УДК 004.056:004.77

Е.И. Неткачёва

Таврический национальный университет им. В.И. Вернадского, Симферополь, Украина

ИССЛЕДОВАНИЕ ДИНАМИКИ ОБНАРУЖЕНИЯ УЯЗВИМОСТЕЙ В ПРОГРАММНЫХ КОМПОНЕНТАХ С ИСПОЛЬЗОВАНИЕМ ЛОГИСТИЧЕСКОЙ КРИВОЙ

Исследуются и сравниваются характеристики безопасности различных программных продуктов на основании информации об уязвимостях, представленной в открытых источниках. С помощью логистической кривой проводится аппроксимация кумулятивного профиля отказов, определяются границы этапов, что позволяет ввести дополнительную метрику, сделать выводы о стадии, на которой находится продукт и прогнозировать ожидаемую интенсивность обнаружения уязвимостей в будущем.

Ключевые слова: *уязвимости, коэффициент уязвимостей, логистическая функция, OTS компоненты.*

Введение

Компонентно-ориентированный COTS-подход [1] используется во многих современных системах, в том числе системах критического применения, используемых в атомной энергетике [2], военных и аэрокосмических комплексах [3]. В связи с этим, оценка надежности и безопасности (C)OTS компонетов является актуальным направлением современных исследований. В последние годы активно развиваются разнообразные ресурсы, направленные на поиск, учет и хранение данных о проблемах в информационной безопасности программного обеспечения. Примерами таким ресурсов являются NVD[4], OSVDB [5], Secunia [6] и др. Их цель – проинформировать разработчиков и пользователей ПО, помочь опознать и разрешить уже известные проблемы надежности и безопасности, а также своевременно узнавать о новых уязвимостях и устранять их на ранних стадиях обнаружения. На основании информации из таких ресурсов имеется возможность провести исследование показателей надежности и безопасности различных OTS компонентов, используя количественную или вероятностную оценку.

Подходы, используемые для такой оценки, а также некоторые полученные результаты представлены в работах [7, 8].

В данной работе автором производится анализ обнаруженных уязвимостей с помощью логистической кривой и сравнение различных программных продуктов на основании полученных результатов.

Построение кумулятивного профиля уязвимостей

Одним из возможных вариантов оценки и сравнения аналогичных по функциональности программных продуктов является оценка общего (кумулятивного) числа найденных в них уязвимостей за определенные промежутки времени. Общее (кумулятивное) число N¬i отказов ПО к моменту времени t=ti·i является одной из базовых характеристик надёжности и вычисляется как:

$$N_i = n_1 + n_2 + \ldots + n_i$$

где n_i – количество отказов ПО на i-том интервале тестирования, t_i – длина временного интервала, t – общее время тестирования. Для рассмотрения были выбраны три ветки Apache: 1.3, 2.0, 2.2 от "Apache Software Foundation" или "Apache Group" поставщиков и три версии IIS: 5.0, 6.0 и 7.0 от компании "Microsoft". На основании анализа уязвимостей, обнаруженных в рассматриваемых версиях ПО, были построены графики $N_i(t)$ кумулятивных профилей отказов (рис. 1).

Визуальный анализ представленного графика позволяет увидеть, что график продукта Apache 2.0 является более крутым. Крутой склон кумулятивного профиля является показателем большей частоты проявления уязвимостей. Следовательно, Apache 2.0 имеет худшие характеристики по частоте обнаружения уязвимостей по отношению к другим версиям продукта Apache. График позволяет проследить темпы роста количества обнаруженных уязвимостей: в ноябре 2001 года на момент выпуска версии Apache 2.0 версия Apache 1.3 уже существовала на рынке, к этому времени в ней было найдено 15 уязвимостей. Однако, в 2006 количество уязвимостей, обнаруженное в Apache 2.0, сравнялось с количеством, найденном в Apache 1.3 и после этого только обгоняло Apache 1.3 по темпам роста. Такая динамика свидетельствует о более низких показателях безотказности версии Apache 2.0 по сравнению с версией Apache 1.3.

