МЕТОДЫ КОНТРОЛЯ И ИДЕНТИФИКАЦИИ СОСТОЯНИЯ КОМПЬЮТЕРНЫХ СИСТЕМ НА ОСНОВЕ BDS-ТЕСТИРОВАНИЯ

Сергей Семенов [1], Светлана Гавриленко [2]

[1] НТУ «ХПИ», Украина, г. Харьков, ул. Фрунзе, 21, *e-mail: s_semenov@ukr.net*

[2] НТУ «ХПИ», Украина, г. Харьков, ул. Фрунзе, 21, *e-mail: gavrilenko@mail.ua*

Резюме: В работе предлагается использовать математический аппарат статистического анализа на основе BDS-тестов для идентификации состояния компьютерной системы в условиях воздействий компьютерных вирусов.

Ключевые слова: BDS-статистика, безопасность компьютерных систем, компьютерный вирус, показатель загрузки центрального процессора, математическая модель.

Введение.

Экономическое и социальное развитие современного государства зависит от выполнения целенаправленной политики широкой информатизации общества. Интенсифи-кация инфокоммуникационных отношений является одним из условий приумножения экономического и духовного потенциала страны, совершенствования самосознания ее граждан.

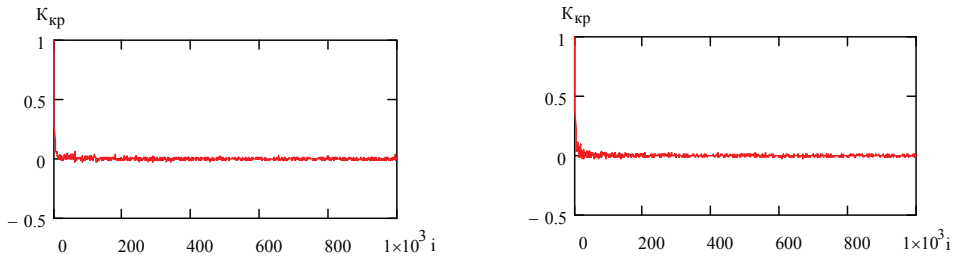
Главными особенностями в решении комплекса поставленных задач является наличие ряда внутренних и внешних факторов, существенно повышающих сложность получения результата. Одним из таких факторов является участившиеся случаи компьютерных атак с помощью злоумышленного программного обеспечения (ЗПО).

Анализ литературы [1-4] показал, что для защиты компьютерных систем от ЗПО в настоящее время существует множество подходов, основанных в основном на сигнатурном и эвристическом анализе и идентификации технических и программных структур. Наиболее сложным и противоречивым при этом остается эвристический подход.

Как показали исследования [1, 2] в основу множества эвристических анализаторов в настоящее время положен принцип вычисления корреляционной размерности. Пример идентификации различных интерактивных сетевых приложений компьютерной системы представлен нам рис. 1.

Вид приведенных графиков в большей степени показывает на отсутствие, каких либо статистических зависимостей в поведении компьютерной системы, чем на их наличие. В связи с этим низкая точность результатов анализа с помощью стандартных методов вычисления корреляционной размерности, снижает эффективность антивирусного анализа и идентификации состояния компьютерных систем.

Проведенные исследования показали, что в настоящее время существует несколько графических методов обработки информации и идентификации состояния, среди которых можно выделить методы, основанные на тестах хаоса (тест Гилмора и др.). Однако, ограничение исследований сложных компьютерных систем только рамками априорных данных и возможностями графических фазовых портретов в условиях компьютерных вирусных атак, может привести к различным ошибкам или неточностям.



а) *б)*

Рис. 1. График автокорреляционной функции FTP и HTTP-трафика

Следовательно, разработка и исследование методов антивирусного контроля и идентификации состояния компьютерных систем, является актуальной научно-прикладной задачей. Ее решение непосредственно связано с обеспечением безопасности современных компьютерных систем и применяемых компьютерных технологий.

