УНИВЕРСИТЕТ ИТМО
НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
июль–август 2016 Том 16 № 4 ISSN 2226-1494 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
July–August 2016 Vol. 16 No 4 ISSN 2226-1494 http://ntv.ifmo.ru/en

IMPLEMENTATION OF SIDE-CHANNEL LEAKAGE DETECTION TECHNIQUE BASED ON NORMALIZED INTER-CLASS VARIANCE METHOD

A.B. Levina[a], P.S. Borisenko[a]

[a] ITMO University, Saint Petersburg, 197101, Russian Federation
Corresponding author: alla_levina@mail.ru

Article info
Received 22.03.16, accepted 31.05.16
doi:10.17586/2226-1494-2016-16-4-697-702
Article in English

For citation: Levina A.B., Borisenko P.S. Implementation of side-channel leakage detection technique based on normalized inter-class variance method. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016, vol. 16, no. 4, pp. 697–702, doi:10.17586/2226-1494-2016-16-4-697-702

Abstract
The paper presents a new mathematical method for parasitic signal analyzing. NICV (Normalized Inter-Class Variance) method allows reducing considerably of computing and time expenditure in carrying out side channel attacks. To analyze NICV efficiency mathematical statistics methods and theory of probability were used. The paper presents an algorithm implementing NICV within a developed software package. The main shortcomings of the existing solutions have been taking into consideration during development of the new software: architecture of the presented software is easily extensible for adding new tools; unified format is used for all processing data. NICV was tested on the first round of 64-bit Data Encryption Standard algorithm. To assess the effectiveness two attacks based on differential power analysis and correlation power analysis have been simulated. Another advantage of the package is flexibility in adding of new methods for processing, saving, both original information and its new statuses in the database after carrying out signal processing. Side-Channel Attacks (SCA) are considered as a serious threat for data protected by cryptographic devices. Therefore such devices must be tested for resistance to these attacks. It should be taken into account that SCA are very powerful tool but they require significant computation capacity, especially in case of countermeasures. Presented software package program can help to analyze cryptographic devices on resistance to SCA and implemented NICV method allows decreasing of time and computation costs.

Keywords
cryptography, side-channel attacks, NICV

Acknowledgements
We are gratefully thankful to Sylvain Guilley as one of the authors of the NICV technique; without his help presented study could not have been initiated and completed.
This work was presented on the Information Security and Protection of Information Technology Conference 2015

УДК 004.056.5

РЕАЛИЗАЦИЯ ТЕХНОЛОГИИ ОБНАРУЖЕНИЯ УТЕЧКИ ИНФОРМАЦИИ ПРИ ПРОВЕДЕНИИ АТАК ПО СТОРОННИМ КАНАЛАМ, ОСНОВАННОЙ НА МЕТОДЕ ВЫЧИСЛЕНИЯ НОРМАЛИЗОВАННОЙ ВНУТРИКЛАССОВОЙ ДИСПЕРСИИ

А.Б. Левина[a], П.С. Борисенко[a]

[a] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
Адрес для переписки: alla_levina@mail.ru

Информация о статье
Поступила в редакцию 22.03.16, принята к печати 31.05.16
doi: 10.17586/2226-1494-2016-16-4-697-702
Язык статьи – английский

Ссылка для цитирования: Левина А.Б., Борисенко П.С. Реализация технологии обнаружения утечки информации при проведении атак по сторонним каналам, основанной на методе вычисления нормализованной внутриклассовой дисперсии // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 4. С. 697–702. doi: 10.17586/2226-1494-2016-16-4-697-792

Аннотация
Предложен алгоритм реализации новейшего математического метода анализа паразитного сигнала. Метод NICV (Normalized Inter-Class Variance) позволяет значительно уменьшить вычислительные и временные затраты при проведении атак по сторонним каналам. Анализ эффективности метода выполнен с использованием инструментов математической статистики и теории вероятности. В работе представлен алгоритм, реализующий NICV на базе

