УДК 004.7

С.Г. Семенов

Национальный технический университет «Харьковский политехнический институт», Харьков

ИССЛЕДОВАНИЕ МЕТОДОВ ИДЕНТИФИКАЦИИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИХ ХАРАКТЕРИСТИК

Проведен анализ основных уязвимостей программного обеспечения (ПО). Определена актуальность обеспечения защиты от неавторизованного изменения данных. Проведены исследования методов идентификации ПО, как одного из путей решения поставленной задачи. Определено, что анализируемые данные, должны обладать специфическими параметрами, позволяющими указать меру распознавания программы. Доказана целесообразность и возможность использования R/S анализа для установления целостности и подлинности ПО, при этом в качестве исследуемой характеристики использовалась частота появления операторов.

***Ключевые слова**: идентификация, защита, программное обеспечение, R/S анализ.*

Введение

Постановка проблемы. Эффективное использование информационных технологий в различных сферах жизнедеятельности современного общества возможно при условии решения задач, связанных с предотвращением кибератак непосредственно на составляющие информационных ресурсов. При этом, рассматривая программное обеспечение (ПО) как активный эксплуатируемый информационный ресурс, можно говорить о том, что безопасность информации в целом во многом определяется безопасностью ПО. Следует заметить, что данная проблематика актуальна на всем протяжении жизненного цикла ПО (формирование требований, проектирование, реализация, тестирование, внедрение, эксплуатация и сопровождение), и требует внесения защитных функций с учетом особенностей каждого из перечисленных этапов. Несложно отметить, что одним из наиболее продолжительных этапов жизненного цикла ПО является эксплуатация и сопровождение.

В этой связи актуальным представляется исследование различных аспектов защиты ПО, присущих именно этому этапу. Анализ кибератак на ПО показал, что одним из наиболее распространенных видов ($\approx 10\%$) является неавторизованное изменение данных (рис. 1).

Данный вид атаки возможен в случаях, когда исходные тексты ПО попадают в руки злоумышленников. При этом безопасность ПО может быть обеспечена с использованием методов идентификации ПО и его характеристик.

Анализ литературы [1, 2] показал, что на практике для идентификации ПО используются следующие методы:

- анализ стиля программирования;
- анализ идентификационных меток;
- анализ программных процедур;
- анализ характеристик ПО.

1. Анализ методов идентификации ПО

Исследования методов анализа стиля программирования показали достаточно большое многообразие различных биометрических, психологических и функциональных особенностей и параметров, присущих программистам, объединив которые, можно составить профиль анализируемого ПО и подтвердить первоисточник разработки. В качестве таких индивидуальных параметров могут выступать характерные ошибки программиста, избыточные тексты, заготовки, наработанные в процессе кодирования, избыточные команды, дополнительные подпрограммы и др.

Однако, следует заметить, что некоторые приведенные индивидуальные особенности могут быть присущи разным людям (особенно одной профессии). В этой связи точность идентификации ПО на основе только стиля программирования низка.

В последнее время в литературе [3, 6] все больше внимания уделяется методам анализа на основе идентификационных меток, а также цифровых водяных знаков. Эффективность использования меток для целей идентификации обусловлена скрытым характером их внедрения в ПО. Программист, использующий идентификационные метки, обладает широким выбором кодов, а также способов их маскировки. Кроме этого, использование криптографических средств защиты идентификационных меток может повысить уровень их скрытности.

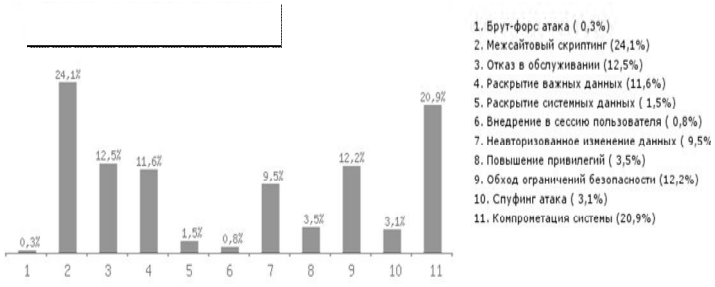


Рис. 1. Уязвимости программного обеспечения

Еще один из подходов формирования идентификационных меток – стеганографический (с помощью цифровых водяних знаков). Однако, в связи с тем, что основные стеганографические методы защиты данных используют природную избыточность контейнеров (а ПО такой избыточностью практически не обладает), внедрение цифровых водяных знаков в большей степени имеет целью защиту среды носителя программных средств.