Целью работы является разработка и исследование метода контроля и идентификации состояния компьютерных систем на основе BDS-тестирования.

Основная часть. Проведенный анализ, показал, что BDS-тесты, предложеные в результате анализа финансовых рынков экономистами Броком, Дечертом и Шейнкманом (*B. Brock, W. Dechert* и *J. Scheinkman*) в 1987 [5], представляют собой эффективные методы выявления зависимостей во временных рядах в рамках их нелинейного анализа. Их цель состоит в том, чтобы различить данные *I.I.D.* и любой вид зависимости – проверить нулевую гипотезу H_0 о независимости и тождественном распределении значений временного ряда $\vec{\xi} = (\xi_1, \xi_2, ..., \xi_N)$, используя для этого критерий значимости. Согласно этому критерию для принятия гипотезы H_0 необходимо выбрать критическую область G_α, удовлетворяющую условию $P(g \in G) = \alpha$, где $g(\xi_1, \xi_2, ..., \xi_N)$ – статистика наблюдения, а α – устанавливаемый уровень значимости.

Из [2-5] известно, что *BDS*-тест основан на статистической величине $w(\vec{\xi})$ (*BDS*-статистике):

$$w_{m,N}(\varepsilon) = \sqrt{N-m+1}\,\frac{C_{m,N}(\varepsilon) - C_{1,N-m}(\varepsilon)^m}{\sigma_{m,N}(\varepsilon)}, \quad (1)$$

где $C_{m,N}(\varepsilon) - C_{1,N-m}(\varepsilon)^m$ – (числитель *BDS*-статистики) определяется корреляционными интегралами $C_{m,N}(\varepsilon)$, $C_{1,N}(\varepsilon)$ для размерности m,

ε – радиус гиперсферы;

$\sigma_{m,N}(\varepsilon)$ – среднеквадратическое отклонение разницы $C_{m,N}(\varepsilon) - C_{1,N-m}(\varepsilon)^m$;

N – число элементов временного ряда.

В ряде работ [2, 3] были предложены «упрощенные» алгоритмы оценки *BDS*-статистики. В них для вычисления $C_{m,N}(\varepsilon)$ $(m > 1)$ необходимо выполнить «вложение» временного ряда в m-мерное псевдофазовое пространство, элементами которого, на основании теоремы Такенса [6], являются точки $\xi_i^m = (\xi_i, \xi_{i+1}, ..., \xi_{i+m})$ с координатами $\{\xi_{i+k}\}_{k=1}^m$ заданными m последовательными значениями исходного временного ряда. Корреляционный интеграл определяет частоту попадания произвольной пары точек фазового пространства в гиперсферы радиуса ε :

$$C_{m,N}(\varepsilon) = \frac{2}{(N-m+1)(N-m)} \sum_{s=m}^{N} \sum_{t=s+1}^{N} \prod_{j=0}^{m-1} I_\varepsilon(\xi_{s-j}^m, \xi_{t-j}^m), \quad (2)$$

$$I_\varepsilon(\xi_i^m, \xi_j^m) = \begin{cases} 1, & \left\| \xi_i^m - \xi_j^m \right\| \le \varepsilon \\ 0, & \left\| \xi_i^m - \xi_j^m \right\| > \varepsilon \end{cases}, \quad (3)$$

$\{\xi_i\}_{i=1}^N \quad 0 \le i \le N$ и $0 \le j \le N$

где $I_\varepsilon(\xi_i^m, \xi_j^m)$ – функция Хевисайда для всех пар значений i и j.

Значение корреляционного интеграла стремиться к определенному пределу по мере уменьшения ε. Анализ работ известных авторов [3, 4] показал, что существует диапазон значений ε, который позволяет провести вычисления с заданным коэффициентом точности. Этот диапазон зависит от числа элементов временного ряда N. Если ε является слишком маленьким, не будет достаточного количества точек для захвата статистической структуры; если ε является слишком большим, точек будет слишком много.