Исследование кумулятивного профиля с помощью логистической кривой

Логистическая функция представляет собой сигмоидальную кривую и является гладкой монотонно возрастающей всюду дифференцируемой S-образной нелинейной функцией.

В качестве аналитического выражения для логистической функции будем использовать следующее выражение:

$$f(t) = \alpha \Big/ \left(1 + \exp\left(-\frac{t-\mu}{s}\right)\right)$$

где α – параметр, отвечающий за общее количество уязвимостей, s – параметр наклона логистической функции, изменение которого позволяет построить функции с различной крутизной, μ – медиана, параметр, определяющий положение кривой по оси абсцисс.

Результат аппроксимации для выбранных программных продуктов с помощью логистической функции представлен на рис. 2.

Аппроксимация кумулятивного профиля отказов с использованием логистической кривой, которую также называют кривой жизненного цикла, позволяет определить, на какой стадии развития находится исследуемый программный продукт. Для этого можно условно выделить три периода эволюционирования продукта: начальная стадия, активное развитие и период зрелости. На начальной стадии рост общего количества обнаруженных уязвимостей примерно соответствует экспоненте (показательной

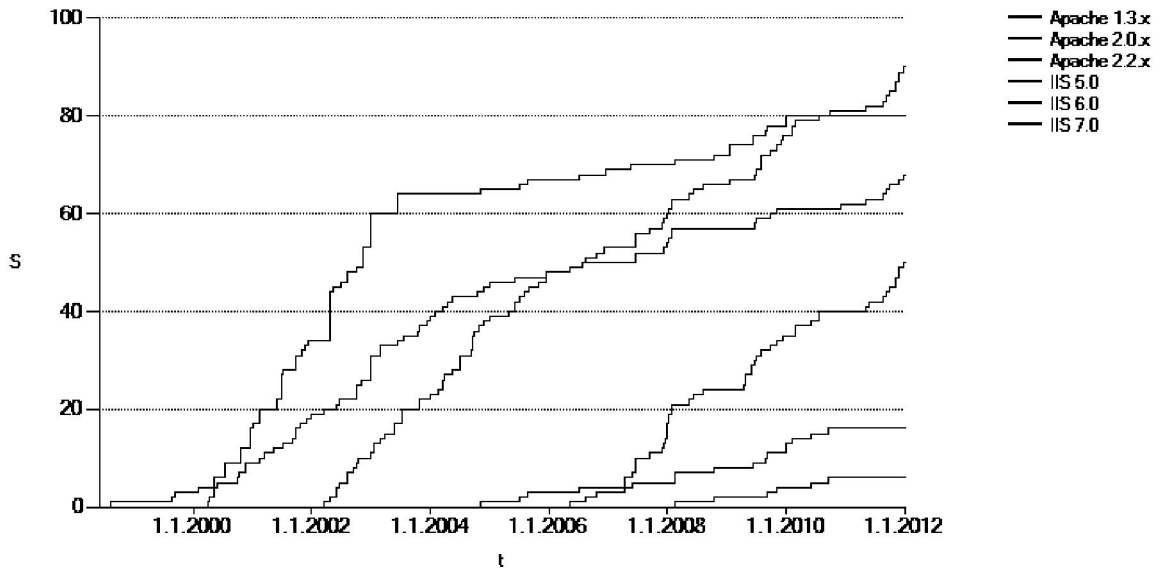

Рис. 1. Кумулятивные профили отказов веб-серверов Apache и IIS

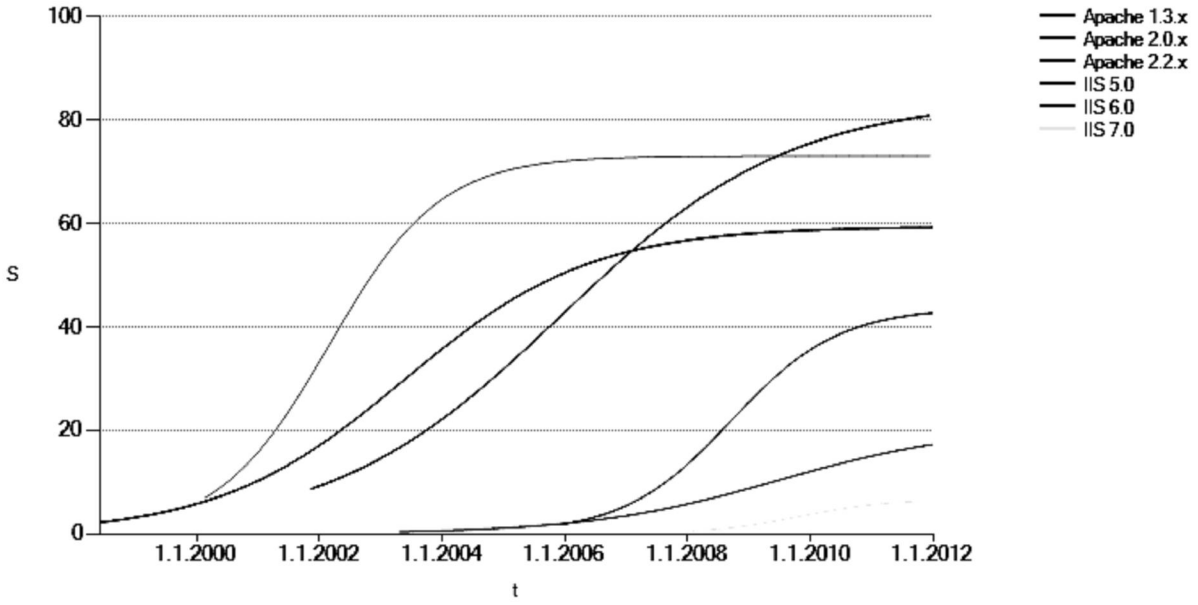

Рис. 2. Аппроксимация кумулятивного профиля отказов с использованием логистической кривой

функции), в период развития рост переходит в линейную фазу, и затем на этапе зрелости продукта практически останавливается.

Исследование логистической функции позволяет обозначить границы этих этапов. Аналитическое выражение первой производной функции представляет собой функцию распределения плотности вероятности с учетом параметра α кумулятивного количества уязвимостей:

$$f'(t) = \frac{\alpha \cdot \exp\left(-\frac{t-\mu}{s}\right)}{\left(1+\exp\left(-\frac{t-\mu}{s}\right)\right)^2 \cdot s}$$

График кумулятивной функции распределения (логистической функции) и соответствующей функции плотности вероятности логистического распределения представлен на рис. 3.

Математическое ожидание μ функции плотности вероятности соответствует значению середины интервала активного развития продукта на S-образной кривой, с границами этого интервала в точках [μ-σ; μ+σ], где σ – среднеквадратичное отклонение. Значение среднеквадратичного отклонения, выражается через параметр s логистической функции f(x) следующим образом:

$$\sigma = \frac{\pi}{\sqrt{3}} s$$

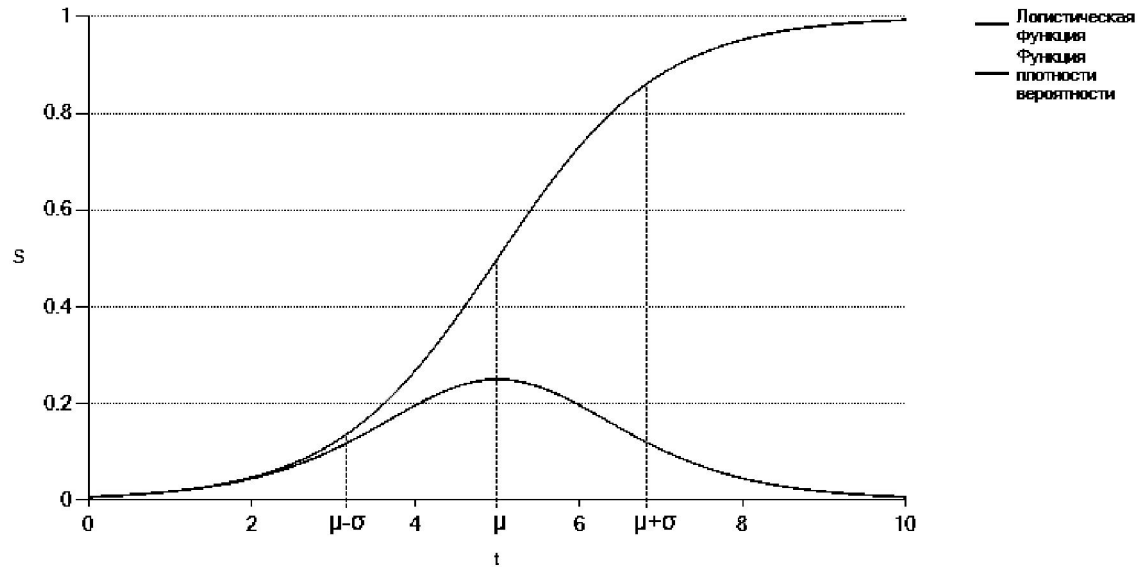

Рис. 3. Вид кумулятивной функции распределения
и функции плотности вероятности логистического распределения

При выборе ПО предпочтение стоит отдавать программным продуктам, которые прошли этапы начального и интенсивного развития и находятся в текущий момент на зрелой стадии. Это позволит минимизировать вероятность обнаружения новых уязвимостей при использовании данного ПО.

С целью формального определения стадии развития, а следовательно, ожидаемой стабильности функционирования продукта по кумулятивному профилю отказов, аппроксимированному с помощью логистической функции, введем дополнительную метрику τ – коэффициент уязвимостей ПО. Значение показателя оценивается следующим выражением:

$$\tau = \frac{t - \mu}{\sigma}$$

где t – момент времени исследования, μ – значение середины интервала активного развития продукта, σ – среднеквадратичное отклонение.

На основании значения показателя τ можно сделать выводы о текущей стадии продукта и возможном дальнейшем развитии:

– $\tau < -1$ характеризует начальную стадию, велика вероятность обнаружения большого количество уязвимостей в будущем

– $\tau \in [-1; 0)$ характеризует первую половину стадии интенсивного развития продукта, на которой будет найдено основное количество уязвимостей, ожидается активный рост количества обнаруженных уязвимостей

– $\tau \in [0; 1]$ характеризует вторую половину стадии интенсивного развития продукта, вероятность обнаружения новых уязвимостей уменьшается

– $\tau > 1$ характеризует стадию зрелости продукта, низкая вероятность обнаружения новых уязвимостей.

Чем выше значение показателя τ, тем более стабильным является программный продукт и тем предпочтительнее его использование по сравнению с другими аналогами. Таким образом, вычисление данной метрики позволяет определить стадию развития, на которой находится исследуемое ПО, прогнозировать уровень обнаружения новых уязвимостей в будущем, а также сравнить аналогичные программные продукты с целью выбора наиболее предпочтительного варианта.

Значения показателя τ, рассчитанного для рассматриваемых Apache и IIS компонентов, представлено в табл. 1.

Рассчитанные значения коэффициентов показывают, что наиболее стабильными из исследуемых веб-серверов на текущий момент являются версии Apache 1.3 и IIS 5.0, которые достаточно давно прошли стадию интенсивного развития и в настоящее время являются зрелыми программными продуктами с низкой вероятностью обнаружения новых уязвимостей. Эти результаты также подтверждаются информацией о количестве уязвимостей, обнаруженных в данных версиях за последние годы.