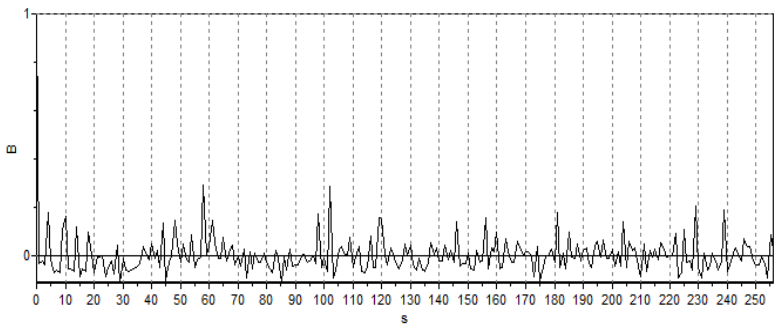


Рис. 2. Автокорреляционная функция характеристики частоты появления операторов в программе для редакторов текста

Анализ приведенных методов формирования идентификационных меток показал, что их реализация наряду с преимуществами имеет и недостатки: слабая производительность, низкая степень скрытия и др. Поэтому использование только идентификационных меток для защиты ПО представляется недостаточным.

Для выявления несанкционированного копирования ряд разработчиков [1, 2, 7] предлагает использовать метод программных процедур. Однако данный метод в настоящее время не получил широкого применения в связи с тем, что определение характерной последовательности признаков и идентификация программ на базе этих признаков затруднительна.

При установлении целостности и подлинности ПО также используется метод анализа характеристик ПО. При этом в качестве показателя целесообразнее всего выбрать тот, который в большей степени, не зависит от маскировок. К таким параметрам может относиться, например, частота использования операторов по тексту программы [1, 2], которую сложно изменить нарушителю, не искажая назначения программы, а подлинность структур может быть выражено как максимум взаимной корреляции функций.

Проведенные исследования показали, что данный критерий не всегда эффективен. Например, возникают проблемы при встраивании отдельного модуля в состав программы, а также анализе исходной программы на языке высокого уровня.

Другая характеристика программы – автокорреляционная функция, определяющая меру соответствия, с которой одни и те же последовательности операторов повторяются в самой программе. Однако точность определения степени подобия с помощью автокорреляционной функции низка, что наглядно иллюстрирует график рис. 2. Поэтому для установления целостности и подлинности ПО в статье предлагается использовать R/S анализ.

2. R/S анализ ПО

Алгоритм R/S анализа подробно описан в работах [4, 5], при этом в качестве основного параметра, характеризующего степень подобия систем чаще всего используется показатель Херста. В то же время исследования процесса идентификации с помощью данного показателя выявили факт невысокой точности результатов исследования. В этой связи для идентификации ПО представляется целесообразным использовать следующие ожидаемые показатели R/S (рис. 3):

$$E(R/S(n)) = \frac{n-0,5}{n} \cdot (n \cdot \frac{\pi}{2})^{-0,5} \cdot \sum_{r=1}^{n-r} \sqrt{\frac{n-r}{n}}, \quad (1)$$

где n – число наблюдений.

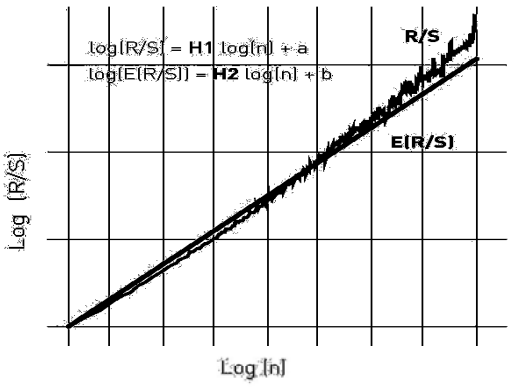


Рис. 3. Пример иллюстрации показателей R/S анализа

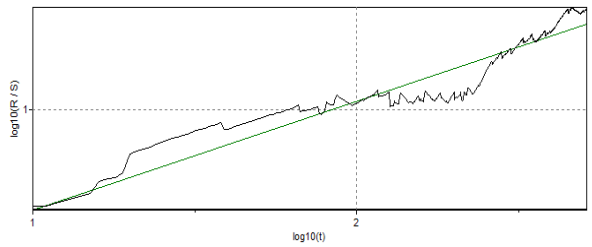
Исходя из выражения (1), находим ожидаемый показатель Херста $E(H)$, и дисперсии $D(H) = 1/T$, где T – количество наблюдений в выборке.

Проведем исследования и определим показатели R/S анализа для частоты появления операторов в программе для редакторов текста. При этом смоделируем ситуацию поглощения основной процедуры, состоящей из 90 операторов другой (нелегитимной) программой. На рис. 4 представлены результаты исследований и расчетов показателя Херста, а также графики зависимости $\log(R/S)_t$ от $\log(t)$ числовой последовательности частоты появления операторов в программе для редакторов текста.

Анализ результатов моделирования и расчетов показал явление персистентности Херста (H>0,5) для представленного примера числового ряда.

Следует заметить, что R/S анализ характеристики частоты появления операторов программы-эталона показал близкие значения рассматриваемого показателя. Данный фактор может быть использован для подтверждения внесения изменений в авторское

ПО либо встраивания его в другой программный продукт.