разработанного программного обеспечения. При разработке учитывались основные недостатки существующих решений: архитектура представленного программного комплекса легко расширяема для добавления новых инструментов, весь функционал основывается на унифицированном формате обрабатываемых данных. Реализация NICV протестирована для первого раунда 64-битного алгоритма Data Encryption Standard. Для оценки эффективности были смоделированы атаки на базе дифференциального и корреляционного анализа мощности. Преимуществом комплекса также является гибкость в задачах добавления новых методов обработки и хранения как оригинальной информации, так и ее новых состояний в базе данных после проведения процедур обработки сигнала. Атаки по сторонним каналам представляют серьезную угрозу для данных, защищаемых с использованием криптографических устройств. В связи с этим криптографические устройства должны проходить тестирование на защищенность от атак по сторонним каналам. При этом следует учитывать, что данные атаки являются очень сильным инструментом, но для их проведения необходимы большие вычислительные мощности, особенно в случае наличия контрмер. Разработанный программный комплекс позволяет анализировать криптоустройства на криптостойкость к атакам по сторонним каналам, а реализованный метод NICV – сократить вычислительные и временные затраты при его использовании.

Ключевые слова
криптография, атаки по сторонним каналам, NICV

Благодарности
Выражаем огромную благодарность Сильвиану Гилле, одному из авторов метода NICV. Без его помощи данная работа не была бы начата и завершена.
Работа была представлена на международной конференции Information Security and Protection of Information Technology Conference 2015.

Introduction

Side-channel attacks (SCA) are an actual trend in the field of information security, as they consider no theoretical models of cryptographic algorithms, but information about physical processes in the encryption units. For example, timing information, power consumption, electromagnetic leaks or even sound can provide an extra source of information that can be exploited to break the system. Some side-channel attacks require technical knowledge of the internal operation of the system in which the cryptography is implemented, although others such as differential power analysis are effective as black-box attacks. Many powerful side-channel attacks are based on statistical methods pioneered by Paul Kocher [1].

There are many different types of side-channel attacks. Each of them involves the processing of a specific data set obtained by parasitic signal retrieved from the device performing cryptographic operations. Implementation of attacks may be directed not only to disclose a secret key, but also for the detection of potential vulnerabilities during the process of devices certification.

There are three main types of approaches used for side-channel attacks: SPA, DPA and CPA. The difference is that SPA (Simple Power Analysis) is a side-channel attack that involves visual examination of graphs of the current used by a device over time, DPA (Differential Power Analysis) looks at differences of category averages for all key guesses and CPA (Correlation Power Analysis) looks at correlation between all key guesses. NICV works in co-ordination with any SCA distinguishers like CPA to enhance their performance.

In the field of tasks related to software and hardware elements of protection systems, determination of optimal and reliable ways for cryptographic protection of information and its transmission channels occupies a separate niche. Data protection with cryptographic tools is one of the most important solutions of security problem.

A large number of mechanisms have been created to protect information: symmetric and asymmetric encryption algorithms, protocols, digital signatures, identification, and authentication schemes [2]. However, more and more new methods of attacks on emerging technologies regularly appear following the improvement of security elements. Despite the negative connotation of this phenomenon, it certainly contributes to the further development of cryptographic solutions and products based on them.

Normalized Inter-Class Variance (NICV) was presented in 2013 in [3]. This method allows detection of interesting time samples, without the need of a profiling stage on a clone device. Hence the SCA traces can be compressed and the analysis could be greatly accelerated. The main characteristics of NICV are:

- NICV operates without the need of a clone device, i.e. it requires no profiling stage and use the same set of traces that are to be analyzed;
- it uses only public information like plaintext or ciphertext;
- the method is leakage model agnostic, it is not an analysis tool but a helper to speed up the analysis;
- it can serve to evaluate the accuracy of various leakage models and choose which is the best applicable.

