BRAZILIAN JOURNAL OF INFORMATION SECURITY AND CRYPTOGRAPHY

VOLUME 2 - ISSUE 1 - Brasília — Brazil - 2015

RENASIC
Rede Nacional em Segurança da Informação e Criptografia

National Network of Information Security and Cryptography

Science and Technology Department

Brazilian Army

Cyber Defense Center

ENIGMA — Brazilian Journal of Information Security and Cryptography

Eduardo Takeo Ueda

Mirela Sechi Moretti Annoni Notare
Rafael Timóteo de Sousa Júnior

VOLUME 2

ISSUE 1

Brasília — Brazil
2015

ENIGMA — Brazilian Journal of Information Security and Cryptography

Editors

Editorial Board

Coordinator of the National Network of Information Security and Cryptography

Antônio Carlos Menna Barreto Monclaro

Chief of the Center of Cyber Defense

General de Divisão Paulo Sergio Melo de Carvalho

Chief of the Science and Technology Department

General de Exército Juarez Aparecido de Paula Cunha

Brazilian Army Commander

General de Exército Eduardo Dias da Costa Villas Boas

Production

Matheus Barbosa, Graphic Designer of Logo
Marcílio Costa Bezerra, Publication Designer

SUMMARY

Editorial 1

Accepted Papers

Proposal of Enhancement for Quartz Digital Signature 3
E. R. Andrade and R. Terada

Untappable Key Distribution System: a One-Time-Pad Booster 16
G. A. Barbosa and J. van de Graaf

Cyber-Attacks Based in Electromagnetic Effects 28
M. B. Perotoni, R. M. Barreto and S. K. Manfrin

Invited Papers

A Modularity and Extensibility Analysis on Authorization Frameworks 36
E. M. Guerra, J. O. Silva and C. T. Fernandes

The Producer-Consumer Collusion Attack in Content-Centric Networks 48
A. Nasserala and I. M. Moraes

SpamBands - a Methodology to Identify Sources of Spam Acting in Concert 56
E. Fazzion, P. H. B. Las-Casas, O. Fonseca, D. Guedes,
W. Meira Jr, C. Hoepers, K. Steding-Jessen and M. H. P. Chaves

ENIGMA – Brazilian Journal of Information Security and Cryptography Volume 2 Issue 1 September 2015

E. T. Ueda, *Editor in Chief,* M. S. M. A. Notare, *Associate Editor in Chief,* and R. T. de Sousa Júnior, *Associate Editor in Chief*

***Abstract*— This is the first issue of Volume 2 of ENIGMA – Brazilian Journal of Information Security and Cryptography. Submissions were accepted in English, Portuguese and Spanish. In this issue, 6 papers are published, of which 3 were peer-reviewed while the other 3 were invited and reviewed by the editorial board of the Journal. In addition, one of the invited papers was the second Best Paper from the conference SBRC'2015 while another was the Best Paper of the conference SBSeg'2014.**

***Keywords*— Brazilian Journal, Cryptography, Information Security.**

I. INTRODUCTION

ENIGMA – Brazilian Journal of Information Security and Cryptography – is a technical-scientific publication that aims at discussing theoretical aspect contributions and practical applications results in information security, cryptography and cyber defense as well as fundamental subjects in support of those issues.

The choice of the name ENIGMA for this publication is related to the ENIGMA cryptography machine. However, the main reason for this choice is to pay tribute to the mathematician and computer scientist Alan Mathison Turing (1912-1954), considered one of the leading scientists in the history of computing.

This journal is directed to academia researchers, industry professionals, members of government and military organizations, and all people that have interest in the area of information security and cryptography in order to disseminate and share their new technologies, scientific discoveries and research contributions.

The creation of this periodical is due the necessity to solve a gap represented by the lack of a technical-scientific Brazilian journal that emphasizes information security and cryptography. In this manner, ENIGMA – Brazilian Journal of Information Security and Cryptography – must provide this demand, publishing papers of high quality within the international state-of-the-art.

Therefore, ENIGMA – Brazilian Journal of Information Security and Cryptography – aims to fulfill this demand, and will publish state-of-art and original research papers and timely review articles on the theory, design, and evaluation of all aspects of information, network and system security.

II. ABOUT VOLUME 2, ISSUE 1 OF ENIGMA

In this issue of Volume 2 of ENIGMA – Brazilian Journal of Information Security and Cryptography – 6 papers are published. This section briefly presents the contribution of each of these papers.

The first selected paper, entitled "Proposal of Enhancement for Quartz Digital Signature", is published in Portuguese. In this paper, the authors propose a new digital signature scheme, based on the Quartz digital signature scheme, which pertains to the class of Hidden Field Equations (HFE), with a special choice of parameters. The presented scheme achieves an estimated security level estimated at 2^{112}, regarding adaptive chosen message attacks that make calls to the random Oracle.

The selected paper "Untappable Key Distribution System: a One-Time-Pad Booster", which is published in English, proposes a solution for the secure sharing and renewal of keys for the One-Time-Pad (OTP) protocol. To provide fast and unlimited renewal of secure keys, the proposed untappable key distribution system utilizes two layers of confidentially protection, based on the physical noise intrinsic to the optical channel and a bit pool of refreshed entropy.

Also in English, the selected paper "Cyber-Attacks Based in Electromagnetic Effects" covers cyber-attacks that take advantage of unintended electromagnetic emanations from the data sources, comprising a survey of some attacks, alongside with measurements that show the basic nature and underlying principles involved.

The first invited paper is "A Modularity and Extensibility Analysis on Authorization Frameworks", published in English, and presenting a comparative analysis between the existing authorization frameworks developed either within the academic and industry environments. This analysis uses a motivating example to present the main industry frameworks and consider the fulfillment of modularity, extensibility and granularity requirements facing its suitability for the existing access control models.

The next two invited papers come from relevant Brazilian conferences, being in this ENIGMA issue published in Portuguese. The first one, "The Producer-Consumer Collusion Attack in Content-Centric Networks", which was considered the 2nd best paper of SBRC´2015, evaluates the impact of a denial-of-service attack in information-centric networks based on the Content Centric Networking (CCN) architecture. In the considered attack, both malicious consumers and producers collude, by generating, publishing, and changing content popularity, thus increasing the content retrieval time.

Closing this ENIGMA issue, the invited paper "SpamBands - a Methodology to Identify Sources of Spam Acting in Concert", the best paper of SBSeg´2014, considers the relationships between the machines used to send spam as the basis for an analysis that could reveal how different machines may be used by a single spammer to spread his messages. Then, this work proposes a methodology to

E. T. Ueda, Institute for Technological Research of the State of São Paulo, edutakeo@usp.br

M. S. M. A. Notare, IEEE Latin America Transactions Editor in Chief, IEEE South Brazil, mirela@ieee.org

R. T. de Sousa Júnior, University of Brasilia, desousa@unb.br

cluster the machines used by spammers, thus identifying different aspects of the spam dissemination process.

III. CONCLUSION

ENIGMA – Brazilian Journal of Information Security and Cryptography – is now in its second year. Adopting since its creation the best practices from IEEE Transactions publications, it is hoped that soon this journal will become a reference among the leading international publication dedicated to information security and cryptography.

With the creation of this journal Brazil makes a considerable step toward the future, because the ENIGMA journal is an important tool for communication and integration of knowledge between universities, research centers, industries, government or military institutions around the world. Moreover, as threats to information security and privacy are risks for any nation, the ENIGMA journal can envision the international community.

ACKNOWLEDGEMENTS

We would like to thank all the authors who contributed with their papers for this issue of the ENIGMA journal, this publication would not exist if not for the dedication to their research. We must also thank all reviewers who worked very hard and in a timely fashion, so that we could select high quality papers. We are grateful to National Network of Information Security and Cryptography (RENASIC) and the Cyber Defense Center (CDCiber) of the Defense Minister of Brazil for their support in the creation of this journal, and the University of Brasilia (UnB) for providing space on one of their servers to host the official website of this journal. Moreover, we wish to thank Coronel Coronel Eduardo Wallier Vianna and Major Helder Vieira Bezerra for their support and in believing in this journal as well as their continuous help to get the printed version of this issue. Last but not least, thanks to Itamar Annoni Notare for his hard work and for his assistance in the revision process as well as on the formatting this journal.

Eduardo Takeo Ueda received the Ph.D. degree in Electrical Engineering in 2012, MSc degree in Computer Science in 2007, both from University of São Paulo (USP), and Specialist degree in Health Informatics in 2014 by Federal University of São Paulo (UNIFESP). He also holds a Mathematics degree by the São Paulo State University (UNESP), year 2000. His research interest includes topics of Cryptographic Algorithms and Protocols, Models of Access Control, and Computational Trust and Reputation. He has been committee member in conferences and reviewer of scientifics journals. Currently, he is Professor in Senac University Center of São Paulo, Master's Thesis Advisor in Institute for Technological Research of the State of São Paulo, member of the National Network of Information Security and Cryptography (RENASIC), and Editor in Chief of ENIGMA – Brazilian Journal of Information Security and Cryptography. http://lattes.cnpq.br/8367973725203446.

Mirela Sechi Moretti Annoni Notare received her Ph.D. and MSc degrees from the Federal University of Santa Catarina (UFSC) and a BSc degree from Passo Fundo University – all the three degrees in Computer Science. Her main research of interest focuses on the proposition of security management solutions for Wireless, Mobile, Sensor and Ad-Hoc Networks. Dra. Mirela Notare published widely in these areas. She also received several awards and citations, such as National Award for Telecommunication Software, British Library, TV Globo, INRIA and Elsevier Science. She served as General Co-chair for the I2TS (International Information and Telecommunication Technologies Symposium) and Program Co-Chair for the IEEE MobiWac (Mobility and Wireless Access Workshop) and IEEE ISCC. She has been a committee member in several scientific conferences, including ACM MSWiM, IEEE/ACM ANSS, IEEE ICC, IEEE IPDPS/WMAN IEEE/SBC SSI, and IEEE Globecom/Ad-Hoc, Sensor and Mesh Networking Symposium. She has been Guest Editor for several international journals, such as JOIN (The International Journal of Interconnection Networks), IJWMC (Journal of Wireless and Mobile Computing), JBCS (Journal of Brazilian Computer Society), Elsevier ScienceJPDC (The International Journal of Parallel and Distributed Computing), Wiley & Sons Journal of Wireless Communications & Mobile Computing, and Wiley InterScience Journal Concurrency & Computation: Practice & Experience. She has some Books and Chapters – Protocol Engineering with LOTOS/ISO (UFSC) and Solutions to Parallel and Distributed Computing Problems (Wiley Inter Science), for instance. She is the current Editor in Chief of IEEE Latin America Transactions magazine and Associate Editor in Chief of ENIGMA – Brazilian Journal of Information Security and Cryptography. She is the founding and president of STS Co, a senior member (21 years) of IEEE, and member of SBrT and SBC societies. http://lattes.cnpq.br/8224632340074096.

Rafael Timóteo de Sousa Júnior, was born in Campina Grande – PB, Brazil, on June 24, 1961. He graduated in Electrical Engineering, from the Federal University of Paraíba – UFPB, Campina Grande – PB, Brazil, 1984, and got his Doctorate Degree in Telecommunications, from the University of Rennes 1, Rennes, France, 1988. He worked as a software and network engineer in the private sector from 1989 to 1996. Since 1996, He is a Network Engineering Professor in the Electrical Engineering Department, at the University of Brasília, Brazil. From 2006 to 2007, supported by the Brazilian R&D Agency CNPq, He took a sabbatical year in the Group for the Security of Information Systems and Networks, at Ecole Superiéure d´Electricité, Rennes, France. He is a member of the Post-Graduate Program on Electrical Engineering (PPGEE) and supervises the Decision Technologies Laboratory (LATITUDE) of the University of Brasília. He is a member of the Brazilian Computer Society (SBC) and member of the National Network of Information Security and Cryptography (RENASIC). His field of study is distributed systems and network management and security. http://lattes.cnpq.br/3196088341529197.

Proposal of Enhancement for Quartz Digital Signature

E. R. Andrade and R. Terada

***Abstract*— Today, we see a large dependence on systems developed with cryptography. Especially in terms of public key cryptosystems, which are widely used on the Internet. However, public key cryptography was threatened and new sources began to be investigated when Shor in 1997 developed a polynomial time algorithm for factoring integers and to compute the discrete logarithm with a quantum computer. In this context, Patarin proposed Hidden Field Equations (HFE), a trapdoor based on $\mathcal{MQ}$ ($\mathcal{M}$ultivariate $\mathcal{Q}$uadratic) and IP (Isomorphism of Polynomials) problems. Such problems are not affected by the Shor algorithm, moreover $\mathcal{MQ}$ Problem was proved by Patarin and Goubin to be NP-complete. Despite the basic HFE has been broken, there are variants that are secure, obtained by a generic modification. The Quartz – digital signature scheme based on HFEv-, with special choice of parameters – is a good example of this resistance to algebraic attacks aimed at the recovery of the private key, because even today it remains secure. Furthermore, it also generates short signatures. However, Joux and Martinet, based on axioms of Birthday Paradox Attack, proved that Quartz is malleable, showing that if the adversary has a valid pair (message, signature), he can get a second signature with 2^{50} computations and 2^{50} calls to the signing oracle, so that the estimated current security standards are at least 2^{112}. Thus, based on Quartz, we present a new digital signature scheme, achieving the adaptive chosen message attacks that make calls to the random oracle, with a security level estimated at 2^{112}. Our cryptosystem also provides an efficiency gain in signature verification algorithm and vector initializations that will be used for signing and verification algorithms. Furthermore we provide an implementation of Original Quartz and Enhanced Quartz in the Java programming language.**

***Keywords*— Post-Quantum Cryptography, $\mathcal{MQ}$ Problem, Digital Signature, Quartz, MPKC.**

I. INTRODUÇÃO

PODEMOS perceber que atualmente, seja conscientemente ou não, uma dependência dos sistemas desenvolvidos sob a seara da criptografia foi instaurada em todos nós. Principalmente no tocante dos sistemas criptográficos de chave pública, que são vastamente utilizados na Internet, incluindo-se aí os esquemas de assinatura digital.

No entanto, desde quando Shor em 1997 desenvolveu um algoritmo de tempo polinomial para fatorar inteiros e para calcular o logaritmo discreto num computador quântico [49] – computador este, proposto por Deutsch em 1985 [15] – a criptografia de chave pública se viu ameaçada e começou a investigar novas fontes de problemas para seus sistemas. Este alarde ocorrera porque, basicamente, os criptossistemas de chave pública usados na atualidade têm sua segurança baseada na intratabilidade dos problemas da fatoração de inteiros, no caso de sistemas RSA, e do logaritmo discreto, em sistemas ElGamal ou de Curvas Elípticas, e tal descoberta tornaria estes sistemas inseguros quando possuíssemos computadores quânticos com a capacidade adequada para implementarmos o algoritmo de Shor.

Acreditamos que a possibilidade de evolução dos computadores quânticos não deveria ser encarada como único fator para a obsolescência dos criptossistemas de chave pública atuais. Pois além de existir incontáveis estudos acerca da segurança destes problemas ditos clássicos, a capacidade computacional aumenta significativamente a cada década, e isto, sem dúvida, torna padrões outrora considerados seguros em inseguros.

Uma interessante proposta para enfrentarmos estes desafios é utilização de sistemas MPKC (acrônimo da nomenclatura em inglês que significa Criptossistema de Chave Pública Multivariável), que se apoiam no Problema $\mathcal{MQ}$(*$\mathcal{M}$ultivariate $\mathcal{Q}$uadratic*) para o desenvolvimento ou aprimoramento de sistemas criptográficos de chave pública seguros.

O Quartz é um esquema de assinatura digital baseado no HFEv-, com escolha especial de parâmetros. Sua versão original proposta por Patarin, Courtois e Goubin em 2001 [44] foi atualizada pelos mesmos autores logo em seguida [14], sendo que desde então adotamos esta última como versão original. Este esquema de assinatura foi submetido e aceito no NESSIE (*New European Schemes for Signatures, Integrity and Encryption*), um projeto de pesquisa desenvolvido com a *Information Societies Technology* (IST) *Programme of the European Commission* para identificar sistemas criptográficos seguros que forneçam – em sentido amplo – confidencialidade e integridade dos dados, além de autenticidade das entidades [38]. De acordo com os relatórios públicos do NESSIE, o principal trunfo deste esquema são suas assinaturas curtas (apenas 128 bits) e a fundamentação em um problema intratável até mesmo em computadores quânticos (o problema $\mathcal{MQ}$) [37].

Contudo, o Quartz não foi selecionado para figurar no portfólio final desse projeto de pesquisa. Isto porque – em linhas gerais – o cálculo de suas chaves secretas foi considerado muito lento (comparando com os demais esquemas submetidos) [37]; por possuir algumas divergências nas especificações de sua implementação (quando confrontado com o requerido pelo NESSIE) [20]; e também por possuir

E. R. Andrade, Laboratório de Arquitetura de Redes de Computadores (LARC), Escola Politécnica, Universidade de São Paulo (Poli-USP), São Paulo, SP, Brasil, ewe@ime.usp.br

R. Terada, Departamento de Ciência da Computação (DCC), Instituto de Matemática e Estatística, Universidade de São Paulo (IME-USP), São Paulo, SP, Brasil, rt@ime.usp.br

uma arquitetura maleável que permite ao adversário obter uma segunda assinatura, caso ele possua um par (mensagem, assinatura) válido, com uma quantidade de cálculos muito menor do que o solicitado pelo projeto [33].

Desta forma, o objetivo de nosso trabalho foi analisar o esquema de assinatura digital Quartz, apresentando, ao final deste estudo, um novo protocolo de assinatura digital baseado nele, porém, com um nível de segurança ainda maior e um algoritmo de verificação mais eficiente. Além disto, foi desenvolvida uma implementação de referência, tanto do modelo original quanto de nosso modelo proposto, para então ser analisada a viabilidade de nosso modelo através da estimativa de segurança e apreciação dos tempos obtidos durante os testes realizados a partir de nossa implementação.

A. Contribuições e organização do trabalho

As principais contribuições deste trabalho são: a apresentação de um novo protocolo de assinatura digital baseado no Quartz, logo, com assinaturas curtas e fundamentado em um problema intratável até mesmo em computadores quânticos; obtenção de um criptossistema resistente a ataques adaptativos que realizem chamadas ao oráculo aleatório, com um nível de segurança estimado em 2^{112}, contra os 2^{50} do protocolo original; constatação de que nosso aprimoramento irá testar até 4.096 vezes menos hipóteses de utilização da chave pública durante a verificação de assinatura, quando comparado com o Quartz Original; implementação do Quartz Original e do Quartz Aprimorado em uma linguagem de programação portável.

Para isto, organizamos este trabalho da seguinte forma. Na seção II, apresentamos as principais notações e definições utilizadas durante o desenvolvimento deste trabalho. Na seção III, descrevemos sucintamente o Problema $\boldsymbol{\mathcal{MQ}}$ e seu uso na criptografia de chave pública, apresentando seu esquema genérico de funcionamento. Neste ponto, também elencamos e descrevemos sucintamente as principais características dos modificadores genéricos aplicáveis as funções $\boldsymbol{\mathcal{MQ}}$ básicas. Na seção IV até a XI colocamos as principais contribuições de nosso trabalho. Nelas, revisamos o HFE e o Quartz Original, levantamos alguns aspectos referente a segurança do modelo original, apresentamos nossa proposta de aprimoramento, analisamos as modificações propostas, estimamos o impacto destas modificações na segurança, e ainda, apresentamos os tempos obtidos durante os testes realizados a partir de nossa implementação de referência. Por fim, na seção XII, apresentamos as considerações finais de nossa pesquisa e propomos novas direções para trabalhos futuros.

II. PRINCIPAIS NOTAÇÕES UTILIZADAS

Com intuito de facilitar a leitura, destacamos na TABELA I os principais parâmetros, bem como algumas definições e terminologias pertinentes ao Quartz. Além disto, frisamos que utilizaremos $\mathbb{F}$ ou $\mathbb{F}_q$ para indicar um Corpo Finito de ordem q, onde q possua característica p, para algum p primo, e $k \in \mathbb{N}$, tal que $q = p^k$. Quando utilizarmos n estaremos tratando sobre o número de variáveis do sistema de equações, e v definirá quantas destas variáveis são do tipo vinagre. Por sua vez, quando empregarmos m estaremos indicando a quantidade de equações utilizadas em nosso sistema, sendo que r denotará quantas destas equações foram removidas, quando for o caso. E ainda, por definição, temos que h representará o grau da extensão do Copo Finito $\mathbb{F}$, ou seja $h \stackrel{\text{def}}{=} n - v$ e $\mathbb{E} = \mathbb{F}_{q^h}$.

TABELA I. DEFINIÇÕES, NOTAÇÃO E TERMINOLOGIA PERTINENTES AO QUARTZ.

Símbolo	Significado
$[\lambda]_{p\to q}$	Dada uma cadeia de bits $\lambda = (\lambda_0, \dots, \lambda_t)$ e dois inteiros p e q tais que $0 \le p \le q \le t$, temos que $[\lambda]_{p\to q} = (\lambda_p, \lambda_{p+1}, \dots, \lambda_{q-1}, \lambda_q)$
$\|\|$	Concatenação
$\lambda\|\|\mu$	Se $\lambda = (\lambda_0, \dots, \lambda_t)$ e $\mu = (\mu_0, \dots, \mu_t)$ são duas cadeias de bits, então $\lambda\|\|\mu = (\lambda_0, \dots, \lambda_t, \mu_0, \dots, \mu_t)$
$p(x_1, \dots, x_n)$	Um polinômio de grau d com n variáveis sobre $\mathbb{F}$
$\mathcal{P} = (p_1, \dots, p_m)$	Um sistema de m polinômios de grau d com n variáveis sobre $\mathbb{F}$

III. CRIPTOSSISTEMAS DE CHAVE PÚBLICA MULTIVARIÁVEL

Uma das motivações (não a única) para a nossa pesquisa é o risco do comprometimento dos atuais sistemas criptográficos de chave pública, que basicamente são fundamentados no problema da fatoração de inteiros e do logaritmo discreto, no caso de computadores quânticos com capacidade de processamento adequada serem desenvolvidos. Dentre as classes de criptossistemas pós-quânticos, os MPKCs se destacam por, principalmente, possibilitar a criação de esquemas de assinatura digital com tamanho de assinatura reduzida [13], que é um objetivo implícito de nosso trabalho.

A. O problema $\boldsymbol{\mathcal{MQ}}$

Seja $\mathbb{F}$ um Corpo Finito de ordem q, $n \in \mathbb{N}$ o número de variáveis, $m \in \mathbb{N}$ o número de equações, $\mathcal{P} = (p_1, \dots, p_m)$ um sistema de m polinômios com grau d e n variáveis sobre $\mathbb{F}$. Temos que o Sistema de Equações Polinomiais Multivariáveis Simultâneas consiste em encontrar $x = (x_1, \dots, x_n) \in \mathbb{F}^n$ tal que $\mathcal{P}(x) = y$, sendo $y = (y_1, \dots, y_m) \in \mathbb{F}^m$, onde y é um vetor de dimensão m:

$$\mathcal{P} = \begin{cases} p_1(x_1, \dots, x_n) = y_1 \\ p_2(x_1, \dots, x_n) = y_2 \\ \vdots \\ p_m(x_1, \dots, x_n) = y_m \end{cases}$$

Assim, quando o grau de $\mathcal{P}$ é maior ou igual a 2, ou seja $d \ge 2$, chamamos, então, este sistema de equações polinomiais de Problema $\boldsymbol{\mathcal{MQ}}$.

O problema $\boldsymbol{\mathcal{MQ}}$ baseia-se no trabalho apresentado por Fraenkel e Yesha em 1979, onde os autores provaram que solucionar sistemas de equações polinomiais multivariáveis sobre $GF(2)$ é NP-difícil [27], sendo este resultado popularizado pelo livro de Garey e Johnson [28]. Entretanto, a demonstração de segurança desta primitiva sendo utilizada na

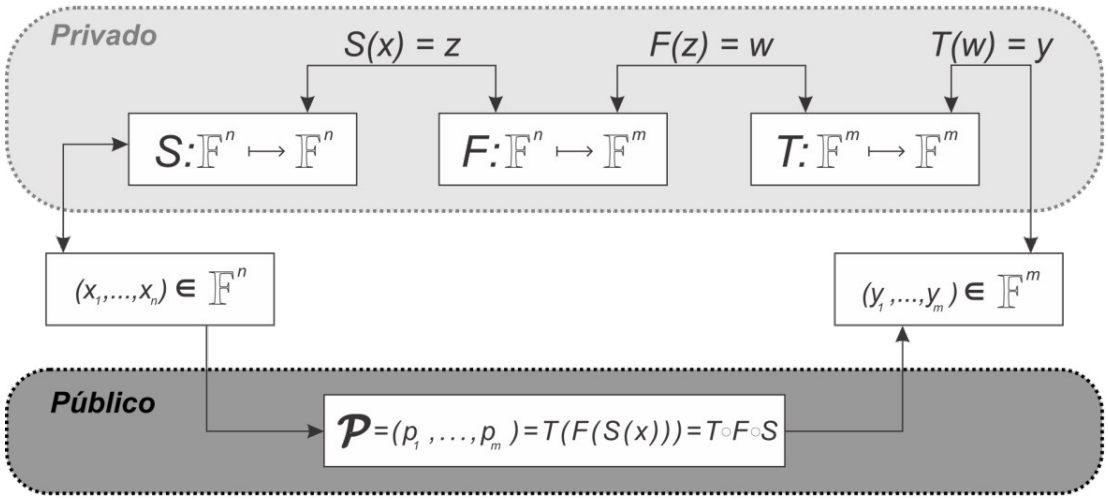


Figura 1. Modelo genérico de MPKC.

concepção de criptossistemas só foi apresentada em 1997. Neste novo trabalho, Patarin e Goubin demonstraram que a *trapdoor* criada a partir do problema $\mathcal{MQ}$ é NP-completo [45], não se conhecendo até hoje nenhum algoritmo, nem mesmo quântico, de tempo polinomial que possa resolver este problema [3], [39].

Precisamos destacar ainda, que ao estabelecer os parâmetros para instanciar uma função $\mathcal{MQ}$, devemos ter $n > m$ e $n \approx m$ para que o problema permaneça intratável. Isto porque, caso $n < m$, teríamos um sistema de equações superdefinido, onde existiriam mais equações do que variáveis, o que tornaria a sua resolução "fácil"' [10]. Por outro lado, caso $n \gg m$, teríamos um sistema de equações com muito mais variáveis do que equações, o que possibilitaria ao adversário utilizar o algoritmo proposto por Courtois *et al.* em 2002; sendo este algoritmo muito mais veloz do que a busca exaustiva [9].

B. Visão geral dos MPKCs

Sabemos que uma função construída a partir do problema $\mathcal{MQ}$ é uma função unidirecional [45], porém, para a construção de um criptossistema de chave pública é necessário que esta função seja também uma função alçapão (*trapdoor*). No entanto, Ding e Yang enfatizam, de forma simples e direta, que não é possível obter uma função deste tipo a partir de uma instância aleatória de função $\mathcal{MQ}$ [3]. Por isso, nos atuais MPKCs, a chave pública $\mathcal{P}$ é criada a partir da composição de duas transformações $S: \mathbb{F}^n \mapsto \mathbb{F}^n$, $T: \mathbb{F}^m \mapsto \mathbb{F}^m$ e um mapeamento central $F: \mathbb{F}^n \mapsto \mathbb{F}^m$, ou seja, $\mathcal{P} = T \circ F \circ S$, onde T, F e S são as chaves privadas [55]. Sendo este o modelo conceitual genérico adotado pelos criptossistemas MPKC, o qual pode ser vislumbrado na Figura 1.

O fato de $\mathcal{P} = T \circ F \circ S$, ou seja, $\mathcal{P}$ ser uma composição de outras funções, faz com que os atuais Criptossistemas de Chave Pública Multivariável não dependam exclusivamente do problema $\mathcal{MQ}$. Esta interdependência entre o problema $\mathcal{MQ}$ e IP (Isomorfismo de Polinômios) gera certa controvérsia, uma vez que mesmo acreditando-se que o problema IP seja difícil, nenhuma prova da intratabilidade deste problema foi formulada até o momento.

Contudo, ressaltamos que o modelo aqui exposto é apenas uma generalização das diversas $\mathcal{MQ}$-*trapdoors* existentes na atualidade, onde a principal diferença entre elas está, principalmente, no formato do mapeamento central F. Sendo que também existem algumas *trapdoors*, como é o caso da UOV (*Unbalanced Oil and Vinegar*), que utilizam apenas uma transformação afim para criar a chave pública $\mathcal{P}$, de modo que $\mathcal{P} = F \circ S$ [34]. Característica que utilizamos para propor o modelo Aprimorado do Quartz (veja Seção VII), pois apesar de fugir ligeiramente do modelo genérico, este tipo de composição é amplamente aceito, já que uma segunda transformação afim não adiciona segurança alguma ao sistema criptográfico [3], [36], uma vez que suas operações são apenas lineares; e a não realização desta transformação melhora a performance do processo de assinatura e geração de chaves.

C. Modificadores Genéricos

Com o passar dos anos, e também devido o aumento dos estudos acerca de MPKC, foi constatado que versões básicas das $\mathcal{MQ}$-*trapdoors* existentes na atualidade são inseguras [3],[55]. Contudo, para que todos estes criptossistemas não fossem descartados, modificadores genéricos (que, a grosso modo, são blocos construtores que alteram algumas estruturas das funções alçapão básicas) foram desenvolvidos para serem aplicados (pelo menos na teoria) em todas estas funções $\mathcal{MQ}$ básicas.

TABELA II. PRINCIPAIS MODIFICADORES GENÉRICOS E ALGUMAS DE SUAS CARACTERÍSTICAS.

Símbolo	Nome	Segurança	Ideia básica	Perda
-	Menos	seguro	descarta alguns polinômios	encriptação mais lenta
+	Mais	maioria sem efeito	adiciona polinômios	assinatura mais lenta
v	Vinagre	pouco mais seguro	variáveis extras são definidas	encriptação mais lenta
p	Pré-fixo ou Pós-fixo	em aberto	força algum $p_l = 0$	assinatura mais lenta
i	Pertubação Interna	em aberto	equivalente a p + v	tudo mais lento
f	Fixador	em aberto	usa algumas variáveis aleatórias	-
m	Mascaramento	em aberto	descarta algumas variáveis	-
s	Esparso	em aberto	usa polinômios esparsos	*speedup* mais lento

A TABELA II expressa de maneira sucinta – sem querer abordar todos os aspectos pertinentes a este tema –, os principais modificadores existentes, seguidos de suas características mais marcantes.

Vale ressaltar, também, que alguns destes modificadores mostraram-se mais eficientes para alguns esquemas do que para outros [55]. Além disto, assim como ocorrera nas versões básicas das $\mathcal{MQ}$-*trapdoors*, alguns modificadores genéricos como: Ramificação ($\perp$), Sub-Corpo ($/$) e Homogeneização(h), também foram considerados inseguros ou sem efeito [7], [24], [29], [41], [42], [54].

Desta forma, podemos pressupor que no momento da melhoria ou desenvolvimento de um criptossistema MPKC, há de se considerar a possibilidade de inserção do modificador genérico. Ponderando sobre qual é o ideal para aquele tipo de $\mathcal{MQ}$-*trapdoors*, considerando suas peculiaridades e contribuições para segurança. Mesmo não sendo esta uma tarefa trivial, que pode, inclusive, gerar vulnerabilidades em vez de melhorias.

IV. REVISANDO O QUARTZ

Nesta seção revisaremos o protocolo de assinatura digital Quartz, apresentado no NESSIE em 2001. Inicialmente, explanaremos sobre os aspectos gerais do HFE e exibiremos como é formada sua função de mapeamento central. Em seguida, explicaremos como funciona o algoritmo Quartz Original.

D. HFE (Hidden Field Equation)

Após ter quebrado a primeira *trapdoor* baseada em sistemas de equações multivariadas que se mostrou viável para os computadores da época, o MIA (*Matsumoto Imai Scheme A*) [41], Patarin – fundamentado nas ideias desta *trapdoor* considerada insegura – desenvolveu uma nova função alçapão denominada *Hidden Field Equations*, ou simplesmente HFE [42]. Esta nova *trapdoor* é uma generalização que modifica a função de mapeamento central F do MIA. Tal generalização tem como principal característica a troca dos monômios, empregados na *trapdoor* quebrada, por polinômios. Porém, apesar desta troca de monômios por polinômios, é mantido o conceito de utilizar uma extensão do corpo $\mathbb{F}_q$, comumente denotado por $\mathbb{E}$, tal que $\mathbb{E} = \mathbb{F}_{q^n}$ (onde q possua característica p, para algum p primo, e $k \in \mathbb{N}$, tal que $q = p^k$), juntamente com o corpo $\mathbb{F}_q$.

Deste modo, sejam: i e j números naturais; ξ_{ij}, ψ_i e μ elementos de $\mathbb{E}$; e θ_{ij}, σ_{ij} e γ_i números inteiros; a função do mapeamento central fica definida como [43]:

$$f(x) = \sum_{i,j}^{d} \xi_{ij} x^{q^{\theta_{ij}} + q^{\sigma_{ij}}} + \sum_{i}^{d} \psi_i x^{q^{\gamma_i}} + \mu \qquad (1)$$

$$onde \begin{cases} \xi_{ij} x^{q^{\theta_{ij}} + q^{\sigma_{ij}}} & são\ os\ termos\ quadráticos, \\ \psi_i x^{q^{\gamma_i}} & são\ os\ termos\ lineares, e \\ \mu & são\ os\ termos\ constantes \end{cases}$$

tal que $f(x)$ seja um polinômio em x sobre $\mathbb{E}_{q^n}$ com grau d, para $0 \leq \theta_{ij}, \sigma_{ij}, \gamma_i \leq d$.

Como $\mathbb{E}$ e $\mathbb{F}$ são isomórficos, podemos representar os elementos de $\mathbb{E} = \mathbb{F}_{q^n}$ numa n-tupla sobre $\mathbb{F}_q$, e a função (1) pode ser representada por polinômios com n variáveis $x_1, x_2, \ldots, x_n$ também sobre $\mathbb{F}_q$ [42].

