Exploring the Intersections of Web Science and Accessibility // The MITRE Corporation Scientific journal. Aug 2019.