This paper presents implementation and testing NICA for analyzing the signal received by side-channel attacks. This complex uses certain format of data presentation, uniform for all processes included in him; it allows analyzing any loaded data. The developed program complex, besides providing access to the available set of methods, considerably saves time spent for search/development of converting.

Implementation of NICV, testing it on real data and developing software package provides a set of tools for manipulating parasitic signal, attacks and finding the most vulnerable signal components.

Software package tools

The developed software uses two reference-attacks for the first round of 64-bit Data Encryption Standard (hereinafter DES), based on the DPA [4] and CPA. The reference-attacks were proposed by Paul Kosher and were based on the principle of markup used in cryptanalysis (partitioning cryptanalysis) [5]. The markup attack used against block algorithms implies separation of possible intermediate states of the cipher in order to detect the most common conditions in a particular set of plaintexts and ciphertexts.

Depending on the configuration of file presented software provides one of two attacks on chosen signal group to search for the desired subkeys, detection of which is sufficient to obtain the secret key encryption.

To speed up the attack NICV has been selected, it is the newest method for the detection of vulnerable components of the signal. The calculation of the encryption key may take a large amount of time (hours to days) when analyzing takes large amounts of data. The main application of NICV is to speed up subsequent search of the key.

NICV calculation is based on the following ideas.

Let us call X one byte of the plaintext or of the ciphertext (that is, the domain of X is $X = \mathbb{F}_2^8$), and $Y \in \mathbb{R}$ the leakage measured by the attacker, in general, Y can be continuous, but X must be discrete (and X must be of finite cardinality). Both random variables are public knowledge. Then, for all leakage prediction function L of the leakage knowing the value of x taken by X, we have:

$$\rho^2[L(X); Y] = \rho^2[L(X); \mathbb{E}[Y|X]] \times \rho^2[\mathbb{E}[Y|X]; Y]. \quad (1)$$

Also, it has been proved in [6]:

$$\rho^2[\mathbb{E}[Y|X]; Y] = \frac{\mathrm{Var}[\mathbb{E}[Y|X]]}{\mathrm{Var}[Y]}. \quad (2)$$

Once combined, equations (1) and (2) yield that for all prediction function $L: \mathbb{F}_8^2 \to \mathbb{R}$, we have:

$$0 \le \rho^2[L(X); Y] \le \frac{\mathrm{Var}[\mathbb{E}[Y|X]]}{\mathrm{Var}[Y]} = \mathrm{NICV} \le 1. \quad (3)$$

Therefore, the NICV is the envelop or maximum of all possible correlations computable from X with Y. There is an equality in (3) if and only if $L(x) = \mathbb{E}[Y|X = x]$, which is the optimal prediction function.

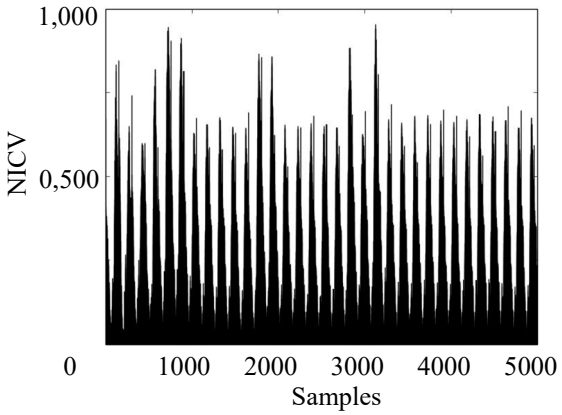

Figure. NICV diagram for a records' group

NICV values above a predetermined threshold, defined as the most vulnerable. On the Figure above, NICV was calculated for records containing 5002 points.

Implementation Tools

Implementing software system that handles large volumes of data (signal records collected for one device may have a volume of several GB) and to maintain an optimal storage of files within the database imposes certain restrictions on choice of the programming language and DBMS. In addition, cost plays another important role, as well as the ability to run your software on multiple platforms.