E. Quartz Original

Como vimos anteriormente, o Quartz é um esquema de assinatura digital baseado no HFEv-. Neste esquema de assinatura, como os próprios modificadores genéricos já sugerem, algumas variáveis extra (chamadas de "variáveis vinagre") são adicionadas, e também, alguns "polinômios de perturbação" são inseridos no local dos polinômios removidos (em algumas fontes estes polinômios são chamados de polinômios secretos). Além disto, Patarin *et al.* destacam que os parâmetros escolhidos para o Quartz são cuidadosamente selecionados para melhorar sua segurança e impedir o funcionamento dos principais ataques conhecidos [44].

1) Parâmetros: Na versão original do Quartz temos definido que: $h = 103$, assim, a extensão do corpo utilizada pelo Quartz fica definida como $\mathbb{F}_{2^{103}} = \mathbb{E}$, mais precisamente, $\mathbb{E} = \mathbb{F}_2[X]/(X^{103} + X^9 + 1)$; $q = 2$; $d = 129$; $v = 4$; $r = 3$; $n = 107$ (pois $n \stackrel{\text{def}}{=} h + v$); $m = 100$ (pois $m \stackrel{\text{def}}{=} h - r$) [13], [14], [44]; e a função pública $\mathcal{P}$ – função *trapdoor* – é um mapeamento de 107 bits para 100 bits, ou seja $\mathbb{F}^{107} \mapsto \mathbb{F}^{100}$ [44].

2) Assinando Mensagens: Seja M uma mensagem representada por uma cadeia de bits, e S a assinatura obtida desta mensagem. Então, os procedimentos necessários à

obtenção de S devem ser realizados conforme segue:

1. Sejam M_0, M_1, M_2 e M_3 quatro cadeias de 160 bits definidas por: $M_0 = SHA\_1(M)$, $M_1 = SHA\_1(M_0||0)$, $M_2 = SHA\_1(M_0||1)$, $M_3 = SHA\_1(M_0||2)$.
2. Sejam H_1, H_2, H_3 e H_4 quatro cadeias de 100 bits definidas por: $H_1 = [M_1]_{0\to99}$, $H_2 = [M_1]_{100\to159}||[M_2]_{0\to39}$, $H_3 = [M_2]_{40\to139}$, $H_4 = [M_2]_{140\to159}||[M_3]_{0\to79}$.
3. Seja $\tilde{S}$ uma cadeia de 100 bits, tal que $\tilde{S}$ seja inicializada com 00 ... 0.
4. Para $i = 1$ até 4, faça:
   a. Calcule a cadeia de 100 bits $Y = H_i \oplus \tilde{S}$.
   b. Calcule a cadeia de 160 bits $W = SHA\_1(Y||\Delta)$.
   c. Obtenha a cadeia de 3 bits $R = [W]_{0\to2}$.
   d. Obtenha a cadeia de 4 bits $V = [W]_{3\to6}$.
   e. Calcule $B = \varphi(t^{-1}(Y||R))$.
   f. Considerando $F_V(Z) = B$ em Z sobre $\mathbb{E}$:
      i. Se a equação $F_V(Z) = B$ não tiver solução, troque W por $SHA\_1(W)$ e retornar ao passo 4c.
      ii. Neste passo a equação $F_V(Z) = B$ tem uma ou mais soluções em $\mathbb{E}$. Logo, temos que $A(1), A(2), \dots, A(\delta)$ são as soluções de $F_V(Z) = B$.
      iii. Se $F_V(Z) = B$ tiver apenas uma solução, defina $A = A(1)$. Caso contrário, aplique a função hash em cada uma das soluções, ou seja $I(j) = SHA\_1(A(j))$. Em seguida escolha o $A(j)$ que resulta no menor $I(j)$, considerando a ordenação *big-endian*.
   g. Calcule a cadeia de 107 bits $X = s^{-1}(\varphi^{-1}(A)||V)$.
   h. Defina um novo valor para $\tilde{S}$ como sendo $\tilde{S} = [X]_{0\to99}$.
   i. Obtenha a cadeia de 7 bits X_i definida por $X_i = [X]_{100\to106}$.
5. A assinatura S é a cadeia de 128 bits definida por $S = \tilde{S}||X_4||X_3||X_2||X_1$.

3) Verificando Assinatura: Dadas uma mensagem M, representada por uma cadeia de bits, e uma assinatura S, que neste caso é uma cadeia de 128 bits. Então, os procedimentos que seguem devem ser realizados para verificar se S é ou não uma assinatura válida para M.

1. Sejam M_0, M_1, M_2 e M_3 quatro cadeias de 160 bits definidas por: $M_0 = SHA\_1(M)$, $M_1 = SHA\_1(M_0||0)$, $M_2 = SHA\_1(M_0||1)$, $M_3 = SHA\_1(M_0||2)$.
2. Sejam H_1, H_2, H_3 e H_4 quatro cadeias de 100 bits definidas por: $H_1 = [M_1]_{0\to99}$, $H_2 = [M_1]_{100\to159}||[M_2]_{0\to39}$, $H_3 = [M_2]_{40\to139}$, $H_4 = [M_2]_{140\to159}||[M_3]_{0\to79}$.
3. Seja $\tilde{S}$ uma cadeia de 100 bits definida por $[S]_{0\to99}$.
4. Sejam X_4, X_3, X_2 e X_1 quatro cadeias de 7 bits definidas por: $X_4 = [S]_{100\to106}$, $X_3 = [S]_{107\to113}$, $X_4 = [S]_{114\to120}$, $X_4 = [S]_{121\to127}$.
5. Seja U uma cadeia de 100 bits, tal que U seja inicializada com $\tilde{S}$
6. Para $i = 4$ até 1, faça:
   a. Calcule a cadeia de 100 bits Y definida por $Y = G(U||X_i)$.
   b. Defina um novo valor para a cadeia de 100 bits U como sendo $U = Y \oplus H_i$.
7. Se U é igual à cadeia 00 ... 0, aceite a assinatura. Caso contrário, rejeite-a.

V. SOBRE A SEGURANÇA DO QUARTZ (HFEv-)

Existem diversos ataques capazes de recuperar a chave privada em criptossistemas desenvolvidos a partir da versão básica do HFE (ou seja, a *trapdoor* HFE sem modificadores aplicados a ela) [4], [11], [17], [19], [21], [22], [30], [32], [35]. Porém, até o momento, poucas criptoanálises foram capazes de sobrepujar a segurança adicionada pelos modificadores genéricos.

Ironicamente, uma das primeiras criptoanálises do HFE com modificadores que fora publicada afirmava decrementar a segurança desta *trapdoor* com, justamente, os modificadores "v" (vinagre), "-" (menos), ou os dois aplicados simultaneamente; além de também atacar sua versão básica [22]. Este trabalho foi desenvolvido por Faugère e Joux em 2003, e nele os autores afirmavam ser possível recuperar a chave privada do Quartz com um esforço muito menor do que 2^{80} triplo-DES através de um ataque algébrico viabilizado pelo algoritmo F5 [22], desenvolvido anteriormente por Faugère (frisando que o algoritmo F5 é baseado na teoria de bases de Gröbner [8]). Todavia, algum tempo após o desenvolvimento deste ataque, Courtois publicou uma versão estendida de [12] onde ele mostra que o ataque desenvolvido por Faugère e Joux possui informações imprecisas ou enganosas, que acabam tornando sua criptoanálise inválida para o HFEv- e Quartz [13]. Sobre este ataque, ainda vale ressaltar, que Wolf (em sua tese de doutorado) alertara sobre os argumentos utilizados por Faugère e Joux, afirmando que era necessário ter cautela antes de adotar tal ataque, uma vez ele ainda não havia sido analisado por outros pesquisadores [54].

Acreditamos que os estudos realizados por Courtois e Wolf ocorreram de maneira independente e em paralelo. Sendo assim, a suspeita de Wolf juntamente com as explicações de Courtois são suficientes para conjecturarmos que o Quartz (HFEv-) é resistente ao ataque algébrico que visa a recuperação da chave privada, desenvolvido por Faugère e Joux.

Posteriormente, visando o HFE com modificadores, alguns outros ataques capazes de recuperar a chave privada foram criados. Porém nenhum que atingisse o Quartz [3], [16], [36].

Lembramos que tentar recuperar a chave privada não é o único objetivo de um adversário que ataca esquemas de assinatura digital. Uma abordagem possível é tentar gerar uma assinatura válida (porém falsificada) sem poder determinar ou modificar a mensagem cuja assinatura foi forjada. Este é o nível de sucesso mais baixo para um adversário, porém, se alcançado, é suficiente para considerarmos um esquema de

assinatura inseguro.

Seguindo esta abordagem, Joux e Martinet – baseados em axiomas do Ataque pelo Paradoxo de Aniversário – provaram que o Quartz é maleável, demonstrando que caso o adversário possua um par (mensagem, assinatura) válido, ele conseguirá obter uma segunda assinatura com $2^{m/2}$ computações e $2^{m/2}$ chamadas ao oráculo de assinatura, com um método que consiste em encontrar a segunda pré-imagem sem se preocupar com a inversão da função pública G [33]. Assim, sob este cenário, estima-se que um adversário do Quartz consegue forjar uma segunda assinatura com somente 2^{50} computações e 2^{50} chamadas ao oráculo aleatório, logo muito inferior aos padrões segurança atuais que são de, no mínimo, 2^{112} [1].

Apesar disto, recentemente Sakumoto *et al.* apresentaram um trabalho sugerindo que os modelos usuais de prova de segurança não devem ser diretamente aplicados a *trapdoors* como o HFE e UOV [48]. Além disto, os autores desenvolveram um novo modelo de prova de segurança para esquemas de assinatura digital baseado em HFE e UOV, frisando que nenhum modelo deste tipo existia até o momento da publicação desse artigo. Neste novo modelo, juntamente com pequenas modificações também propostas neste trabalho, as assinaturas passam a ser uniformemente distribuídas [48], e com esta nova distribuição, Sakumoto *et al.* afirmam que os esquemas de assinatura baseados em HFE e UOV podem atingir um nível de segurança resistente a falsificação existencial através de um ataque adaptativo de mensagem escolhida, e fornecem um teorema para calcular a segurança provável destes esquemas modificados [48].

Assim, fundamentados nas criptoanálises aqui citadas e principalmente na nova prova de segurança e modificações desenvolvidas pro Sakumoto *et al.* em 2011, seguiremos propondo nosso aprimoramento para o Quartz nas seções subsequentes.

VI. A QUESTÃO DO SHA-1

Sabemos que algoritmos de assinatura e verificação são mais rápidos quando aplicados sobre o resultado de uma função hash (y) do que quando aplicado diretamente sobre sua entrada (x), isto porque y é relativamente muito mais curto que x [50]. Motivado por este atributo das funções hash, o Quartz emprega o SHA-1 em diversos pontos do seu algoritmo de assinatura e, consequentemente, no de verificação também. Mais precisamente, o uso do SHA-1 pode ser vislumbrado nos passos 1, 4.b, 4.f.i e 4.f.iii do algoritmo de assinatura, e no passo 1 do algoritmo de verificação.

Todavia, a cada ano que passa a resistência a colisões do SHA-1 tem sido consideravelmente reduzida. Por exemplo, em 2005, Wang *et al.* publicaram um algoritmo para colisões do SHA-1 reduzido a 58 iterações com a complexidade 2^{33}, sendo este o principal marco no declínio do SHA-1 [51].

Além disto, Joux publicou em 2004 um trabalho averiguando a segurança em funções hash iteradas (situação presente no passo 1 dos algoritmos de assinatura e verificação). Neste trabalho o autor demonstra que concatenando-se os resultados de funções hash a resistência a colisões é de apenas $\mathcal{O}(n2^{n/2})$ e não $\mathcal{O}(2^n)$ como esperava-se.

Desta forma, acreditamos ser latente a necessidade de atualização do Quartz (e qualquer outro esquema de assinatura digital que utilize o SHA-1) quanto ao emprego desta função de hash em suas rotinas.

VII. QUARTZ APRIMORADO

Agora que já estudamos os detalhes pertinentes ao Quartz, podemos iniciar a explicação acerca do nosso modelo aprimorado. Iniciaremos esta seção abordando quais foram os parâmetros escolhidos para o Quartz Aprimorado, pois a escolha de parâmetros além de definir o *trade-off* entre segurança e performance também garante a resistência contra as criptoanálises conhecidas, quando feita corretamente.

Em seguida, descreveremos os algoritmos de Geração de Chaves, Assinatura e Verificação. Nestes algoritmos será possível notar três grandes mudanças. A primeira delas está no fato de utilizarmos somente uma transformação afim no processo de assinatura das mensagens. Outra mudança está na substituição do SHA-1 pelo SHA-3 no momento de inicializar os vetores de bits que serão utilizados no processo de assinatura e verificação de assinaturas, bem como no momento de gerar as variáveis R e V do algoritmo de assinatura. A terceira grande mudança esta no fato de concatenarmos um *salt* Γ à mensagem M antes de empregarmos a função de hash nesta mensagem.

A. Parâmetros

Em nossa versão aprimorada do Quartz temos definido que: $h = 229$, assim, a extensão do corpo utilizada pelo Quartz Aprimorado fica definida como $\mathbb{F}_{2^{229}} = \mathbb{E}$, mais precisamente, $\mathbb{E} = \mathbb{F}_2[X]/(X^{229} + X^9 + X^6 + X^5 + X^2 + X + 1)$; $q = 2$; $d = 129$; $v = 2$; $r = 5$; $n = 231$ (pois $n \stackrel{\text{def}}{=} h + v$); $m = 224$ (pois $m \stackrel{\text{def}}{=} h - r$); e a função pública $\mathcal{P}$ – função *trapdoor* – é um mapeamento de 231 bits para 224 bits, ou seja $\mathbb{F}^{231} \mapsto \mathbb{F}^{224}$.

Temos definido, ainda, um parâmetro adicional g, onde g expressa o tamanho do *salt* aleatório Γ que será concatenado à mensagem antes dela servir como entrada para a função hash, ou seja, $g = |\Gamma|$. Lembramos que Sakumoto *et al.*, em seu novo modelo de prova, propuseram a utilização deste *salt* aleatório para uniformizar as assinaturas em esquemas de assinatura digital baseados no HFE [48], sendo estimado um tamanho aproximado de $\log(q_{assina}(q_{hash} + q_{assina}))$ bits para que o esquema de assinatura seja considerado seguro [48]. Sabemos que q_{hash} e q_{assina} correspondem, respectivamente, à quantidade de consultas aos oráculos de hash e assinatura; e que em provas de esquemas de assinatura digital normalmente são considerados $q_{assina} = 2^{30}$ e $q_{hash} = 2^{60}$ [2]. Assim, segue que $g = 96$.

B. Assinando Mensagens

Seja M uma mensagem representada por uma cadeia de bits, e S a assinatura obtida desta mensagem. Então, em nosso esquema aprimorado, os procedimentos necessários a obtenção de S devem ser realizados conforme segue:

1. Seja Γ uma cadeia de 96 bits, tal que $\Gamma \in_R \{0,1\}^{96}$.
2. Seja M_0 uma cadeia de 512 bits definida por $M_0 = SHA\_3(M|\Gamma)$.
3. Sejam H_1 e H_2 duas cadeias de 224 bits definidas por: $H_1 = [M_0]_{0\to223}$, $H_2 = [M_0]_{224\to447}$.
4. Seja $\tilde{S}$ uma cadeia de 224 bits, tal que $\tilde{S}$ seja inicializada com 00 ... 0.
5. Para $i = 1$ até 2, faça:
   a. Calcule a cadeia de 224 bits $Y = H_i \oplus \tilde{S}$.
   b. Calcule a cadeia de 512 bits $W = SHA\_3(Y||\Delta)$.
   c. Obtenha a cadeia de 5 bits $R = [W]_{0\to4}$.
   d. Obtenha a cadeia de 2 bits $V = [W]_{5\to6}$.
   e. Considerando $F_V(Z) = B$ em Z sobre $\mathbb{E}$:
      i. Se a equação $F_V(Z) = B$ não tiver solução, troque W por $SHA\_3(W)$ e retornar ao passo 5c.
      ii. Neste passo a equação $F_V(Z) = B$ tem uma ou mais soluções em $\mathbb{E}$. Logo, temos que $A(1), A(2), \dots, A(\delta)$ são as soluções de $F_V(Z) = B$.
      iii. Se $F_V(Z) = B$ tiver apenas uma solução, defina $A = A(1)$. Caso contrário, aplique a função hash em cada uma das soluções, ou seja $I(j) = SHA\_3(A(j))$. Em seguida escolha o $A(j)$ que resulta no menor $I(j)$, considerando a ordenação *big-endian*.
   f. Calcule a cadeia de 231 bits $X = s^{-1}(\varphi^{-1}(A)||V)$.
   g. Defina um novo valor para $\tilde{S}$ como sendo $\tilde{S} = [X]_{0\to223}$.
   h. Obtenha a cadeia de 7 bits X_i definida por $X_i = [X]_{224\to230}$.
6. A assinatura S é a cadeia de 334 bits definida por $S = \tilde{S}||X_2||X_1||\Gamma$.

C. Assinando Mensagens

Dadas uma mensagem M, representada por uma cadeia de bits, e uma assinatura S, que neste caso é uma cadeia de 334 bits. Então, no Quartz Aprimorado, os procedimentos que seguem devem ser realizados para verificar se S é ou não uma assinatura válida para M.

1. Seja $\tilde{S}$ uma cadeia de 224 bits definida por $\tilde{S} = [S]_{0\to223}$.
2. Sejam X_2 e X_1 duas cadeias de 7 bits definidas por: $X_2 = [S]_{224\to230}$, $X_1 = [S]_{231\to237}$.
3. Seja Γ uma cadeia de 96 bits definida por $\Gamma = [S]_{238\to334}$.
4. Seja M_0 uma cadeia de 512 bits definida por $M_0 = SHA\_3(M|\Gamma)$.
5. Sejam H_1 e H_2 duas cadeias de 224 bits definidas por: $H_1 = [M_0]_{0\to223}$, $H_2 = [M_0]_{224\to447}$.
6. Seja U uma cadeia de 224 bits, tal que U seja inicializada com $\tilde{S}$.
7. Para $i = 2$ até 1, faça:
   a. Calcule a cadeia de 224 bits Y definida por $Y = G(U||X_i)$.
   b. Defina um novo valor para a cadeia de 224 bits U como sendo $U = Y \oplus H_i$.
8. Se U é igual à cadeia 00 ... 0, aceite a assinatura. Caso contrário, rejeite-a.

VIII. ANÁLISE DA PROPOSTA

Para a análise de nosso protocolo Quartz Aprimorado vamos inicialmente explicar porque foram escolhidos os parâmetros elencados na Seção VII-A, para isto, descreveremos os benefícios de adotar tais parâmetros, justificando o impacto que tal escolha gera na segurança, abordando também a possível perda ou ganho de eficiência que tal modificação pode gerar, quando comparada ao esquema original. Explanaremos também, sobre a substituição da função de hash SHA-1 (utilizada no modelo original) pela função de hash SHA-3 (escolhida para nosso aprimoramento). E ainda, justificaremos porque é interessante adotarmos as modificações propostas por Sakumoto *et al.* em seu novo modelo de prova para esquemas de assinaturas baseados em HFE [48], mesmo sabendo que tais modificações acarretam um aumento de 96 bits no tamanho final da assinatura.

Como acreditamos ser dedutível que uma segunda rodada de operações lineares não adicione segurança alguma a um esquema baseado na intratabilidade de equações multivariáveis quadráticas, não nos preocuparemos em dar maiores detalhes sobre a não utilização de duas transformações afim em nosso aprimoramento. Sendo que o leitor pode consultar maiores detalhes sobre o tema em [3], [6], [16] e [34].

A. Escolha de Parâmetros

Sabemos que no Quartz, caso o adversário possua um par (mensagem, assinatura) válido, é possível que este adversário obtenha uma segunda assinatura válida com $2^{m/2}$ computações e $2^{m/}$ chamadas ao oráculo aleatório [33]. Como nosso aprimoramento não modifica a estrutura geral do Quartz, apenas a adapta; podemos deduzir que tal ataque também é válido para ele. Desta forma, temos que $m \geq 224$ para obtermos um nível de segurança de no mínimo 2^{112} (padrão mínimo exigido para sistemas criptográficos atuais [1]).

Todavia, ao definirmos nossos parâmetros não nos preocupamos somente com o tamanho de m. Isto porque ao estabelecermos parâmetros para instanciar uma função $\boldsymbol{\mathcal{MQ}}$, devemos ter $n > m$ e $n \approx m$ para que o problema permaneça intratável. Também existe o fato de Ding e Schmidt em 2005 terem demonstrado ser possível quebrar o HFEv quando $v = 1$ [18], assim, tomamos o cuidado de escolher um $v > 1$, porém não excessivamente maior. Além disto, escolhemos cuidadosamente o valor de h para que o mesmo fosse primo (fato que também ocorre no modelo original [14], [44]); isto porque até hoje não foi apresentada nenhuma criptoanálise que atinja MPKCs que utilizem extensões de corpos com característica igual a um número primo [13], [23], [36], [54].

Desta forma, lembramos que os parâmetros de nosso

modelo aprimorado são: $m = 224$, $n = 231$, $h = 229$, $v = 2$, $r = 5$, $q = 2$, $d = 129$ e $g = 96$.

Com estes parâmetros, constatamos dois inconvenientes em nosso aprimoramento. O primeiro deles está no aumento das chaves de nosso criptossistema, pois a chave privada aumenta de 3 Kbytes no Quartz Original para 8 Kbytes em nosso protocolo, sendo que a chave pública salta de 71 Kbytes para 739 Kbytes. O outro inconveniente de nossa escolha de parâmetros está na perda de eficiência dos algoritmos de Geração de Chaves e Assinatura (perda que pode ser constatada nas Tabelas IV e V da Seção XI). Em linhas gerais, acreditamos que tal ineficiência se deu devido ao aumento significativo na quantidade de objetos e instâncias a serem manipuladas por nossa implementação, sendo que melhorias como paralelismo ou instruções de máquina possam facilmente ser incorporadas a futuras implementações que visem à melhoria deste quesito.

Apesar desta perda de eficiência ocorrer, acreditamos que ela não seja tão grave, uma vez que a geração é efetuada apenas uma vez para cada usuário, dentro de um prazo longo de validade das chaves. Além disto, em determinados cenários, a ineficiência do processo de assinatura não é grave se supormos que o ato de assinar é realizado apenas uma vez para cada mensagem (ou cada Certificado Digital), enquanto a verificação é efetuada por muitos receptores da mensagem (ou do Certificado Digital).

Contudo, aliado ao expressivo ganho no nível de segurança (para maiores detalhes consulte a seção IX), a escolha de parâmetros do nosso esquema aprimorado proporcionou, também, uma melhoria no algoritmo de Verificação de Assinatura. Isto porque testaremos até 4.096 vezes menos hipóteses de utilização da chave pública no momento da resolução da função G, enquanto estamos verificando a validade de uma assinatura. Para sermos mais específicos, consideremos as seguintes características: (a) o mapeamento central F_V de $\mathbb{E}$ para $\mathbb{E}$ utilizado tanto na assinatura de mensagens quanto na composição da chave pública terá seu número de possibilidades para V variando conforme a quantidade de variáveis vinagre utilizadas pelo criptossistema, pois $(F_V)_{V\in\{0,1\}^v}$; (b) estas possibilidades representam um total de 2^v chaves públicas a serem geradas pelo algoritmo de Geração de Chaves; (c) como o Quartz (tanto o original quanto nosso modelo aprimorado) realiza operações iteradas para assinar as mensagens, e antecipadamente não podemos definir quais das 2^v possibilidades de variáveis vinagres foram utilizadas no processo de assinatura, temos que, seja K o número de iterações do algoritmo, então todas as $(2^v)^K$ combinações devem ser testadas para que o algoritmo de verificação negue uma assinatura falsa. Deste modo $(2^4)^4 = 2^{16}$ possibilidades devem ser testadas durante a resolução de G no Quartz Original e $(2^2)^2 = 2^4$ no Quartz Aprimorado, o que representa um intervalo de possibilidades $2^{12} = 4.096$ vezes menor em nosso aprimoramento.

B. Substituição do SHA-1 pelo SHA-3

Sabemos que algoritmos de assinatura (Assina) e verificação (Verifica) são mais rápidos quando aplicados sobre o resultado de uma função hash (y) do que quando aplicado diretamente sobre sua entrada (x), isto porque y é relativamente mais curto que x [50]. Logo, caso a função hash $H()$ não seja resistente a colisões, um adversário poderia obter uma mesma assinatura para duas mensagens distintas. Ou seja, dado um par de legíveis x_1 e x_2, onde $x_1 \neq x_2$ acarrete $H(x_1) = H(x_2)$, então $\sigma = \text{Assina}_{sk}\big(H(x_1)\big) = \text{Assina}_{sk}\big(H(x_2)\big)$. Caso isto ocorra, teríamos ainda que $\text{Verifica}_{pk}\left(M, \text{Assina}_{sk}\big(H(x_1)\big)\right) = \text{Verifica}_{pk}\left(M, \text{Assina}_{sk}\big(H(x_2)\big)\right) = 1$, ou seja, o algoritmo de verificação aceitaria as assinaturas, ferindo, também, os princípios de autenticidade e irretratabilidade, essenciais a esquemas de assinaturas digital.

Para ilustrar o quão prejudicial pode ser utilizar uma função hash inadequada, pensemos em nosso aprimoramento. Suponha que em vez do SHA-3 de 512 bits (que tem sua segurança estimada em 2^{256} [31]) utilizássemos o SHA-1. Desta forma, um adversário de nosso criptossistema poderia forjar uma segunda assinatura com aproximadamente 2^{33} operações [51] sem atacar diretamente o nosso protocolo.

Portanto, acreditamos que substituir a função SHA-1 pelo SHA-3 seja imprescindível para mantermos nosso esquema aprimorado dentro da segurança estimada. Sendo que, conforme pode ser vislumbrado nas TABELA IV e TABELA V, a adesão ao SHA-3 além de ajudar na segurança do Quartz Aprimorado também proporciona um ganho de eficiência no momento de inicializarmos os vetores. Tal melhoria ocorre em virtude da não realização de operações iteradas e também por não concatenar suas saídas para obter vetores do tamanho estabelecido pelo algoritmo de assinatura.

C. Modificação proposta por Sakumoto et al. em 2011

Inicialmente, no artigo que submetia o Quartz ao NESSIE [14], [44], seus autores apenas justificavam a dificuldade de quebrar seu protocolo, não fornecendo nenhuma prova matemática devido a falta de modelo apropriado para demonstrar a segurança de criptossistemas baseados em equações multivariáveis quadráticas. Alguns anos depois, o primeiro modelo de prova para MPKCs foi formalizado [13], neste trabalho, Courtois apresentava um modelo de prova que estimava a segurança de criptossistemas baseados nesta primitiva quando colocados sob ataques em que o adversário possua apenas a chave pública da vítima [13]. Apesar de ser uma prova matemática, este modelo é considerado "fraco", visto que em cenários reais o adversário facilmente pode obter (através de interceptações, por exemplo) muito mais informações do que somente a chave pública da vítima.

Felizmente, há pouco tempo Sakumoto *et al.* desenvolveram um novo modelo de prova, demonstrando que criptossistemas baseados nas *trapdoors* HFE e UOV podem ser existencialmente seguros contra ataques adaptativos de mensagem escolhida [48]. Este novo modelo de prova é considerado "forte", pois dá amplos poderes ao adversário e exige o seu nível de sucesso mais baixo. No entanto, para utilizarmos este modelo na demonstração de segurança de um criptossistema baseado no HFE, como é o caso do Quartz

Aprimorado, necessitamos concatenar um *salt* a mensagem antes de empregarmos a função de hash nesta mesma mensagem afim de obtermos assinaturas uniformemente distribuídas [48]. Esta modificação acarreta um aumento no tamanho final da assinatura. Em nosso modelo aprimorado este aumento é de 96 bits, porém, mesmo com este aumento no tamanho final da assinatura, consideramos viável a adesão desta modificação já que nosso esquema aprimorado poderá ser provado como sendo "fortemente infalsificável" em vez de somente "infalsificável", como ocorre no Quartz Original.

IX. ESTIMATIVA DE SEGURANÇA

Nesta seção, iremos inicialmente olhar para a probabilidade de recuperar a chave privada do Quartz (HFEv-) através do melhor ataque conhecido na atualidade. Em seguida, como a modificação proposta por Sakumoto *et al.* foi aderida por nosso aprimoramento, buscaremos determinar a segurança exata de nosso criptossistema de acordo com o novo modelo de prova proposto pelos mesmos autores. Por fim, visando demonstrar que o ataque de Joux e Martinet é computacionalmente inviável em nosso aprimoramento, calcularemos quantas computações e quantas chamadas ao oráculo aleatório serão necessárias para que um adversário derive uma segunda assinatura caso ele possua um par (mensagem, assinatura) válido.

A. Melhor ataque ao Quartz (HFEv-)

Discutimos anteriormente na Seção V que até o momento não foi desenvolvido nenhum ataque ao Quartz capaz de recuperar a chave privada (inverter a função G) com um esforço menor do que o Ataque por Força Bruta, ou seja, a *trapdoor* HFEv- permanece segura [3], [16], [36].

Além disto, em 2010, utilizando Unidades de Processamento Gráfico (GPUs – *Graphics Processing Unit*) Bouillaguet *et al.* apresentaram um algoritmo para resolução de equações polinomiais em $\mathbb{F}_2$ que pode ser empregado para solucionar qualquer instância de problema $\boldsymbol{MQ}$ onde, evidentemente, $q = 2$. Neste algoritmo, em vez de utilizarem bases de Gröbner (principalmente devido a necessidade exponencial de memória para sua implementação) Bouillaguet *et al.* resolveram utilizar uma abordagem diferente, baseando-se no algoritmo padrão de Busca Exaustiva [5]. Com este novo método os autores deste trabalho demonstraram ser possível encontrar todos os zeros de um polinômio de grau d com n variáveis, em corpos $\mathbb{F}_2$, efetuando apenas $d.2^n$ operações binárias [5].

Como este é o algoritmo para resolução de equações polinomiais através da Busca Exaustiva mais rápido da atualidade, então, podemos conjecturar que um adversário conseguirá inverter a função pública G com uma probabilidade estimada em $\epsilon' \leq 1/(d.2^n)$. Desta forma, com os parâmetros adotados em nossa proposta de aprimoramento, temos que $\epsilon' \leq 1/(129.2^{231}) \approx 2^{-238}$.

B. Estimativa de segurança segundo Sakumoto et al.

No apêndice A de [48], onde consta a demonstração dos Teoremas formulados por Sakumoto *et al.*, os autores mostram que a probabilidade do algoritmo de inversão, simulado pelo oráculo aleatório, encontrar a inversa da assinatura S utilizando chamadas a este oráculo aleatório e a chave pública G é de aproximadamente $\epsilon(1 - (q_{hash} + q_{assina})q_{assina}2^{-g})/(q_{hash} + q_{assina} + 1)$. Onde q_{hash} e q_{assina} correspondem, respectivamente, a quantidade de consultas aos oráculos de hash e assinatura. E g corresponde ao tamanho do *salt* aleatório inserido na inicialização de vetores do algoritmo proposto, lembrando que g = $\log(q_{assina}(q_{hash} + q_{assina}))$-bits.

Como, em provas de esquemas de assinatura digital normalmente são considerados $q_{assina} = 2^{30}$ e $q_{hash} = 2^{60}$ [2]. Desta forma, obtemos $\epsilon = \epsilon'(q_{hash} + q_{assina} + 1)/(1 - (q_{hash} + q_{assina})q_{assina}2^{-g}) \approx \epsilon'.2^{60}$ [48].

Assim, utilizando este Teorema que compõe o novo modelo de prova de segurança proposto por Sakumoto *et al.* e o ϵ' calculado anteriormente, temos que a probabilidade de um adversário recuperar a chave privada do Quartz Aprimorado utilizando um oráculo aleatório é de $\epsilon \approx 2^{-178}$.

C. Estimativa de esforço para o ataque de Joux e Martinet

Joux e Martinet em 2003 desenvolveram um poderoso ataque ao Quartz; neste trabalho, os autores – baseados em axiomas do Ataque pelo Paradoxo de Aniversário – provaram que este criptossistema é maleável, demonstrando que caso o adversário possua um par (mensagem, assinatura) válido, ele conseguirá obter uma segunda assinatura com $2^{m/2}$ computações e $2^{m/2}$ chamadas ao oráculo de assinatura, com um método que consiste em encontrar a segunda pré-imagem sem se preocupar com a inversão da função pública G [33].

Desta forma, com os parâmetros adotados em nossa proposta de aprimoramento, temos que através deste ataque o adversário terá que realizar $2^{224/2} = 2^{112}$ computações e $2^{224/2} = 2^{112}$ chamadas ao oráculo de assinatura.

Portanto, para recuperar a chave privada do Quartz Aprimorado será necessário um esforço maior do que 2^{178}, e para derivar uma segunda assinatura através do ataque de Joux e Martinet o adversário terá que efetuar mais de 2^{112} operações e chamadas ao oráculo de assinatura.

X. COMPARANDO O QUARTZ APRIMORADO COM OUTROS PROTOCOLOS

Na TABELA III, listamos o tamanho das assinaturas de alguns outros esquemas baseados no problema $\boldsymbol{MQ}$. Além disto, para também haver uma comparação com esquemas de assinatura mais convencionais e já padronizados, incluímos o comprimento das assinaturas do ECDSA e do RSA. Frisamos que as informações listados na referida tabela tem como referência um nível de segurança de aproximadamente 2^{100}, sendo tal segurança calculada de acordo com suas fontes, que também encontram-se listadas na TABELA III.

TABELA III. TAMANHO DAS ASSINATURAS DE ALGUNS CRIPTOSSISTEMAS.

	Criptossistema	q	d	m	n	Tamanho da Assinatura (em bits)	Referência
Pós-Quântico	CyclicUOV	256	256	77	77	624	[46]
	Rainbow	16	30	58	58	352	[47]
	NC-Rainbow	256	17	26	26	672	[56]
	CyclicRainbow	256	17	26	26	344	[46]
	Quartz Aprimorado	2	129	224	231	334	Nosso
Quântico	ECDSA					400	[40]
	RSA					2048	[1]

Compreendemos que o tamanho da assinatura não é a única métrica a ser avaliada no momento da adesão de um esquema de assinatura digital. Porém, entendemos que este é um quesito importante para a economia de tráfego de rede, além de ser um ponto extremamente forte de nosso aprimoramento.

XI. IMPLEMENTAÇÃO DE REFERÊNCIA E TESTES

Nossa implementação tem como objetivo comparar o tempo gasto pelo Quartz Original e pelo Quartz Aprimorado durante a inicialização de vetores, geração de chaves, assinatura de mensagens, verificação de assinatura verdadeira e a verificação de uma assinatura falsa.