В работах [2-4] ε рекомендовано выбирать таким, что $\varepsilon = 0.5\sigma \div 2\sigma$, где σ – среднеквадратическое отклонение процесса $\{\xi_i\}_{i=1}^N$. В соответствии с теорией статистики, зависимость корреляционного интеграла от ε имеет вид:

$$C_{m,N}(\varepsilon) \sim \varepsilon^{D_c},$$

где D_c – корреляционная размерность временного ряда.

Для $m = 1$ имеем:

$$C_{1,N}(\varepsilon) = \frac{2}{N(N-1)} \sum_{s=1}^{N} \sum_{t=s+1}^{N} I_\varepsilon(\xi_s, \xi_t).$$

Поведенные исследования показали, что при $N \to \infty$, корреляционный интеграл $C_{m,N}(\varepsilon) \Rightarrow C_{1,N}(\varepsilon)^m$, а величина $\left(C_{m,N}(\varepsilon) - \left(C_{1,N}(\varepsilon)\right)^m\right) \cdot \sqrt{N-m+1}$ является случайной асимптотически нормально распределенной величиной с нулевым средним и среднеквадратическим отклонением $\sigma_{m,N}(\varepsilon)$, которое определяется как:

$$\sigma_{m,N}(\varepsilon) = 2\sqrt{\begin{array}{l} k^m + 2\sum_{j=1}^{m-1} k^{m-j} \cdot (C_{1,N}(\varepsilon)^{2j} + (m-1)^2 \times \\ \times (C_{1,N}(\varepsilon)^{2m} - m^2 k (C_{1,N}(\varepsilon)^{2m-2} \end{array}} \qquad (4)$$

где

$$k = \frac{1}{(N-1)(N-2)N}\left\{\sum_{t=1}^{N}\left[\sum_{s=1}^{N} I_\varepsilon(\xi_t, \xi_s)\right]^2 - 3\sum_{s=1}^{N}\sum_{t=s+1}^{N} I_\varepsilon(\xi_t, \xi_s) + 2N\right\}$$

BDS-статистика $w(\vec{\xi})$ является нормально распределенной случайной величиной при условии, что оценка $\hat{\sigma}_{m,N}(\varepsilon)$ близка к ее теоретическому значению $\sigma_{m,N}(\varepsilon)$.

Задача обнаружения хаотического сигнала рассматривается как непараметрическая проверка одной из двух гипотез:

1) H_0 – наблюдаемые данные (информационный трафик) $\vec{\xi} = (\xi_1, \xi_2, ..., \xi_N)$ независимы и одинаково распределены, т.е. плотность (функция) распределения факторизуется $F_N(\xi_1, \xi_2, ..., \xi_N) = \prod_{i=1}^{N} F(\xi_i)$.

2) H_1 – полученные в результате эксперимента данные (информационный трафик) имеют определенную зависимость (процесс структурирован).

Согласно гипотезе H_0 статистика $w(\vec{\xi})$ асимптотически распределена как $N(0,1)$, если число наблюдений асимптотически стремится к бесконечности. В ряде работ [2-4] обосновывается гипотеза о необходимости проведения экспериментального исследования объемом более 500 наблюдений. Такое количество экспериментов позволит утверждать о достоверности полученных результатов.

Исследования показали, что критерием достоверности гипотезы H_0 (об отсутствии в информационном трафике каких либо зависимостей) является неравенство:

$$\left|w_{m,N}(\varepsilon)\right| \le 1{,}96 \cdot \qquad (5)$$

В ходе исследования была выдвинута гипотеза об уменьшении значения *BDS*-теста в случаях злоумышленных внешних воздействий на систему. На примере случая воздействия на компьютерную систему *Dos*-атаки проведена оценка значений *BDS*-теста, результаты которой показали снижение этого показателя до 40 раз, что наглядно показывают данные в табл. 1.

Для выявления возможности и особенностей идентификации состояния компьютерной системы в условиях воздействия злоумышленного программного обеспечения с помощью BDS-теста была разработана имитационная модель, при этом ее входными данными стали показатели загрузки центрального процессора.