Analysis of the existing instruments based on the criteria described above, led to the choice of a combination of the Python programming language and PostgreSQL DBMS.

Python [7] – high-level programming language, focused on improving the readability of the code, and developer productivity. It is well ported to almost all known programming languages. The standard library and additional modules include a large amount of useful features, like the ones optimized to perform statistical calculations, signal processing, etc.

PostgreSQL [8] – free object-relational database management system implemented for a large number of UNIX-platform, as well as Microsoft Windows. One of the strengths of PostgreSQL is support of virtually unlimited database size, the library interface with many programming languages, and the ability to access efficiently the stored files if they are small.

Signal Format

The operation of each element of software, working with the signal should be based on pre-regimented format (or a number of formats) of data within the package. The main criteria in the current implementation are minimalistic meta-data downloaded files, as well as a compact representation of records. Both factors significantly affect the performance of software at processing.

The implemented software system supports Agilent Binary Format, created by Agilent Technologies, the manufacturer of measuring, medical equipment and equipment for chemical analysis. Files formed on its basis are represented by the following structure (the basic elements of metadata are described):

File Header.
Cookie –AG bytes, which represent the format.
Version – two bytes representing version.
File Size – file size in bytes.
Number of Waveforms – number of signal records (it is possible to store more than one).
Waveform Header – header of a signal record.
Header Size – in bytes.
Waveform Type – decimal number which can be:
1 = Raw
2 = Average
3 = Horizontal Histogram
4 = Vertical Histogram
5 = Versus
6 = Interpolate
8 = Color Grade
10 = Peak Detect
Date – date of the record creation. Default value – 27 DEC 1996.
Time – time of the record creation. Default value – 01:00:00:00.
Frame – serial number and model number of oscilloscope in form of MODEL#:SERIAL#.
Waveform Data Header – particular record information.
Waveform Data Header Size – header size in bytes.
Buffer Type – record data type:
0 = Unknown data
1 = Normal 32 bit float data
2 = Maximum float data
3 = Minimum float data
4 = Time float data
5 = Counts 32 bit float data
6 = Digital unsigned 8 bit char data
Buffer Size – record size in bytes.

Software package commands

The software package provides the user interface to perform operations on groups of signal records. Appeal to the methods performed through the OS terminal running startup file scapack.py with a certain set of parameters (see Table 1).

Command	Parameters	Description
-nicv	<group name>	Applies NICV to all signal records of the chosen group.
-reference	<group name>	Applies reference attack to all signal records of the chosen group. Save the results in TRACE table. Parameters of the attack based on config file.
-load	<group name> <path to the folder containing signal records in Agilent format>	Load new signal records to the database.
-kinds		Provides the list of all signal groups in the database.
-delete	<group name>	Delete all records in the chosen group.
-upload	<group name> <path to output folder>	Upload all records of the chosen group to the output folder.
-help/null		Provides information about all supported commands.
-config		Goes to the editor of config file.

Table 1. Commands description

Tests

During the software testing, loads, processing and attacks of two types have been applied to parasite signals, published during the DPA Contest, organized by the French University of Télécom ParisTech. The purpose of this contest is to give to different teams the possibility to check their attacks and compare the results with each other based on final statistics [9].

Total records volume is 11 GB. Records are divided into three groups, depending on the accessories to the device with the corresponding private key. They were loaded into the software package in the form of three groups: des_first (81089 records), des_second (81,569 entries) and des_third (67,753 entries). The results of the attack on each group are given in Table 2.

Group	Key detection time	Records processed	Key result	Attack type
des_first	2000–7000 s	8000–30000	Correct	DPA
des_first	3000–4500 s	9000–12000	Correct	CPA
des_second	3000–8000 s	10000–32000	Correct	DPA
des_second	3000–6000 s	10000–14000	Correct	CPA
des_third	~17000 s	67753	Incorrect	DPA
des_third	~17000 s	67753	Incorrect	CPA

Table 2. The attacks without NICV pre-processing

In order to speed up the process NICV was calculated for each group. After that attacks were carried out only the data items that have been marked for the group as the most vulnerable. NICV designed to optimize attacks such as CPA, obtained statistics of this attack is shown in Table 3.