Para não realizarmos muito re-trabalho, baseamos nossa implementação no Projeto QuartzLight, feito em Java e que tem como autor Christopher Wolf [53]. Este projeto implementa uma versão mais "frágil" do Quartz [52], porém, mesmo com esta vulnerabilidade, é interessante para nós devido seus diversos métodos e classes que podem ser re-utilizados para fatorarmos polinômios sobre Corpos Finitos e executarmos operações sobre $GF(2^n)$ e matrizes binárias. Além disto, utilizamos implementações em Python das funções de hash SHA-1 [25] e SHA-3 [26], também para não temos uma carga de re-trabalho excessiva.

Lembrando que o código fonte desta implementação está disponível em nossa página pessoal (http://www.ime.usp.br/~ewe/QuartzAprimorado/) para que o mesmo possa ser baixado, alterado e/ou executado conforme a necessidade e interesse de cada usuário.

A. Testes realizados

Utilizando nossa implementação do Quartz em sua versão original e também em sua versão aprimorada – proposta por este trabalho –, realizamos os testes necessários para a coleta dos tempos de execução dos algoritmos de geração de chaves, inicialização de vetores, assinatura e verificação (tanto de assinaturas legítimas quanto de assinaturas falsas).

Para isso, utilizamos dois computadores distintos. Sendo eles:

Brucutu: processador Intel Xeon E5645 de 2,4 GHz × 24, com 128 GB de memória RAM, utilizando o Sistema Operacional Linux Debian 7.0 (wheezy), OpenJDK 1.6.0_27 IcedTea e Python 2.7.3;

PC: processador Intel Core i7-2670QM de 2,2 GHz, com 8 GB de memória RAM, utilizando o Sistema Operacional Linux Ubuntu 12.10 (quantal), Java 1.7.0_25 da Oracle e Python 2.7.3.

Nesses computadores, rodamos os testes em apenas uma linha de execução (*thread*) e efetuamos um total de 1000 repetições para cada um dos quesitos analisados, coletando o tempo de barreira (horário em que a função encerrou seu processamento menos o horário inicial da chamada da função) em cada uma destas repetições.

B. Tempos obtidos nos testes

Apresentamos nas TABELA IV e TABELA V o intervalo, a média e o desvio padrão de tempo gasto pelo Quartz Original e pelo Quartz Aprimorado.

Acreditamos que a discussão levantada na Seção VIII – durante a análise da proposta de nosso aprimoramento – seja suficiente para justificar os tempos obtidos em nossos testes. Ou seja, a substituição do SHA-1 pelo SHA-3 além de ajudar na segurança do Quartz Aprimorado também proporcionou um ganho de eficiência de aproximadamente 75% no momento da Inicialização dos Vetores; sendo que tal melhoria ocorreu principalmente devido a não realização de operações iteradas e também por não concatenar suas saídas para obter vetores do tamanho estabelecido pelo algoritmo de assinatura.

Além disso, observamos que os algoritmos de Geração de Chaves e Assinatura têm sua eficiência prejudicada devido a nossa escolha de parâmetros (cerca de 360% acrescido no tempo de Geração de Chaves e 240% na Assinatura), isto porque tal escolha ocasionou um aumento significativo na quantidade de objetos e instâncias a serem manipuladas por nossa implementação.

TABELA IV. TEMPOS OBTIDOS DURANTE A REALIZAÇÃO DOS TESTES NO BRUCUTU.

			Quartz Original	Quartz Aprimorado
Inicialização dos Vetores	**SHA-1**	**Média (ms)**	158	-
		Desvio Padrão (ms)	16	-
		Intervalo (ms)	121 - 236	-
	SHA-3	**Média (ms)**	-	40
		Desvio Padrão (ms)	-	7
		Intervalo (ms)	-	34 - 57
Geração de Chaves		**Média (s)**	16,9	75,1
		Desvio Padrão (s)	0,2	0,4
		Intervalo (s)	16,5 - 17,7	74,2 - 77,8
Assinatura		**Média (s)**	5,2	19,1
		Desvio Padrão (s)	0,7	0,2
		Intervalo (s)	4,4 - 27,2	18,9 - 20,0
Verificação de Assinatura		**Média (ms)**	3814	18
		Desvio Padrão (ms)	233	3
		Intervalo (ms)	4 - 3927	17 - 40
Verificação de Assinatura Falsa		**Média (ms)**	60074	180
		Desvio Padrão (ms)	959	14
		Intervalo (ms)	52067 - 62258	159 - 194

TABELA V. TEMPOS OBTIDOS DURANTE A REALIZAÇÃO DOS TESTES NO PC.

			Quartz Original	Quartz Aprimorado
Inicialização dos Vetores	**SHA-1**	**Média (ms)**	62	-
		Desvio Padrão (ms)	15	-
		Intervalo (ms)	47 - 130	-
	SHA-3	**Média (ms)**	-	15
		Desvio Padrão (ms)	-	4
		Intervalo (ms)	-	12 - 44
Geração de Chaves		**Média (s)**	18,5	87,0
		Desvio Padrão (s)	1,7	10,4
		Intervalo (s)	15,6 - 26,8	72,2 - 108,3
Assinatura		**Média (s)**	5,4	16,6
		Desvio Padrão (s)	5,4	2,9
		Intervalo (s)	4,3 - 169,2	16,5 - 25,6
Verificação de Assinatura		**Média (ms)**	164	35
		Desvio Padrão (ms)	89	4
		Intervalo (ms)	136 - 2447	33 - 53
Verificação de Assinatura Falsa		**Média (ms)**	43248	99
		Desvio Padrão (ms)	3488	14
		Intervalo (ms)	36197 - 54591	96 - 145

Contudo, a Verificação de Assinatura de nosso aprimoramento mostrou-se significativamente melhor do que a do modelo original, sendo possível observar um ganho de $99{,}5\,\%$ (no Brucutu) e cerca de $78\,\%$ (no PC) durante a verificação de assinaturas legítimas e aproximadamente $99{,}7\,\%$ quando trata-se de assinaturas falsas; lembrando que no Quartz Aprimorado, durante a resolução da função G, são testadas até 4.096 vezes menos hipóteses de utilização da chave pública do que no Quartz Original.

Frisa-se que os tempos expostos acima foram obtidos a partir de nossa implementação de referência: feita em Java, sem paralelismo ou qualquer conjunto de instruções avançadas. Ou seja, todas as alterações no desempenho ocorreram em virtude dos parâmetros escolhidos e do novo design proposto em nosso aprimoramento.

XII. CONSIDERAÇÕES FINAIS

Nossa principal contribuição nesta pesquisa foi à apresentação de um novo protocolo de assinatura digital, baseado em sistemas polinomiais multivariados quadráticos. Como resultado, obtivemos um criptossistema com um nível de segurança estimado em 2^{112}, contra os 2^{50} do protocolo original. Nossa proposta apresenta, ainda, um ganho de eficiência na inicialização dos vetores que serão utilizados pelo algoritmo de assinatura; para ser mais específico, através de nossos testes – realizados a partir de nossa implementação de referência – constatamos um ganho de aproximadamente 75%. Além disto, mostramos que no Quartz Aprimorado, durante a resolução da função G, testaremos até 4.096 vezes menos hipóteses de utilização da chave pública, quando comparado com o Quartz Original.

Todavia, observamos que devido os parâmetros escolhidos para nosso criptossistema, houve um aumento significativo no tamanho das chaves, fato que também acarretou uma perda de eficiência nos algoritmos de geração de chaves e assinatura.

Em virtude do tamanho das chaves de nosso aprimoramento, acreditamos que uma possível extensão para nosso trabalho seria pesquisar uma maneira de reluzi-las. Outra possível extensão de nosso trabalho seria buscar uma prova de segurança mais eficiente (*tight*) no modelo do

oráculo aleatório para protocolos baseados na *trapdoor* HFE. Isto porque, conforme vimos durante a estimativa de segurança do Quartz Aprimorado, com o modelo de prova proposto por Sakumoto *et al.* [48] perdemos em média 60 bits de segurança.

AGRADECIMENTOS

Agradecemos à CAPES pelo apoio financeiro concedido.

REFERÊNCIAS

[1] Elaine Barker and Allen Roginsky. NIST Special Publication 800-131a - Transitions: Recommendation for Transitioning the Use of Cryptographic Algorithms and Key Lengths. Technical report, National Institute of Standards and Technology, NIST, U.S. Department of Commerce, Washington DC. http://csrc.nist.gov/publications/nistpubs/800-131A/sp800-131A.pdf, 2011. Último acesso em 09/07/2013.

[2] Mihir Bellare and Phillip Rogaway. The exact security of digital signatures: How to sign with RSA and Rabin. In U. Maurer, editor, *Advances in Cryptology - EUROCRYPT 96 Proceedings*, volume 1070 of *Lecture Notes in Computer Science*, pages 399–416. Springer-Verlag, 1996.

[3] Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors. *Post-Quantum Cryptography*. Springer, 2009.

[4] Luk Bettale, Jean-Charles Faugère, and Ludovic Perret. Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. *Designs, Codes and Cryptography*, pages 1–52, 2012.

[5] Charles Bouillaguet, Hsieh-Chung Chen, Chen-Mou Cheng, Tung Chou, Ruben Niederhagen, Adi Shamir, and Bo-Yin Yang. Fast Exhaustive Search for Polynomial Systems in $\mathbb{F}_2$. In Stefan Mangard and François-Xavier Standaert, editors, *Cryptographic Hardware and Embedded Systems, CHES 2010*, volume 6225 of *Lecture Notes in Computer Science*, pages 203–218. Springer Berlin Heidelberg, 2010.

[6] Charles Bouillaguet, Jean-Charles Faugère, Pierre-Alain Fouque, and Ludovic Perret. Practical Cryptanalysis of the Identification Scheme Based on the Isomorphism of Polynomial with One Secret Problem. In Dario Catalano, Nelly Fazio, Rosario Gennaro, and Antonio Nicolosi, editors, *Public Key Cryptography - PKC 2011*, volume 6571 of *Lecture Notes in Computer Science*, pages 473–493. Springer Berlin Heidelberg, 2011.

[7] An Braeken, Christopher Wolf, and Bart Preneel. A study of the security of Unbalanced Oil and Vinegar signature schemes. Cryptology ePrint Archive, Report 2004/222. http://eprint.iacr.org/2004/222, 2004. Último acesso em 11/06/2013.

[8] Jean charles Faugère. A new efficient algorithm for computing Gröbner Bases without reduction to zero (F5). In *ISSAC 02: Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, pages 75–83, 2002.

[9] Nicolas Courtois, Louis Goubin, Willi Meier, and Jean-Daniel Tacier. Solving underdefined systems of multivariate quadratic equations. In David Naccache and Pascal Paillier, editors, *Public Key Cryptography*, volume 2274 of *Lecture Notes in Computer Science*, pages 211–227. Springer Berlin Heidelberg, 2002.

[10] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In Bart Preneel, editor, Advances in *Cryptology - EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 392–407. Springer Berlin Heidelberg, 2000.

[11] Nicolas T. Courtois. The security of Hidden Field Equations (HFE). In David Naccache, editor, *Topics in Cryptology - CT-RSA 2001*, volume 2020 of *Lecture Notes in Computer Science*, pages 266–281. Springer Berlin Heidelberg, 2001.

[12] Nicolas T. Courtois. Generic Attacks and the Security of Quartz. In YvoG. Desmedt, editor, *Public Key Cryptography - PKC 2003*, volume 2567 of *Lecture Notes in Computer Science*, pages 351–364. Springer Berlin Heidelberg, 2002.

[13] Nicolas T. Courtois. Short signatures, provable security, generic attacks and computational security of multivariate polynomial schemes such as HFE, QUARTZ and SFLASH. Cryptology ePrint Archive, Report 2004/143. http://eprint.iacr.org/2004/143, 2004. Versão extendida e revista do artigo *Generic Attacks and the Security of Quartz* publicado no PKC 2003. Último acesso em 12/06/2013.

[14] Nicolas T. Courtois, Louis Goubin, and Jacques Patarin. Quartz, na asymmetric signature scheme for short signatures on PC. Primitive specification and supporting documentation (second revised version), 2001.

[15] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London Ser. A*, A400:97–117, 1985.

[16] Jintai Ding, Jason E. Gower, and Dieter Schmidt. *Multivariate public key cryptosystems*, volume 25 of *Advances in information security*. Springer, 2006.

[17] Jintai Ding and TimothyJ. Hodges. Inverting HFE systems is quasi-polynomial for all fields. In Phillip Rogaway, editor, *Advances in Cryptology - CRYPTO 2011*, volume 6841 of *Lecture Notes in Computer Science*, pages 724–742. Springer Berlin Heidelberg, 2011.

[18] Jintai Ding and Dieter Schmidt. Cryptanalysis of HFEv and Internal Perturbation of HFE. In Serge Vaudenay, editor, *Public Key Cryptography - PKC 2005*, volume 3386 of *Lecture Notes in Computer Science*, pages 288–301. Springer Berlin Heidelberg, 2005.

[19] Jintai Ding, Dieter Schmidt, and Fabian Werner. Algebraic attack on HFE revisited. In Tzong-Chen Wu, Chin-Laung Lei, Vincent Rijmen, and Der-Tsai Lee, editors, *Information Security*, volume 5222 of *Lecture Notes in Computer Science*, pages 215–227. Springer Berlin Heidelberg, 2008.

[20] Emmanuelle Dottax and École Normale Supérieure. Tweak reviews: ES-IGN, RSA-PSS, QUARTZ and SFLASH. NES/DOC/ENS/WP1/018/1. Technical report, Commission of the European Communities, out 2002. Último acesso em 04/07/2013.

[21] Jean-Charles Faugère. Algebraic cryptanalysis of HFE using gröbner bases, February 2003.

[22] Jean-Charles Faugère and Antoine Joux. Algebraic cryptanalysis of Hidden Field Equation (HFE) cryptosystems using gröbner bases. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 44–60. Springer Berlin Heidelberg, 2003.

[23] Adam Thomas Feldmann. A survey of attacks on Multivariate Cryptosystems. Master's thesis, University of Waterloo, Ontario, Canada, 2005.

[24] Patrick Felke. On the Affine Transformations of HFE-Cryptosystems and Systems with Branches. Cryptology ePrint Archive, Report 2004/367. http://eprint.iacr.org/2004/367, 2004. Último acesso em 03/07/2013.

[25] Python Software Foundation. 14.1. hashlib – Secure hashes and message digests. https://docs.python.org/2/library/hashlib.html, 2015. Último acesso em 31/01/2015.

[26] Python Software Foundation. pysha3 0.3. https://pypi.python.org/pypi/pysha3/, 2015. Último acesso em 31/01/2015.

[27] A.S. Fraenkel and Y. Yesha. Complexity of problems in games, graphs and algebraic equations. *Discrete Applied Mathematics*, 1(1–2):15–30, September 1979.

[28] Michael R. A. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. A Series of Books in the Mathematical Sciences. W. H. Freeman, 1979.

[29] Henri Gilbert and Marine Minier. Cryptanalysis of SFLASH. In LarsR. Knudsen, editor, *Advances in Cryptology - EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 288–298. Springer Berlin Heidelberg, 2002.

[30] Louis Granboulan, Antoine Joux, and Jacques Stern. Inverting HFE is Quasipolynomial. In Cynthia Dwork, editor, *Advances in Cryptology - CRYPTO 2006*, volume 4117 of *Lecture Notes in Computer Science*, pages 345–356. Springer Berlin Heidelberg, 2006.

[31] Shu jen Chang, Ray Perlner, William E. Burr, Meltem Sönmez Turan, John M. Kelsey, Souradyuti Paul, and Lawrence E. Bassham. NIST Interagency or Internal Reports 7896: Third-Round Report of the SHA-3 Cryptographic Hash Algorithm Competition. Technical report, National Institute of Standards and Technology, NIST, U.S. Department of Commerce, Washington DC. http://nvlpubs.nist.gov/nistpubs/ir/2012/NIST.IR.7896.pdf, 2012. Último acesso em 09/07/2013.

[32] Xin Jiang, Jintai Ding, and Lei Hu. Kipnis-shamir attack on HFE revisited. In Dingyi Pei, Moti Yung, Dongdai Lin, and Chuankun Wu, editors, *Information Security and Cryptology*, volume 4990 of *Lecture*

Notes in Computer Science, pages 399–411. Springer Berlin Heidelberg, 2008.

[33] Antoine Joux and Gwenaëlle Martinet. Some weaknesses in Quartz Signature Scheme. NES/DOC/ENS/WP5/026/1. Technical report, Commission of the European Communities, jan 2003. Último acesso em 12/06/2013.

[34] Aviad Kipnis, Jacques Patarin, and Louis Goubin. Unbalanced Oil and Vinegar signature schemes. In Jacques Stern, editor, *Advances in Cryptology - EUROCRYPT 99*, volume 1592 of *Lecture Notes in Computer Science*, pages 206–222. Springer Berlin Heidelberg, 1999.

[35] Aviad Kipnis and Adi Shamir. Cryptanalysis of the HFE Public Key Cryptosystem by Relinearization. In *CRYPTO 99: Proceedings of the 19th Annual International Cryptology Conference on Advances in Cryptology*, pages 19–30, London, UK, 1999. Springer-Verlag.

[36] Dongdai Lin, Jean-Charles Faugère, Ludovic Perret, and Tianze Wang. On enumeration of polynomial equivalence classes and their application to MPKC. Cryptology ePrint Archive, Report 2011/055. http://eprint.iacr.org/2011/055, 2011. Último acesso em 29/06/2013.

[37] Gwenaëlle Martinet and École Normale Supérieure. QUARTZ, FLASH and SFLASH. NES/DOC/ENS/WP3/006/2. Technical report, Commission of the European Communities, mar 2001. Último acesso em 14/06/2013.

[38] NESSIE. Final report of European project IST-1999-12324: New European Schemes for Signatures, Integrity, and Encryption (NESSIE), (Abril de 2004). https://www.cosic.esat.kuleuven.be/nessie/Bookv015.pdf. Technical report, Commission of the European Communities, Abril 2004. Último acesso em 10/06/2013.

[39] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[40] NIST. FIPS 186-3: Digital Signature Standard (DSS). Technical report, National Institute of Standards and Technology, NIST, U.S. Department of Commerce, Washington DC. http://csrc.nist.gov/publications/fips/fips186-3/fips_186-3.pdf, 2009. Último acesso em 16/07/2013.

[41] Jacques Patarin. Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88. In Don Coppersmith, editor, *Advances in Cryptology - CRYPT0 95*, volume 963 of *Lecture Notes in Computer Science*, chapter 20, pages 248–261. Springer Berlin / Heidelberg, Berlin, Heidelberg, July 1995.

[42] Jacques Patarin. Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms. In Ueli Maurer, editor, *Advances in Cryptology - EUROCRYPT 96*, volume 1070 of *Lecture Notes in Computer Science*, pages 33–48. Springer-Verlag, 12–16 May 1996.

[43] Jacques Patarin. Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms - Extended Version, 1996.

[44] Jacques Patarin, Nicolas T. Courtois, and Louis Goubin. QUARTZ, 128-bit Long Digital Signatures. In David Naccache, editor, *Topics in Cryptology - CT-RSA 2001*, volume 2020 of *Lecture Notes in Computer Science*, pages 282–297. Springer Berlin Heidelberg, 2001.

[45] Jacques Patarin and Louis Goubin. Trapdoor one-way permutations and Multivariate Polynomials - Extended Version. In *Proc. of ICICS 97*, *LNCS 1334*, pages 356–368. Springer, 1997.

[46] Albrecht Petzoldt, Stanislav Bulygin, and Johannes Buchmann. CyclicRainbow – A Multivariate Signature Scheme with a Partially Cyclic Public Key. In Guang Gong and KishanChand Gupta, editors, *Progress in Cryptology - INDOCRYPT 2010*, volume 6498 of *Lecture Notes in Computer Science*, pages 33–48. Springer Berlin Heidelberg, 2010.

[47] Albrecht Petzoldt, Stanislav Bulygin, and Johannes Buchmann. Selecting parameters for the Rainbow Signature Scheme. In Nicolas Sendrier, editor, *Post-Quantum Cryptography*, volume 6061 of *Lecture Notes in Computer Science*, pages 218–240. Springer Berlin Heidelberg, 2010.

[48] Koichi Sakumoto, Taizo Shirai, and Harunaga Hiwatari. On provable security of UOV and HFE Signature Schemes against Chosen-Message Attack. In Bo-Yin Yang, editor, *Post-Quantum Cryptography*, volume 7071 of *Lecture Notes in Computer Science*, pages 68–82. Springer Berlin Heidelberg, 2011.

[49] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[50] Routo Terada. *Segurança de dados: Criptografia em redes de computador*. Blucher, 2ª revisada e ampliada edition, 2008.

[51] Xiaoyun Wang, Yiqun Lisa Yin, and Hongbo Yu. Finding collisions in the full SHA-1. In *Advances in Cryptology - CRYPTO 2005*: 25th Annual International Cryptology Conference, Santa Barbara, California, USA, August 14-18, 2005, Proceedings, volume 3621 of *Lecture Notes in Computer Science*, pages 17–36. Springer, 2005.

[52] Christopher Wolf. Implementing QUARTZ in java. Draft for the 3rd NESSIE Workshop. http://www.christopher-wolf.de/ql/quartzJava.pdf, 2002. Último acesso em 05/07/2013.

[53] Christopher Wolf. QuartzLight in Java. http://www.christopher-wolf.de/ql/, 2002. Último acesso em 17/07/2013.

[54] Christopher Wolf. *Multivariate Quadratic Polynomials in Public Key Cryptography*. PhD thesis, Katholieke Universiteit Leuven – Faculteit Ingenieurswetenschappen - Departement Elektrotechniek (ESAT), 2005.

[55] Christopher Wolf and Bart Preneel. Taxonomy of Public Key Schemes based on the problem of Multivariate Quadratic equations. Cryptology ePrint Archive, Report 2005/077. http://eprint.iacr.org/2005/077, 2005. Último acesso em 28/06/2013.

[56] Takanori Yasuda, Kouichi Sakurai, and Tsuyoshi Takagi. Reducing the Key Size of Rainbow using non-commutative rings. In Orr Dunkelman, editor, *Topics in Cryptology – CT-RSA 2012*, volume 7178 of *Lecture Notes in Computer Science*, pages 68–83. Springer Berlin Heidelberg, 2012.

Ewerton R. Andrade possui graduação em Sistemas de Informação pelo Centro Universitário Luterano de Ji-Paraná (2009), também é graduado em Matemática pela Universidade Federal de Rondônia (2011), mestrado em Ciência da Computação pelo Instituto de Matemática e Estatística da Universidade de São Paulo (2013), e atualmente é doutorando em Engenharia de Computação pela Escola Politécnica da Universidade de São Paulo. Tem experiência na área de Ciência da Computação, com ênfase em Segurança de Dados / Criptografia, atuando principalmente nos seguintes temas: segurança de dados, criptografia, assinaturas digitais baseadas em polinômios multivariados quadráticos, funções de derivação de chaves e esquemas de hash de senhas.

Routo Terada possui graduação em Engenharia Elétrica Eletrônica pela Universidade de São Paulo (1970), mestrado em Matemática Aplicada pela Universidade de São Paulo (1975) e doutorado em Computer Science - University of Wisconsin - Madison (1979). Atualmente é professor titular da Universidade de São Paulo, avaliador de artigos do International Journal of Information Security e do Journal of the Brazilian Computer Society. Tem experiência na área de Ciência da Computação, com ênfase em Criptografia, atuando principalmente nos seguintes temas: segurança de dados, criptografia, algoritmos, criptossistemas.

Untappable Key Distribution System: a One-Time-Pad Booster

G. A. Barbosa and J. van de Graaf

***Abstract*—The One-Time-Pad (OTP) protocol gives unconditional security for the information being encrypted. Correctly implemented, not even an adversary with a quantum computer can crack it. However, the need of sharing in a secure way supplies of symmetric random keys turned the method almost obsolete as a stand-alone method for fast and large volume telecommunication. Basically, this secure sharing of keys and their renewal, once exhausted, had to be done through couriers, in a slow and costly process. This paper presents a solution for this problem providing a fast and unlimited renewal of secure keys: An *untappable key distribution system* is presented and detailed. This *fast* key distribution system utilizes two layers of confidentially protection: 1) Physical noise intrinsic to the optical channel that turn the coded signals into *stealth* signals and 2) Privacy amplification using a bit pool of refreshed entropy run after run, to eliminate any residual information. The resulting level of security is rigorously calculated and demonstrates that the level of information an eavesdropper could obtain is negligible. The random bit sequences, fast and securely distributed, can be used to encrypt text, data or voice.**

***Keywords*—Random, Physical processes, Cryptography, Privacy amplification.**

I. Introduction

A *key distribution* system that uses the intrinsic light noise of an optical carrier to forbid an attacker E (or Eve) to extract clean signals from the transmitted ones was described in Refs. [1] and [2]. The basic characteristics of that system is that the legitimate users, A (or Alice) and B (or Bob), are *not* affected in the same way as E by the channel's noise. This asymmetry is caused by a starting *information* shared by A and B but not by E – it produces a measurement *advantage* for A and B over Eve: Signals buried under noise for Eve and clear signals for A or B.

This work stresses basic theoretical and practical aspects of that physical system (Section III) and enhances its security by an explicit privacy amplification protocol (PA) (Sections VI and VII). The security level due to these two protection layers, physical and computational, is calculated and discussed. The physical implementation will be presented in a following work.

The use of optical noise to secure *encryption* of signals in telecommunication channels was analyzed in [3] (alpha-eta *encryption* system). This system was tested in secure networks in US. Also in land-air tests (Optix/NuCrypt), reached market applications (NuCrypt LCC) and produced independent developments in Japan [4]. The system discussed here is akin to the alpha-eta (or $\alpha\eta$) system in the use of optical noise but is specific as a *key distribution* system. Furthermore, it uses true random bit generators instead of linear feedback-shift-registers (LFSR) used in the alpha-eta systems.

The resulting securely shared random sequences of bits can be used for encryption in arbitrary communication channels. It can be used for bit-to-bit encryption as a "one-time-pad" system, with constantly renewed keys in a fast process.

The discussed system has a basic connection in principle with Ref. [5], where cryptography using continuous variables with coherent states was proposed: the use of the *optical noise* is also at the core of that scheme. However, several important differences exist between these systems. Among them, optical quantum demolition measurements are not necessary, neither quadrature measurements. Therefore, the system discussed in this paper is widely different and, as such, no need for comparisons exist.

The key distribution system presented here (as in [1]) uses light's noise for protection by modulating each signal representing a random bit by another randomly chosen (secret) signal representing a physical basis. The superposition of "noise" signals, "basis modulation" and "bit" signals frustrates an attacker trying to obtain either the basis used or the transmitted bit. A privacy amplification protocol (PA) operates in a *bit-pool* constantly renewed in entropy that enhances the security of the system. The overall security achieved is calculated giving the users a guaranteed security level.

This system is designed such that its ultimate security should depend only on the secure transmission of the bits and their safe storage. The system can be fully understood and signals openly acessed by the adversary and yet full security for A and B resides just on the keys (Kerckhoff's principle).

One among the possible uses for this continuously renewable and fast one time process, is the protection of energy infrastructures (generation, distribution and their control interconnected by smart-grids). The proposed system provides fast and secure sharing of keys between end-points connected by an optical channel as well as a fast one-time-pad encryption.

This paper is roughly divided in two parts, one dealing with the physical noise and other with privacy amplification aspects. Although the subjects are different, they are intrinsically connected by the architecture of the key distribution system and are essential for a full understanding of this system.

II. Phase modulation and optical noise

One of the simplest ways to implement the physical part of the scheme is using optical phase modulation of a laser beam. This modulation is achieved by fastly modifying the refractive index of an optical medium in the beam path before the transmission channel (fiber optics or any continuous media).

Bits could be represented by a given phase, say $\phi_1 = 0^\circ$ for bit 1 and bit 0 by phase $\phi_0 = 180^\circ$. A *basis* is represented by an extra phase ϕ_b, within a manyfold of M possible values, that is added to ϕ_1 or ϕ_0 producing a different resulting phase: $\phi_{b,j} = \phi_j + \phi_b$, $(j = 1, 0)$ and $(b = 1, 2, 3 \ldots M)$. Both bit and basis are unknown to the adversary. Fig. 1 explain these ideas. It represents a uniformly spaced set of bases constituted of M

G. A. Barbosa, PhD, is the CEO of QuantaSec – Consulting and Projects in Physical Cryptography Ltda, Av. Portugal 1558, Belo Horizonte MG 31550-000 Brazil. Phone: +11 55 (31) 3441–4121, e-mail: GeraldoABarbosa@gmail.com

J. van de Graaf, PhD, is Professor, Instituto de Ciências Exatas, Dept. de Ciência da Computação, Av. Antônio Carlos 6.627, Belo Horizonte MG 31270-901 Brazil. E-mail: jvdg@dcc.ufmg.br

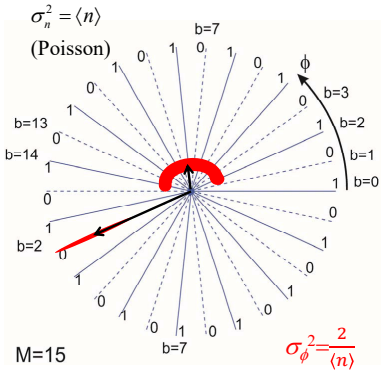


Figure 1. Wheel of phases representing encryption bases for bits. Bits 0 and 1 are represented at extremes of a basis and separated by π. Encryption bases are separated by π/M. A bit signal represented by an amplitude and phase has an intrinsic phase noise (given by $\sigma_\phi^2 = 2/\langle n \rangle$) that may cover adjacent bases and do not allows an attacker to identify the signal being sent. While a strong amplitude signal allows easy identification of the bases and bit (e.g., see signal at basis $b = 2$), a weak signal does not allow such identification (e.g., see signal around basis 7 or 8). It should be emphasized that a signal representing one bit is sent just once and never repeated.

physical phase bases ($M = 15 : b = 0, 1, 2, \ldots, 14$ in this example, closest bases are separated by an angle of $12°$). Each basis is represented as a *single* line made up of a solid line continued by a dashed one (a given phase value and this same value $+180°$ represents a single basis).

The fundamental characteristics of these bases is that bits are represented in an alternate order in neighboring bases. For example, if Alice wants to send a bit 1 she may pick one basis, say $b = 2$ (without Eve's knowledge). In this basis, bit 1 is represented along a solid line ($\phi_2 = 24°$). If she had picked $b = 3$, bit 1 would have been placed at the phase $\phi = (12 \times 3 + 180)°$. In any of these cases, the closest bits in neighboring bases would have been opposite.

The distance from the center, along any basis lines in Fig. 1, gives the *amplitude* of the light field that carries the bit signal while phases are represented around a circle as indicated. The uncertainty in the signal to be measured (e.g., by measuring Stokes parameters) can be represented in Fig. 1 by a smeared figure representing uncertainties in amplitude and phase (see red features in figure).

A difference between a strong and a weak coherent signal is that the phase uncertainty over the average signal level with n photons ($1/\sqrt{\langle n \rangle}$) for the weak signal is larger.

The phase uncertainty can be calculated in a similar way as done in the polarization uncertainty obtained in [1] (Eq. (2)): Assume that a laser beam in a coherent state $|\Psi_0\rangle = |\alpha\rangle$ passes through an optical modulator that produces a phase difference ϕ between its two physical axis (say x and y) for an incoming polarization state. One should recall that for a coherent state $|\alpha|^2 = \langle n \rangle$ [13].

The optical modulator is a two-port device for an incoming state. The modulator transforms the state $|\alpha\rangle$ according to the angular momentum rotation operator J_z for two states (or two modes). These states are represented by photon annihilation operators a_x and b_y: $J_z = (1/2)(a_x^+ a_x - b_y^+ b_y)$ [7]. The transformation produces

$$|\Psi(\phi)\rangle = e^{-iJ_z\phi}|\Psi_0\rangle = |\frac{\alpha}{\sqrt{2}}e^{-i\phi/2}\rangle_x|\frac{\alpha}{\sqrt{2}}e^{i\phi/2}\rangle_y \,. \quad (1)$$

In $|\Psi(\phi)\rangle$ a phase is established due to a phase difference between two orthogonal components x and y. From this result the overlap of two states $|\Psi(\phi)\rangle$ and $|\Psi(\phi')\rangle$ can be obtained:

$$\langle\Psi(\phi)|\Psi(\phi')\rangle = e^{-|\alpha|^2\left[1-\cos(\frac{\phi-\phi'}{2})\right]} \,. \quad (2)$$

This overlap is a measure of the "indistinguishability" degree between the two states. For mesoscopic states $|\alpha|^2 \gg 1$ (but not intense) and the exponential term gives a vanishing contribution unless $\phi - \phi'$ is small. Considering $\Delta\phi \equiv \phi - \phi' \ll 1$ one has the Gaussian distribution

$$\langle\Psi(\phi)|\Psi(\phi')\rangle \to e^{-|\alpha|^2(\Delta\phi)^2/2} \,. \quad (3)$$

The probability for indistinguishability between ϕ and ϕ' is then given by

$$|\langle\Psi(\phi)|\Psi(\phi')\rangle|^2 \to e^{-|\alpha|^2(\Delta\phi)^2} \equiv e^{-(\Delta\phi)^2/(2{\sigma_\phi}^2)} \,, \quad (4)$$

where $\sigma_\phi = \sqrt{2/\langle n \rangle}$.

This shows that a strong signal (large $\langle n \rangle$) has a reduced phase uncertainty. For example, the large amplitude signal representing a bit 0 in basis $b = 2$ could be easily identifiable (see Fig. 1) by either A, B or E because the phase uncertainty is small and no confusion is possible with a neighboring bit. Differently, if the phase uncertainty is such (weak signal) that the obtained signal overlaps neighboring bases (see uncertainty around basis $b = 8$ in Fig. 1), the information of *which* basis is being used is not available. Consequently, the bit sent cannot be identified without a large probability of error. This noisy channel can be referred as the $\alpha\eta$ channel.

However, if the legitimate users know which basis was used, there is no ambiguity in bit identification. For them, bit identification is just a question of identifying if the signal is *around* a given phase value ϕ or at $\phi + 180°$, not between closest bases where identification is not allowed due to the phase noise.

If the basis information is not available to Eve, her measurements will produce errors in the bases or bit estimation (for example, signals around bases 7 and 8). In other words, as the separation between closest bases is π/M, the resolution needed for bit or basis identification has to be better than this value. However, the noise is tailored by the legitimate users to produce an uncertainty much larger than π/M and the attacker has no way to reduce this noise.