Модель предусматривает вариацию количества значений временного ряда N. Значение

Таблица 1 – Значение BDS-статистики для различного вида трафиков при N=500

	m=6		m=5		m=4	
	ε=0.5	ε=0.25	ε=0.5	ε=0.25	ε=0.5	ε=0.25
IP-телефония	17.512	26.893	16.431	22.455	15.709	18.691
Торрент услуги	45.876	67.028	40.727	49.076	35.145	41.392
Потоковое видео	52.329	117.954	38.824	83.730	30.032	54.371
DoS-атака	1.391	4.298	1.692	4.939	1.847	5.231

загрузки центрального процессора сканируется посекундно и сохраняются в файле.

Полученные значения загрузки центрального процессора делаться на выборки по 500 значений и подаются на вход модуля анализа, который подвергается дальнейшей обработке и анализу с помощью BDS-статистики.

Результатом работы программной модели является N/500 значений BDS-теста.

На рис. 2 представлены графики значений BDS-теста при обычной загрузке процессора в режиме «Пользователь».

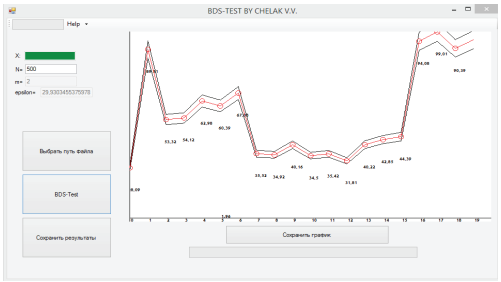


Рис.2 – Графики значений BDS-теста при обычной загрузке процессора в режиме «Пользователь»

Как видно из графика минимальное значение BDS-теста близко 28, а максимальное значение возрастает до 99. Таким образом, максимальное значение BDS-теста превышает минимальное более чем в три раза, джиттер (разброс) значений достигает 70%.

На рис. 3 представлены графики значений BDS-теста в условиях заражения компьютера вирусом Svchost.exe (svchost.exe - это безопасный системный процесс Microsoft Windows, который называется "Generic Host Process".

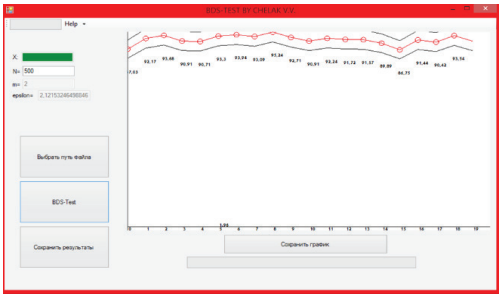


Рис. 3 – Графики значений BDS-теста в условиях заражения компьютера вирусом Svchost.exe

Как видно из графика минимальное значение BDS-теста близко 77, а максимальное значение возрастает до 95. Таким образом, джиттер значений сокращается до 19%. Данный факт может стать сигналом о возможном заражении компьютерной системы злоумышленным программным обеспечением.

Аналогичные результаты получены во время эксперимента в условиях заражения компьютера вирусом *KillProc* (рис. 4) и *Kb657048.crs* (рис. 5).

В первом случае джиттер между максимальным и минимальным значением также не превышает 20.%, во втором - 19%.

Проведенные исследования показали, что ряд вирусов приводят к существенному увеличению загрузки процессора. Для подтверждения этой гипотезы было создано злоумышленное программное обеспечение, имитирующее работу форк-бомбы для ОС Windows.