Group	Key detection time	Records processed	Key result	Attack type
des_first	2000–3500 s	8000–11000	Correct	CPA
des_second	2000–4000 s	9000–13000	Correct	CPA
des_third	~12000 s	67753	Incorrect	CPA

Table 3. The attacks after NICV pre-processing

NICV allows reducing the amount of input data for future attacks and, consequently, reducing the time for their implementation. However, the calculations required to use this method in themselves require hardware resources. At the average execution time of ~ 700 s, CPA-attacks were decreased for 1000–1500 s that is an indicator of performance improvements from pre-record processing by NICV method.

The tests method were carried out on parasitic signals from the device that performs AES encryption with each record containing 7 000 000 components. NICV compression was significantly higher (7000 times) with such volumes of data so it can be successfully applied before the relevant attacks execution.

Conclusion

This publication presents implementation of side-channel leakage detection technique based on NICV.

The developed software package allows pre-processing of signal records in order to identify the most vulnerable components of data, provides two types of attacks, as well as the tools for management of collected signal records.

The work of this package is based on the specific format of data common to all the processes included in it that allows you to apply any of them to properly loaded information. Commonality of this kind makes it necessary to transform the format of the input data required in case of differences, but as a result of the latter, users get access to all methods of the software package. Where the tools are scattered and not integrated in the general solution, users are forced to perform format conversion for each specific instance of the software. Thus, the developed software package, in addition to providing access to the existing set of methods saves time spent on search / development of tools for such conversion.

The use of the package is possible during analysis of cryptographic devices, as well as in research related to side-channel attacks. The test results of the package may be useful in assessing the effectiveness of other existing signal analyzing methods.

References

1. Kocher P., Jaffe J., Jun B. Introduction to Differential Power Analysis and Related Attacks. 1998.
2. Левина А.Б. Моделирование криптосистем. СПб.: НИУ ИТМО, 2013. 82 с.
3. Bhasin S., Danger J.-L., Guilley S., Najm Z. NICV: normalized inter-class variance for detection of side-channel leakage // IEEE International Symposium on Electromagnetic Compatibility. Tokyo, 2013. V. 3. P. 310–313.
4. Kocher P., Jaffe J., Jun B. Differential power analysis // Lecture Notes in Computer Science. 1999. V. 1666. P. 388–397.

5. Genkin D., Pipman I., Tromer E. Get your hands off my laptop: physical side-channel key-extraction attacks on PCs // Lecture Notes in Computer Science. 2014. V. 8731. P. 242–260.
6. Prouff E., Rivain M., Bevan R. Statistical analysis of second order differential power analysis // IEEE Transactions on Computers. 2009. V. 58. P. 799–811. doi: 10.1109/TC.2009.15
7. Python Documentation and Downloads. Режим доступа: https://www.python.org, своб. Яз. англ. (дата обращения 10.06.2016)
8. PostgreSQL: the world's most advanced open source database. Режим доступа: http://www.postgresql.org, своб. Яз. англ. (дата обращения 10.06.2016)
9. Danger J.-L., Duc G., Guilley S., Sauvage L. Education and open benchmarking on side-channel analysis with the DPA contests. NIST, USA, 2011. V. 2. 7 p.

Alla B. Levina — PhD, Associate professor, ITMO University, Saint Petersburg, 107101, Russian Federation, alla_levina@mail.ru

Pavel S. Borisenko — student, ITMO University, Saint Petersburg, 107101, Russian Federation, borisenkopp@yandex.ru

Левина Алла Борисовна — кандидат физико-математических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, alla_levina@mail.ru

Борисенко Павел Сергеевич — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, borisenkopp@yandex.ru