DOI: 10.51790/2712-9942-2022-3-4-05

МЕТОДОЛОГИЯ КОСВЕННОГО МОНИТОРИНГА НЕСАНКЦИОНИРОВАННОЙ АКТИВНОСТИ В ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

К. И. Бушмелева, А. В. Гавриленко, А. В. Никифоров[a]
Сургутский государственный университет, г. Сургут, Российская Федерация
[a] ✍ *nikiforov_av@surgu.ru*

Аннотация: в статье описаны недостатки существующих средств противодействия несанкционированной активности и возможность их устранения посредством методологии косвенного мониторинга несанкционированной активности в вычислительных системах. Разобраны проблемы применения математического ожидания, дисперсии и стандартного отклонения при расчете граничных значений косвенных параметров и предложено решение, которое основано на использовании выборочного среднего, выборочной дисперсии и выборочного стандартного отклонения для проведения данных расчетов. Помимо расчета граничных значений косвенных параметров, в статье описывается метод прогнозирования объема выборки, который основывается на частоте сбора косвенных параметров вычислительной системы и времени ее работы. Граничные значения косвенных параметров, предельные значения косвенных параметров, прогнозируемый объем выборки и другие расчетные показатели, предусмотренные методологией, образуют множество показателей, описывающих нормальную работу вычислительной системы. Наличие данного множества позволяет определить кусочную функцию проверки косвенных параметров вычислительной системы на нормальное состояние. Следовательно, возможно представить вычислительную систему в виде предиката. Данный предикат и множество показателей лежат в основе шаблона, описывающего выбранную вычислительную систему. Полученный шаблон и сценарии его применения являются базой для архитектуры системы косвенного мониторинга несанкционированной активности в вычислительных системах.

Ключевые слова: методология, мониторинг, несанкционированная активность, вычислительные системы, косвенные параметры.

Для цитирования: Бушмелева К. И., Гавриленко А. В., Никифоров А. В. Методология косвенного мониторинга несанкционированной активности в вычислительных системах. *Успехи кибернетики*. 2022;3(4):41–45. DOI: 10.51790/2712-9942-2022-3-4-05.

INDIRECT MONITORING OF SUSPICIOUS ACTIVITY ON COMPUTER SYSTEMS

K. I. Bushmeleva, A. V. Gavrilenko, A.V. Nikiforov[a]
Surgut State University, Surgut, Russian Federation
[a] ✍ *nikiforov_av@surgu.ru*

Abstract: this study conspires the drawbacks of the existing fraud detection tools and offers a solution: indirect monitoring of suspicious activity on computer systems. We applied the expected value, variance, and standard deviation concepts to estimate the thresholds of indirect indicators of compromise and derived a solution based on the selective mean, selective variance, and selective standard deviation. The paper also describes a sample size estimation procedure from the computer system's indirect indicator sampling rate and runtime. The thresholds of the indirect indicators, the estimated sample size, and other proposed indicators describe the normal operation of the computer system. With this set of indicators, we can define a piecewise function used to check the indirect indicators against the normal operation conditions. Consequently, the computer system can be represented as a predicate. The predicate and the set of indicators are a template describing the computer system. The resulting template and its application scenarios provide a foundation for developing the architecture of an indirect suspicious activity monitoring tool.

Keywords: methodology, monitoring, suspicious activity, computer systems, indirect indicators of compromise.

Cite this article: Bushmeleva K. I., Gavrilenko A. V., Nikiforov A.V. Indirect Monitoring of Suspicious Activity on Computer Systems. *Russian Journal of Cybernetics*. 2022;3(4):41–45. DOI: 10.51790/2712-9942-2022-3-4-05.

Введение

Одной из проблем современных вычислительных систем является несанкционированная активность. Причины появления несанкционированной активности в вычислительных системах имеют различную природу возникновения [1–6]. Определение наличия несанкционированной активности рядовыми пользователями в вычислительных системах без специальных средств невозможно, т. к. для этого необходимо обладать достаточно высокой квалификацией. Существующие специальные средства по большей части нацелены на противодействие вредоносному программному обеспечению. Уязвимостям и неисправностям аппаратного обеспечения вычислительных систем уделяется намного меньше внимания.