With a proper choice of a separation $\Delta\phi$ between bases and average number of photons $\langle n \rangle$, not only the separation can be set $\Delta\phi < \sigma(= 1/\sqrt{2\langle n \rangle})$ but also the probability of an error by Eve, P_e^E can be set arbitrarily close to 1/2 (see Fig. 3 in [1] and discussions therein). The derivation of P_e using POVM (Positive Operator Valued Measurement) can be found in Ref. [1]. Appendix B discuss the effect of noise in the channel in an alternate way, by means of Poincaré measurements. This way, the reader has different contexts for comparisons.

The different amount of information between E and the legitimate users produce the different results between E and B in a transmitted sequence of bits. This is made possible by information shared beforehand in each transmission round by A and B in the form of a secret stream with the information of the bases being used and about which E has no information.

We will also see that this process can be continued without limitation, without any need for A and B to meet after the first contact.

XOR encoding?– Some questions may be asked, such as "why this basis encoding cannot be replaced by a simple XOR of *basis*(=*single bit*) with the *bit*, especially in the model where the eavesdropper gets a perfect copy of the transmitted state?".

The level of signals used in this implementation is such that the term "perfect copy" does not apply. Whereas classical signals may admit the concept of a perfect copy (apart from technical noises), *any* "copy" of a signal in the mesoscopic range produces a distinct output due to the inherent noise in the channel. In other words, use of signals where the signal-to-noise ration S/N is very high (="classical" signals) produces undistinguishable copies. Of course, an XOR of classical signals produces well-defined signals and the attacker's task will be solely cryptanalysis of perfectly defined signals - a purely mathematical task. On the other hand, when dealing with an intrinsically noisy channel, the first task of the adversary task is to make sense of the signals being transmitted. Moreover, when these signals are distributed among different physical states, the adversary task to obtain even the sequence of signals to a posterior cryptanalysis is shown impossible. Access to the $\alpha\eta$ channel gives the adversary very little information on the signals. The amount of information available for cryptanalysis in the two cases is vastly different.

III. The transmission and reception system

Alice has a physical random bit generator (PhRBG) that produces true random bits continuously. One of the basic questions to be answered in this paper is: *Given that A and B start sharing a secret random sequence of length c_0, what is a secure length of bits to be extracted in the successive rounds of this system after c_0 have been used*?

The original transmission protocol described in [1] and [2] indicate the need for a PA protocol to be applied at the transmitter and receiver stations to eliminate any information that the attacker could have obtained in these attempts but no specific procedure was proposed. Furthermore, in the present paper, instead of assuming some repetition of bases as done in [1], the idea of a "*bit pool*" is used leading to a substantial improvement on the overall security of the system. This will be explained in Section VII.

Basically the idea is use to PA to reduce the amount of information accessed by E to a negligible amount while giving A and B access to a refreshed bit sequence to be used. This frustrates even a-posteriori attacks using known-plaintext by an adversary trying to recover past bit sequences based on the bit sequences obtained from the plaintext used. Before explaining the protocol within the bit-pool, a description will be given of the physical system where the protocols will operate.

Fig. 2 sketches the main parts of the key distribution system.

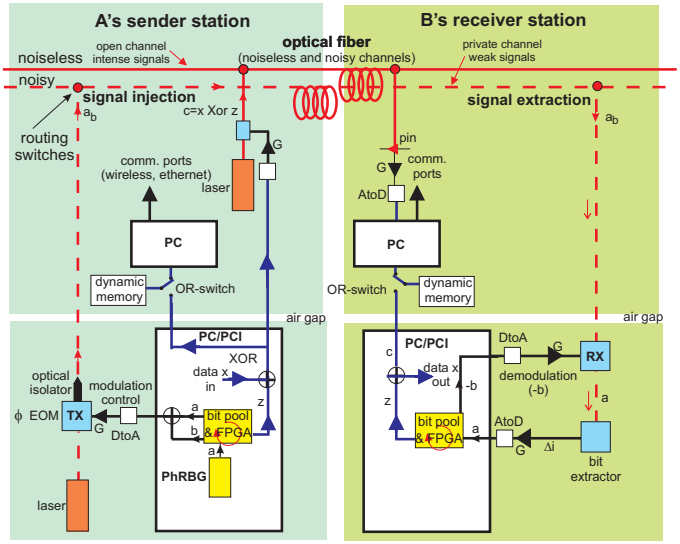


Figure 2. A simplified sketch of the sender and receiver stations used by the legitimate users Alice and B. The fiber channel may be a single fiber with a noiseless channel and a noisy channel. The noisy channel is used to distribute secret random bits between Alice and Bob while the noiseless channel is used to transmit encrypted information. These channels have to be spectrally separated to avoid spill-overs from the intense noiseless channel to the noisy channel. The distributed secret bits are treated and privacy amplified in a bit pool in Alice's and Bob's stations. The distilled secure sequence of bits, z, is used for encryption of text, image or voice. Signals are sent just once and never repeated. Actually, sender and receiver systems may be contained in both A and B stations to simultaneously offer sender/receiver capabilities.

A single optical fiber contains an optical noisy channel ($\alpha\eta$) and a noiseless channel, *both* fully accessible to Alice, Bob and Eve. Signals from the laser (carrier) are modulated at station A and demodulated at station B. A physical random bit generator (PhRBG) [8] feeds a control station composed of a computer and electronics to perform required functions such as digital to analog (DtoA) conversion, analog to digital (AtoD), XOR, and PA operations on a bit pool. A final stream z of secure bits is extracted from the pool to encrypt bit-by-bit any desired data x ("message") ($c = x \oplus z$) to be sent from A to B or from B to A by the public channel. The PhRBG continuously generates a fast stream of random bits a that feeds the control station.

Briefly described, A and B secretely shared $c_0 = m\,n_0$ bits to create n_0 modulation bases to encode n_0 fresh bits generated by the PhRBG (discussed in Sections IV and VII). m is the number of bits necessary to specify each basis. Eve has full access to both channels, noisy and noiseless. The optical signals are created by phase modulation of a laser beam to create the information signals transmitted by the optical fiber. Signals in the classical channel have a high signal to noise ratio (SNR), while the noisy channel has a relatively small SNR. The noisy channel is wavelength separated from the classical channel such that the wavelength separation avoids overlap of the wings of the strong signals in the classical channel with the weak signal carried by the noisy channel. Fig. 3 (taken from [9]) exemplifies signal and noise levels taken at some point in a fiber network that include optical amplifiers. The mesoscopic signals used in the key distribution system being discussed are above the QKD level (single photon level), but well below an intense (=classi-

cal) signal. We define a classical signal as a signal that can be perfectly copied or only subjected to technical noises, that could be eliminated by improved techniques.

Wavelength Division Multiplexing (WDM, dense or coarse)

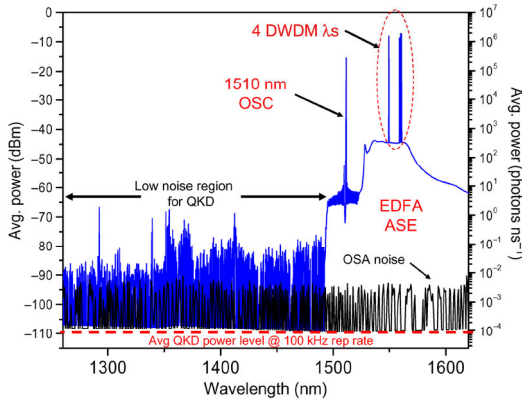


Figure 3. Different signal and noise levels due to different processes in a network. The minimum noise level is the OSA noise due to the spectrum analyzer being used while the maximum signal is due to amplified signal that also produces a relatively high amplified spontaneous emission signal (EDFA ASE) that acts as a noise for some communication channels. Signal from an Optical Supervisory Channel (OSC) is also shown.

technology can be used to set distinct channels in the same single fiber around 1553nm. WDM wavelengths are standardized with 100GHz spacing in optical frequencies, with a reference fixed at 1552.52nm (193.10 THz). DWDM can use 50 GHz channel spacing or even 25GHz spacing for up to 160 channel operation. For a small size network, where no optical amplifier is needed and only an Optical Supervisory Channel is used, the amount of noise is much less, simplifying the setting of the wavelengths to avoid cross-talk with the noisy channel carrying the bits for the key distribution protocols.

The communication protocol is presented in the next section and the operations performed by the control station will be discussed in Section VII. It is assumed that bits can be sent in runs of size $n(i)$, where i is the run index. It is important to realize

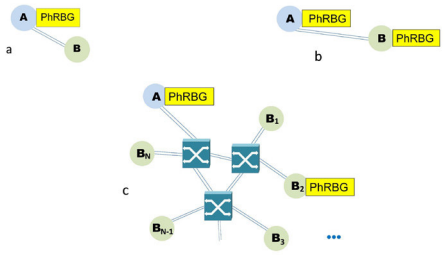


Figure 4. Some networking possibilities for the key distribution platform. a) is a configuration where only A possess a PhRBG whereas in b) both A and B have equal capabilities. In c), several receiving stations can be set for a single key-sender.

that the separate sender and receiver station capabilities shown in Fig. 2 and Fig. 4-(a) could also be set in a same station, where both A or B have emission and receiver capabilities Fig. 4-(b). In this case, there is no change in the logical procedures used, all arguments and explanations remain valid. Networking with more stations is also possible; see Fig. 4-(c).

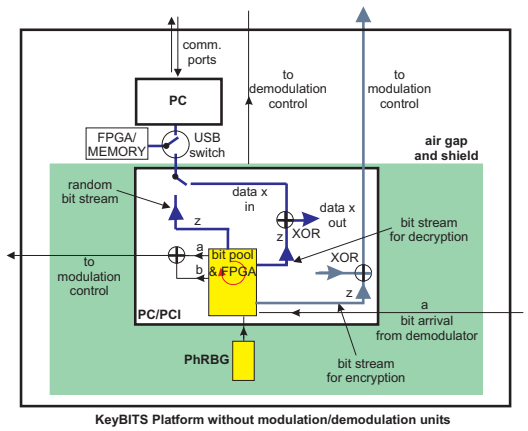


Figure 5. KeyBITS modular secure communications platform. See Fig. 2 for more details.

In Fig. 2 an "air-gap" is indcated that separates a region (bottom part) assumed free of undesired interferences of any kind, including reception or emission of electromagnetic and acoustical signals. All information going to the secluded region by the air-gap has to be monitored so that only controlled electronic signals are allowed in and out the air-gap. Details are not of interest at this moment but the main objective of the air-gap is to block a large number of attacks that could be launched against an open system.

In fact, separated emitter and reception stations in Fig. 2 may be together in both stations A and B, so that each one is autonomous, in a modular unit. Fig. 5 shows a modulus of this platform, without the modulation and demodulation units.

IV. The Physical Protocol

As shown in [1] and [2], and further discussed in Section III, the noise in the channel combined with the use of closely separated bases reduce enormously the probability of success of Eve. The signals she obtains do not allow her to obtain reliable bit sequences to be analyzed. This is the physical protection level in the key distribution scheme. However, this is not the only level of difficult existing in the system.

A first round of sending bits will be described to establish the basic ideas. A PhRBG continuously generates random bits a that can be processed in a bit pool with operations fastly processed by a FPGA or ASIC (Application-Specific Integrated Circuit).

Initially, this bit pool starts with the shared c_0 random bits, constituted of $c_0 = m\, n_0$ bits to create n modulation bases. For the sender, the total number of bits in the beginning of the process is then $n_0 + n_0 m$. It should be emphasized that despite the need of m bits to create a basis for modulate *one* bit, the process has been demonstrated to be very fast in hardware.

The choice of m depends on the physical choice of the bases to be used and the intensity of light, or average photon number $\langle n \rangle$ in the noisy optical channel (see [1] for explanations). For example, if optical *phase* values are used in a circle of 2π values, a choice of M values implies a distance of π/M between bases. A bit 1 could be represented by a phase value $\phi_1 = \pi$

while a bit 0 is given $\phi_0 = 0$ in one given basis n_M. In the other closest bases $n_M \pm 1$, the opposite choice is adopted, alternating ones and zeros.

The idea [1] is to set the light's noise such that it overlaps several physical bases. The choice of the number of bases M is based on a POVM –Positive Operator Valued Measure– which defines the probability of error given to the attacker (See Section V of [1]) and it will be directly connected to the average number of photons $\langle n \rangle$. Once M is defined, physical signals are generated creating a modulation (say, a phase ϕ) upon the laser beam. This physical modulation is being called an encryption basis for a fresh bit. To create *each* basis, m bits are necessary, $2^m = M$ or $m = \log_2 M$:

$$\begin{aligned} &\text{phase basis number(for } \phi_b = 0\frac{\pi}{M}, 1\frac{\pi}{M}, \cdots (M-1)\frac{\pi}{M}) \\ &\equiv b(m)2^{m-1} + b(m-1)2^{m-2} + \cdots b(1)2^0 \,. \end{aligned} \tag{5}$$

All bits of a generated by the PhRBG are represented by phase values ϕ_a (either ϕ_1 or ϕ_0) and added to the corresponding phase ϕ_b associated with the basis being used: $\phi_a + \phi_b$ and sent from A to B.

With $n\,m$ bits, n modulation bases b are created. They modulate the n fresh bits of a_0: $a_0 \oplus b$. As this sequence b is known to A and B, B could used it to demodulate the received sequence, extracting $a_0 = b \oplus (a_0 \oplus b)$.

As a brief comment, in the BB84 protocol, two bases are defined to send one bit that is carried by a single photon. The adversary must not know in which of the two bases the bit was encoded. In a parallel way, for the key distribution protocol discussed in this paper, the optical noise must protect against attempts by the adversary to know which basis was used.

Now, A and B share the sequence a_0. Eve may have obtained some *statistical* information t on these bits and A and B task is to eliminate t by PA – therefore, calculation of t is essential. This is shown ahead and with a numerical example in Fig. 6.

Another level of difficulty, computational, will be added – usually, this level of difficulty is used as a stand-alone protection level and may be sufficient by most of the cases even with noiseless signals. This mathematical level of protection will be discussed as well as the effect of combining these two protection levels.

Physically leaked bits – Assume that Alice sends n ($n =$ Length$[n_0]$) uniform random bits to B. Eve has complete access to the transmission channel, close to the sender, where the signal is maximum, and disposes of the ideal equipment, subjected to the laws of Physics, to measure and record all emitted signals. No one monitors Eve's intrusion and will not constrain her endeavor in anyway. However, Alice will not send bits in a repeated way; every bit information is sent *only once*. This way, Eve obtains noisy signals representing the sent bits and will treat them individually or collectively, as she pleases.

As shown by the POVM calculation in Section V of [1], there is a minimum probability of error P_e for Eve when measuring any bit due to the inherent noise in the optical channel and the M-ry bases used. P_e is a function of the average photon number $\langle n \rangle$ in the signal representing a bit and M, the number of bases used in the M-ry communication protocol. The POVM calculation utilizes the wavefunction or density matrix representing *all* information about the transmitted bit. This is the maximum amount of information available about a physical system. The result of the calculation indicates the best Eve could obtain, even with an ideal measuring system and analyzing capabilities.

For numerical examples of these results, see Fig. 3 in [1]. The probability of having a correct bit assignment by Eve is $P_r = 1 - P_e$. Therefore, Eve is able to *statistically* assign correctly or "extract" $t \equiv t_{\text{bit}} = P_r - 0.5 = 0.5 - P_e$ of each bit.
Therefore, t_{bit} is a extraction rate of Eve (or, leak per bit).

Fig. 6 exemplifies the behavior of $\log_{10} t_{\text{bit}}$ as a function of M. As will be shown in Section VII, Eves probability to obtain

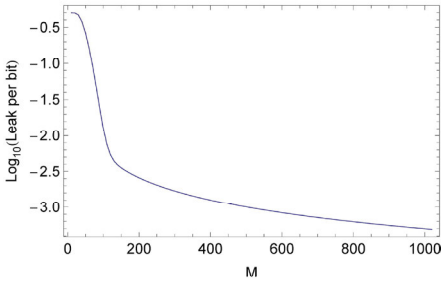


Figure 6. Logarithm of the loss per bit to Eve, t_{bit}, as given by the POVM calculation that provides P_e. Here $\langle n \rangle = 1000$.

all bits in a sequence is completely negligible.

V. Signal Modulation and Demodulation

The left side of Fig. 2 shows a modulation system (at Alice's station) that injects signals in the "noisy" channel of the optical fiber. At the right side of the same figure (at Bob's station) a demodulation system extracts the signals sent by Alice, erasing the signals representing the encoding bases that produce the indistinguishability of the signals to the attacker.

The modulation and demodulations systems are discussed in [10] and will not be discussed here. The final signal $\Delta i = i_e - i_f$ representing the bits as extracted by Bob, come from the two pin detectors in the demodulation system and are proportional to the streams of photons

$$\langle n_e \rangle = -\frac{1}{4}|\alpha|^2\left[\sqrt{3}\sin\varphi\cos^2\left(\frac{\Delta}{2}\right) + \sqrt{3}\cos\varphi\sin\Delta - 2\right] \tag{6}$$

$$\langle n_f \rangle = \frac{1}{4}|\alpha|^2\left[\sqrt{3}\sin\varphi\cos^2\left(\frac{\Delta}{2}\right) + \sqrt{3}\cos\varphi\sin\Delta + 2\right], \tag{7}$$

where Δ is the path phase difference between the two arms of a fiber Michelson interferometer in the bit extractor (see Fig. 2).

Fig. 7 shows plots for the direct currents i_e and i_f for a given laser intensity (arbitrarily taken at $|\alpha| = 10$, and $G = 1$, $\eta_d = 1$ and a unitary time interval. Δ, set by the piezoelectric driver, is set at $\Delta = \pi/2$); G is the detector's electronic *gain* and η_d is the detector's *efficiency* in the photon-to-electron conversion. It is seen that the best resolution for bits 0 and 1 is obtained from the difference of the two currents, or $\Delta i = i_f - i_e$ and not from either current outputs i_e or i_f alone.

As was shown, physical noise can create a physical barrier to the attacker making it impossible for him/her to extract clean bit signals from the channel. At the same time the legitimate users

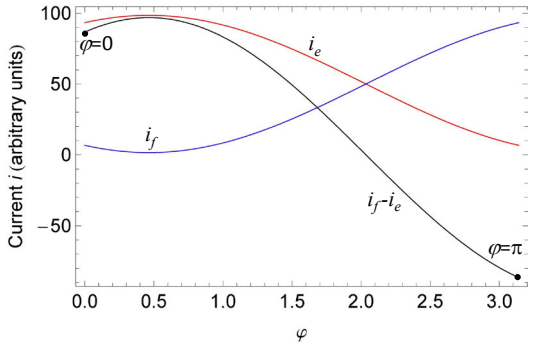


Figure 7. Photon currents i_e and i_f obtained from the pin diodes and the final difference current $\Delta i = i_f - i_e$ as a function of the input phase ϕ. Particular values are indicated with $\phi = 0$ and $\phi = \pi$ that represent bits 0 and 1.

extract clean signals and obtain the bits sent from the current at the demodulator, as given by Fig. 7.

These different results for the attacker and for the legitimate user can be understood by the overlap between two states, as given by Eq. (2). For the legitimate user he has only to distinguish between a bit 0 and a bit 1 encoded by a known basis value b, within M possible values. For him, $\Delta\phi = \pi$ and, therefore

$$\langle \Psi(\phi) | \Psi(\phi + \pi) \rangle = e^{-2|\alpha|^2} \to 0\,. \tag{8}$$

This result shows a comparison between two almost orthogonal states, of easy identification, with negligible overlap. On the contrary, not knowing the basis value the adversary has a complex measurement problem that limits his/her knowledge according to what was shown by the POVM calculation in [1].

VI. Increasing the protection level

The security of the proposed key distribution system does not stop at this physical barrier but has its security further strengthened by Privacy Amplification (PA). The final security then rests on a combination of physical and mathematical protections. Each of these aspects can be calculated as well as the amount of information that the attacker might have obtained about the final bit sequences being shared by A and B. This provides a rigorous proof of the security of the system. The adopted PA protocol will be discussed in Section VII.

In order to protect the communication on the classical channel from tampering, the system also uses a Message Authentication Code (MAC) to guarantee that the messages sent between Alice and Bob are authentic, i.e. where not sent by someone else. The used MAC is Galois Counter Mode. In conventional counter mode (CTR), pseudo-random blocks are generated by incrementing a counter and encrypting each result. The resulting blocks are then used as a one-time pad key, so each ciphertext block is the xor of a key block and a message block. Galois Counter Mode is an extension of counter mode which consists of computing a cryptographic CRC of 128 bits by performing an xor and a multiplication over $GF(2_{128})$ (hence the name) of the ciphertext block and some constant derived from authentication information. Obviously, nothing in GCM prevents us from substituting blocks pseudo-randomly generated from some block ciphered by keyblocks originating from a true random process; GCM also works with the one-time pad.

TABLE I

Privacy Amplification protocol for the KeyBITS platform

PA protocol		
INITIALIZATION: A and B share c_0 of size and entropy $m\,s$.		
ALICE		
#	ACTION	COMMENT
1a	$a_i =$ GetString(PhRBG)	get bitstring from PhRBG
1b	$b_i = c_{i-1}[1, ms]$	extract $m\,s$ from pool for bases b
1c	Code&Send(a_i, b_i)	send over $\alpha\eta$ channel
2	SendCC(f)	send instance of universal hash f over classical channel
3a	$c_i = f(c_{i-1}\|\|a_i)$	Alice applies PA from $ms + s$ bits to $ms + s - t - \lambda$
3b	$c_i = f(c_{i-1}\|\|a_i)$	Alice uses $\overline{s} = s - t - \lambda$ bits from pool as the key stream z. The remaining $m\,s$ bits form the bases' bits for next round.
BOB		
1a		no matching step to Alice's
1b	$b_i = c_{i-1}[1, ms]$	get bases bits from initial pool value
1c	$a_i =$Receive&Decode(b_i)	receive bits from $\alpha\eta$ channel
2	ReceiveCC(f)	receive instance of universal hash f
3a	$c_i = f(c_{i-1}\|\|a_i)$	Bob applies PA from $ms + s$ bits to $ms + s - t - \lambda$
3b	$z_i = c_i[ms+1, ms+s-t-\lambda]$	Bob uses $\overline{s} = s - t - \lambda$ bits from pool as the key stream z. The remaining $m\,s$ bits form the bases' bits for next round.

VII. Privacy Amplification protocol

Before discussing the PA protocol, is should be emphasized that although the physical protocol uses a somewhat larger number of bits ($\log_2 M$) to encode *one* bit, the process is continuously sustained in rounds of s bits, in an unlimited way. This process has been shown to be very fast in hardware. The Privacy Amplification protocol adopted uses a bit pool of constantly renewed random bits. For details, see [6]. Fig. 2 can be used as a reference for description of the PA protocol. Before discussing the level of security, a summary of the PA protocol steps is given in Table I. After this summary, conditions for its applicability will be discussed.

The first round of the protocol will be described in words:
INITIALIZATION: Alice and Bob share a starting sequence c_0 of secret random bits. The sequence has size $c_0 = m\,s$, where m is the number of bits necessary to describe one of the M basis and s is the size of the first fresh sequence of random bits to be shared between A and B.

Alice first steps –

A1a: Alice gets a random bitstring of length s in the bit pool fed from the PhRBG.

A1b: Alice gets the shared starting sequence c_0 and partition it in s parts with m bits each. Each subsequence of length m randomly specifies one basis among the M bases.

A1c: Alice encodes each bit in s with the corresponding basis and sends the signal to Bob over the noisy channel. See Section IV for a description of the physical modulation to be used. Be-

forehand, Alice and Bob had agreed on the PA's security parameter λ and calculated the statistical fraction t of a bit ($t \equiv t_{\text{bit}}$) leaked to Eve, see Fig. 6 and Eqs. (7) to (12) in Ref.[1].
A2: Alice sends an instance of a universal hash function $f(\in \mathcal{F})$ to Bob over a noiseless channel with public acess.
A3a: Using f Alice applies PA and reduce the total number of bits $s+ms$ to $\overline{s} = s+ms-t-\lambda$. The eavesdropper has *no* knowledge on $\overline{s}$ or on the modified ms sequence. See [6] for details.
A3b: The reduced sequence $\overline{s}$ is the distilled fresh random sequence z to be used for OTP encryption. The remaining fresh random sequence $m\,s$ will form the bases for the next run.
Bob first steps –
B1a: There is no corresponding step to Alice's 1a.
B1b: Bob gets the shared starting sequence c_0 and partitions it in s parts with m bits each. Each subsequence of length m randomly specifies one basis among the M bases. Bob and Alice are then using the same set of bases.
B1c: Bob receives the physical signals sent by Alice, demodulates them (See Fig. 2) and obtain the random bit stream coded with the sm bases. As Bob knows the bases coding, he decodes the random stream and obtains the stream s sent by Alice.
B2: Bob receives f over the classical channel.
B3a: Using f Bob applies PA and reduce the total number of bits $s+ms$ to $\overline{s} = s+ms-t-\lambda$.
B3b: The reduced sequence $\overline{s}$ is the distilled fresh random sequence z to be used for OTP encryption. The remaining fresh random sequence $m \times s$ will form the bases for the next run. Therefore, both Bob and Alice share a secure sequence of random bits $\overline{s}$ to be used as OTP.

This means that the generated stream from the PhRBG at Alice's station was transferred to Bob and a distilled secure sequence of random bits and base bits is obtained.
The protocols proceeds to next similar runs. After n runs, Alice and Bob share $n\overline{s}$ bits.

Preliminary conditions for the PA protocol – The protocol for Privacy Amplification [14] (or PA) offers a powerful tool to decrease the amount of information an adversary (E) might have acquired on a bit string transmitted from one legitimate user (A) to a second one (B).

In this paper, A sends n random bits to B, from which E is able to statistically gain t_{bit} of information per pulse sent. The amount of gained information by the adversary over a string of n randomly distributed bits is t^n_{bit}.

Among quantities or conditions that the PA protocol need to be applicable are the statistical gain t_{bit} as well as a bound on the second-order conditional Rényi entropy, $R_2(n|V=v)$, as seen by the adversary E. Here V designate the variable under control of E that has some degree of correlation to n.

Some preliminary steps may help to calculate $R_2(n|V=v)$. The "collision probability" for a variable X specified by a probability distribution$P_X(x)$ can be defined (see [14] for details) as

$$P_c(X) = \sum_{x\epsilon\chi} P_X(x)^2 \ , \tag{9}$$

from which the Rényi entropy of X can be calculated:

$$R(X) = -\log_2 P_c(X)\,. \tag{10}$$

The *entropy* in binary digits is also given in "bit" units, and may be fractionary, differently from the physical bits 0 or 1 (encoded physical signals). Given an event $\mathcal{E}$ on X, with conditional probability $P_{X|\mathcal{E}}$, one may directly write the collision probability $P_c(X|\mathcal{E})$ and the conditional Rényi entropy

$$R(X|\mathcal{E}) = -\log_2 P^2_{X|\mathcal{E}}\,. \tag{11}$$

The variable X of interest should represent the bit stream transmitted from A to B and the event $\mathcal{E}$ represents Eves access to that stream.

The mapping of $P_{X|\mathcal{E}}$ in terms of the physical processes gives

$$P_{X|\mathcal{E}} \to P_r = 1 - P_e\,, \tag{12}$$

and

$$R(X|\mathcal{E}) = -\log_2 P^2_{X|\mathcal{E}} = -\log_2 (1-P_e)^2 \tag{13}$$

Fig. 8 shows an example of the collision Rényi entropy given by Eq. (13) for the case of bits being transmitted with an average number of photons $\langle n\rangle = 1000$ per bit as a function of the number M of bases used. The asymptotic limit for Rényi entropy is

$$R(X|\mathcal{E}) = -\log_2 (1-P_e)^2 \to 2 \ /\text{per bit}\ , \tag{14}$$

for $P_e \to 1/2$. This result can be interpreted as follows: For large M, Eve succeeds in obtain *one* collision, or statistical-success, in every two trials. Due to the uniformity of the random

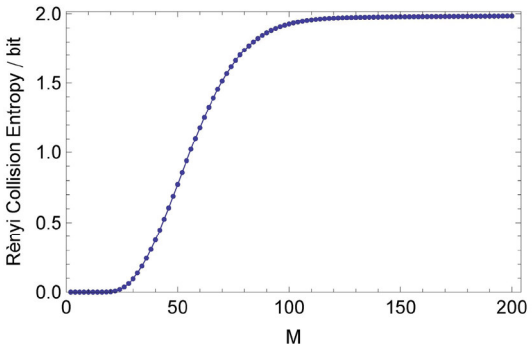


Figure 8. Example of the Conditional Rényi collision entropy of order two for bits being transmitted with $\langle n\rangle = 100$/bit signal as a function of the number of bases M.

sequences, for a stream of n bits the Rényi entropy will give the corresponding limit $R(n|\mathcal{E}) \to 2\times n \equiv c$.

The PA protocol using a compression function G within a universal class of hash functions maps the received stream $\{1,0\}^n$ onto $\{1,0\}^r$. Assuming that A and B uses $z \equiv \{1,0\}^r$ as their secret stream of bits, it is known that [14]

$$H\,(z|G,V=v) \geq r - \frac{2^{r-c}}{\log_e 2}\,. \tag{15}$$

Therefore, as $r<n$ and $c=2n$, then $r<c$. Eve's entropy on the keys is

$$H\,(z) - H\,(z|G,V=v) \simeq \frac{2^{r-2n}}{\log_e 2}\,, \tag{16}$$

and goes exponentially to zero as n increases. In conclusion, the string of r random bits can be protected by the PA protocol.

Fig. 9 exemplifies Eq. (16) for a range of n and r values.

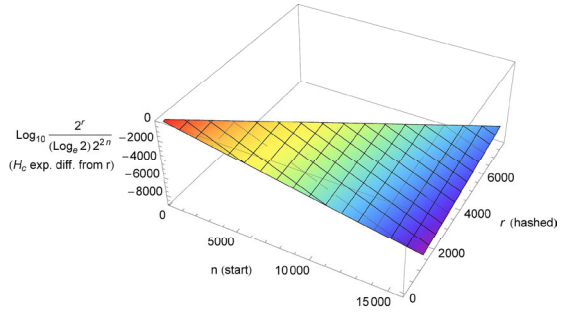


Figure 9. An example that shows that Eve's amount of information on the hashed key stream is negligible (see Eq. (16)).

A. Overall security

Having demonstrated the possibility for application of the PA protocol [14] for this key distribution scheme, one may invoke corollary 5 of the PA theorem. In words, the expected information of Eve about the secret key (assume length r_t) is given by the mutual information I on the secret key given the information t acquired by Eve when Alice and Bob use a randomly chosen function from a universal class of hash functions:

$$I \leq \frac{1}{\ln 2 \times 2^{\lambda}} \ , \tag{17}$$

where λ is a security parameter $\lambda < r_t - t$. This way, by eliminating t bits of r_t, Eve's information decreases exponentially while Alice and Bob knows $\overline{s} - 1/(\ln 2 \times 2^{\lambda})$ bits.

In a sequence of s bits sent, two factors will work against Eve. The first one is the effect of the noisy channel on her measurements and the second one is the result of applying the Privacy Amplification protocol.

Using the probability of an error by Eve, P_e, as shown in Section V of Reference [1], to correctly guess a particular bit sent through the noisy channel, the "hit" probability t_1 is $t_1 = 1 - P_e$. This says that Eve's probability of obtaining all s bits is $t_s = (1 - P_e)^s$, because the keys are uncorrelated as well as the physical signals that carry them. This probability gives Eve a negligible chance of success.

The legitimate users may adopt the strategy of defining the key sequence length s such that after sending *all* of them, statistically the adversary could have gained less than one bit, that is to say $s(1 - P_e) < 1$. In other words, the legitimate users choose $s < 1/(1 - P_e)$. This says that the amount of $t + \lambda$ bits to be reduced from $s + ms$ (see Step 3a in Table I) would be $t + \lambda \simeq \lambda$.

The physical noise in the channel basically reduces enormously the amount of information that Eve could obtain from the channel. On the other way, if A and B use long sequences such that $s(1 - P_e) \gg 1$, a larger number of bits have to be added to λ to make effective the PA protocol. Therefore, from now on Eq. (A) will be adopted as a condition to establish the lengths of the key sequence runs.

After A and B have applied the PA protocol, together with the effects of the noisy channel, the number of bits is reduced from s to r distilled bits. The information obtained by Eve is given by

$$I_r \simeq \frac{1}{\ln 2 \times 2^{\lambda}} \ . \tag{18}$$

Fig. 10 gives an example of Eq. (18).

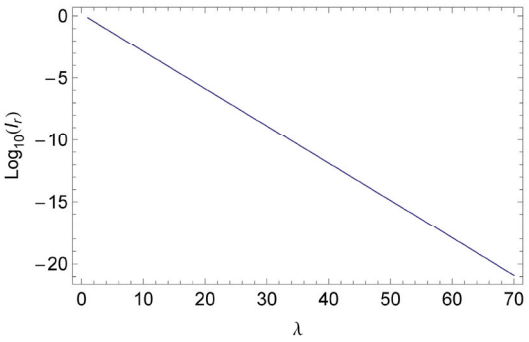


Figure 10. Eves information on r bits after Alice and Bob applies the PA protocol on the bit stream obtained from the noisy $\alpha\eta$ channel under the condition $s(1 - P_e) < 1$.

VIII. Adversary with information on the bit pool?

It may be argued that even if the adversary tries to obtain the basis that has encoded a bit and fails, some exclusive information on the bases' wheel is gained. This information may be seen as a set of bases to be *excluded* from the bit pool and, therefore, will simplify a posterior analysis. Therefore, the question *"Will this gained information increases Eve's knowledge on the bit pool for posterior analysis on a reduced set of unknowns?"*

This question may be answered with the Mutual Information between B (or A) and E, $I(B;E)$. More specifically, assuming that A nd B utilized (secretly) a given basis ϕ_b to encode a bit b, 0 or π, what will be the Mutual Information $I(\phi_b;\phi_E)$, where ϕ_E is Eve's estimated value obtained from an arbitrary measurement?

First of all, the Mutual Information will be calculated to reveal the amount of information the adversary could obtain from a bit *only* considering the optical noise effect on the mesoscopic signal. This absolute measure could be compared with Minimum Probability of Error by Eve P_e^E already calculated in [1]. As discussed, any small amount of information leaked by the channel can be privacy amplified.