Экстраполяция результатов оценки

С целью решения задачи прогнозирования количества уязвимостей и дефектов, имеющихся в программном продукте, предлагается использовать статистические методы.

При прогнозировании показателей делается предположение, что закономерности и тенденции, установленные на основе анализа уязвимостей, найденных в программных компонентах в прошлом, сохранятся и в прогнозируемом будущем. На основании статистических данных проводится анализ различных показателей надежности и безопасности ПО и строятся графики временных зависимостей. Далее, происходит распространение выявленных в прошлом закономерностей на будущий период времени.

С целью демонстрации данного подхода предлагается выбрать в качестве примера несколько версий веб-серверов, которые имеют невысокое значение коэффициента τ и соответственно представляют интерес в плане прогнозирования количества уязвимостей в будущем. Результат аппроксимации и экстраполяции, проведенной с помощью логистической функции, представлен на рис. 4.

Заключение

В работе было проведено исследование динамики обнаружения уязвимостей в программных компонентах на основе данных, полученных из открытых источников уязвимостей. Построены кумулятивные профили обнаруженных уязвимостей ПО, проведена аппроксимация и экстраполяция профилей с помощью логистической функции. Введен новый показатель уязвимостей, позволяющий формально определить стадию развития программного продукта, сделать выводы о наличие в нем уязвимостей и дальнейшей динамике их обнаружения. Предложенный подход позволяет производить сравнение различных программных продуктов, а также прогнозировать количество уязвимостей, которые будут обнаружены в будущем.

Таблица 1

Коэффициенты уязвимостей, вычисленные для веб-серверов Apache и IIS

Программный продукт	Apache 1.3	Apache 2.0	Apache 2.2	IIS 5.0	IIS 6.0	IIS 7.0
Коэффициент уязвимостей, τ	3.2	1.8	2.1	6.1	0.9	1.7

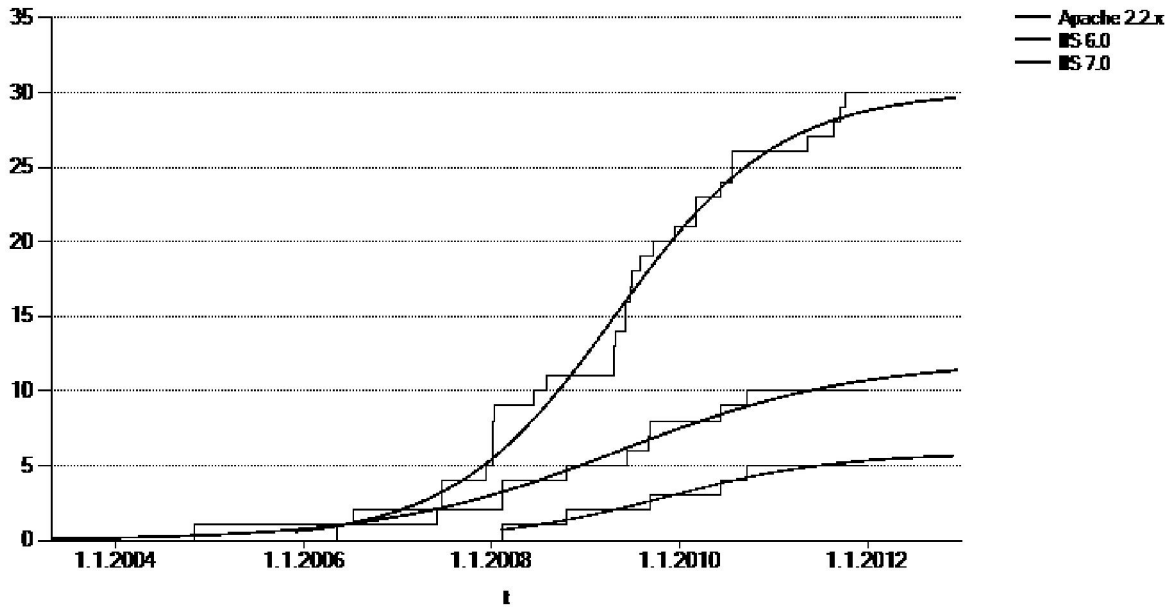

Рис. 4. Аппроксимация и экстраполяция кумулятивного профиля уязвимостей Apache и IIS веб-серверов с помощью логистической функции