Методология косвенного мониторинга несанкционированной активности в вычислительных системах, описанная в данной статье, направлена на решение проблем, связанных с несанкционированной активностью вне зависимости от природы ее возникновения. Эта методология опирается на мониторинг не прямых, а косвенных параметров работы вычислительной системы [7]. Прямыми являются частота процессора, количество ядер процессора, объем кеша процессора и т. д. Косвенными параметрами вычислительной системы являются температура процессора, скорость вращения вентилятора системы охлаждения, уровень электромагнитного поля и т. д. Существующие специальные средства работают именно с прямыми параметрами вычислительной системы, но все чаще злоумышленники подделывают их с целью сокрытия несанкционированной активности. Использование описанной в статье методологии позволяет избежать проблем, связанных с достоверностью прямых параметров работы вычислительной системы.

Методология косвенного мониторинга несанкционированной активности в вычислительных системах

Суть описанной методологии заключается в наблюдении за параметрами, которые оказывают косвенное воздействие на вычислительные процессы, происходящие в вычислительных системах. Каждый элемент вычислительной системы имеет определенное количество косвенных параметров, которые можно использовать при оценивании состояния вычислительной системы.

Обозначим множество всех значений некоторого косвенного параметра любого элемента вычислительной системы буквой P [8].

$$P = \{x_1, \cdots, x_n\}, \qquad x_n \in R,\ 1 \le n \le m, \tag{1}$$

где m — максимально возможное количество значений косвенного параметра.

Из определения множества P, описанного формулой (1), следует, что оно является генеральной совокупностью некоторого косвенного параметра. Следовательно, можно рассчитать следующие базовые математические и статистические показатели, которые помогут оценить состояние вычислительной системы [9]:

1) минимальное значение;
2) максимальное значение;
3) математическое ожидание;
4) дисперсия;
5) стандартное отклонение.

$$\begin{cases} \varepsilon_{left} = \mu - c_\sigma * \sigma \\ \varepsilon_{right} = \mu + c_\sigma * \sigma \end{cases}, \quad c_\sigma \in R,\ \varepsilon_{left} \in R,\ \varepsilon_{right} \in R, \tag{2}$$

где ε_{left} — это левое граничное значение некоторого косвенного параметра;
ε_{right} — это правое граничное значение некоторого косвенного параметра;
μ — это математическое ожидание некоторого косвенного параметра;
c_σ — это математическое ожидание количества стандартных отклонений, при котором вычислительная система находится в нормальном состоянии;
σ — это стандартное отклонение некоторого косвенного параметра.

Имея значения данных показателей и воспользовавшись формулой (2), можно рассчитать на множестве P граничные значения, при которых вычислительная система функционирует в нормальном состоянии.

$$\begin{cases} \varepsilon_{left\ lim} = \min P \\ \varepsilon_{right\ lim} = \max P \end{cases}, \quad \varepsilon_{left\ lim} \in P,\ \varepsilon_{right\ lim} \in P, \tag{3}$$

где $\varepsilon_{left\ lim}$ — это предел левого граничного значения некоторого косвенного параметра;
$\varepsilon_{right\ lim}$ — это предел правого граничного значения некоторого косвенного параметра.

Также, воспользовавшись формулой (3), можно найти пределы левого и правого граничных значений.

$$\begin{cases} \varepsilon_{left\ norm} = \varepsilon_{left} * 10^n \\ \varepsilon_{right\ norm} = \varepsilon_{right} * 10^n \\ \varepsilon_{left\ lim\ norm} = \varepsilon_{left\ lim} * 10^n \\ \varepsilon_{right\ lim\ norm} = \varepsilon_{right\ lim} * 10^n \end{cases}, \quad n \in N,\ 1 \le n \le m, \tag{4}$$

где m — максимальное количество разрядов, которое переносится в целую часть числа.

Отдельного внимания требуют косвенные параметры, значения которых никогда не достигают целой части. Например, напряжение питания процессора или чипсета. Поэтому граничные значения таких косвенных параметров нормализуются при помощи формулы (4).

Процедура нормализации значений косвенных параметров требуется для увеличения производительности расчетов на вычислительных системах, т. к. вычисления с использованием целых чисел осуществляются быстрее. В остальных случаях процедура нормализации является опциональной.

Однако описанный выше метод, основанный на использовании граничных значений косвенного параметра, имеет две существенные проблемы:

1) по причине сложных взаимосвязей между косвенными и прямыми параметрами вычислительной системы невозможно получить все значения множества P и, как следствие, рассчитать базовые статистические показатели для реальной вычислительной системы;
2) отсутствие базовых статистических показателей делает невозможным расчет c_σ, т. к. базовые статистические показатели лежат в основе процедуры расчета.