A. Mutual Information

In order to write the Mutual Information

$$I(X;Y) = H(X) - H(X|Y) \tag{19}$$

on the desired variables, one may start with the relationships

$$H(X|Y) = \sum_{x,y} p(x|y) \log_2 \frac{1}{p(x|y)} \tag{20}$$

$$H(X) = \sum_x p(x) \log_2 \frac{1}{p(x)} \to \sum_{k=0}^{M-1} \left(\frac{1}{M}\right) \log_2 \frac{1}{\left(\frac{1}{M}\right)} = \log_2 M \,. \quad (21)$$

Therefore,

$$I(X;Y) = \log_2 M - \sum_{k,k_E} p(k|k_E) \log_2 \frac{1}{p(k|k_E)} \,. \quad (22)$$

Eq. (2) gives the un-normalized Conditional Probability

$$p(k|k_E) = e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]} \,. \quad (23)$$

The notation designates an angle set by the legitimate users as $\phi = k\pi/M$ and an angle set by the adversary as $\phi_E = k_E\pi/M$. The normalized form of the Conditional Probability will be written

$$p_{\text{norm}}(k|k_E) = \frac{e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]}}{\sum_{k=0}^{M-1}\sum_{k_E=0}^{M-1} e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]}} \quad (24)$$

The Mutual Information can be written, in normalized form, for specific values of the phases ϕ and ϕ_E, using $p_{\text{norm}}(k|k_E)$. One obtains

$$I(X=\phi;Y=\phi_E) = \frac{1}{M}\log_2 M - \frac{e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]}}{\sum_{k=0}^{M-1}\sum_{k_E=0}^{M-1} e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]}} \times \log_2 \left[e^{|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]} \times \sum_{k=0}^{M-1}\sum_{k_E=0}^{M-1} e^{-|\alpha|^2[1-\cos[(\pi/M)(k-k_E)]]} \right] \,. \quad (25)$$

One may interpret $I(X=\phi;Y=\phi_E)$ as the average reduction in uncertainty about ϕ when Eve in some way learns ϕ_E. As the k values are uniformly distributed the entropy of ϕ (or k) is

$$H(k) = \frac{1}{M}\log_2 \frac{1}{\frac{1}{M}} = \frac{1}{M}\log_2 M \,, \quad (26)$$

The relative reduction in uncertainty obtained by Eve can be quantified by

$$r_{I/H} \equiv \frac{I(k;k_E)}{H(k)} \quad \left(0 \le r_{I/H} \le 1\right) \,. \quad (27)$$

Fig. 11 shows $I(k=20;k_E)/H(k=20)$ for $|\alpha|^2 = \langle n\rangle = 100$ and two set of bases, $M = 100$ and $M = 200$. The value 20 is arbitrary, as any other value gives similar results.

Going back to the question *"Will this gained information increases Eve's knowledge on the basis or bit sent?"*, one concludes that she gains some information around the basis value k used by the legitimate users but no knowledge for distant values from k. Therefore, the small amount of information gained by Eve just by excluding bases numbers distant from the basis value set, will be *irrelevant* after PA methods reduce this leaked information to a negligible level. Therefore, the amount of Eve's useful information about the bit pool is negligible.

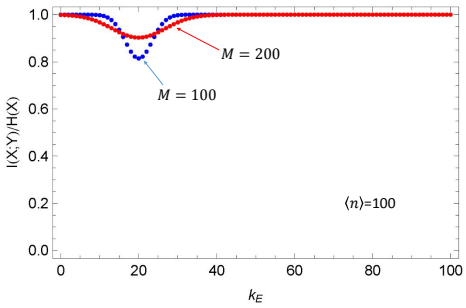


Figure 11. Example of the relative information $I(k;k_E)/H(k)$ gained by Eve on $H(k)$ when she learns k_E from her measurement. When $I(k=20;k_E)/H(k=20) \to 1$ she gains no information on x, what occurs for most of the bases values. Only when k_E is set close to k she acquires some information. However, due to the physical noise that produces the physical uncertainty on k_E (or ϕ_E), *even* with $k_E = k$ her knowledge does not give her the desired $I(k=20;k_E)/H(k=20) \to 0$ but only a limited gain. Her overall probability of success is given by $(1-P_e)$.

IX. Conclusions

The key distribution system introduced in [1] was revisited and improved with inclusion of a specific protocol for the key distribution that includes both the noise protected step and a PA protocol. The PA protocol uses the bit pool shown in Fig. 2.

It was shown that starting with a shared sequence of $n_0\, m$ random bits to form physical bases, A and B can distribute in a secure way an *unlimited* number of secure bits generated by a PhRBG generator. The overall security of the key distribution depends on signal to noise ratio in the transmitted signals, the number of bases M used and the shuffling produced by the PA protocol.

The key distribution process is a "one-time-pad booster" which allow users to use bit-by-bit encryption for *top security level* applications and, at the same time, allows fast encryption of large volumes of information. When working in fiber-optic channels, the system demands, besides the use of a true physical random bit generator working at high speeds, analog-to-digital, digital-to-analog converters, optical modulators and separate channels to avoid perturbation from ordinary signal channels. In optical channels with signals of mesoscopic intensities, the system presents two layers of protection, physical noise and computational difficulties, such as the one exemplified by PA with universal hash functions. The system can also be used with classical signals in generic channels using only the computational difficulty guaranteed by the PA protocol when the protection given by the optical noise is not present (noiseless channels). When used in optical channels both security levels are present.

X. Acknowledgements

We acknowledge the support of Ministério da Ciência, Tecnologia e Inovação (MCTI)-Finep(0276/12)-Fundep(19658)-Comando do Exército(DCT)-RENASIC.

APPENDICES

A. PRIVACY AMPLIFICATION - TOEPLITZ MATRICES

Privacy Amplification can be used to increase the security level already given by the physical noise in the channel. The PA procedure establishes [14] that once A have sent to B a sequence a of bits ($n = \text{Length}[n_0]$), if Eve obtains an estimated number of st bits, A and B could reduce the amount of Eve's information.

To achieve this end, A and B need to agree on a procedure that results in a shorter secure string a' on which Eve has an exponentially vanishing knowledge. This procedure demands that the number initial bits n has to be reduced by st and even further by a security parameter λ (in bits) to guarantee that Eve can obtain at most $1/\ln(2)2^{\lambda}$ bits (see Section IV of [14]) on a'. These operations can be performed on the bit pool shown in Fig. 2. The procedure to reduce a to a' may use a hash function $\mathcal{H}$: $\mathcal{H}a = a'$, where $\mathcal{H}$ is a matrix with random elements.

Among the several possible PA choices and for $\mathcal{H}$ - of random elements, one could choose a matrix where all elements are randomly and independently chosen or even constructed with a starting set of randomly chosen elements. This matrix can be renewed at each distribution bit round for maximum security.

Just to explain the PA process with renewed matrix elements for increased security, an example using a Toeplitz matrix will be presented. A Toeplitz matrix has a simple structure of form

$$\mathcal{H} = \begin{pmatrix} r_1 & c_2 & c_3 & c_4 & \dots \\ r_2 & r_1 & c_2 & c_3 & \dots \\ r_3 & r_2 & r_1 & c_2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \tag{28}$$

where r_i and c_j are binary random digits taken from the PhRBG. The number of columns should be equal to the number of fresh bits n to be transmitted plus the number of bits $n\,m$ (secretely shared by A and B) to generate the modulation bases for the transmitted bits. The number of rows is equal the $n + n \times m - nt_{\text{leak}} - \lambda$. The number of bits in a column is the same number of bits in a. This way, an input bit stream, or a vector with $n(m+1)$ components (bits) gives an output of $n(m+1) - nt_{\text{leak}} - \lambda$ bits, which is the desired reduction in the number of bits for A and B such that Eve has a negligible knowledge on them.

What is the number of secure bits finally available for OTP? – What was just described was a PA protocol applied for a first run starting with $n_0(m+1)$ bits. The output number of bits n_0 was reduced to $n_1(m+1) - n_1 t_{\text{leak}} - \lambda$ bits.

As shown in the PA protocol (Section VII), after the first round both Alice and Bob share a distilled sequence of $\overline{s}$ secure bits to be used as OTP and still have a shared fresh sequence of bases bits $s\,m$.

The process is *unlimited* in number of runs and will be as fast as the current technology allows because there are no fundamental physical limitations in the bit generation occurring in the PhRBG.

B. STOKES PARAMETERS OF A NOISY FIELD

The phase modulation specified by Eq. (1)

$$|\Psi(\phi)\rangle = e^{-iJ_z\phi}|\Psi_0\rangle = |\frac{\alpha}{\sqrt{2}}e^{-i\phi/2}\rangle_x|\frac{\alpha}{\sqrt{2}}e^{i\phi/2}\rangle_y\,, \tag{29}$$

is imposed as a phase difference between two orthogonal polarization components represented by annihilation operators a and b, representing fields of equal intensity. This form is due to the optical modulator being considered. As the imposed electric field (assumed of weak intensity to avoid non-linear effects) travels along the optical fiber, it undergoes randomized polarization fluctuations in direction caused by several somewhat localized effects that modifies the dielectric constants of the supporting glass medium. These effects include thermal fluctuations, acoustic modes, Mie scattering, mechanical stresses,.... The demodulation system represented at the right in Fig. 2 subtracts the base modulation effects regardless these random contributions, by operating on two arbitrary polarization components.

One may question if the adversary, Eve, would be able to perform phase measurements close to the emitter, such that these complicating perturbations have not taken an appreciable effect yet. Her goal is to extract precise phase information such that she could *resolve* the angular separation $\Delta\phi_1$ between two closest bases k and $k+1$. In general, writing $\Delta\phi$ in terms of base indexes k, k', one has $\Delta\phi = (\pi/M)(k-k')$. The question is "*what is the effect on the inherent optical shot-noise on her measurements?*". An equivalent but more precise question would be "*what is the maximum resolution on $k-k'$ possible to be achievable by Eve given an average number of photons $\langle n\rangle$ and a number of bases M?*"

In order to answer this question, one may start recalling some adequate tools such as Stokes parameters and the Poincaré sphere. In order to understand transformations of an optical medium or a device on any incoming light mode, it is useful to depict the Poincaré sphere of polarizations (See Fig. 12) and to write the incoming polarized electric field in terms of the variables for this sphere. A polarization state is represented by

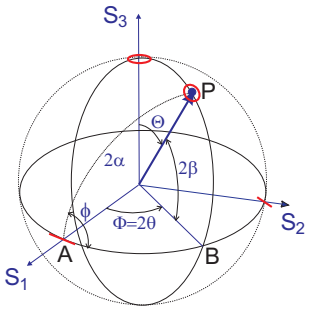


Figure 12. Θ and Φ are the polar and azimuthal angles that indicate a point P on the sphere. Polarizations on the equator of this sphere represent linearly polarized states with different inclination angles. For example, a point on $\Phi = 0, \Theta = \pi/2$ represent a linearly polarized state along x, for light propagating along z, while its antipodal point $\Phi = \pi, \Theta = \pi/2$ represents a linearly polarized state along y. Similarly, a point $\Theta = 0$ represents a (+) circularly polarized light state and point $\Theta = \pi$ a (-) circularly polarized light state. Arbitrary points, like P, represent elliptical polarization states.

a point on the Poincaré sphere given by the two coordinates Θ (polar angle) and Φ (azimuthal angle). The Poincaré sphere has

its radius defined by the Stokes' parameter s_0 (equal to the *intensity* of the polarized light) and the projections of the point (Θ, Φ) on the orthogonal axes S_1, S_2 and S_3. These projections have values s_1, s_2 and s_3 and are known as Stokes' parameters (Stokes, 1852) [16]. Therefore, the Stokes parameters describe a general polarized light state.

Physical analyzers such as the crystal axis of a polarization beam splitter or wave-plates produce field projections onto the Poincaré's sphere axes and allow for photon number or intensity measurements, leading to s_0, s_1, s_2, s_3.

Operations and operators describing these measurements are known either in the classical or quantum domain. In the quantum domain these parameters are given by the expectation values of the Hermitian operators of the number operator $\widehat{N}$ and of the total angular momentum of light as described by Schwinger in terms of two bosonic modes given by annihilation operators $\widehat{a}$ and $\widehat{b}$

$$\widehat{s}_0 = \widehat{a}^\dagger\widehat{a} + \widehat{b}^\dagger\widehat{b} = \widehat{N} \tag{30}$$
$$\widehat{s}_1 = \widehat{a}^\dagger\widehat{a} - \widehat{b}^\dagger\widehat{b} = 2\widehat{J}_z = \widehat{\sigma}_z \tag{31}$$
$$\widehat{s}_2 = \widehat{a}^\dagger\widehat{b} + \widehat{b}^\dagger\widehat{a} = 2\widehat{J}_x = \widehat{\sigma}_x \tag{32}$$
$$\widehat{s}_3 = \frac{1}{i}(\widehat{a}^\dagger\widehat{b} - \widehat{b}^\dagger\widehat{a}) = 2\widehat{J}_y = \widehat{\sigma}_y \ , \tag{33}$$

from which $\langle\widehat{s}_0\rangle = s_0$, $\langle\widehat{s}_1\rangle = s_1$, $\langle\widehat{s}_2\rangle = s_2$, $\langle\widehat{s}_3\rangle = s_3$.

At this point it is interesting to note that Eqs. (30) to (33) define operators J_i, $(i = x, y, z)$ in terms of boson operators a and b (hats will be ignored from now on) and that they obey the same commutation properties as the ones connected with angular momentum: $[J_i, J_k] = i\epsilon_{ijk}J_k$. This leads to the conservation of the total number of photons n as they go through a lossless optical device: $n = n_a + n_b = a^\dagger a + b^\dagger b$.

Standard procedures to perform these measurements have been well established [16] and, from the experimental side, even automated measuring systems can be found to perform these tasks. For example, designating an intensity by I and by x the horizontal axis H and by y the vertical axis V, and by the indexes R and L, circular states of light, one could write

$$s_0 = I_H + I_V = a^2 + b^2 \tag{34}$$
$$s_1 = I_H - I_V = a^2 - b^2 = s_0\cos(2\beta)\cos(2\theta) \tag{35}$$
$$s_2 = I_{45^0} - I_{-45^0} = 2ab\cos\phi = s_0\cos(2\beta)\sin(2\theta) \tag{36}$$
$$s_3 = I_R - I_L = 2ab\sin\phi = s_0\sin(2\beta) \ . \tag{37}$$

See Fig. 13 for definitions. Light noise associated to a source

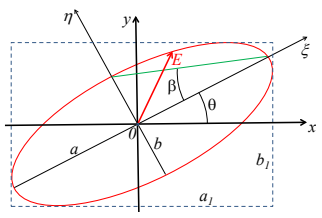


Figure 13. A polarization state of light represented by an ellipse, with principal axes along ξ and η, depicted with the main axes x and y. β can be either positive or negative, giving the senses in which the ellipse may be described (See Fig. 12, top or bottom half hemisphere).

of light cannot be eliminated and noise-to-signal ratio cannot be be rendered arbitrarily negligible. The influence of the noise on the propagated signals can be explored in several ways, including for cryptographic purposes. The associated error in these measurements are by far less established and belong to the quantum research realm ([17], [18]). Generalized quantum measurements have been applied to separate deterministically two nonorthogonal quantum states and add a necessary set of inconclusive results [19]. The subject is a permanent area of research [20].

To see the effect of noise, one may calculate $\langle\widehat{J}_z\rangle$, $\langle\widehat{J}_x\rangle$, $\langle\widehat{J}_y\rangle$ and the associated variances $\sigma_z^2 = \langle\left(\widehat{J}_z - \langle\widehat{J}_z\rangle\right)^2\rangle$, $\sigma_x^2 = \langle\left(\widehat{J}_x - \langle\widehat{J}_x\rangle\right)^2\rangle$, $\sigma_y^2 = \langle\left(\widehat{J}_y - \langle\widehat{J}_y\rangle\right)^2\rangle$.

Observing that $\langle(J_i - \langle J_i\rangle)\,(J_k - \langle J_k\rangle)\rangle = \langle J_iJ_k\rangle - \langle J_i\rangle\langle J_k\rangle$, the quantities $\langle J_iJ_k\rangle$ and $\langle J_i\rangle$ have to be calculated for the x, y, z components. Expansion of the J_iJ_k products in normal order and application of the operators on the wave-function given by Eq. (1) is straightforward. One obtains

$$\langle\psi|J_xJ_x|\psi\rangle = \frac{\langle n\rangle}{8}\left[2 + \langle n\rangle\left(1 + \cos(2\phi)\right)\right] \tag{38}$$
$$\langle\psi|J_xJ_y|\psi\rangle = \frac{\langle n\rangle}{8}\sin(2\phi) \tag{39}$$
$$\langle\psi|J_xJ_z|\psi\rangle = -i\frac{\langle n\rangle}{4}\sin\phi \tag{40}$$
$$\langle\psi|J_yJ_x|\psi\rangle = \frac{\langle n\rangle}{8}\sin(2\phi) \tag{41}$$
$$\langle\psi|J_yJ_y|\psi\rangle = \frac{\langle n\rangle}{8}\left[2 + \langle n\rangle\left(1 - \cos(2\phi)\right)\right] \tag{42}$$
$$\langle\psi|J_yJ_z|\psi\rangle = i\frac{\langle n\rangle}{4}\cos\phi \tag{43}$$
$$\langle\psi|J_zJ_x|\psi\rangle = i\frac{\langle n\rangle}{4}\sin\phi \tag{44}$$
$$\langle\psi|J_zJ_y|\psi\rangle = -i\frac{\langle n\rangle}{4}\cos\phi \tag{45}$$
$$\langle\psi|J_zJ_z|\psi\rangle = \langle n\rangle\,, \tag{46}$$
$$\langle\psi|J_x|\psi\rangle = \frac{\langle n\rangle}{2}\cos\phi, \langle\psi|J_y|\psi\rangle = \frac{\langle n\rangle}{2}\sin\phi, \langle\psi|J_z|\psi\rangle = 0 \tag{47}$$

The ratio $\langle\psi|J_y|\psi\rangle/\langle\psi|J_x|\psi\rangle$ of expected values of the Hermitian operators would give a measure of $\tan\phi$ and, therefore, of ϕ – if not for the deviations produced by the light noise. Considering these deviations one has

$$\tan(\phi \pm \Delta\phi) = \frac{\langle\psi|J_y|\psi\rangle \pm \sigma_y}{\langle\psi|J_x|\psi\rangle \pm \sigma_x} = \frac{\sin\left[\frac{\pi}{M}k\right] \pm \frac{1}{\sqrt{\langle n\rangle}}}{\cos\left[\frac{\pi}{M}k\right] \pm \frac{1}{\sqrt{\langle n\rangle}}}\,, \tag{48}$$

where ϕ was written in the discrete set of M k phase values. In order to get the extrema of $\tan(\phi \pm \Delta\phi)$ one writes:

$$\tan\phi_{\text{Max}} = \frac{\sin\left[\frac{\pi}{M}k\right] + \frac{1}{\sqrt{\langle n\rangle}}}{\cos\left[\frac{\pi}{M}k\right] - \frac{1}{\sqrt{\langle n\rangle}}}, \tan\phi_{\text{min}} = \frac{\sin\left[\frac{\pi}{M}k\right] - \frac{1}{\sqrt{\langle n\rangle}}}{\cos\left[\frac{\pi}{M}k\right] + \frac{1}{\sqrt{\langle n\rangle}}}$$

Fig. 14 shows Eqs. 49 and 49 in a range of values. Δk represents the irreducible uncertainty due to the phase noise. In this example $\Delta k \gg 1$. Fig. 15 and Fig. 16 show the Δk for a differ-

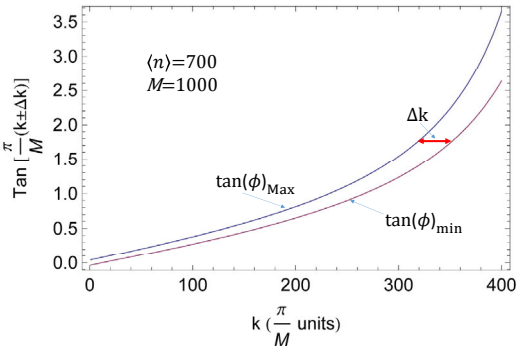


Figure 14. A sample of extrema for $\tan\left(\frac{k}{M} \pm \Delta k\right)$ versus k.

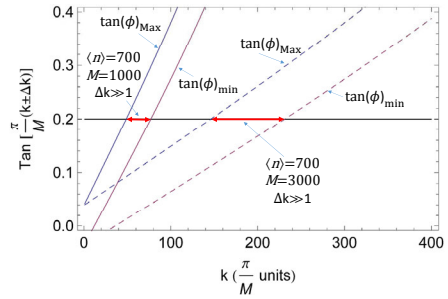


Figure 15. Extrema for $\tan\left(\frac{k}{M} \pm \Delta k\right)$ versus k.

ent set of values $\langle n \rangle$ and M. Fig. 15 shows that for a fixed $\langle n \rangle$ the uncertainty Δk (or $\Delta\phi$) increases with the number of bases M used. Fig. 16 shows that for an intense field $\langle n \rangle \gg 1$ the uncertainty Δk can be reduced giving the resolution $\Delta k \ll 1$ or $\Delta\phi < \pi/M$. In this condition of intense fields, the adversary could identify any basis used and, therefore, obtain the bit sent from A to B. This shows that A and B can frustrate the adversary by choosing $\langle n \rangle$ and M such that the adversary cannot distinguish which basis was used in every emission. The POVM calculation shown in [1] details this in a complementary and quite general way.

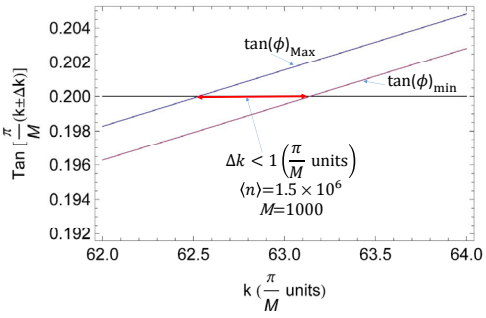


Figure 16. Extrema for $\tan\left(\frac{k}{M} \pm \Delta k\right)$ versus k.

References

[1] G. A. Barbosa, Physical Review A **68**, 052307 (2003).
[2] G. A. Barbosa, Physical Review A **71**, 062333(2005).
[3] G. A. Barbosa, E. Corndorf, P. Kumar, and H. P. Yuen, Physical Review Letters **90**, 227901-1 (2003).
[4] O. Hirota, K. Ohhata, M. Honda, S. Akutsu, K. Harasawa, and K. Yamashita, SPIE Newsroom 10.1117/2.1200909.1679 (2009).
[5] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).
[6] G. A. Barbosa and J. van de Graaf, "Untappable communication channels over optical fibers from quantum-optical noise", http://eprint.iacr.org/2014/146 .
[7] A. Messiah, "Mécanique Quantique" Vol. II, Chapter XIII (Dunod 1965).
[8] There are true physical random generators in the market. The one planned for this key distribution system was specially designed to operate at a fast speed compatible with optical communication rates. It is discussed in G. A. Barbosa, ENIGMA - Brazilian J. of Information Security and Cryptography, Vol. 01 (2013)– to be published.
[9] T. E. Chapuran, P. Toliver, N. A. Peters, J. Jackel, M. S. Goodman, R. J. Runser, S. R. McNown, N. Dallmann, R. J. Hughes, K. P. McCabe, J. E. Nordholt, C. G. Peterson, K. T. Tyagi, L. Mercer and H. Dardy, New Journal of Physics **11**, 105001 (2009).
[10] G. A. Barbosa, QuantaSEC - Technical Notes – Modulation and Demodulation systems, Feb. 2104.
[11] T. Iwata and K. Kurosawa, FIPS PUB 198 (2002) (Federal Information Processing Standards Publication), The Keyed- Hash Message Authentication Code (HMAC), and Natl. Inst. Stand. Technol. Spec. Publ. 800-38B (May 2005) CODEN: NSPUE2, in http://csrc.nist.gov/publications/nistpubs/index.html#sp800-38B.
[12] M. A Nielsen and I. L. Chuang, "Quantum Computation and Quantum Information", Cambridge University Press 2005.
[13] R. J. Glauber, Physical Review **131**, 2766-2788 (1963).
[14] C.H. Bennett, G. Brassard, C. Crepeau, U.M. Maurer, IEEE Transactions on Information Theory **41**, 1915 - 1923 (1995).
[15] G. A. Barbosa, QuantaSEC Technical Notes, KeyBITS Platform 2104.
[16] M. Born and E. Wolf, "Principles of Optics" (Pergamon Press, Sixth Edition, 1980), Section 10.8.3.
[17] P. Carruthers and M. M. Nieto, Reviews of Modern Physics **40**, 411 (1968).
[18] R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, Phys. Rev. A **63**, 040305(R)(2001). S. M. Barnett and E. Riis, J. Mod. Optics **44**, 1061 (1997).
[19] B. Huttner, A. Muller, J. D. Gautier, H. Zbinden, and N. Gisin, Phys. Rev. A **54**, 3783 (1996).
[20] L. L. Sánchez-Soto, A B Klimov, P. de la Hoz and G. Leuchs, arXiv:1306.0351v1 [quant-ph] 3 Jun 2013.

G. A. Barbosa – PhD (Physics)/University of Southern California, 1974. Areas of work: Quantum Optics, Condensed Matter (Theory and Experiment), Physical Cryptography. Full Professor, Universidade Federal de Minas Gerais/MG/Brazil (up to 1995)/Northwestern University (2000/2012), and CEO, QuantaSec Consultoria e Projetos em Criptografia Física Ltda /Brazil.

J. van de Graaf – PhD (Informatics)/Université de Montréal, 1998. Areas of work: Cryptography – theoretical and the applied aspects of cryptographic protocols. Assistant Professor at the Universidade Federal de Ouro Preto (August 2008/January 2011). Professor at the Universidade Federal de Minas Gerais (March 2011 up to date).

Cyber-Attacks Based in Electromagnetic Effects

M. B. Perotoni, R. M. Barreto and S. K. Manfrin

***Abstract*— This article covers eavesdropping on computer and auxiliary data communication equipment by means of hardware, namely unintended electromagnetic emanations. The physical basis that underlies the process is covered, alongside with a canonical electromagnetic simulation. Some known cases of these exploits are covered, and real world examples of a leaking coaxial cable and a shielding conductive sheet are measured in the laboratory, with results relate to the data protection and its implications. The measured shielding effectiveness of the sheet proved to comply with usual Tempest requirements.**

***Keywords*— Shielding, Electromagnetic Compatibility (EMC), Information Security.**

I. INTRODUCTION

SECURE communication is a theme that concerns almost everyone who uses emails, makes phone calls or even browses news on Internet on a regular basis. The confidentiality of these mundane activities, almost taken for granted in the beginning of the Internet era, is a foregone recall. Back then, the only concern for the majority of ordinary users was credit card passwords theft, when shopping online – nowadays much more sophisticated schemes are being used to steal and decode information, not only from individual hackers, but also by official government agencies. Emails, phone calls, history of the visited pages– all these data can be gathered and analyzed by third parties elsewhere, making daunting the confidential exchange of information.

Though these eavesdropping activities are usually related to software (Trojan horses, back door exploits, suspicious emails hiding executables, etc), the existing hardware and infrastructure also provides information windows from where data can be stolen. These infrastructure and hardware exploits, so far, require more sophisticated means to be implemented, therefore are not accessible to delinquent teenagers living next door. Though harder to implement, due to the both higher technical expertise required and more expensive pieces of equipment necessary, they are at the same time harder to detect and to prevent, partially due to the fact they are yet unbeknownst to most people and corporation IT sectors. These complicated requirements make cyber-attacks based on hardware and infrastructure more focused on corporate targets, rather than ordinary internet users. Usually they focus on valuable information theft (IP, intellectual property), kept in confidentiality due to the relevant financial or industrial impact, for instance.

This article deals with some ways by which attacks based on infrastructure and hardware are implemented, and the solutions that can be employed to prevent it. Naturally most information is kept disclosed from the public literature; therefore much of it cannot be checked or tested. The field is still on its development, with many tools yet to be widespread. Regardless, the means to implement such attacks are already available in the market.

M. B. Perotoni, UFABC, Santo André, SP, Brasil, marcelo.perotoni@ufabc.edu.br
R. M. Barreto, QEMC Consulting, Rio de Janeiro, RJ, Brasil, roberto.menna@qemc.com.br
S. K. Manfrin, UFABC, Santo André, SP, Brasil, stilante.manfrin@ufabc.edu.br

II. PHYSICAL BACKGROUND

Any data or signal that is carried by a guided medium (for instance a wire, cable or PCB – Printed Circuit Board- trace) acts as source of unintended electromagnetic emission.

Though a single wire can be seen as self-contained element that is used to transport an electric current between two points, it also operates as a radiator (antenna). Electromagnetic (EM) waves arise whenever there are accelerated charges [1,2], according to the Larmor Formula [3] . As Fig. 1 shows, an infinite wire is not radiating, but if it is made finite (by truncating one of its ends), radiation arises, due to the fact that the carrier velocities drop abruptly to zero at the finite end. By the same token, a change of direction on the wire results on acceleration as well (since acceleration is a vector quantity) – the more acute the angle the more efficient is the radiation, as Fig. 1 illustrates. The rationale for implementing a cyber-attack based on this effect is that the radiated signal bears a relevant resemblance with the original signal, therefore by remotely capturing this EM wave one is able to reconstruct the original data.

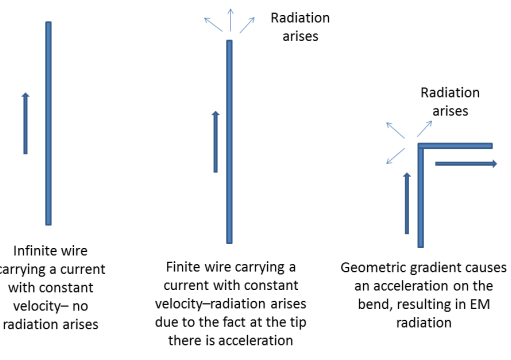


Figure 1. Pictorial examples of radiation arising from accelerated charges.

Real world circuitry and equipment are more complicated than the straight wires from Fig. 1. Fig. 2 depicts EM simulations carried on with CST MICROWAVE STUDIO® [4], a 3D field solver which here employed its TLM (Transmission Line Method). It consists on a simple trace printed on PCB (FR-4 material, 1 mm thick, dielectric constant 4.9), excited by a generator and terminated with a 50 Ω load. Two scenarios were evaluated: the bare PCB and with the Plexiglas (chosen only for the sake of similarity with common used materials) enclosure.

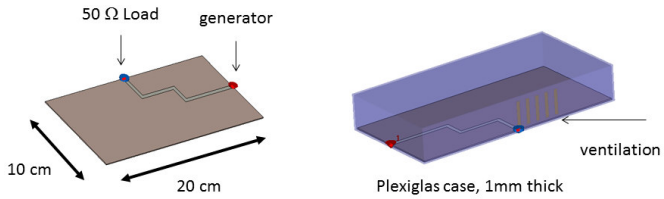


Figure 2. Simulated models to evaluate the radiation from a copper trace (bare board and with a case).

The system shown in Fig. 2 was excited with a rectangular pulse in the generator end (Fig. 3). Fig. 3 shows additionally the electric field captured at 3 meters from the board (the curves had their amplitudes normalized to ease the visualization).

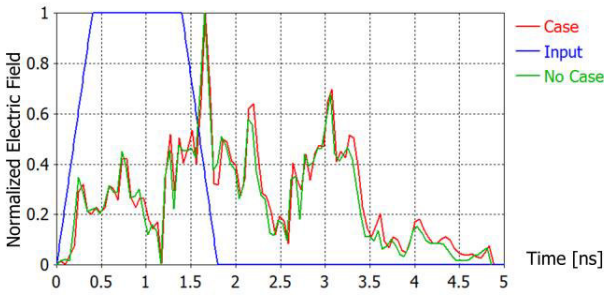


Figure 3. Normalized time response of the input signal on the trace (blue curve) and the electric fields measured at 3 meters distance from the equipment under test – for both scenarios.

Some conclusions can be drawn from the numerical simulations presented in Fig. 2:

- There is no significant difference whether a plastic case is used or not; their maximum amplitudes differ only 12% in the worst case. That means, the plastic case does not provide a relevant attenuation to the radiated signal. Other materials could be used, for instance plastic mixed with Carbon Black, which presents higher attenuation for electromagnetic waves [5];
- near field coupling effects (and consequently radiation) increase with the frequency, therefore signals with shorter rise times and higher repetition rates (which are generally associated with high data rates) are more prone to radiate and get captured elsewhere. It can be seen on the fields computed at 3 meters from the system; both in time (Fig.3) and frequency (Fig 4) domains;
- the comparison between the original and radiated signals power spectrums is shown in Fig. 4. It can be seen that the original signal has energy spread up to 2 GHz (limit where the power fell by approximately 40 dB). And the resemblance among the spectra is larger as the frequency increases – lower frequencies are poorly radiated, therefore the information is lost. However a remote pulse reconstruction based on the captured electric field could be implemented using filters (as a first order analysis a kind of low pass filter) to compensate for the channel frequency response.

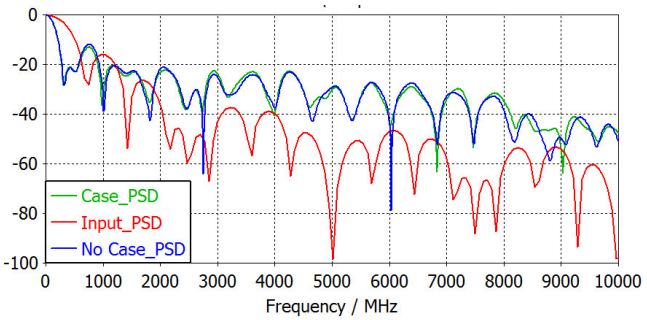


Figure 4. Power spectrum densities (PSD's) of the input and radiated signals.

III. A BRIEF INTRODUCTION TO ATTACKS

This section covers some known examples where attacks are performed exploring infra-structure fragilities. It is not an exhaustive list, but it pinpoints some typical examples where these physical fragilities can be explored.

A. Laser Bouncing on Windows

If a meeting is taking place in some closed space, the sound waves resonate and propagate inside the area. Upon hitting the windows, the same sound waves make the glass structure vibrate, similarly to loudspeakers. Though these mechanical vibration on the glass presents very low amplitudes, if a laser bounces the window its amplitude will be modulated with the same time pattern as the one from the conversation (as Fig. 5 shows). A receiver located outside needs only to demodulated the backscattered beam to recover the information. Since lasers can be made invisible, the technique is almost impossible to be detected.