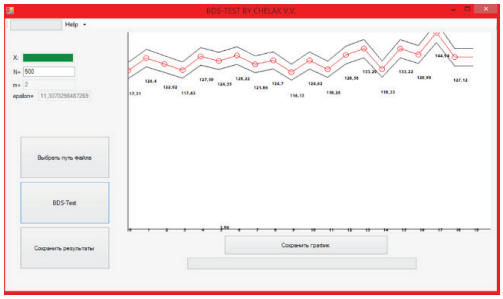


Рис.4 – Значение BDS-теста после заражения компьютера вирусом KillProc

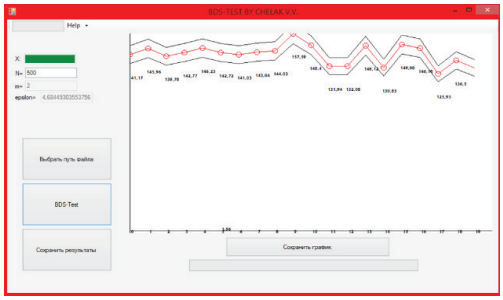


Рис.5 – Значение BDS-теста после заражения компьютера вирусом Kb657048.crs

Полученные временные характеристики процессора были обработаны с помощью BDS-теста. Результаты обработки представлены на рис. 6.

Как видно из результатов эксперимента в случае заражения компьютерной системы форк-бомбой значения BDS-теста в большинстве тестовых случаев стремится к бесконечности, что может также быть сигналом заражения компьютерной системы злоумышленным программным

обеспечением.

N	Значения BDS-теста при зараженими fork-бомбой
0	23,38
1	22,45
2	Стремится к бесконечности
3	Стремится к бесконечности
4	Стремится к бесконечности
5	Стремится к бесконечности
6	22,45
7	Стремится к бесконечности
8	Стремится к бесконечности
9	Стремится к бесконечности
10	Стремится к бесконечности
11	Стремится к бесконечности
12	Стремится к бесконечности
13	Стремится к бесконечности
14	Стремится к бесконечности
15	Стремится к бесконечности
16	Стремится к бесконечности
17	Стремится к бесконечности
18	22,55
19	22,45

Рис. 6 – Результаты анализа при заражении линейной программой

Выводы

В статье разработан метод контроля и идентификации состояния компьютерных систем на основе BDS-тестирования.

Результаты исследования показали:

1. В условиях воздействия на компьютерную систему злоумышленного программного обеспечения изменяются статистические показатели основных параметров функционирования системы (загрузка центрального процессора, оперативной памяти и т.д.).

2. Воздействия ряда вирусов на компьютерную систему приводит к резкому снижению джиттера значений BDS-теста до 19-20%, что может послужить сигналом аномального поведения системы.

3. Ряд видов злоумышленного программного обеспечения (например форк-бомбы), имеющего своей целью вывод компьютерной системы из строя могут быть обнаружены по факту резкого возрастания значений BDS-теста стремящихся к бесконечности.

4. Проведенные экспериментальные исследования подтверждают возможность использования аппарата BDS-тестирования как дополнительного средства для выявления вирусных атак, об общей системе обнаружения вредоносного программного обеспечения.

Литература

[1] **К. Касперский.** Записки исследователя компьютерных вирусов. / К. Касперский. – СПб.: Питер, 2006. – 316 с.

[2] **С. Г. Семенов.** Защита данных в компьютеризированных управляющих системах (монография) / С.Г. Семенов, В.В. Давыдов, С.Ю. Гавриленко // Изд. «LAP LAMBERT ACADEMIC PUBLISHING» Германия, 2014.– 236 с.

[3] **С. М. Порошин.** Разработка и исследования математической модели компьютеризированной информационно-измерительной управляющей системы критического применения с учетом фактора внешних воздействий / С.М. Порошин, С.Г. Семенов // Системи обробки інформації. – Х.: ХУ ПС. – 2013. – Вип. 2(110). – С. 208-210.

[4] **С. Г. Семенов.** Динамическая модель информационной системы на основе наблюдаемого структурно-информационного портрета / С.Г. Семенов, В.В. Давыдов // Вісник Національного технічного університету «Харківський політехнічний інститут».– Х.:НТУ «ХПІ». – 2011. –№36. – С. 156-163.

[6] **W. A. Brock.** Test for independence based on the correlation dimension / W. Brock. W. Dechert and J. Scheinkman. // Working Paper, University of Wisconsin, 1987.