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \mu \tag{5}$$

Решением данных проблем является переход от базовых статистических показателей к эквивалентным, а именно выборочному среднему, выборочной дисперсии и выборочному стандартному отклонению. Возможность эквивалентного перехода доступна благодаря пределу, описанному в формуле (5) [9].

$$\begin{cases} \varepsilon_{left} = \overline{x} - c_{sd} * sd \\ \varepsilon_{right} = \overline{x} + c_{sd} * sd \end{cases}, c_{sd} \in R \tag{6}$$

Функция, предел которой равен математическому ожиданию, является ни чем иным, как формулой расчета выборочного среднего. Следовательно, граничные значения рассчитываются по обновленной формуле (6).

$$c_{sd} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\overline{x} - x_i|}{sd}, \tag{7}$$

где n — количество элементов в выборке.

Значение переменной c_{sd}, которая является заменой c_σ, рассчитывается по формуле (7).

$$\begin{cases} \varepsilon_{left} = \overline{x} - \frac{1}{n}\sum_{i=1}^{n}\frac{|\overline{x}-x_i|}{sd} \\ \varepsilon_{right} = \overline{x} + \frac{1}{n}\sum_{i=1}^{n}\frac{|\overline{x}-x_i|}{sd} \end{cases} \tag{8}$$

Если объединить формулы (6) и (7), получится формула (8).

$$n = \sum_{i=1}^{m}\nu_i * T_i, \quad n \in N, \quad \nu_i \in N, \quad T_i \in N, \quad 1 \le n \le k, \tag{9}$$

где ν_i — частота сбора значений косвенного параметра вычислительной системы;
T_i — период работы вычислительной системы;
k — максимальное количество периодов, на которых проводится сбор значений косвенного параметра вычислительной системы.

Опираясь на формулы (5) и (8), можно сделать вывод, что расчетные граничные значения будут наиболее близкими к действительным, если максимизировать количество элементов в выборке. Поэтому, воспользовавшись формулой (9), можно рассчитать количество элементов в выборке за определенный период времени работы вычислительной системы.

Формула (9) позволяет не только рассчитать, но и спрогнозировать количество элементов в выборке. Пополнение количества элементов в выборке осуществляется при помощи датчиков, отслеживающих значения косвенных параметров. Датчики, применяемые при отслеживании значений косвенных параметров, бывают встроенные и внешние. Большинство современных элементов, из которых состоят вычислительные системы, имеет встроенные датчики (датчики измерения температуры и скорости вращения вентиляторов системы охлаждения). Использование внешних датчиков при пополнении выборки значениями является минимальным.

$$M = \left\{n,\overline{x},sd,c_{sd},\varepsilon_{left},\varepsilon_{right},\varepsilon_{left\ lim},\varepsilon_{right\ lim}\right\} \tag{10}$$

Все полученные ранее расчетные показатели образуют множество M, которое позволяет оценить некоторый косвенный параметр. Данное множество описывается формулой (10).

$$\varepsilon_{state} = vlim(x) = \begin{cases} 0,\ \varepsilon_{left} \le x \le \varepsilon_{right} \\ 1,\ \varepsilon_{left} < x \le \varepsilon_{left\ lim} \\ 1,\ \varepsilon_{right} < x \le \varepsilon_{right\ lim} \end{cases}, x \in P,\ \varepsilon_{state} \in B \tag{11}$$

В итоге множества P и M позволяют описать следующую кусочную функцию из формулы (11).

$$\varepsilon_{state}^{(1)} \wedge \varepsilon_{state}^{(2)} \wedge \cdots \wedge \varepsilon_{state}^{(n)},\quad 1 \le n \le m, \tag{12}$$

где m — максимальное количество косвенных параметров элементов, из которых состоит вычислительная система.

Эта кусочная функция показывает, находится некоторый косвенный параметр в нормальном состоянии или нет. Опираясь на вышесказанное, можно произвести описание вычислительной системы в виде предиката, который позволит определять состояние системы на основе состояния косвенных параметров ее элементов. Описанный выше предикат представлен в формуле (12).

Наличие предиката, описывающего вычислительную систему и множества M, делает возможным создание шаблона любой вычислительной системы. Этот шаблон позволяет осуществлять мониторинг вычислительной системы в процессе ее работы. Следовательно, становится возможным определить наличие несанкционированной активности в вычислительной системе, что позволит своевременно принимать меры по противодействию.