A simple way to avoid this technique is the use of curtains, possible made of heavy cloth, as to damp the acoustic waves before they hit the window. This technique was allegedly used to discover first the new Popes name, Francis, back in 2013.

Figure 5. An incident constant amplitude laser beam is modulated as it backscatters the glass window.

B. *Electronic Bugs*

The use of electronic bugs is known since the beginning of the electronic era. They basically have the ambient sound modulating a carrier which in turn is wirelessly transmitted. They might be hidden in telephone sets, portraits, small gifts received by the subject etc. Since electronics became more and more advanced and present in everyday gadgets, so does increase the possibilities to tap conversations. Nowadays it is virtually impossible to guarantee that a conversation is kept secret, since electronics can be even implanted inside human bodies [6,7].

Though in the past conversations carry a lot of relevant information, nowadays access to data stored in computers or transmitted by email is much richer – it can convey blueprints of industrial projects, strategic or financial decisions. Therefore the use of electronic bugs became less important in the arsenal of the contemporary espionage.

Electronic bugs can be traced and located by searching for their wireless signal (carrier). Since the signal amplitude becomes stronger closer to the bug, a directional antenna plugged into a tuned field strength meter can help find them. Other alternative is the use of jamming, where a high amplitude broadband noise source overpowers eventual hidden transmitters in an area. It is the principle behind Mobile phone jammers, commercially available.

C. *Hardware with hidden exploits*

Microprocessors are present in almost any mass consumer equipment. Currently, due to their lower production prices, most of the commercial microelectronics manufactures are based in Asia, whereas the design is still partially kept on Western countries [8,9]. Their small sizes and high complexity make almost impossible to get them checked after the purchase, so the consumers and final users have to rely on their internal content. Indeed they are nice targets to be used as a vector to steal information by means of Trojans and Backdoors implemented inside them [10]. One documented case is presented by a group from Cambridge [11] which found backdoors in a commercial FPGA (field-programmable gate array) chip specifically targeted for military and industrial applications. Ironically, the chip manufacturer states that "*low power flash devices are unique in being reprogrammable and having inherent resistance to both invasive and noninvasive attacks on valuable IP*" [12].

The use of counterfeit chips is hard to be repressed, due to their large volume and small sizes. In 2010 two men were caught with 13,000 fake chips imported from China, having common brand names on it (like Intel, AMD and National Instruments) [8]. One of the purchasers of such batch was the US Navy.

The New York Times reported that during an Israeli air raid on Syrian nuclear laboratories their anti-aircraft defenses were temporarily disabled by a "kill-switch" feature that had been surreptitiously introduced by Israeli intelligence [13].

Given the extreme sensitivity that microprocessors impose into the data security realm, chip design and manufacture is considered a national security issue for critical applications. Brazil has its own secure microprocessor designed and produced so that it can be used in highly secure and critical applications, like defense, financial transactions and also powering the electronic voting machines [14,15]. Naturally, the complete design of a microprocessor is a costly and long endeavor, which demands continuous investment and a highly specialized workforce, not easily achieved by any country.

D. *Infrastructure and Data Center exploits (TEMPEST)*

As stated before, any current circulating in a conductor generates magnetic fields. Though common Kirchhoff laws do not cover electromagnetic emanations [16], there is always an electromagnetic wave associated with time-varying electric currents (accelerated charges produce electromagnetic radiation). Therefore the information encoded in an electric current running in wires or printed circuits is not fully guided and contained by the medium, it is also free to propagate as a wave. That means that normal wires, cables and printed circuit traces can also be seen as antennas, normally ineffective but still radiating part of the energy. By gathering this emanation it might be possible to reconstruct the original data content.

In order to illustrate the leaking problem, Fig. 6 shows the measurement setup (inside an anechoic chamber) of 1.2 m long standard RG-58 cable, terminated with a matched 50 Ω as to avoid reflections and stationary waves. The cable was excited by a signal generator (from 20 MHz to 1 GHz), with 10 dBm power, and a broadband antenna captures the leaked electric field at a distance of 2.7 m. Fig. 7 shows the received power for two scenarios: with both ends grounded (connection between the shield to the ground plane) and floating. It can be seen that the simple act of connecting the metallic shields from both cable ends to the ground plane helps reduce the emissions in most frequencies across the band.

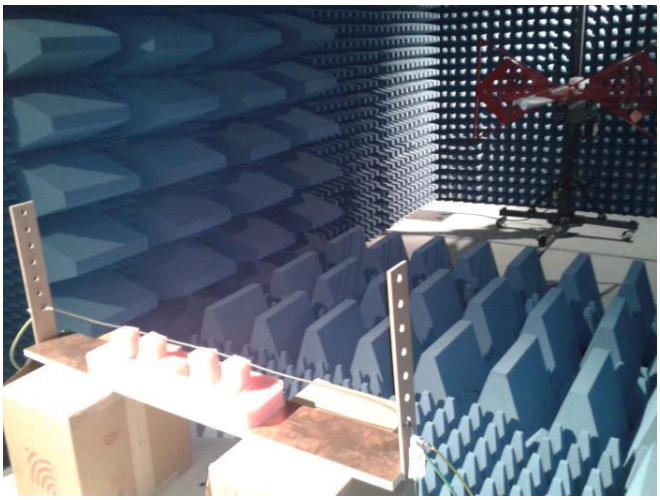


Figure 6. RG 58 over a ground plane, matched. The electric field that leaks from it is captured at distance, with a broadband antenna.

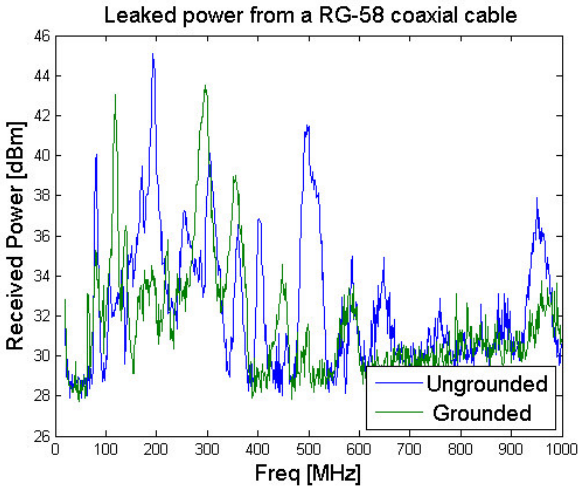


Figure 7. Received power on the antenna, for two different cases: with and without connection to the ground plane.

TEMPEST is a codename established by the US Department of Defense (DoD) Agency in 1974 to address eavesdropping based on leaking emanations, unintentional radio or electrical signals [17]. The code stands for "Telecommunication and Electronic Material Protected from Emanating Spurious Transmission".

Cathode ray tubes (CRT's) were once commonplace as computer screens. They operate based on an electron beam made to sweep a phosphor-based screen. The electron beam is deflected, focused and directed by means of high voltages and magnets (yokes). In order to synchronize transmission and reception, the frequencies by which the beam was made sweep the screen was fixed (15.734 kHz for TV). Since there were high amplitudes involved in the accelerated beam voltage (typically 20 kV), proportional high fields were also generated. By means of decoding these fields (both at fundamental or harmonics) it was possible to reconstruct, at distance, the same image seen on the eavesdropped TV set (as shows Fig. 8). Van Eck [18] in 1985 published a comprehensive and seminal paper on such apparatus, where with a directional antenna, a TV receiver and tuned circuits housed in a van he was able to gather and reconstruct signals from a nearby TV set located inside a building. Modern similar techniques were also applied to normal flat screen monitors, based on the same principle of focusing on the fixed frequency synchronization signals [19]. Common contemporary monitors have adopted the raster scan method, where each line is swept at a time. Therefore the picked signal should be the analog RGB information, not the digital complex image transmitted from the graphic board to the monitor – an approach that required only a broadband scanner and a dipole antenna to reconstruct the image generated by a video signal.

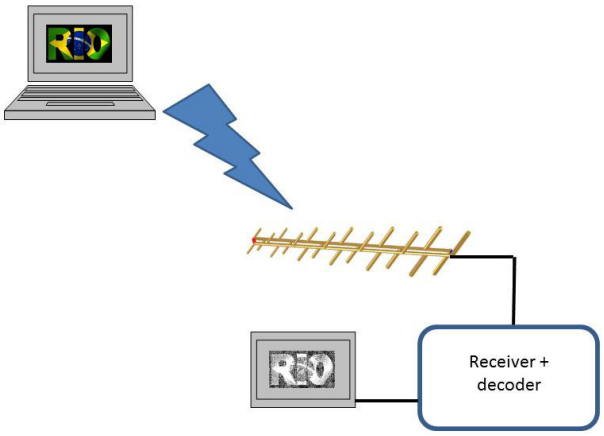


Figure 8. Basic elements to decode the leaked information from the monitor.

In addition to monitors, keyboards can also have their keys stroke remotely detected, since to each one is assigned a different code [21]. One found countermeasure consists on assigning each key a code defined by a cryptographic routine [22].

One more sophisticated way to countermeasure the information gathered through the monitor electromagnetic leakage is jamming. In [23] a circuit was enclosed in a box connected to a laptop by the USB port. It receives tracing information from the monitor RGB signals and adds jamming signal in frequencies where the emanation occurs. The artificially polluted signal is then back-fed into the laptop as a common mode voltage, and it turns out radiated by the laptop with the same mechanism as the normal information, making the eavesdropping harder.

Since most equipment is fed by AC mains, the connection to power chords is a viable vehicle to where information can be extracted. In this case, the information is not radiated, but conducted [16]. Power analysis is the name given to techniques that try to guess cryptographic keys by statistically observing power magnitude fluctuations from the target computer [24]. Magnitude fluctuations can be analyzed both in single and in differential mode, the latter conveying much richer information about the binary transitions inside the CPU [25].

The standard measure to tackle unintended leaked emissions from data equipment is shielding [26]. Operating

similarly to Faraday cages, a metallic enclosure will block radiation from sources within its volume. Though simple in concept, enclosing a complete data center in a metallic box is a daunting task. Even a single computer is hard to be completely enclosed in a metallic case, since it requires cabling inlets and ventilation holes. Therefore the shielding has to be used in a smart way, in such a way it provides an adequate shielding to the fields. The shielding the metallic layer provides prevents electric fields from propagating, but low frequency magnetic fields are still allowed to go through.

Filtering can also be used to reduce the emission levels [27]. By allowing the passage of only a limited frequency range, the detection of harmonics whose higher frequencies are more effectively radiated from cables and circuits can be made harder.

So in order to reduce the unintended radiation from data centers, the following procedures can be taken:

- use of metallic layers to block the EM waves generated by screens, keyboards and other elements;
- use of power chords with filters and shields;
- cables, connectors and jacks made with higher isolation rates.

An adequate protection of a data center demands a thorough analysis. To evaluate the performance of the shielding, there are three different levels of Tempest protection, according to the respective NATO regulation [28]; A, B and C, with decreasing levels of shielding from both conducted and emitted radiation. Solutions provided from industries refer to these aforementioned different levels as parameters.

E. HIRF (High Intensity Radiated Fields)

Historically, after the first nuclear detonation test in 1940 it was noticed that the monitoring electronic apparatuses based on semiconductors (then a brand new technology) were destroyed, not due to the blast, but because of the intense electromagnetic wave originated after the explosion [16], phenomena currently named HIRF. It aroused the interest on the unintended radiation effects on systems that may not properly operate. In 1978 the first open conference focused on the theme took place (named Nuclear EMP Meeting), as well as a publication followed on the subject [29]. Particularly electromagnetic effects due to nuclear explosions are treated by the acronym NEMP (Nuclear Electromagnetic Pulse), and nowadays, studies regarding it are carried out mainly based on simulations, considering that the last high altitude nuclear test detonations took place in 1963. Waveforms of typical pulses caused by NEMPs are known, some in the open literature, some to restricted applications [30], and they are similar to the ones used by lightning discharges, though with spectra reaching higher frequencies (NEMPs around 300 MHz whereas lightning reaching barely 10 MHz). The very first analytical waveform concerning a NEMP was proposed in 1963 by Bell Labs , whose rise time was 4.6 ns and maximum electric field amplitude of 50 kV/m [31], presented in Fig. 9.

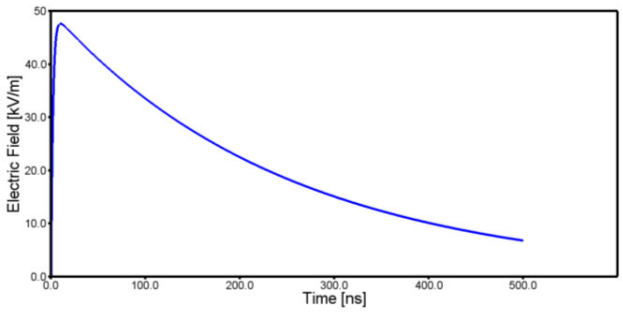


Figure 9. Analytical waveform of the Electric Field developed after a nuclear explosion in high altitude, according to Bell Labs [31].

Physically, the high amplitude electromagnetic fields generated by the nuclear blast propagate at the speed of light, and reach electronic circuitry and systems almost instantly. Induced voltages may eventually burn integrated circuits, effects that can even make planes in nearby areas crash due to lack of electronic flight control. The major defense against HEMP and HIRF is shielding and radiation hardening components, mainly integrated circuits. High impedance devices, such as CMOS, of widespread use due to their low cost and good performance in integrated circuits, are particularly sensitive to this incoming high amplitude electric fields, and should therefore go through a hardening process if they are intended to be made robust against NEMPs (for instance in missiles or jet planes supposed to fly over combat zones).

Though NEMPS and HIRFS are not meant to explicitly steal data and eavesdropping, they can block and turn inoperative large communication areas, including power distribution and transmission systems, causing havoc by destroying the existing infrastructure.

IV. SHIELDING CLOTH EVALUATION

The main counter measure against attacks based on Tempest or other similar emanation-based techniques is shielding. There has been considerable investigation in sheets and cloth operating as shielding materials, with carbon often used as a component, in the format of graphite [32], Carbon nanotubes and polymer composites [33] and also graphene-based sheets [34]. Copper wires inserted into a cloth made of Polypropylene and fiberglass is also used [35], also with evaporated silver and polypyrrole –both conductive – into a plastic fabric [36]. Shielding a room requires metallic sheets on the walls, floor and ceiling, with appropriate soldering and seaming, which is an expensive and complicated task that demands trained workforce and sophisticated instrumentation to certify its performance. An alternative is the shielding cloth or sheet – fabrics with metallic parts embedded, that provide an easily deployed and light weigh surface, able to be quickly deployed on normal and temporary spaces (such as field data centers).

A sample of such shielding cloth was investigated, from Soliani EMC s.r.l. [37]. It is a polyester sheet impregnated with metallic particles, such as Nickel, and does not present visually any difference to other types of clothing. The fabric was placed between two rectangular loop wire antennas, each with 17 cm length. Fig. 10 shows the antenna in front of the fabric and Fig. 11 the antenna measured reflection loss in dB (S11 parameter). This method is based on the IEEE Std. 299-1997, specifically aimed to Shielding Effectiveness Measurement of Enclosures[38].

Figure 10. Quad Loop Antenna used to address the shielding effectiveness of the fabric, seen on the background.

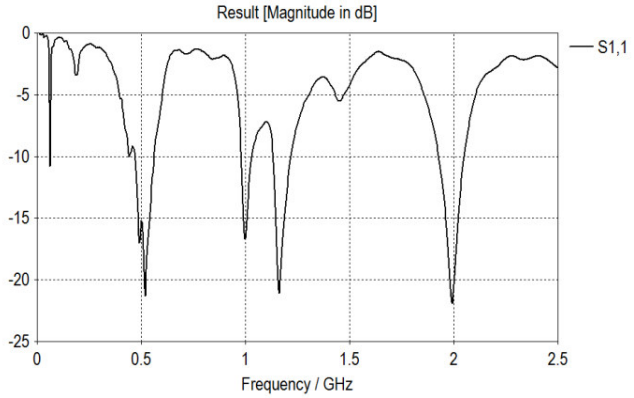


Figure 11. Quad Loop Antenna S11 measured response.

This specific antenna is chosen to investigate the cloth shielding around 500 MHz, which according to Fig. 11 is the resonant frequency of the antenna. The available sheet had its performance indicated for frequencies up to 1 GHz, therefore well within the antenna operating range.

The two similar quad loop antennas were placed 70 cm apart, and their transmission scattering parameter S21 (equivalent to the power transmission factor of both) was measured in two conditions: with and without a shielding sheet placed at mid distance, as Fig. 12 shows. The measurement was performed outside of an anechoic chamber, so there were reflections on the pieces of furniture, ceiling and further objects, and the covering was not perfect (due to the limited sample fabric width).

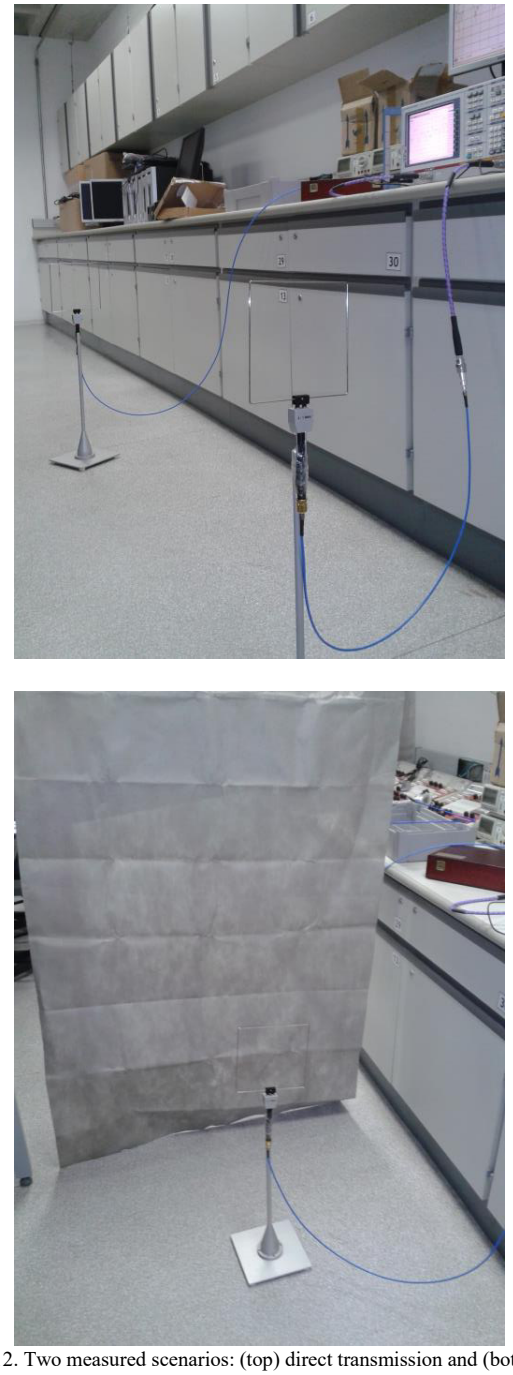


Figure 12. Two measured scenarios: (top) direct transmission and (bottom) with the conductive sheet.

Fig. 13 shows the measured shielding effectiveness of the sheet, based on the difference of the S21 parameter for the two scenarios. It was considered the frequency range of the 300 MHz to 700 MHz, around the center resonant antenna frequency. Since the measurement covers a broad band,

eventual emissions from a computer, for instance, are covered by the measurement. It can be seen that at the frequency where the antenna resonates the nominal attenuation is larger than 50 dB, a high value but yet insufficient to be used as protection against Tempest attacks – the usual assumed safe Shielding Effectiveness is 100 dB, according to the NSA-65-6 [39], value normally achieved only with large solid metal plates. The cloth provides, though, a quick and transportable way to offer protection to sites temporarily located in areas where a complete shielded room is not available or too costly.

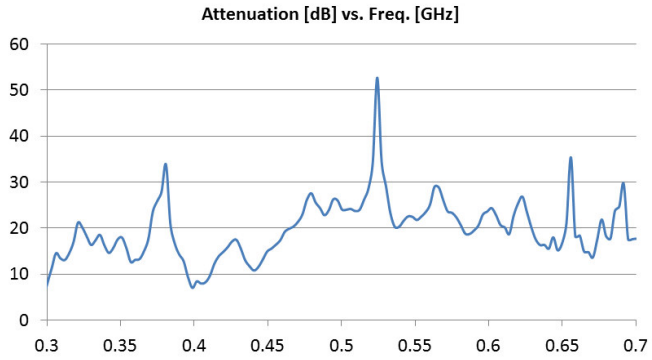


Figure 13. Measured attenuation of the conductive shielding sheet.

V. CONCLUSIONS

The article covered some formats of Cyber-Attacks based on hardware, mostly the ones that are based on eavesdropping taking advantage of unintended electromagnetic emanations from the data sources, as well as the effects caused by nuclear bombs detonation at high altitudes. A survey of some attacks are presented, alongside with simulations of measurements that show the basic nature and underlying principles involved. A sample of shielding sheet is measured and had its performance evaluated, which proved to provide a simple and light way to prevent this sort of attacks.

VI. ACKNOWLEDGEMENTS

The authors thank Mr. Maurizio Rizzati, from Soliani EMC, for kindly providing the sheet samples that were subjected to the measurement.

This article is included within the larger RENASIC Project, this one particularly named IT Critical Installation Security, and it has been conducted by Roberto M. Barreto, from QEMC, Engineering, Quality and Electromagnetic Compatibility Ltda, and Prof. Marcelo B. Perotoni.

REFERENCES

[1] Z. Popovic, B. D. Popovic, "Introductory Electromagnetics",Ed. Prentice Hall, New Jersey, 2000.

[2] D. Halliday, R. Resnick, J. Walker, "Fundamentals of Physics", Ed. Wiley, 9th Edition, 2010.

[3] J. D. Jackson, "Classical Electrodynamics", Ed. John Wiley, 2nd Edition, New York, 1975.

[4] CST STUDIO SUITE EM simulation software, v.2015, www.cst.com.

[5] Q. H and M. S. Kim, "Electromagnetic Interference Shielding Properties of CO_2 Activated Carbon Black Filled Polymer Coating Materials", Carbon Letters, Vol. 9, No.4, Dec. 2008, pp. 298-302.

[6] K. Y. Yazdandoost , R. Kohno, "Wireless Communications for Body Implanted Medical Device", Proceedings of Asia-Pacific Microwave Conference 2007, pp.1-4.

[7] M. Balouchestani, K. Raahemifar , S. Krishnan, "Wireless Body Area Networks with Compressed Sensing Theory", Proceedings of 20121 CME International Conference on Complex Medical Engineering, pp. 364-369.

[8] M. Bilzor, T. Huffmire, C. Irvine, T. Levin, "Security Checkers: Detecting processor malicious inclusions at runtime", Hardware-Oriented Security and Trust (HOST), 2011 IEEE International Symposium on, 2011.

[9] M. M. Farag, L. W. Lerner, C. D. Patterson," Interacting with Hardware Trojans Over a Network", 2012 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), pp.69-74.

[10] M. Tehranipoor, F. Koushanfar, "A survey on hardware Trojan taxonomy and detection", IEEE Design and Test of Computers, 2010.

[11] S. Skorobogatov, C. Woods, "Breakthrough silicon scanning discovers backdoor in military chip", Cryptographic Hardware and Embedded Systems – CHES 2012, Lecture Notes in Computer Science, Vol. 7428, 2012, pp. 23-40.

[12] Military ProASIC3/EL FPGA Fabric User's Guide. Microsemi, 2011, http://www.actel.com/documents/Mil PA3 EL UG.pdf.

[13] J. Markoff. Old Trick Threatens Newest Weapons. New York Times, October 2009.

[14] R. Gallo, H.Kawakami and R. Dahab, "SCuP – Secure Cryptographic Microprocessor", Proceedings of the XI Brazilian Symposium of Information Security and Computational Systems SBSEG 2011, 2011.

[15] R. Gallo et al, "T-DRE: a hardware trusted computing base for direct recording electronic vote machines", Proceedings of the 26th Annual Computer Security Applications Conference, 2010, pp. 191-198.

[16] C. R. Paul, "Introduction to Electromagnetic Compatibility", 2nd Edition, Ed.Wiley, 2006.

[17] A. Auddy and S. Sahy, "Tempest: Magnitude of threat and mitigation techniques", Electromagnetic Interference & Compatibility, 2008. INCEMIC 2008. 10th International Conference on, 2008.

[18] W. van Eck: "Electromagnetic Radiation from Video DisplayUnits: An Eavesdropping Risk?", Computers & Security, Vol. 4, pp. 269–286, 1985.

[19] C. Xiang and J. Xi, "A Method to Extract the Synchronous Characters in the Electromagnetic Information Leaked by a Computer", 2011 4th International Congress on Image and Signal Processing.

[20] H. Sekiguchi and S. Seto, "Study on Maximum Receivable Distance for Radiated Emission of Information Technology Equipment Causing Information Leakage", IEEE Transactions on Electromagnetic Compatibility, vol.55, No. 3, June 2013, pp. 547-554.

[21] M. Kinugawa, Y. Hayashi, T. Mizuki and H. Sone, "The effects of PS/2 keyboard setup on a conductive table on electromagnetic information leakages", Proceedings of SICE Annual Conference, 2012, pp. 60-63.

[22] J. A. Ross and M. G. Kuhn, "Soft tempest – an opportunity for NATO", Protecting NATO Information Systems in the 21st century (1999).

[23] Y. Suzuki and Y. Akyiama, "Jamming Technique to Prevent Information Leakage Caused by Unintentional Emissions of PC Video Signals", 2010 IEEE International Symposium on Electromagnetic Compatibility (EMC), 2010, pp. 132-137.

[24] A. Arora, J. A. Ambrose, J. Peddersen and S. Parameswaran, "A Double-width Algorithmic Balancing to prevent Power Analysis Side Channel Attacks in AES", 2013 IEEE Computer Society Annual Symposium on VLSI, pp. 76-83.

[25] P. Kocher, J. Jaffe, and B. Jun, "Introduction to differential power analysis and related attacks", Technical Report, 1998.

[26] S. Pennesi and S. Sebastiani, "Information security and emissions control", 2005 International Symposium on Electromagnetic Compatibility, pp. 777-781.

[27] S. Sebastiani, "Characterization to a TEMPEST testing laboratory and methodology for control to compromising emanation", 1998 IEEE International Symposium on Electromagnetic Compatibility, 1998, pp. 165-170.

[28] NATO SDIP-27 Standard.
[29] K. S .H .Lee, "EMP Interaction: Principles, Techniques and Reference Data", New York: Hemisphere, 1980.
[30] W. A. Radasky, "Review of unclassified HEMP calculations and analytic waveforms", NEM 1990 Record, p. 71.
[31] EMP Engineering and Design Principles, Bell Telephone Labs, Whippany, NJ, 1975.
[32] D. D. L. Chung, "Electromagnetic interference shielding effectiveness of carbon materials", Carbon, vol. 39, No. 2, 2001, 279-285.
[33] M.H. Al-Saleh and U. Sundararaj, "Electromagnetic interference shielding mechanisms of CNT/Polymer composites", Carbon, vol.47, No.7, 2009, pp.1738-1746.
[34] J. Liang et al, "Electromagnetic interference shielding of graphene/epoxy composites", Carbon, vol..47, No. 3, 2009, pp. 922-925.
[35] K. B. Cheng, S. Ramakrishna and K. C. Lee, "Electromagnetic Shielding effectiveness of copper/glass fiber knitted fabric reinforced polypropylene composites", Composites Part A: Applied Science and Manufacturing, vol. 31, No. 10, 2000, pp. 1039-1045.
[36] Y. K. Hong et al, "Electromagnetic interference shielding characteristics of fabric complexes coated with conductive polypyrrole and thermally evaporated Ag", Current Applied Physics, vol. 1, No. 6, 2001, pp. 439-442.
[37] Soliani EMC s.r.l., http://www.solianiemc.com/
[38] IEEE Std. 299-1997.
[39] MIL-HDBK-1195 Military Handbook Radio Frequency Shielded Enclosures, 1988.

Marcelo Bender Perotoni is from Porto Alegre, RS, Brazil. He holds an Electrical Engineer degree from UFRGS, Porto Alegre, 1995; a Masters (2001) and PhD (2005) degrees in Electrical Engineering, both from Escola Politécnica da USP, São Paulo, SP. He was visiting researcher at the Colorado University, Boulder , US, 2004 and a postdoc researcher at TEMF Institute, Darmstadt, Germany, 2006. He currently is Professor at UFABC, Santo André, SP, Brazil.

Roberto Menna Barreto holds an Electrical Enginner degree from IME, Rio de Janeiro, (1976), and a masters (1979) from Philips International Institute. He is currently General Manager of QEMC, located in Rio de Janeiro, focused on consulting and training in EMC-related areas. He is member of the "dB Society" and also " ssociation of Old Crows", both from US.

Stilante Koch Manfrin is from Santo André, SP, and holds an Electrical Engineering degree from FEI, São Bernardo do Campo (1990), a MsC (1995) and PhD (2003) degrees from USP São Carlos, both in Electrical Engineering. He currently is Professor at UFABC, Santo André, SP, Brazil.

A Modularity and Extensibility Analysis on Authorization Frameworks

E. M. Guerra, J. O. Silva, and C. T. Fernandes

Abstract — **Authorization in its most basic form can be reduced to a simple question: "May a subject X access an object Y?" The attempt to implement an adequate response to this authorization question has produced many access control models and mechanisms. The development of the authorization mechanisms usually employs frameworks, which usually implements one access control model, as a way of reusing larger portions of software. However, some authorization requirements, present on recent applications, have demanded for software systems to be able to handle security policies of multiple access control models. Industry has resolved this problem in a pragmatic way, by using the framework to solve part of the problem, and mingling business and the remaining authorization concerns into the code. The main goal of this paper is to present a comparative analysis between the existing frameworks developed either within the academic and industry environments. This analysis uses a motivating example to present the main industry frameworks and consider the fulfillment of modularity, extensibility and granularity requirements facing its suitability for the existing access control models. This analysis included the Esfinge Guardian framework, which is an open source framework developed by the authors that provides mechanisms that allows its extension to implement and combine different authorization models.**

Keywords— **Software architecture, Access control, Authorization, Metadata-based frameworks, Decoupling, Metadata, Security, Software development, Software engineering.**

I. Introduction

ACCESS control is usually referred to as a broader term that includes authentication and authorization procedures. The former can be defined as a procedure that confirms if the subject is who it claims to be. The latter can be defined as a procedure that verifies if the subject has the right privileges to access a certain object. Even though both types of access control procedures are interesting to investigate, the focus of this work is on the analysis of authorization mechanisms architectures.

During our research, we noticed that many of the existing access control mechanisms used for developing industry applications tend to offer more features for authentication, limiting the authorization procedures to fewer ones, usually bounded to authorization based on roles. While it is understandable for a great amount of effort to be used to prevent the entrance of intruders into a system, it is still very important to control the authorization concerns. Analogously, a person can be allowed (authenticated) to enter a building, but it is still very important to control which floors or rooms this person is allowed (authorized) to go and under what circumstances.

Online systems can be cited as an environment in which the importance of access control has greatly increased. Software has been increasingly being made available through web services, requiring the control of authorization aspects of how these services are going to be consumed.

Since the very early days of Software Engineering, mechanisms have been developed on software systems to provide effective authorization procedures. However, mingling the implementation of the authorization rules with business concerns has proven to be ineffective regarding some software design principles, such as modularity, extensibility, reuse, cohesion, code readability, and testability. The use of the traditional object-oriented paradigm alone does not solve the issue adequately, mainly because authorization has a crosscutting nature.

The use of authorization frameworks are one of the current academic and industry answers to this issue, because they allow different levels of code reuse, extensibility, and modularity. For an object-oriented developer, the usage of these authorization frameworks implies in a learning process of how to use each one of them. From the point of view of the framework developer, it is vital that the application developer can use their features without major complications, being able to focus mainly on business tasks.

Additionally, the choice of an authorization framework ordinarily implies that it will be bound to some specific access control model, and once this choice is made, it becomes difficult to incorporate new authorization requirements belonging to others access control models. In frameworks that uses a more granular access control mechanism, another problem happen in parts of the system where more simple models could be used, because it is bounded to an access control model that is more complex than necessary

In this matter, the existing authorization frameworks underachieve requirements of separating business from authorization concerns appropriately. As a consequence, developers have to improvise and craft solutions for more

E. M. Guerra, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP, Brasil, eduardo.guerra@inpe.br

J. O. Silva, Pontificia Universidade Católica de São Paulo (PUC-SP), São Paulo, SP, Brasil, silva.o.jefferson@gmail.com

C. T. Fernandes, Instituto Tecnológico da Aeronáutica (ITA), São José dos Campos, SP, Brasil, clovis@ita.br

complex authorization scenarios [1], or to use complex rules definition in simple scenarios making its management unnecessarily hard. In practice, it leads to applications where the authorization is coupled to the business code, a clearly undesirable situation [1]. Additionally, the existing authorization mechanisms leave developed applications coupled to their architecture, technology, or access control models, making the resulting application – once coded – more difficult to evolve [10].

The goal of this paper is to present a comparative analysis on existing authorization frameworks, focusing on modularity, granularity and extensibility requirements. As part of this work contribution, it presents development-guiding principles that can be used to perform the design and analysis of such kind of framework. Additionally, this paper presents a solution implemented by the authors called Esfinge Guardian, which is a framework that allows its extension to implement different access control models and particular authorization requirements, not coupled to any particular technology or architecture.

This research work can be considered an expansion of one previous work of ours [30], where Esfinge Guardian was presented, however this one has a strong focus on the frameworks comparison analysis. More specifically: (*i*) we provide more information on the theoretic background presenting the access control models $UCON_{ABC}$ and RAdAC; (*ii*) we establish development guiding principles for extensible and decoupled authorization models; (*iii*) we expand the Section on the use of the Esfinge Guardian; (*iv*) we propose a development guiding principles to provide a baseline for comparison analyses among authorization frameworks, specifically on extensibility and decoupling features.

This work is organized as follows: Section II provides the theoretic background on access control models and frameworks. Section III formalizes the problems of the current authorization framework implementations. Section IV schematizes a motivating authorization scenario. Section V presents how each of the main authorization solutions implements the access control policies of Section IV. Section VI makes a deeper analysis on academic and industry authorization frameworks, highlighting differences related to modularity and extensibility. Section VII concludes the paper, highlighting the main points and contributions of this work.