Список литературы

1. *Харченко В.С. COTS- и CrOTS-подходы к повышению эффективности критических и коммерческих IT-проектов / В.С.Харченко, К.В. Харченко // Системи обробки інформації: Зб. наук. праць. Вип. 2(18). – Х.: ХВУ, 2002. – С. 214 – 217.*

2. *Афанасьев Н.В. Обеспечение и оценка безопасности информационно-вычислительной системы энергоблока АЭС с реактором ВВЭР-1000 / Н.В. Афанасьев, О.М. Белохин и др. // Ядерная и радиационная безопасность. – № 4. – 2002. – С. 87 – 104.*

3. *Лупян Е.А. Технология построения автоматизированных информационных систем сбора, обработки, хранения и распространения спутниковых данных для решения научных и прикладных задач / Е.А. Лупян, А.А. Мазуров, Р.Р. Назиров и др. // Тр. Всероссийской конференции "Современные проблемы дистанционного зондирования Земли из космоса". – М., 2003. – С. 62 – 79.*

4. *National Vulnerability Database [Электронный ресурс]. – Режим доступа: http://nvd.nist.gov.*

5. *The Open Source Vulnerability Database [Электронный ресурс]. – Режим доступа: http://osvdb.org.*

6. *Secunia – The Leading provider of vulnerability management and vulnerability intelligence solutions [Электронный ресурс]. – Режим доступа: http://secunia.com.*

7. *Lobachova K.I. Assessing Software Vulnerabilities and Recovery Time: Elements Of Technique And Results / K.I. Lobachova, V.S. Kharchenko // Radioelectronic and Computer Systems. – 2007. – № 8 (27). – P. 61 – 65.*

8. *Лобачева Е.И. Исследование показателей безотказности веб серверов Apache и IIS / Е.И. Лобачева, В.С. Харченко // Информационные технологии и информационная безопасность в науке, технике и образовании «ИНФОТЕХ - 2007». Международная НПК. Тез. докл. – Севастополь, 2007. – С. 45.*

Поступила в редколлегию 16.02.2012

Рецензент: д-р техн. наук, проф. В.С. Харченко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.

ДОСЛІДЖЕННЯ ДИНАМІКИ ВИЯВЛЕННЯ УРАЗЛИВОСТЕЙ У ПРОГРАМНИХ КОМПОНЕНТАХ З ВИКОРИСТАННЯМ ЛОГІСТИЧНОЇ КРИВОЇ

К.І. Неткачова

Досліджуються та порівнюються характеристики безпеки різних програмних продуктів на підставі інформації про уразливості, наданої у відкритих джерелах. За допомогою логістичної кривої проводиться апроксимація кумулятивного профілю відмов, визначаються межі етапів, що дозволяє ввести додаткову метрику, зробити виводи про стадію, на якій знаходиться продукт, і прогнозувати очікувану інтенсивність виявлення уразливостей у майбутньому.

Ключові слова: *уразливості, коефіцієнт уразливостей, логістична функція, OTS компоненти.*

ANALYZING THE DYNAMICS OF SOFTWARE VULNERABILITY DETECTION USING A LOGISTIC CURVE

K.I. Netkachova

Security characteristics of different software products are analyzed and compared based on the data collected from public vulnerability databases. An approximation of the cumulative failure distribution by a logistic function is presented, the boundaries of different stages are outlined, which makes it possible to introduce a new metric, determine the current stage of the product and predict the vulnerability detection rate expected in the future.

Keywords*: vulnerabilities, coefficient of vulnerabilities, logistic function, OTS components.*