Рассмотрим применение методологии косвенного мониторинга несанкционированной активности в вычислительных системах для всех возможных сценариев:

1. Вычислительная система не имеет шаблона, и база данных шаблонов пустая. Тогда производится первичный сбор статистики для косвенных параметров вычислительной системы. Далее на основе полученной статистики создается шаблон, который записывается в базу данных. После чего запускается процесс мониторинга вычислительной системы.

2. Вычислительная система не имеет шаблона, и база данных шаблонов не пустая. Тогда производится поиск наиболее подходящего временного шаблона из базы данных. Далее запускается процесс мониторинга вычислительной системы. После чего параллельно процессу мониторинга запускается сбор статистики для косвенных параметров вычислительной системы с целью создания шаблона данной вычислительной системы.

3. Вычислительная система имеет шаблон, и база данных шаблонов не пустая. Тогда производится поиск шаблона данной вычислительной системы в базе данных. После чего запускается процесс мониторинга вычислительной системы.

Все описанные выше сценарии сводятся к запуску процесса мониторинга вычислительной системы вне зависимости от различий на предыдущих этапах. Однако общим для всех сценариев является периодическая корректировка шаблонов на основе собираемой статистики, сбор которой продолжается параллельно процессу мониторинга. Это делается с целью создания наиболее точного шаблона, описывающего вычислительную систему.

Описанная выше методология, основанная на создании шаблона и сценариях его применения, легла в основу архитектуры системы косвенного мониторинга несанкционированной активности в вычислительных системах [10].

Заключение

Представленная методология косвенного мониторинга несанкционированной активности в вычислительных системах позволяет обнаруживать несанкционированную активность вне зависимости от природы ее возникновения. Созданная на основе методологии система значительно облегчает обнаружение несанкционированной активности как специалистами, так и рядовыми пользователями. Данный факт позволяет своевременно принять все необходимые меры по противодействию и увеличивает шансы сохранения вычислительной системы в рабочем состоянии.

В дальнейшем планируется апробация полученной методологии и системы на вычислительных машинах действующей организации или предприятия для получения отзывов рядовых пользователей и специалистов. Апробация позволит выявить недочеты предложенной методологии и недоработки созданной системы.

ЛИТЕРАТУРА

1. Киселев А. Н. Подход к обнаружению вредоносного программного обеспечения web-shell на основе анализа сетевого трафика web-инфраструктуры. *Труды военно-космической академии имени А.Ф. Можайского.* 2021;677:143–152.
2. Саенко М. А., Шепель Д. П. Аппаратная безопасность, уязвимости и атаки: всеобъемлющая таксономия. *Вопросы устойчивого развития общества.* 2022;6:1294–1302.
3. Спиридонов С. Б., Чертилин А. А., Черненький М. В. и др. Аппаратные уязвимости современных процессоров, вызванные спекулятивным исполнением инструкций, и методы их исправления. *Естественные и технические науки.* 2018;5:270–273.
4. Бабенко Л. К., Кириллов А. С. Разработка автоматизированной системы обнаружения вредоносного программного обеспечения. *Известия ЮФУ. Технические науки.* 2021;7:153–167.
5. Сахно В. В., Проказова Ж. В. Анализ вредоносного программного обеспечения. *Modern Science.* 2021;9(2):226–229.
6. Pham D.-P., Marion D., Mastio M. et al. *Obfuscation Revealed: Leveraging Electromagnetic Signals for Obfuscated Malware Classification.* France: HAL archives ouvertes; 2021. 14 p. DOI: 10.1145/3485832.3485894.
7. Бушмелева К. И., Гавриленко А. В., Никифоров А. В. *Использование инверсии управления и внедрения зависимостей в архитектуре системы косвенного мониторинга несанкционированной активности.* М: Ассоциация выпускников и сотрудников ВВИА имени профессора Н.Е. Жуковского содействия сохранению исторического и научного наследия ВВИА имени профессора Н.Е. Жуковского; 2020. 472 с.
8. Вечтомов Е. М., Широков Д. В. *Математика: логика, теория множеств и комбинаторика.* Юрайт; 2022.
9. Гмурман В. Е. *Теория вероятностей и математическая статистика.* Юрайт; 2022.
10. Бушмелева К. И., Гавриленко А. В., Никифоров А. В. Информационная система косвенного мониторинга несанкционированной активности в компьютерных системах. *Вестник кибернетики.* 2021;4:16–21.