II. Authorization in Access Control Frameworks

One way to understand authorization is as the accurate management of three parameters: subject, resource, and privileges [1]. That is, for a resource to be accessed, a subject must have the right privileges. In fact, authorization is the process that ensures that resources are only made available to authorized subjects, and the selection of the privileges that each subject should have brings the notion of access control policies [3].

Access control policies are a key concept in the construction of an access control mechanism. They are here defined as "high-level requirements that specify how access is managed and who may access information under what circumstances" [2]. One example is: "Only account managers can credit money into a client's account". Clearly, computers cannot understand policies as high-level requirements. When translated into a format that programs can understand, authorization policies become digital policies [8].

The literature makes a distinction about the moments of creation and use of authorization policies. Privilege management is the process that creates and manages attributes and policies that are used by the access control [2]. Access control is the process responsible for the enforcement of policies and rules [4].

The effective enforcement of policies usually requires a mechanism, due to the complexity involved. Mechanisms must enforce system policies for every subject request to protected resources [16]. However, in many scenarios, in order to implement a mechanism, it is necessary to firstly design a model [1]. It is reasonable to think that a mechanism depend on a model. More precisely, access control models are mathematical formalizations of the security properties of a system, which are used to describe and, in some cases to prove these properties [1]. Models enter to bridge the gap between policies and mechanisms [2].

In application development, authorization mechanisms are many times implemented as frameworks, which can be considered as incomplete pieces of software with some special points that can be specialized to add application-specific behavior [5]. Frameworks' extension points are called hot spots [6], which are the points that applications use for customization. Each kind of behavior that a framework can execute is called variability [6].

Frameworks that base their logic decisions on the class metadata that they are working with are called metadata-based frameworks [5]. In this type of framework, a class needs to contain additional metadata so that the framework can consume, process, and make the decision for which variability to follow. In the context of object-oriented programming, metadata is information about the program structure itself such as classes, methods, and attributes [5].

Metadata-based frameworks' decoupled approach has represented an important facet in the reduction of coupling between the framework and business application concerns. As a general rule, it can be said that the more decoupled an approach the more general the types of algorithms it can execute [5]. This kind of framework can be applied for crosscutting functionality [28], such as authorization.

The main variability that access control frameworks need to handle is what rules should be enforced in each point of the system. The access control mechanism should also be able to prevent the execution of the functionality when the authorization is not confirmed. Some frameworks use metadata as an approach to configure the security rules related to a code element, such as a class or a method [6].

Different needs and contexts have led the development of many access control models and a plethora of access control mechanisms [16]. For brevity, this work only discusses some of the classical methods of authorization, but the architectural

model here presented is general enough to contain other authorization methods. Authorization methods refer to both access control mechanisms and access control models [2]. The following sub-sections provide a brief view about the classical access control methods implemented by the main security frameworks.

A. Identity-based Access Control (IBAC)

Due to the immense diversity of access control models, some works condense many access models into a category called IBAC [2][3][4], which despite essentially different among themselves, they share in common that privileges are somehow associated to the identity of the subjects.

In terms of policy enforcement, IBAC mechanisms tend to be relatively simple, as long as they handle simple policies [2]. Their drawback is when the number of resources grows too much [9], because it became problematic to privilege management. In a company of thousands of employees, it is difficult to centrally manage the creation and attribution of privileges for this huge number of resources. Scalability problems, like the previous example, were among the main reasons why it was advocated for the adoption of RBAC worldwide [7].

However, this access control model is suitable for applications where the privilege management is distributed among the users. For instance, when users own resources and can control the access to them. Nevertheless, nowadays social networks' access control model fit into this category, where a user can define who is allowed to access each of her resources [24][25], such as files, photos and information.

B. Role-based Access Control (RBAC)

RBAC introduces the concept of accessing resources mediated by roles. A role is a set of related privileges, normally equivalent to a function performed by someone in an enterprise organization. Instead of having the privileges bounded directly to subjects, they are attributed to roles [10]. Roles are attributed to subjects. The inclusion of this level of indirection immensely facilitates privilege attribution, for all a privilege administrator has to do is to set a person up with a role.

Although very efficient in the representation of hierarchies, such as companies and organizations, RBAC presents difficulties in the representation of other contexts. As an example, consider a global organization with branches in many countries. It could be necessary to divide the Information Technology team into multiple sub-teams, each team in one country administrating local resources. The creation of the one role Administrator for all sub-teams would not be adequate since each sub-team must only have access to their local resources. This is known as the least privilege principle [8]. One solution adopted by companies is to create as many roles as the number of sub-teams. However, this practice may lead into the role explosion problem in some cases, when the number of roles to be created is too numerous [9].

Another issue is that RBAC is not much adaptable to situations that demand change according to dynamic factors. The case of a hospital system illustrates the point. Information about patients is confidential by law and ethical reasons. Only the designated medical doctor must have access to the patient information. However, there are cases in which other doctors must attend to the patient for a matter of urgency. In these cases, doctors must have access to the patient information, but RBAC does not inherently deals with contextual and dynamic authorizations.

In fact, the literature does document RBAC variations built for dealing with dynamic situations such as Rule-based Access Control model (RuBAC) [3]. RuBAC is essentially RBAC that makes use of rules to create and manage roles. However, there are those who consider that these types of adaptation change the essence of RBAC, turning it into ABAC in disguise [2].

These sorts of situations – beyond the proposed scope of RBAC – are normally resolved in industry by embedding the additional access control policies into the application code [11][12][13].

C. Attribute-based Access Control (ABAC)

ABAC introduces the notion of access control based on the attributes of the subject, environment, and resources [14]. It still does not have a formal definition and its description can differ in the access control literature. For our purposes, ABAC will be defined as "access control based on attributes and policies. Attributes are distinguishable characteristics of users or resources, conditions defined by an authority, or aspects of the environment, and policies specify how to use attributes to determine whether to grant or deny an access request" [2].

Because it is based on the attributes of authorization entities, ABAC is generally said to be a fine-grained access control. It also includes the environment as part of the authorization, allowing rules to depend on other factors. It is worth mentioning that whatever access control can be defined with IBAC or RBAC can also be defined with ABAC [2].

Consider a hypothetical fine-grained policy: "Only account managers of level 2 can give credit to their clients during the working time". The problem with fine-grained policies is that they do not fit into the categories of IBAC and RBAC models.

ABAC mechanisms do not need to know the subject identity to authorize an operation. Instead, they rely on the attributes that the request proves to have [15]. In the case of the previous example policy, the request to give credit operation must contain that the requester is a level-2-manager, and the request time to be within the working time.

The ABAC's downside is that its fine-grained management increases considerably the complexity in the management of authorization policies [15], demanding a great effort to define and maintain the semantics of attributes in the enterprise.

D. Policy-based Access Control (PBAC)

Although PBAC [3] is often cited as a different access control model, it is essentially ABAC with a few differences [16]. The question about why it was necessary to create a slightly different model than ABAC naturally arises. The reason is that using ABAC in its pure form does not offer any

means of standardization in the communication of the attributes [18].

Consider an attribute called organization-name. It may happen that in one company the value of this attribute is "Aeronautical Institute of Technology", while in the other it could be "A.I.T.". Another issue is when enterprises use the same attribute name for different things, introducing the problem of name collapse [3].

In order to standardize the communication of attributes, OASIS' has created a standard named eXtensible Access Control Markup Language (XACML) [17], which is a general purpose language in XML for the declaration and communication of digital access control policies. Since then, XACML has become the de facto standard for writing fine-grained access control applications [18].

E. $UCON_{ABC}$

Traditionally, access control models focus on protecting resources on the server side and do not deal with client-side controls for locally stored digital information. Additionally, the advent of public-key infrastructure has allowed the authorization of subjects using models categorized as trust management [35]. In many cases, trust management utilizes subject properties for authorization in the form of digital credentials or certificates.

Usage Control (UCON) is a notion, a conceptual framework, introduced to be comprehensive enough to encompass traditional access control, trust management, and DRM [34] The term has some connotations, which reference [33] present them: "*In the DRM context, it conveys the sense that digital content is provided for use of the end-user's system, but the provider would like to retain some control over what the user does with the bits. In the privacy context the situation is reversed. It is the end-user who often provides personal information to a service provider, and would like to control how the service provider can use that information. Sometimes the personal information is provided by a third-party originator, say a health-care provider, but the individual, called 'identifiee', to whom it pertains, would nevertheless like to exercise control over its use. Usage also has a connotation of duration, so the access may continue for some time.*"

We can see some new concerns for authorizations. One example for access control in the DRM context is the re-distribution of a music file (e.g.: MP3) once it has been bought from a service provider. The service provider may be able to retain the right of distribution from the end user. One example for access control with duration, suppose an application that must control the use of prepaid mobile phone. In this case, even if the subject (user) is authorized to complete the phone call, the application must continuously check if the subject still has the credits for continuing the call.

Park and Sandhu [33] not only use the concept of Authorization (A), but also introduce oBligations (B), and Conditions (C), integrating them into the conceptual framework, forming the $UCON_{ABC}$ access control model. Obligations are requirements that have to be fulfilled for allowing access. Conditions are environmental or system requirements – related to resources – that have to be satisfied for access.

F. Risk-Adaptive Access Control (RAdAC)

RAdAC is an emerging access control model that takes into account risks to grant resources, being used basically in contexts that demand large-scale computing. The RAdAC model represents the cutting-edge model envisioned for the new contexts of grid and cloud computing [31].

In a world that is each day more interconnected, a differentiation in the access control must be made beyond roles, attributes and identities [32]. Risks must be taken into account. An example is the netbanking services that we do customarily. The risks of accessing the netbanking services from a trusted PC are different from the ones we take on accessing the same services from an untrusted PC.

RAdAC is still a very recent model that needs much research on it. Hu *et al* [3] have proposed a formal framework – at a policy layer – in terms of components and their interactions to develop abstract models for RAdAC.

G. Hybrid Authorization Models

Despite each model focus on solving the authorization problem for a given scenario, real applications can have needs that are not solved by a single model. In these cases, it is important to combine models in order to fulfill the authorization requirements.

For instance, imagine a military application where each user has a role in a military organization, but the documents also have a sensitivity that requires a certain privilege level from the subject. In this scenario, in order to access a given document, the user should have the appropriate role, the document should be related to the organization where he is allocated and he should have the minimum privilege level. Based on this example, it is possible to see that different models can be appropriate for different authorization requirements.

In such hybrid scenarios, by using a more restrictive model, such as RBAC, it does not cover all the authorization requirements. However, a more general model, such as ABAC, can be hard to manage for rules that fit better on other models. A possible solution in such scenarios is to combine authorization models, using each ones for the scope where it is more appropriate. In [29] there is an example of an authorization model that combines characteristics of RBAC and ABAC, creating what it calls a Contextual Authorization Model.

III. Problems in Existing Authorization Frameworks

The basic premise of access control mechanisms is that authorizations can be enforced in terms of subjects accessing protected resources in a particular environment. In the application development world, access control mechanisms are many times implemented as frameworks. It is noteworthy that the main security framework developers already provide

them as metadata-based frameworks.

Although there is not much debate about the importance of authorization, there is not still a general solution that decouples business from authorization concerns, except for simpler authorization policies. For more complex ones, the existing security frameworks fall short on offering tools that can be used declaratively, necessarily forcing developers to craft solutions tangling business with authorization codes.

The existing authorization frameworks offer rudimentary coding tools – or none at all – for software customization such as Spring Framework, Java EE, and Axiomatics XACML. These frameworks are each restricted, in the best cases, to a few access control models – usually RBAC –, but still far from providing means for extending to other authorization models. In other words, they have little or non-existent extensibility. The work of building an exhaustively complete access control mechanism that comprised all the possible user needs would be an impossible one. Therefore, extensibility must have a high priority in the design of an authorization framework architectural model.

To our best knowledge, there are not works that research why access control developers do not invest more in ABAC systems. However, a NIST report mentions that it is because its many-to-many relationships are difficult to represent [2]. It also states that the lack of more complex mechanisms maintains enterprises using RBAC solutions, leaving the ABAC ones on the horizon for most organizations.

Another issue about some current authorization solutions is in the technology dependence for its instantiation. Usually, these frameworks are coupled to some specific architecture or other frameworks. For instance, Java EE authorization solution can only be used in application containers and Spring Security can only be applied to objects managed by the Spring Frameworks, which can limit its application to a small set of applications.

An general architectural model for authorization frameworks is important because it provides ways to adequately separate concerns such as code tangling and technology dependence, representing an important step in the move from programmatic solutions to declarative ones. If a framework can include other framework solutions as its own, we say that such a framework is extensible. If a framework can be plugged in applications independently from its architecture and from the frameworks that it uses, we say that such a framework is technology independent.

IV. A MOTIVATING AUTHORIZATION SCENARIO

This Section schematizes a reasonable access control scenario and it aims to show – at the next Section – how each access control mechanism implements access control policies. This approach helps to create a common baseline for comparison of access control solutions and to create a more concrete view of each implementation.

A. The Scenario Access Control Policy

Consider the following hypothetical access control policy:

"*Any management position can oversee the operations performed by any of its subordinates, but must be restrained of overseeing the operations of their peers, their peers subordinates, and any superior position in the bank management hierarchy. In addition, the oversee operations function must only be accessed from within the perimeters of the bank facilities*."

B. Bank Career Hierarchy

For a richer comparison scenario, Fig. 1 defines a career bank hierarchy. This organizational chart is hypothetical but we consider it to be within the limits of reasonable.

According to the bank access control policy, a manager can oversee operations of Clerks and Officers as long as they are their direct subordinates, but cannot oversee operations of other Managers, Senior Managers or any other position above in the hierarchy. Also, all accesses must be made within the bank facilities.

C. Access Control Policy Rationale

This authorization scenario is composed of elements belonging to different access control models. For instance, the organizational chart is made of roles, each having a set of operations that they can execute. This indicates the use of RBAC.

In addition, the access control policy mentions the oversee operation, which has hierarchical features, which makes the policy in compliance with the MAC model [23] or at least with some type of hierarchical RBAC.

Finally, by limiting the execution of oversee the subordinate operations to the inside of the bank facilities, the access control policy uses elements of the ABAC/PBAC model, because it depends on the access context and not only on the subject and object. The requirement that restrict this authorization to its subordinates also is related to this access control model.

In this fashion, despite the apparent simplicity, the access control policy can be considered a complex one, from the point of view of the modularization of authorization concerns.

V. HOW EACH AUTHORIZATION FRAMEWORK IMPLEMENTS THE SCENARIO ACCESS CONTROL POLICY?

The objective of this Section is to present how each of the main existing solutions implements the access control policy presented in Section IV. Since our focus is on solutions ready to be used on industry projects, we are going to focus on mature solutions that are accessible to use.

We have selected three of the main security industry frameworks, that is: Java EE Security Framework; Spring Security; and Axiomatics XACML. For each one of them we present a possible implementation of the security policy of Section IV. The Esfinge Guardian framework is also included for comparison. Despite Esfinge Guardian is proposed by the authors of this work, it is important to highlight that it is open-source, have a comprehensive tutorial and a good automated test coverage, being ready to be used on real applications.

For better visualization, as a convention, we stress the authorization related code in bold on the code examples.

Java EE Security

The Java EE platform is the current industry standard for building enterprise Java applications. It defines an API that aims to simplify enterprise code development and in the meantime to robust Authentication, Authorization, Confidentiality, Non-Repudiation, Auditing, and Quality of Service. The Java EE security model has the means to secure the Web Tier, Enterprise JavaBean (EJB) Tier, and the Enterprise Information System (EIS) Tier. For the sake of comparison, this research only analyzes authorization in service tiers (i.e. EJB tier).

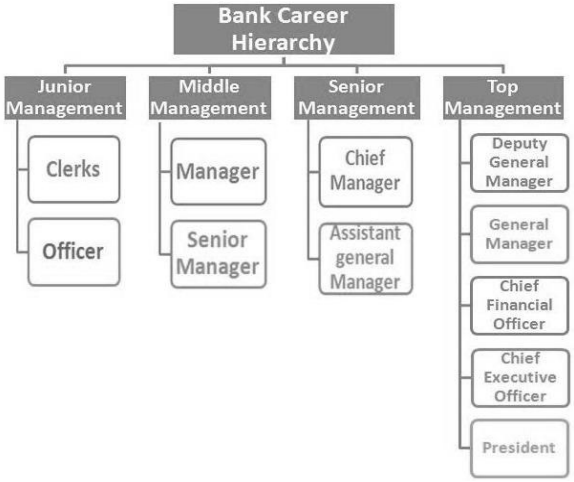


Figure 1. A hypothetical bank career hierarchy

The Java EE platform provides two ways of securing an application: declaratively and programmatically. There is in fact a recommendation in their tutorial for the declarative security, which can be done via XML descriptors files or via framework annotations.

The Java EE reference access control model is the RBAC. However, for anything with more complexity, they offer programmatic security. Listing 1 exemplifies a possible implementation of the bank authorization policy of Section IV. Java EE offers the `@DeclareRoles` annotation for specifying the management roles.

In the actual implementation, this line contains all the roles of the bank hierarchy. However, the only method made available by the platform is the `isCallerInRole()`, that the developer can use when s/he wants to discover is a certain subject belong to a known role. Since the authorization of the calling subject is dynamically relative to its own role, this method does not solve the issue alone. That is the reason why another proprietary method called `canOverseeRole()` is used. This method does not only has the responsibility of discovering which is the role of the calling subject, but also to determine its own role is equal, lower, or higher than the object's role (the other employee).

```
@Stateless
@DeclareRoles("{MANAGER,CLERK,OFFICER,…}")
public class EmployeeServImpl implements EmployeeServ{

  @Resource SessionContext ctx;

  public EmplOps[] overseeAllOps(
          EmplInfo info, CallerLocation cl) {
    boolean canOversee = canOverseeRole(ctx,info);
    boolean isSub =
       info.isSubordinate(ctx.getCallerPrincipal());
    if (!canOversee || !isSub || !cl.isInside()){
      throw new SecurityException(…);
    }
    EmplOps[] empops = //logic for retrieving data
    return empops;
  }

  private boolean canOverseeRole(
      SessionContext ctx, EmplInfo info) {
    //find user role using ctx.isCallerInRole(String)
    //return if this role can oversee the employee role
  }
}
```

Listing 1. A possible authorization policy implementation in Java EE 6.

Another verification that needs to be performed according to the requirements is if the employee is a subordinate from the current user. In Java EE platform, the method `getCallerPrincipal()` can be used to retrieve the current user registered in the session. This information can be used as a parameter to perform this check, which also needs to be done declaratively.

Spring Security

Spring Security is a popular security framework that has the same goals as Java EE, except that it is much more modular and lighter. Spring Security is designed to handle authentication and authorization requirements. In their tutorial [20], they cite four types of security concerns that the framework addresses: (*i*) authentication; (*ii*) web request security; (*iii*) service layer; and (*iv*) domain object security. In this research we focus on (*iii*) and (*iv*) security concern.

Considering the extensibility and modularity capabilities of each framework, Spring Security 3.X represents an enhancement when compared to Java EE 6. Although heavily based on the RBAC model, the Spring framework offers the possibility of accessing the application beans declaratively, through the use of authorization annotations.

As can be seen in Listing 2, by implementing the spring interface `PermissionEvaluator` it is possible to decouple authorization code from business code, except for the framework annotation declared on the business method, such as shown in Listing 3. There is also some configuration to bind these classes in the XML descriptors that we have omitted.

```
public class BankPermissionEvaluator implements
                              PermissionEvaluator {
  @Override
  public boolean hasPermission(Authentication auth,
      Object uid, Object pid) {
    return isHierarchyCompliant(auth, uid, pid)
      && isWithinFacilities(auth, uid)
      && isSubordinate(auth, uid, pid);
  }
  private boolean isHierarchyCompliant(
      Authentication auth, Object uid, Object pid) {
    allow = /* allow based on hierarchy */
    return allow;
  }
```

```
  private boolean isWithinFacilities(Authentication auth,
      Object uid) {
    boolean allow = /*allow based on location*/
    return allow;
  }
  private boolean isSubordinate(Authentication auth,
      Object uid, Object pid) {
    allow = /* allow if he is a subordinate */
    return allow;
  }
}
```

Listing 2. A possible implementation of the authorization policy using Spring Security.

```
@PreAuthorize("hasPermission(#subject, #target) and
        hasPermission(#subject, #cl)")
public EmplOps[] overseeAllOps(EmplInfo target,
      CallerLocation cl)
  EmplOps[] empops = //logic for retrieving data
  return empops;
}
```

Listing 3. Business method protected using Spring Security.

This is how the authorization would work. When business method `overseeAllOps()` is invoked, the framework would intercept the operation using aspect-orientation [26], and redirect the flow to the code presented in Listing 3. The access would be granted only if the method `hasPermission()` present on `BankPermissionEvaluator` return true.

Axiomatics XACML

Axiomatics XACML is the current most popular XACML access control platform, and it significantly facilitates the development of fine-grained applications [37]. It standardizes three essential aspects of the authorization process: policy language; XACML request/response protocol; and reference architecture. Axiomatics XACML can be seen as an implementation of the ABAC model, or more precisely, of the PBAC model.

One of the main advantages of this mechanism is the structured standardized use of external authorization. By providing a standardized language for writing authorization policies, it is possible to apply the mechanism into multiple tiers, having only one authorization policy description. This increases the separation of concerns, therefore augmenting flexibility.

The XACML language can be very fine-grained, being able to express a significant amount of authorization scenarios and access control models. One example is that the language can express hierarchical relationships between roles – unlike the other previous presented solutions. Listing 4 presents a code snippet that uses Axiomatics XACML to send a request to evaluate an authorization rule of the defined policy.

```
public boolean isEmployeeAllowed(EmplIinfo info,
    CallerLocation cl) throw SecurityException {
  try{
    //Create the connection to the service;
    ConnectionInterface pep = new MetroPEPModule();
    Properties config = new Properties();
    config.load(new FileInputStream(
      new File("connection.properties")));
    pep.setupConnection(config);

    //create XACML request
    SimpleRequestWrapper r = new SimpleRequestWrapper(4);
    r.addSubjectAttribute(URI.create("location", cl));
    r.addSubjectAttribute(URI.create("role",
      subj.role()));
    r.addActionAttribute(URI.create("action-id",
      "read"));
    r.addResourceAttribute(URI.create("resource",
      info.role()));

    //Send the request and handle response
    SimpleResponseWrapper resp = pep.evaluate(r);
    return resp.isPermit();
  } catch(Exception e) {
    throw new SecurityException(e);
  }
}
```

Listing 4. Business method protected using Spring Security.

Listing 5 shows one possible implementation of the business method, using solely Axiomatics XACML. It is important to note that this solution does not define how the authorization mechanism is plugged in the application.

```
public EmplOps[] overseeAllOps( EmplInfo info,
    CallerLocation cl) throw SecurityException {

  boolean allow = isEmployeeAllowed(info, cl);
  if(allow){
    EmplOps[] empops = //logic for retrieving data
    return empops;
  } else {
    throw new SecurityException("Access Denied");
  }
}
```

Listing 5. A possible implementation of the business method using Axiomatics XACML.

Esfinge Guardian

The Esfinge Guardian framework is an extensible authorization framework, fully capable of being used in the development of any business application. Among its benefits we can include the complete separation of business and authorization code.

The Esfinge Guardian framework can be seen from at least two perspectives. Since the framework completely separates business from authorization concerns, the implementation of the authorization logic can be delegated to experienced developers, usually the ones with technology and business domain background for creating and implementing business security rules. On the other hand, once the framework has been extended, it can be used by the other members of the development team, which can be composed of less experienced people.

Esfinge Guardian provides the application developer with tools for attacking traditional development problems without compromising its simplicity. Two design decisions are responsible for the simplicity of the framework: (*i*) Esfinge Guardian is a metadata-based framework that allows metadata schema extension, fully capable of adapting its internal algorithm based on the declared metadata associated with the protected operations [30]; and (*ii*) the use of Domain Annotations [21][22] allows the abstraction of complex

authorization policies, factoring them with business domain terminology.

The Esfinge Guardian framework contains ready-to-use metadata elements and component implementations for some of the classical access control models: RBAC, ABAC, and MAC. These implementations can be used and combined to represent an expressive number of authorization scenarios. Nevertheless, the framework can be easily extended to implement new authorization models that can be plugged in an application through custom metadata elements, working the same way as the existing implementations.

As with the previous authorization frameworks, we start showing the class that contains the authorization code, which can be seen in Listing 6. Esfinge Guardian links an authorization annotation, `@RespectHierarchy`, to an Authorizer class, `HierarchyAuthorizer`. This binding is done by using the annotation `@AuthorizerClass` in the definition of the authorization annotation, as presented in Listing 7.

```
public class HierarchyAuthorizer implements
    Authorizer<RespectHierarchy> {
  public Boolean authorize( AuthorizationContext ctx,
      RespectHierarchy rh) {
    Set<String> roles = ctx.subject("roles");
    //retrieve other relevant information from ctx
    return //hierarchy authorization logic;
  }
}
```

Listing 6. A possible implementation of the hierarchy authorization policy using Esfinge Guardian.

```
// Retention and ElementType suppressed
@AuthorizerClass(HierarchyAuthorizer.class)
public @interface RespectHierarchy {
}
```

Listing 7. Binding authorization annotation with the respective implementation.

A similar structure composed by an annotation and an authorizer class can be created to define the other rules from the bank authorization policy. For brevity, the code for the other annotations, `@WithingHQ` and `@SubodinateOnly`, and their respective authorizers are omitted. These three authorizer annotations can be added to a business method it is compliant with the authorization policy. Listing 8 presents a possible implementation of the business code.

```
@RespectHierarchy
@WithinHQ
@SubodinateOnly
public EmplOps[] overseeAllOps(EmplInfo info,
    CallerLocation cl) {
  EmplOps[] empops = //logic for retrieving data
  return empops;
}
```

Listing 8. A possible implementation of the business method using Esfinge Guardian.

In the described context, what Esfinge Guardian does is: (*i*) to intercept transparently the `overseeAllOps()` method call; *(ii)* to recognize that `@RespectHierarchy`, `@WithinHQ` and `@SubodinateOnly` are authorization annotations; (*iii*) to populate the authorization context with subject, object and environment information; (i*v*) to execute the authorization logic to verify if the security conditions to execute the method are satisfied; and, finally, *(v)* proceed or not with the method execution according to the result. The authorization context populator, the authorization logic and its representation on annotations are framework hot spots, meaning that they are extensible and can be adapted according to the software system needs.

Further details on how to use Esfinge Guardian framework, or on its respective architectural model, which lays the theoretical foundations for the framework, can be found in our previous works [30] [36].

VI. ANALYZING MODARILARITY AND EXTENSIBILITY FEATURES OF AUTHORIZATION FRAMEWORKS

The purpose of this section is to offer a deeper analysis on academic and industry authorization frameworks, highlighting the main differences between them, specially related to modularity and extensibility.

Subsection A proposes a comparison baseline for comparing extensibility and modularity of authorization frameworks. Subsection B focuses on the analysis of industry frameworks and, finally, Subsection C, of academic frameworks.

A. Common Comparison Baseline

There are a considerable number of authorization solutions in the industry and in the academic world. We have selected three of the main industry and academic authorization frameworks for the Java platform. However, for other languages and platforms there are other solutions with a similar approach.

One issue is how to establish a baseline for a proper comparison among authorization frameworks. This is important because authorization frameworks are designed with different purposes, making necessary to establish the points of comparison in order to reach a fair conclusion.

Table I proposes some requirements that must be taken into account in the authorization frameworks' design. They intend to establish requirements of extensibility and modularity that are desirable on authorization frameworks [5]. It is important to highlight that this analysis focus on extensibility and modularity aspects only, disregarding other equally important quality attributes, necessary for choosing security frameworks in real projects.

TABLE I
FRAMEWORKS DESIGN REQUIREMENTS

Req Id	Description
REQ01	Authorization frameworks must provide a way for granting authorization in a fine-grained level, considering the three basic authorization entities: subject; resource; environment
REQ02	The authorization mechanism must be able to transparently intercept the subject's requests to the resources
REQ03	Authorization concerns must be completely modularized into specific isolated units
REQ04	Authorization rules cannot depend on the location of the access data for authorization
REQ05	The intersection of authorization and business concerns must be declarative, and related to the business domain

B. Industry Authorization Frameworks

This subsection presents the analysis of each industry framework presented before guided by the requirements presented on table I. The code examples presented on Section V are referenced to exemplify some points.

Java EE

The Java EE Security framework specification does make a recommendation for declarative security instead of the programmatic approach [19]. But, it is not possible to avoid the programmatic approach, except for systems with classic RBAC authorization policies. The scenario authorization policy illustrates this point: it presents a hierarchical form of RBAC combined with geolocation based component.

We present a more systematic analysis as follows:

- REQ01 is not satisfied because Java EE does not inherently offer any means for authorizations to be in the fine-grained level.
- REQ02 is not satisfied because Java EE does not provide transparent interception. Instead, the developer has to explicitly call the security procedure, or to embed in the business code. Listing 1 is an example.
- REQ03 is not satisfied because the authorization code mingles with business code in non-trivial scenarios. Listing 1 is an example.
- REQ04 is not satisfied because there is no structured way of obtaining authorization context data in different places such as files, databases, transactions objects, session objects etc. The developer has to embed this logic into the business code as well.
- REQ05 is not satisfied because all the authorization code that is written into the business code adds nothing to the business domain itself. This is also a result of coupling the framework authorization code to the business class.

Spring Security

Spring Security offer means to express authorization policies for the ABAC model, due to the use of expression languages in its annotations. More precisely, the use of expression languages allows a form of RuBAC model [3], which is essentially a simplified version of ABAC.

We present a more systematic analysis as follows:

- REQ01 is partially satisfied because even though Spring Security is able to work in the fine-grained level theoretically, the mechanism does not scale well when the amount of individual items to be protected is large [45].
- REQ02 is satisfied because Spring Security understands that the target method in the resource is protected, freeing the developer from having to explicitly call the mechanism.
- REQ03 is partially satisfied because Spring Security authorization annotations are related to the framework, not to the business itself. A complex authorization policy could make the annotations hard to read and maintain.
- REQ04 is partially satisfied. Spring Security has the concept of `Evaluators` interface, which are called in custom authorization implementations. In the best scenario, a concrete implementation of this interface could be used to search for authorization context data, which can even be injected by the Spring framework. Especially for data passed as a parameter for the method, it would imply that the search for authorization context data would mingle with authorization logic itself.
- REQ05 is satisfied because there are a considerable number of scenarios in which the Spring Security authorization annotations can carry meaning to the business, adopting some best practices on how to write the annotation.

Axiomatics XACML

Although this architecture is good for handling fine-grained authorizations, when it comes to modularity there are not much tools beyond those already provided by object-oriented programming [37].

- REQ01 is fully satisfied. XACML is currently the best mechanism for fine-grained authorization nowadays.
- REQ02 is not satisfied because the developer has to explicitly make the authorization verification request.
- REQ03 is not satisfied because in the best scenario the authorization code would be modularized into a separate method, but an explicit call would have to be made to it for the authorization to take place.
- REQ04 is partially satisfied. The authorization data can be obtained from multiple places, however the software that are instantiating the framework is responsible to retrieve this data.
- REQ05 is not satisfied because the intersection point between authorization and business code is an embedded method call to the authorization code. This requirement is about declarative, cohesive intersection with business data such as by using domain annotations.

Esfinge Guardian

Esfinge Guardian has been developed from the start with these requirements as a guide.

We present a more systematic analysis as follows:

- REQ01 is satisfied because Esfinge Guardian is capable of operating in the fine-grained level in multiple scales.
- REQ02 is satisfied because Esfinge Guardian offers not only a mechanism for transparent interception, but can also be extended to use a different one.
- REQ03 is satisfied because authorization and business concerns are completely modularized.
- REQ04 is satisfied because it has a component type called `Populator`, which allow authorization data to be retrieved by the framework from anywhere. The framework provides some implementations of this kind of component, but it can also be extended by the application to implement custom populators.
- REQ05 is satisfied because of its support for domain annotations, which should be created with business meaning. Exemplifying this practice, the annotations `@RespectHierarchy`, `@SubordinateOnly` and `@WithinHQ` created on the example are not related to Esfinge Guardian, but to the application business.

Table II summarizes the analysis.

TABLE II
SUMMARY OF AUTHORIZATION FRAMEWORK COMPARISON

Req Id	Related Topic	Esfinge Guardian	Java EE 6	Spring Security	Axio-matics XACML
REQ01	Fine-grained capability	Complete	None	Partial	Complete
REQ02	Transparent interception	Complete	None	Complete	None
REQ03	Modularization of authorization concerns	Complete	None	Partial	None
REQ04	Data location independence	Complete	None	Partial	Partial
REQ05	Cohesion with the business domain	Complete	None	Complete	None

Industry authorization frameworks

C. Academic Authorization Frameworks

Sirbi and Kulkarni [38] present a discussion on the modularization of security concerns combining the Aspect-Oriented Programming (AOP) paradigm [26][27] with the Spring Security framework. Even though the authors recognize the importance of separation of concerns in their work, they focus on showing techniques on the implementation level, detailing how to combine AOP with Spring Security. However, their approach is representative of other solutions based on AOP [39][40], which covers modularization of crosscutting concerns (REQ03), transparent interception mechanism (REQ02), and it also offers a simple form of RuBAC (REQ01). The other architectural requirements presented in Table I are not covered. AOP's interception mechanism is based on the selection of joins points by the pointcuts. Join points are well defined points in the execution of a system such as method execution, method call, attribute read, and attribute write. In the case of Spring AOP, a join point is always equivalent to a method execution. Pointcut is the mechanism that specifies which join points will link aspects and classes. The Spring AOP interception mechanism in the business layer is inherently fragile because it is mostly based on method signatures [41]. That means, if a business method signature X changes, there is no feedback mechanism informing that the modularized crosscutting concern is not being considered in X anymore [30]. Another point in this work is that it is coupled to the Spring Security framework.

Camargo [43] propose an implementation of authentication and authorization concerns in AspectJ, aiming to make them reusable for web applications based on the MVC pattern and the Struts framework. The authors have implemented various levels of authorization: class, method, and attribute level, implementing the RuBAC model. Basically, the same arguments presented for the research of Sirbi and Kulkarni [38] apply for the work of Camargo [43], being coupled to the Apache Struts and restricted to web applications.

