Теорема 2. *Пусть* $\mathfrak{N} = \mathfrak{pq}$ *— модуль RSA-криптосистемы в абстрактном числовом кольце* R. *Предположим, что существуют различные простые числа* r, s *и положительные целые числа* k, l *такие, что* $\varphi(\mathfrak{p}) = rk$, $\varphi(\mathfrak{q}) = sl$, *и числа* $r-1$, $s-1$ *имеют различные простые делители* r_1, s_1 *соответственно. Пусть* y *и* e *— независимые равномерно распределенные случайные величины со значениями в* $R_{\mathfrak{N}}$ *и* $\mathbb{Z}^*_{\varphi(\mathfrak{N})}$ *соответственно. Обозначим* $m_{e,y}$ *– минимальное* $m \in \mathbb{N}$ *такое, что* $y_m = y$. *Тогда выполняется неравенство*

$$\mathbb{P}(m_{e,y} \geq r_1 s_1) \geq (1 - r^{-1})(1 - s^{-1})(1 - r_1^{-1})(1 - s_1^{-1}),$$

где $m_{e,y}$ *это наименьшее натуральное число такое, что*

$$y^{e^{m_{e,y}}} \equiv y (mod\ \mathfrak{N}).$$

Литература

1. Кондратёнок Н. В. *Анализ RSA-криптосистемы в абстрактных числовых кольцах.* Журнал Белорусского государственного университета. Математика. Информатика, 2020. С. 13–21.

2. Харин Ю. С., Агиевич С. В., Васильев Д. В., Матвеев Г. В. *Криптология.* Минск: БГУ , 2013.

КЛАСТЕРИЗАЦИЯ ВРЕДОНОСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Абрамович М.С.[1], **Круглик К.С.**[1]

[1]НИИ прикладных проблем математики и информатики, БГУ, Независимости 4, 220050 Минск, Беларусь
abramovichms@bsu.by, kiraverlieren@yandex.by

Вредоносное программное обеспечение (или вредоносное ПО) – это программное обеспечение, используемое с целью нарушения политики безопасности компьютерной системы в отношении конфиденциальности, целостности или доступности.

Обнаружение вредоносных файлов для уже известных семейств вредоносного ПО происходит путем классификации, что упускает из вида образцы новых семейств. Для того, чтобы успешно обнаруживать образцы неизвестных ранее семейств вредоносного ПО, требуется выполнить кластеризацию, которая относится к классу задач обучения без учителя.

В данной работе рассмотрена выборка исполняемых файлов расширения EXE, которая содержит класс легитимных файлов (110 экземпляров) и класс вредоносных файлов (205 экземпляров). Файлы рассматривались как последовательность байтов. В качестве признаков использованы относительные частоты байтов, TF-IDF байтов [1, с. 358] и энтропия. TF-IDF — это статистическая мера, используемая для оценки важности байта в контексте файла, являющегося частью выборки файлов. Вес определенного байта пропорционален частоте встречаемости этого байта в файле и обратно пропорционален частоте встречаемости байта во всех файлах выборки.

Применены методы кластеризации [1, с. 185], а также методы поиска аномалий [2, 3] для кластеризации на два класса — вредоносное и легитимное ПО.

В качестве меры близости полученной кластеризации и исходных классов использовалась F-мера [1, с. 305] для каждого из классов и взвешенное значение F-меры для всей выборки.

Результаты эксперимента приведены в таблице 1.

Таблица 1 — Результаты сравнения кластеризации выборки и исходных классов

Метод	F-мера (вредоносные файлы)	F-мера (легитимные файлы)	F-мера взвешенная (вся выборка)
Метод k-средних	0.94	0.88	0.92
Мини-пакетный метод k-средних	0.90	0.90	0.94
Агломеративная кластеризация	0.99	0.98	0.98
BIRCH	0.93	0.85	0.90
Одноклассовый метод опорных векторов	0.78	0.23	0.59
Изоляционный лес	0.97	0.94	0.96
Эллиптический конверт	0.82	0.63	0.75

Как следует из таблицы 1, лучшая кластеризация получена при помощи иерархического агломеративного метода, для которого значение взвешенной F-меры равно 0.98.

Литература

1. Мюллер А., Гвидо С. *Введение в машинное обучение с помощью Python* // Руководство для специалистов по работе с данными. М.: Альфа-книга. 2017. Т. 487.

2. Liu F.T., Ting K.M., Zhou Z.H. *Isolation-based anomaly detection* // ACM Transactions on Knowledge Discovery from Data (TKDD). 2012. V. 6. № 1 P. 1-39.

3. Rousseeuw P.J., Driessen K.V. *A fast algorithm for the minimum covariance determinant estimator* // Technometrics. 1999. V. 41. № 3. P. 212-223.

ФИЛЬТРАЦИЯ УСЛОВНО-ГАУССОВСКИХ ПРОЦЕССОВ С ЗАДАННОЙ СПЕКТРАЛЬНОЙ ПЛОТНОСТЬЮ

Лобач В.И.[1], **Бернацкая Я.И.**[2]

[1]Белгосуниверситет, факультет прикладной математики и информатики
Независимости 4, 220050 Минск, Беларусь lobach@bsu.by
[2]Белгосуниверситет, факультет прикладной математики и информатики
Независимости 4, 220050 Минск, Беларусь yana.bernachkaya@gmail.com

Задача фильтрации [1] состоит в оценивании наилучшим образом изучаемого процесса по изменениям некоторых его характеристик, часто косвенных и содержащих «шумы» (погрешности). Это необходимо для получения более точных данных и для принятия правильных решений по наблюдению за исследуемым процессом. Многие модели временных рядов после ряда преобразований могут быть сведены к следующей схеме. Задана частично наблюдаемая случайная последовательность на вероятностном пространстве (Ω, F, P) $(\theta, \xi) = (\theta_t, \xi_t)$, $t = 0, 1, \ldots;$ θ_t и ξ_t определяются рекуррентными уравнениями:

$$\begin{aligned} \theta_{t+1} &= \alpha_0(t,\xi) + \alpha_1(t,\xi)\theta_t + b_1(t,\xi)\varepsilon_1(t+1) + b_2(t,\xi)\varepsilon_2(t+1), \\ \xi_{t+1} &= A_0(t,\xi) + A_1(t,\xi)\theta_t + B_1(t,\xi)\varepsilon_1(t+1) + B_2(t,\xi)\varepsilon_2(t+1). \end{aligned} \tag{1}$$

Задача фильтрации последовательности (1) заключается в оценивании ненаблюдаемой части процесса θ_t по наблюдаемым значениям $F_t^{\xi} = \{\xi_s, s = 0, \ldots, t\}$. Известно [2], что оптимальная в среднеквадратическом смысле оценка и величина ошибки оценивания определяются условными моментами:

$$m_t = M\left(\theta_t | F_t^{\xi}\right), \quad \gamma_t = M\left[(\theta_t - m_t)(\theta_t - m_t)^T | F_t^{\xi}\right]. \tag{2}$$