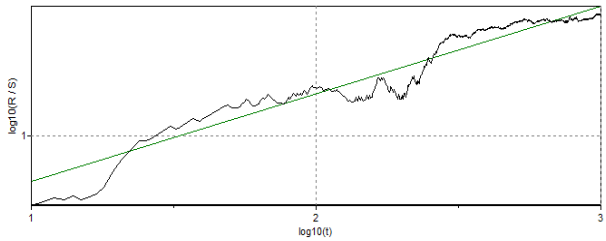


Показатель Херста $H = 0{,}6343 \pm 0{,}2081$,
фрактальная размерность $D = 2\text{-}H = 1{,}3657 \pm 0{,}2081$

Рис. 4. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда частоты появления операторов в программе для редакторов текста

Для подтверждения позитивных тенденций результатов исследования указанного выше фактора для программ-эталонов разного предназначения проведем R/S анализ ПО информационно-поисковой системы.

На рис. 5 представлены результаты исследований Херста, а также графики зависимости $\log(R/S)_t$ от $\log(t)$ числовой последовательности частоты появления операторов в информационно-поисковой системе.



Показатель Херста $H = 0{,}6311 \pm 0{,}1769$,
фрактальная размерность $D = 1{,}3689 \pm 0{,}1769$

Рис. 5. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда частоты появления операторов в информационно-поисковой системе

Выводы

Проведенные исследования показали возможность использования R/S анализа для идентификации ПО, при этом в качестве исследуемой характеристики использовалась частота появления операторов.

Следует заметить, что ПО имеет целую иерархию структур, которые могут быть выявлены, измерены и использованы в качестве характеристик последовательности данных. При этом в ходе тестирования, измерения не должны зависеть от типа данных, хотя данные, имеющие структуру программы, должны обладать специфическими параметрами, позволяющими указать меру распознавания программы.

Список литературы

1. Защита программного обеспечения / Под ред. Д. Гроувера. – М.: Мир, 1992. – 288 с.

2. Казарин О.В. Безопасность программного обеспечения компьютерных систем: монография / О.В. Казарин. – М.: МГУЛ, 2003. – 212 с.

3. Кузнецов О.О. Методи обробки сигналів даних та зображень / О.О. Кузнецов, Г.А. Кучук, С.Г. Семенов. – Х.: НТУ «ХПІ», 2011. – 292 с.

4. Петерс Э. Фрактальный анализ финансовых рынков. Применение теории хаоса в инвестициях и экономике / Э. Петерс. – М.: «Интернет-трейдинг», 2004. – 304 с.

5. Семенов С.Г. Оценка статистических свойств информационного трафика на основе метода нормированного размаха / С.Г. Семенов, Р.В. Королев, О.В. Петров // Системи обробки інформації. – Х.: ХУ ПС, 2011. – Вип. 8(98). – С. 237-240.

6. Семенов С.Г. Анализ и синтез защищенных компьютерных систем и сетей / С.Г. Семенов, А.А. Подорожняк, А.И. Баленко. – Х.: НТУ«ХПІ», 2012. – 204 с.

7. Positive Technologies: аналитика уязвимости веб-приложений. [Электронный ресурс]. – Режим доступа к ресурсу: http://www.comss.info/page.php?al=analitika_uyazvimosti_veb_prilozhenij.

Поступила в редколлегию 28.10.2015

Рецензент: д-р техн. наук, проф. А.А. Можаев, Национальный технический университет «ХПИ», Харьков.

ДОСЛІДЖЕННЯ МЕТОДІВ ІДЕНТИФІКАЦІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ І ЇХ ХАРАКТЕРИСТИК

С.Г. Семенов

Проведений аналіз основних уязвимостей програмного забезпечення (ПЗ). Визначена актуальність забезпечення захисту від неавторизованої зміни даних. Проведені дослідження методів ідентифікації ПЗ, як одного з шляхів рішення поставленої задачі. Визначено, що аналізовані дані повинні володіти специфічними параметрами, що дозволяють вказати міру розпізнавання програми. Доведена доцільність і можливість використання R/S аналізу для встановлення цілісності і достовірності ПЗ, при цьому як досліджувана характеристика використовувалася частота появи операторів.

***Ключові слова**: ідентифікація, захист, програмне забезпечення, R/S аналіз.*

RESEARCH OF SOFTWARE AUTHENTICATION METHODS AND THEIR DESCRIPTIONS

S.G. Semenov

The analysis of basic vulnerabilities of software is conducted. Actuality of providing of protecting from the unauthorized change of information is certain. Researches of methods of authentication of software are conducted, as one of ways of decision of the put task. It is certain that analysable information must possess specific parameters, allowing to specify the measure of recognition of the program. Expedience and possibility of the use of R/S of analysis for establishment of integrity and authenticity of software is well-proven, here as the probed description frequency of appearance of operators was utillized.

***Keywords**: authentication, defence, software, R/S analysis.*