[7] **F. Takens.** Detecting strange attractors in turbulence / F. Takens // Lecture Notes in Math. – 1981.– v.898. – P.366–381.

Данные по авторам:

Семенов Сергей Геннадьевич, кандидат технических наук (2007), доктор технических наук (2013), старший научный сотрудник (2012), заведующий кафедрой «Вычислительная техника и программирование» Национального технического университета «ХПИ». Научные интересы: защита информации.

Гавриленко Светлана Юрьевна, кандидат технических наук (2001), доцент (2005), профессор кафедры «Вычислительная техника и программирование» Национального технического университета «ХПИ». Научные интересы: защита информации, цифровые автоматы.

METHODS OF CONTROL AND IDENTIFICATION OF COMPUTER SYSTEMS STATE BASED ON BDS-TESTING

Sergey Semenov [1], Svetlana Gavrilenko [2]

[1] NTU "KhPI", Frunze street, 21, Kharkov, 61002, Ukraine, *e-mail: s_semenov@ukr.net.*

[2] NTU "KhPI", Frunze street, 21, Kharkov, 61002, Ukraine, *e-mail: gavrilenko@mail.ua.*

Abstract: In this paper we propose to use the mathematical apparatus of statistical analysis based on the BDS-test to identify the state of the computer system under the effect of computer viruses.

Key-Words: BDS-statistics, computer security, computer virus, the CPU load indicator, mathematical model.

References:

[1] **K. Kasperskii.** Zapiski issledovatelya kompyuternih virusov. / K. Kasperskii. – SPb._ Piter_ 2006. – 316 s.

[2] **S.G Semenov.** Zaschita dannih v kompyuterizirovannih upravlyayuschih sistemah _monografiya, / S.G. Semenov_ V.V. Davidov_ S.Yu. Gavrilenko // Izd. «LAP LAMBERT ACADEMIC PUBLISHING» Germaniya_ 2014.

[3] **S.M Poroshin.** Razrabotka i issledovaniya matematicheskoi modeli kompyuterizirovannoi informacionno_izmeritelnoi upravlyayuschei sistemi kriticheskogo primeneniya s uchetom faktora vneshnih vozdeistvii / S.M. Poroshin_ S.G. Semenov // Sistemi obrobki informacií. – H._ HU PS. – 2013. – Vip. 2_110,. – S. 208.

[4] **S.G. Semenov.** Dinamicheskaya model informacionnoi sistemi na osnove nablyudaemogo strukturno_informacionnogo portreta / S.G. Semenov_ V.V. Davidov // Visnik Nacionalnogo tehnichnogo universitetu «Harkivskii politehnichnii institut».– H._NTU «HPI». – 2011. –№36. – S. 156_163.

[6] **W. A. Brock.** Test for independence based on the correlation dimension / W. Brock. W. Dechert and J. Scheinkman. // Working Paper, University of Wisconsin, 1987.

[7] **F. Takens.** Detecting strange attractors in turbulence / F. Takens // Lecture Notes in Math. – 1981.– v.898. – P.366–381.

МЕТОДИ ЗА КОНТРОЛ И ИДЕНТИФИКАЦИЯ НА СЪСТОЯНИЕТО НА КОМПЮТЪРНИТЕ СИСТЕМИ НА ОСНОВАТА НА BDS-ТЕСТОВЕ

Сергей Семенов [1], Светлана Гавриленко [2]

[1] НТУ «ХПИ», Украйна, г. Харков, ул. Фрунзе, 21, *e-mail: s_semenov@ukr.net*

[2] НТУ «ХПИ», Украйна, г. Харков, ул. Фрунзе, 21, *e-mail: gavrilenko@mail.ua*

Резюме: В доклада се предлага да се използва математически апарат на статистически анализ, базиран на BDS-тестове за идентифициране на състоянието на компютърната система в условията на въздействията на компютърни вируси.

Ключови думи: BDS-статистика, безопасност на компютърни системи, компютърен вирус, показател на натоварване на централния процесор, математически модел