Welch and Stroud [42] propose an architectural model for modularizing security concerns using reflective security architecture for distributed computing. They compare a third-party application secured through inheritance and the proxy pattern with a re-engineered version that uses bytecode manipulation, obtaining a code reduction and a degree of separation of concerns that is not complete. They do not provide an access control model, but focus on presenting the technique they used for the separation. However, it should have the granularity of ABAC (REQ01). This inference is necessary because we had not access to their code. Extensibility aspects are not considered, nor cohesion with business domain (REQ05). An interesting point is that they critique the use of the Proxy Pattern, which is one of the interception mechanism used by Esfinge Guardian. The authors argue that applications that rely on this pattern for interception are subject to the bypass problem, which is a variant of the confinement problem [44]. In a complex application, it is always possible that an instance of a proxied class returned by a method invocation might not be replaced with an instance of its proxy. The unwrapped instance would bypass the proxy.

VII. CONCLUSION

This paper is an extension of the one previous work of ours [30]. We provide some theoretical background, discussing the main access control models in use nowadays. In this research, we add a discussion on the RAdAC and $UCON_{ABC}$ models. In addition, we present a discussion about the current problems in the existing authorization frameworks.

A motivating authorization scenario is proposed as a baseline for the comparisons on the rest of the work. Despite contrived, we believe that the proposed authorization scenario is a reasonable one for the comparisons.

We propose an implementation of the authorization policy for each one of the main authorization industry frameworks along with Esfinge Guardian. For each framework, we tried to use the best resources made available. In the case of other approaches for implementation, an extension of this analysis can be made considering the same requirements.

We reserve a Section for analyzing the implementation decisions: strengths and shortcomings, focusing on extensibility and modularity aspects. For a fairer comparison, we propose some development guiding requirements, which must be taken into account in the development of authorization frameworks. Some academic authorization frameworks are also analyzed.

The overall development time and authorization management effort might potentially be reduced, because of the complexity reduction in the use of the authorization rules constructs, and due to the increased semantic cohesion created by the use of domain annotations.

This paper is useful for software architects, framework developers, and software developers in general, by allowing the creation of more decoupled and extensible authorization solutions. Software architects could benefit from the Esfinge Guardian Architectural Model by instantiating a version of the architecture suitable for the enterprise needs. Framework developers could benefit by extending or re-creating the Esfinge Guardian in another language or platform. Finally, software developers in general could benefit from the understanding of the techniques involved in a framework development.

REFERENCES

[1] E. BERTINO; B. CATANIA; E. FERRARI; P. PERLASCA, "A logical framework for reasoning about access control models." *ACM Transactions on Information and System Security*, v. 6, no. 1, pp. 71-127, 2003.

[2] PRIVILEGE MANAGEMENT CONFERENCE COLLABORATION TEAM. A report on the privilege (access) management workshop. Washington, DC: NIST, 2010. (NIST-IR-7657).

[3] Hu, V. C., Ferraiolo, D. F., Kuhn D. R.: Assessment of Access Control (NIST-IR-7316). Gaithersburg, MD (2006)

[4] Hu, V. C., Scarfone, K.: Guidelines for Access Control System Evaluation Metrics NIST-IR-7874. Gaithersburg, MD (2012)

[5] Eduardo Guerra, Felipe Alves, Uirá Kulesza, Clovis Fernandes, A reference architecture for organizing the internal structure of metadata-based frameworks, Journal of Systems and Software, Volume 86, Issue 5, May 2013, Pages 1239-1256.

[6] Fayad, M., Schmidt, D. C., Johnson, R. E.: Building application frameworks: object-oriented foundations of framework design. In: Building application frameworks: object-oriented foundations of framework design, New York, Wiley, 55-83 (1999)

[7] Ferraiolo, D., Kuhn R., Chandramoulli, R.: Role-based access control. Artech House (2007)

[8] Ferraiolo, D., Kuhn, R.: Role-based Access Controls. In: Proceedings of 15th NIST-NCSC National Computer Security Conference, Baltimore, MD, 554-563 (1992).

[9] Elliott, A. A., Knight, G. S.: Role Explosion: Acknowledging the Problem. In: Proceedings of the 2010 International Conference on Software Engineering Research & Practice. (2010)

[10] Sandhu, R., Ferraiolo, D.F., Kuhn, D.R.: The NIST Model for Role-Based Access Control: Toward a Unified Standard. In: 5th ACM Workshop Role-Based Access Control. pp. 47–63. (2000).

[11] Probst, S., Kung, J.: The need for declarative security mechanisms. In: Proceedings of 30th Euromicro Conference, pp. 526- 531 (2004)

[12] Merz, M.: Enabling declarative security through the use of Java Data Objects. In: Journal of Science of Computer Programming, V. 70, n. 2-3, pp. 208-220 (2008)

[13] Bartsch, S.: Authorization Enforcement Usability Case Study. In: ESSoS'11: Proceedings of the Third international conference on Engineering secure software and systems, pp. 209-220 (2011)

[14] Hai-bo, S., Fan, H.: An Attribute-Based Access Control Model for Web Services. In: PDCAT '06. Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies, pp.74-79 (2006)

[15] Peng, J., Yang, F.: Description Logic Modeling of Temporal Attribute-Based Access Control. In: ICCE '06. First International Conference on Communications and Electronics, pp.414-418 (2006)

[16] Hsieh, G., Foster, K., Emamali, G., Patrick, G., Marvel, L.: Using XACML for Embedded and Fine-Grained Access Control Policy. In: ARES '09 International Conference, pp.462-468 (2009)

[17] XACML: eXtensible Access Control Markup Language (XACML), Version 3.0, Committee Specification 01. http://docs.oasisopen.org/xacml/3.0/xacml-3.0-core-spec-cs-01-en.pdf (2011)

[18] Bo, L, Nan, Z., Kun, G., Kai, C.: An XACML Policy Generating Method Based on Policy View. ICPCA 2008: 3rd International Confer. on Pervasive Computing and Applications, v.1, pp.295-301 (2008)

[19] Java EE: Java Enterprise Edition Tutorial 6. http://docs.oracle.com/javaee/6/tutorial/doc/bnbyl.html (2013).

[20] Spring Framework: Spring Source Community. http://www.springsource.org/ (2013)

[21] Perillo, J., Guerra, E., Silva, J., Silveira, F., Fernandes, C.: Metadata Modularization Using Domain Annotations. In: Workshop on Assessment of Contemporary Modularization Techniques. V. 3, Orlando (2009)

[22] Perillo, J., Guerra, E., Fernandes, C.: Daileon-A Tool for Enabling Domain Annotations. In: RAM-SE '09: Proceedings of the Workshop on AOP and Meta-Data for Software Evolution, n. 7 (2009)

[23] Trusted Computer System Evaluation Criteria (Orange Book), Department of Defense. http://csrc.nist.gov/publications/history/dod85.pdf (1985)

[24] Sayaf, R., Clarke D.: Access Control Models for Online Social Networks. In: Social Network Engineering for Secure Web Data and Services, (2012)

[25] R. Sayaf. Access control for online social networks - research summary. In: For your eyes only conference. Brussels. (2012)

[26] Ribeiro, M., Dosea, M., Bonifácio, R., Neto, A. C., Borba, P., Soares, S.: Analyzing Class and Crosscutting Modularity Structure Matrixes. In Proceedings of the 21th Brazilian Symposium on Software Engineering (SBES) (2007)

[27] Neto, A. C., Ribeiro, M., Dósea, M., Bonifácio, R., Borba, P., Soares, S.: Semantic Dependencies and Modularity of Aspect-Oriented Software. In: Workshop on Assessment of Contemporary Modularization Techniques (2007)

[28] Guerra, Eduardo, Buarque, Eduardo, Fernandes, Clovis, Silveira, Fábio (2013) A Flexible Model for Crosscutting Metadata-Based Frameworks. Computational Science and Its Applications – ICCSA 2013, Lecture Notes in Computer Science, V 7972, 391-407.

[29] Motta, G.H.M.B.; Furuie, S.S., "A contextual role-based access control authorization model for electronic patient record," *Information Technology in Biomedicine, IEEE Transactions on* , vol.7, no.3, pp.202,207, Sept. 2003

[30] Silva, J., Guerra, E., Fernandes, C.: An Extensible and Decoupled Architectural Model for Authorization Frameworks. In: Murgante, B., Misra, S., Carlini, M., Torre, C.M., Nguyen, H.-Q., Taniar, D., Apduhan, B.O., Gervasi, O. (eds.) ICCSA 2013, Part IV. LNCS, vol. 7974, pp. 614–628. Springer, Heidelberg (2013)

[31] Kandala, S.; Sandhu, R.; Bhamidipati, V., "An Attribute Based Framework for Risk-Adaptive Access Control Models," Availability, Reliability and Security (ARES), 2011 Sixth International Conference on , vol., no., pp.236,241, 22-26 Aug. 2011

[32] Ferreira, A.; Chadwick, D.; Farinha, P.; Correia, R.; Gansen Zao; Chilro, R.; Antunes, L., "How to Securely Break into RBAC: The BTG-RBAC Model," Computer Security Applications Conference, 2009. ACSAC '09. Annual , vol., no., pp.23,31, 7-11 Dec. 2009

[33] PARK, J.; SANDHU, R. The $UCON_{ABC}$ usage control model. ACM Transactions on Information System Security, v. 0, n. 0, February, 2004.

[34] Yonggang Ding; Junhua Zou, "DRM Application in $UCON_{ABC}$," Advanced Software Engineering and Its Applications, 2008. ASEA 2008 , vol., no., pp.182,185, 13-15 Dec. 2008

[35] Srijith K. Nair, Andrew S. Tanenbaum, Gabriela Gheorghe, and Bruno Crispo. 2008. Enforcing DRM policies across applications. In Proceedings of the 8th ACM workshop on Digital rights management (DRM '08). ACM, New York, NY, USA, 87-94.

[36] Silva, J. O. An Architectural Model for Access Control Frameworks Extensible for Different Authorization. São José dos Campos, 2013. Master's Thesis 114f.

[37] Rissanen E, Brossard D, Slabbert A Distributed access control management—a xacml-based approach. In: ICSOC-servicewave. Springer, Berlin, 2009

[38] Sirbi, K.; Kulkarni, P. J. Modularization of enterprise application security through Spring AOP. International Journal of Computer Science & Communication, v. 1, n. 2, p. 227-231, 2010.

[39] Fernandez, L. L.; Carrillo, M. G.; Pelaez, J.; Fernandez, F. A declarative authentication and authorization framework for convergent IMS/Web application servers based on aspect oriented code injection. In: IMSAA INTERNATIONAL CONFERENCE ON INTERNET MULTIMEDIA SERVICES ARCHITECTURE AND APPLICATIONS, 2, 2008, Bangalore. Proceedings… Bangalore: IMSAA, 2008. p. 1-6.

[40] HAI-BO, S. A semantic and attribute-based framework for web services access control. In: ISA INTERNATIONAL WORKSHOP ON INTELLIGENT SYSTEMS AND APPLICATIONS, 2, 2010, Wuhan. Proceedings… Wuhan: ISA, 2010, p.1-4.

[41] Silva, J. Frameworks orientados a aspectos baseados em metadados. São José dos Campos: Aeronautics Institute of Technology (ITA), 2008.

[42] Welch, I. S.; Stroud, R. J. Re-engineering security as a crosscutting concern. The Computer Journal, v. 46, n. 5, p. 578-589, 2003.

[43] Camargo, V. V. Frameworks transversais: definições, classificações, arquitetura e utilização em um processo de desenvolvimento de software. 2006. PhD's Thesis in Computing Science – University of São Paulo, São Carlos, 2006.

[44] Lampson, B. W. A note on the confinement problem. Communications of ACM. v. 16, n. 10, p. 613–615, October, 1973.

[45] LU, Peng; YIN, Zhao-lin. Analysis and extension of authentication and authorization of Acegi security framework on spring [J]. Computer Engineering and Design, v. 6, p. 022, 2007.

Eduardo Martins Guerra received a PhD in Computer and Electronic Engineering from Aeronautics Institute of Technology - ITA in 2010. He has been working for National Institute for Space Research since 2013 where he is an Associate Researcher. His current research interests include Software Engineering, Framework Development and Application Security.

Jefferson O. Silva received a Master's degree from Aeronautics Institute of Technology - ITA in 2013. He has been working for Pontifícia Universidade Católica de São Paulo - PUC-SP since 2011. He is currently a doctoral student in Instituto de Matemática e Estatística in Universidade de São Paulo - IME-USP. His current research interests include Social Computing, Software Engineering, and Computer Security.

Clovis Torres Fernandes received a PhD in Computer Science from Pontifícia Universidade Católica of Rio de Janeiro – PUC/Rio in 1992. He has been working for Instituto Tecnológico de Aeronáutica – ITA since 1980 where he is an Associate Professor and Director of LAI – Learning and Interaction Laboratory in the Computer Science Department. His current research interests include Software Engineering, Computers and Education and Computer Security.

The Producer-Consumer Collusion Attack in Content-Centric Networks

A. Nasserala and I. M. Moraes

***Abstract*— This paper evaluates a denial-of-service attack in information-centric networks based on the Content Centric Networking (CCN) architecture. This attack aims at increasing the content retrieval time. In this attack, both malicious consumers and producers collude, by generating, publishing, and changing content popularity. Malicious contents are stored by intermediate nodes and occupy the cache space that should be occupied by legitimate content. Thus, the probability of a legitimate consumer retrieves content directly from the producer increases as well as the content retrieval time. We evaluate the impact of the attack by varying the number of consumers and producers in collusion, the interest packets rate, and the way malicious contents are requested. Results show if 20% of consumers are malicious and send 500 interests/s each, the content retrieval time experienced by legitimate users increases by 20 times, which shows the effectiveness of the attack.**

***Keywords*— Future Internet, CCN, Security, Denial of Service.**

I. INTRODUÇÃO

AS REDES Orientadas a Conteúdo são um novo paradigma de comunicação para a Internet [3]. O principal objetivo dessas redes é a entrega de conteúdo para os usuários independentemente da localização desse conteúdo, ao contrário da arquitetura TCP/IP, cujo objetivo é a comunicação entre sistemas finais. Diversas arquiteturas foram propostas para esse novo paradigma de comunicação e uma das arquiteturas com maior destaque na literatura é a *Content Centric Networking* (CCN) [8,13]. Entre as principais características da CCN estão o roteamento através de nomes de conteúdo, o armazenamento de conteúdo em nós intermediários da rede e a capacidade de auto-certificar o conteúdo, aplicando a segurança diretamente aos pacotes de dados [8].

Uma das mais vantajosas características da CCN é o consumo indireto de conteúdo, ou seja, qualquer nó da rede que ao receber uma solicitação de conteúdo e possua esse conteúdo em *cache* pode enviar tal conteúdo para o nó solicitante. Na CCN, o nó que solicita o conteúdo é chamado de consumidor e o nó que disponibiliza o conteúdo é chamado de produtor. O produtor é a fonte de um conteúdo. Na CCN, é possível que um determinado nó, que esteja mais próximo do consumidor, consiga responder à solicitação de um conteúdo sem que o consumidor seja obrigado a recuperar esse conteúdo diretamente do produtor, que pode estar mais distante. Assim, o tempo de recuperação de conteúdos pode ser reduzido. Além disso, o armazenamento de conteúdo em *cache* aumenta a disponibilidade de conteúdos e pode reduzir o consumo de banda, uma vez que o conteúdo é encaminhado por menos saltos.

Outra característica da CCN é que a segurança é aplicada diretamente aos conteúdos, diferentemente da arquitetura TCP/IP, na qual a segurança é aplicada ao canal de comunicação entre os sistemas finais [13]. Um pacote de dados CCN é auto-certificado, isto é, ele contém a assinatura digital do pacote e a chave pública do produtor [8]. Portanto, é possível verificar a integridade do pacote e se ele foi gerado pelo produtor que possui tal chave pública. O uso de assinaturas gera sobrecarga tanto para o produtor assinar o conteúdo quanto para os consumidores verificarem a assinatura. Além disso, a CCN é mais robusta a ataques de negação de serviço (*Denial of Service* - DoS) comuns na Internet atual, como o de esgotamento de banda e o de reflexão, em virtude do uso de *cache* pelos nós intermediários e da agregação de solicitações de conteúdo [5], como será discutido na Seção II.

Um ataque de negação de serviço particular, chamado de conluio produtor-consumidor, entretanto, pode ser efetivo porque os mecanismos nativos empregados pela CCN não são suficientes para inibi-lo. Não foi encontrado na literatura uma avaliação do ataque no qual produtores e consumidores agem em conluio na CCN. Este trabalho avalia o impacto desse ataque em redes CCN, cujo objetivo é aumentar o tempo de recuperação de conteúdos. Nesse ataque, consumidores maliciosos solicitam conteúdos que são disponibilizados apenas por produtores maliciosos a uma alta taxa. Isso aumenta o tempo de recuperação de conteúdos legítimos, em virtude do aumento da taxa de erro do *cache* (*cache miss*) para esses conteúdos e, consequentemente, da necessidade de nós legítimos terem que recuperar o conteúdo diretamente do produtor. Os mecanismos de segurança da CCN padrão são ineficazes na detecção do ataque em conluio, pois, do ponto de vista da rede, as solicitações e os conteúdos são legítimos. São enviados pacotes de interesse para conteúdos que existem e que são disponibilizados por produtores. O conteúdo é malicioso porque torna popular um conteúdo que não é de interesse de usuários legítimos. Como o produtor malicioso assina os conteúdo de acordo com a política definida pela CCN, os consumidores maliciosos podem solicitá-los sem risco de que esses conteúdos sejam descartados por mecanismos de verificação de assinaturas e chaves. Esse ataque é possível, porque a CCN emprega políticas de substituição do *cache* baseadas, em sua maioria, na popularidade dos conteúdos. Assim, se um determinado conteúdo não é solicitado com frequência ou não foi solicitado

A. Nasserala, Universidade Federal do Acre (UFAC), Rio Branco, AC, Brasil, anasserala@ic.uff.br

I. M. Moraes, Universidade Federal Fluminense (UFF), Niterói, RJ, Brasil, igor@ic.uff.br

recentemente pelos consumidores, ele é considerado menos popular. Dessa forma, esse conteúdo terá prioridade de descarte quando houver necessidade de armazenar novos conteúdos. Vários nós maliciosos podem, então, solicitar um conjunto específico de conteúdos produzidos maliciosamente e em taxas altas de envio de interesse para manipular a política de *cache*. Assim, dependendo da forma como os conteúdos maliciosos são solicitados, é possível até remover conteúdos legítimos do *cache*.

A avaliação do ataque de conluio consumidor-produtor é feita através de simulações para diferentes configurações, nas quais se variam o número de consumidores e produtores em conluio, a taxa de pacotes de interesse e o padrão de solicitações de conteúdos maliciosos. As métricas empregadas são o tempo de recuperação de conteúdos legítimos, o percentual de ocupação maliciosa do *cache*, o percentual da taxa de erros de *cache* de conteúdos legítimos e o percentual de conteúdos legítimos recuperados do produtor. Os resultados mostram que o ataque compromete uma das maiores vantagens das CCN que é a redução do tempo de recuperação de conteúdos pelo uso do *cache* nos nós intermediários. Conclui-se que se 20% dos nós consumidores são maliciosos e enviam 500 solicitações de interesses por segundo cada um, o tempo de recuperação de conteúdos por usuários legítimos aumenta em cerca de 20 vezes na topologia avaliada. Além disso, observa-se que até 67% dos conteúdos legítimos são recuperados diretamente do produtor nas configurações analisadas.

O restante do artigo está organizado da seguinte forma. Na Seção II uma revisão sobre o funcionamento e aspectos de segurança da CCN é apresentada. Na Seção III os trabalhos relacionados são discutidos. Na Seção IV, o ataque de negação de serviço em conluio consumidor-produtor é descrito. Na Seção V, é definido o cenário de avaliação usado nas simulações. Na Seção VI, os resultados dos experimentos são analisados e discutidos. Por fim, na Seção VII, as conclusões são apresentadas.

II. A ARQUITETURA CCN: FUNCIONAMENTO E ASPECTOS DE SEGURANÇA

A arquitetura CCN tem como objetivos aumentar a disponibilidade e reduzir o tempo de recuperação de conteúdos. Na CCN, os nós da rede possuem um *cache* para armazenar conteúdos recebidos previamente. Consequentemente, qualquer nó pode responder a um pedido, se o conteúdo solicitado está disponível em seu *cache*, conhecido como *Content Store* (CS). Os consumidores são os nós que solicitam um conteúdo. Quanto mais nós armazenam um conteúdo na rede, maior a disponibilidade desse conteúdo e maior a probabilidade de consumidores recuperarem esse conteúdo de um nó mais próximo. Essa é uma das vantagens da CCN em comparação com a arquitetura atual da Internet.

A CCN emprega dois tipos de pacotes: interesse e dados. Consumidores enviam pacotes de interesse para solicitar um conteúdo. Os produtores ou qualquer outro nó que possua o conteúdo em seu CS respondem aos interesses com pacotes de dados, que carregam o conteúdo em si ou pedaços de conteúdo [4]. Os nós encaminham tanto pacotes de interesse quanto pacotes de dados com base no próprio nome do conteúdo, ao invés do endereço de destino do nó que possui o conteúdo. Para realizar o encaminhamento de pacotes, cada nó CCN tem duas estruturas de dados: a *Pending Interest Table* (PIT) e a *Forwarding Information Base* (FIB). A PIT guarda o estado de cada pacote de interesse encaminhado por um nó que ainda não recebeu uma resposta, ou seja, os interesses que esperam por um pacote de dados. Cada entrada da PIT também armazena a interface de recepção de um pacote de interesse. É importante ressaltar que o tamanho da PIT é limitado e, dessa forma, novos interesses que chegam enquanto a tabela está cheia não são encaminhados. Esse fato é explorado por usuários maliciosos, como detalhado nos próximos parágrafos.

A FIB atua como uma tabela de encaminhamento para pacotes de interesse. Essa tabela contém uma lista de entradas, cada uma contendo o prefixo de um nome e uma lista de interfaces de saída para as quais os pacotes de interesse com nomes de mesmo prefixo devem ser encaminhados.

Quando um nó CCN recebe um pacote de interesse, ele verifica seu CS para encontrar uma cópia do conteúdo solicitado, cujo nome está no cabeçalho do pacote de interesse. Se o conteúdo está armazenado em *cache*, o nó envia um pacote de dados para o consumidor. Caso contrário, o nó verifica a sua PIT. Se houver uma entrada na PIT para o mesmo conteúdo, o nó atualiza a lista de interfaces de entrada e descarta o pacote de interesse. Esse procedimento é chamado de agregação de pacotes de interesse e torna a CCN mais robusta contra ataques de DoS, como discutido a seguir. Caso contrário, o nó cria uma nova entrada na PIT e, então, consulta a FIB para determinar a interface de saída para encaminhar o pacote de interesse. Se não houver nenhuma entrada na FIB relacionada com o nome do conteúdo, o pacote de interesse é descartado. Os nós repetem este processo de encaminhamento para cada pacote de interesse recebido. Os pacotes de dados seguem o caminho reverso percorrido pelos pacotes de interesse porque a PIT armazena a lista de interfaces com interesses a serem atendidos [15].

Ataques de negação de serviço (DoS) são uma ameaça na Internet atual. A arquitetura CCN, entretanto, é mais robusta a esse tipo de ataque do que a pilha TCP/IP em virtude de duas características: o armazenamento de conteúdo pelos nós intermediários e a agregação de pacotes de interesse [6]. Ataques de esgotamento de banda e de reflexão, por exemplo, são pouco eficientes na CCN.

Ataques de esgotamento de banda inundam a vítima com requisições de serviço para esgotar seus recursos. Neste ataque, os pacotes devem chegar à vítima para que o ataque seja efetivo. Na CCN, no entanto, os pacotes não possuem o endereço de destino e os consumidores não podem garantir que os pacotes de interesse alcancem a origem do conteúdo, ou seja, o produtor e nesse a caso a vítima, porque qualquer nó pode responder ao interesse. Portanto, os consumidores maliciosos podem gerar uma quantidade enorme de pacotes de interesse para um dado conjunto de conteúdos, mas nenhum ou poucos desses pacotes alcançarão o produtor.

É importante ressaltar também que nós no caminho reverso entre o consumidor e o produtor armazenam conteúdos em *cache*. Assim, nós intermediários entre consumidor e produtor provavelmente irão satisfazer novos pacotes de interesse, que dificilmente alcançarão o produtor, dependendo da popularidade destes conteúdos. Além disso, a CCN reduz o número de pacotes de interesse transmitidos. Um nó só envia um pacote de interesse que não corresponde a uma entrada PIT. Caso contrário, o nó atualiza a lista de interfaces e descarta o pacote, como descrito nos parágrafos anteriores.

Ataques de reflexão são baseados na técnica de falsificação de endereços IP (IP *spoofing*) e visam atacar vítimas diferentes simultaneamente. Na CCN, esses ataques são menos eficazes porque os pacotes de dados são sempre encaminhados para o consumidor através do caminho reverso percorrido pelo pacote de interesse. Consumidores também não podem garantir que os pacotes de interesse cheguem às vítimas intermediárias ou finais devido ao *cache* nos nós intermediários. Nós CCN, porém, enviam pacotes de interesse em todas as suas interfaces, se não houver nenhuma entrada na FIB para um prefixo de nome solicitado. Portanto, se o atacante e a vítima estão na mesma sub-rede do ataque, a reflexão pode ser eficaz [6]. Neste cenário, o atacante pode enviar pacotes de interesse através de todas as suas interfaces com os endereços da camada MAC falsificados. Assim, múltiplas cópias do conteúdo são enviadas para a vítima. Para evitar isso, nós CCN não transmitem o mesmo conteúdo mais de uma vez no mesmo domínio de difusão (*broadcast*) [6].

Apesar de ser mais robusta do que a arquitetura TCP/IP aos ataques de DoS atuais, a arquitetura CCN possui ataques e vulnerabilidades identificados em trabalhos recentes [6,12], que são discutidos na Seção III.

III. TRABALHOS RELACIONADOS

Os ataques de negação de serviço em CCN são classificados em dois tipos: ataques por inundação de interesses ou envenenamento de *cache* [12].

O objetivo dos ataques de inundação de interesses é sobrecarregar a PIT com solicitações de conteúdo enviadas por um nó malicioso a uma alta taxa [7]. Os pacotes de interesse maliciosos, em geral, solicitam conteúdos inexistentes, o que mantém por mais tempo a informação sobre esses interesses na PIT de um nó. A informação sobre um interesse pendente só é removida após o estouro de um temporizador. Enquanto aguarda pelo pacote de dados, o nó receberá novos interesses para outros conteúdos inexistentes. No pior caso, com a PIT cheia, um nó afetado não atenderá interesses legítimos, o que leva à queda de desempenho da rede.

Gasti *et al.* [6] definem o ataque de inundação de interesses e propõem um mecanismo de *push-back* como contramedida. Esse mecanismo monitora a ocupação da PIT e identifica quando uma determinada interface está próxima de atingir seu número máximo de entradas na PIT. Assim, o mecanismo controla o fluxo de pacotes de interesse que contém os mesmos prefixos de nome. Além disso, a contramedida envia uma notificação na interface supostamente atacada que será recebida por um nó vizinho. Esse nó, por sua vez, deve propagar tal informação no sentido das interfaces atacadas e, ao mesmo tempo, limitar a taxa de interesses encaminhados que contenham o prefixo sob ataque. Portanto, o objetivo da contramedida é empurrar o ataque para o caminho de volta até o atacante, ou pelo menos para um nó no qual seja detectado [6]. A principal característica dessa contramedida é não modificar a arquitetura padrão proposta para a CCN. O ponto fraco do trabalho de Gasti *et al.* é que nem o impacto do ataque e nem a contramedida proposta são avaliados por simulação ou experimentos práticos.

Choi *et al.* [5], por outro lado, avaliam através de simulações a efetividade do ataque de inundação de interesses. Os autores mostram que em uma rede com poucos nós, o desempenho é comprometido. Conclui-se que a vazão de dados total de consumidores legítimos diminuiu cerca de 65%. Da mesma forma, observa-se que o tempo médio de recuperação de conteúdos aumenta rapidamente, logo após o início do ataque.

Afanasyev *et al.* [1] também avaliam o ataque de inundação de interesses através de simulações, porém consideram diferentes cenários e uma rede de maior escala do que a usada no trabalho anterior. Os autores também avaliam a contramedida baseada em um mecanismo de *push-back* proposta por Gasti *et al*. Os resultados mostram que essa contramedida é eficiente, pois isola por completo os atacantes de modo que eles causem pouco ou nenhum impacto no desempenho percebido por usuários legítimos.

Diferentemente dos ataques de inundação de interesses, o objetivo do ataque de envenenamento de *cache* é ocupar o *cache* dos nós com conteúdo poluído. Esse conteúdo é enviado por consumidores maliciosos para fazer com que nós armazenem um conteúdo que possua uma assinatura válida, porém corrompido ou aumentem a popularidade de conteúdos menos populares. No primeiro caso, o objetivo é reduzir o espaço disponível em *cache* para armazenar conteúdos legítimos e fazer com que consumidores recebam conteúdos corrompidos. No segundo, o objetivo é remover do *cache* conteúdos legítimos assumindo que uma política de substituição baseada na popularidade dos conteúdos é usada. Uma contramedida ao ataque de envenenamento de *cache* é a verificação da assinatura contida nos pacotes de dados [13]. Por padrão, a assinatura dos conteúdos é verificada apenas pelos nós de borda, ou seja, os consumidores, e não pelos nós intermediários da rede. Essa característica garante que os consumidores não recebam pacotes de dados contendo conteúdo malicioso. Nesse caso, o serviço da CCN é negado se os consumidores sempre receberem conteúdos inválidos. A solução de obrigar a verificação da assinatura de todos os conteúdos em todos os nós implica sobrecarga de processamento e, por isso, é de difícil adoção prática [6].

Ribeiro *et al.* [10,11] propõem um mecanismo de verificação probabilística de assinaturas. O mecanismo proposto é eficiente, porém se mostrou dependente da topologia de rede utilizada. Quanto maior o número de saltos, maior a probabilidade do conteúdo poluído ser descartado ao longo do caminho. Outra proposta similar, chamada

CacheShield [14], também usa dados estatísticos para verificar se um conteúdo é poluído ou não, porém tem as mesmas limitações do trabalho de Ribeiro *et al.*

Kim *et al.* [9] investigam o impacto de fluxos de conteúdo de longa duração na CCN. A presença de fluxos de longa duração pode ter efeito similar ao ataque de envenenamento de *cache*. Se fluxos de longa duração ocuparem temporariamente um *cache* de um nó por determinado conteúdo, eles podem expulsar pedaços de conteúdos populares do *cache*. Consequentemente, reduz-se a taxa de acertos do *cache* (*cache hit*). Os resultados das simulações mostram que há degradação da taxa de acertos do *cache* quanto maior é o número de fluxos de longa duração.

Todos os trabalhos descritos anteriormente que avaliam e/ou propõem contramedidas para os ataques de inundação de interesses e envenenamento de *cache* não consideram a possibilidade de ataques em que consumidores e produtores maliciosos agem em conluio para gerar, disponibilizar e manipular a popularidade de conteúdos. Avaliar tal ataque, detalhado na Seção IV, é o principal objetivo deste trabalho.

IV. ATAQUE DE NEGAÇÃO DE SERVIÇO POR CONLUIO CONSUMIDOR-PRODUTOR

No ataque de conluio existem pelo menos dois atores: o produtor malicioso e o consumidor malicioso, como ilustrados na Fig. 1. O produtor malicioso é responsável por produzir conteúdo malicioso conforme a demanda do consumidor malicioso. Esse conteúdo tem as mesmas características do conteúdo legítimo e, portanto, não terá tratamento diferenciado por nenhum nó CCN. Isso quer dizer que esse tipo de conteúdo ocupará o *cache* dos nós com o mesmo tratamento dos demais conteúdos que trafegam na rede. Os nomes dos conteúdos maliciosos também seguem as especificações da CCN. O consumidor malicioso, por sua vez, solicita conteúdos maliciosos em altas taxas. Na Fig. 1, os nós R1 e R2 armazenam conteúdos maliciosos em *cache*, uma vez que estão no caminho entre o consumidor e o produtor malicioso.

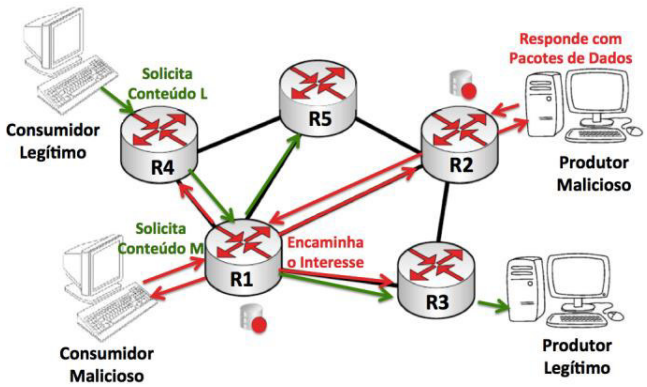


Figura 1. O ataque em conluio consumidor-produtor: nós legítimos e maliciosos em ação.

Com o ataque em conluio, o objetivo é prejudicar o consumo indireto de conteúdos, isto é, obrigar um consumidor legítimo a recuperar o conteúdo desejado diretamente do produtor. Esse objetivo é alcançado através da manipulação da popularidade dos conteúdos armazenados em *cache*.

Consumidores maliciosos enviam pacotes de interesse para um grupo de conteúdos que existem e que são respondidos pelo produtor malicioso. Assim, se solicitado com frequência, um conteúdo se torna popular, apesar de não ter sido solicitado por usuários legítimos. Por isso, o conteúdo é dito malicioso. Esse ataque é possível, porque a CCN emprega políticas de substituição do *cache* baseadas, em sua maioria, na popularidade dos conteúdos. Assim, se um determinado conteúdo não é solicitado com frequência ou não foi solicitado recentemente pelos consumidores, ele é considerado menos popular. Dessa forma, terá prioridade de descarte quando houver necessidade de armazenar novos conteúdos. Ao solicitar um conjunto específico de conteúdos e em taxas altas, os nós maliciosos manipulam a política de *cache*. Com mais conteúdos maliciosos em *cache*, maior a taxa de erro para os conteúdos legítimos e, consequentemente, a necessidade de nós legítimos terem que recuperar o conteúdo diretamente do seu produtor.

Mesmo que os consumidores legítimos não tenham que consumir diretamente do produtor, eles terão seus interesses encaminhados por mais saltos até conseguir o conteúdo desejado. No exemplo da Fig. 1, o consumidor legítimo pode ter que recuperar o conteúdo solicitado diretamente do produtor legítimo, uma vez que o *cache* do nó R1 que está no caminho entre os dois, pode estar sobrecarregado com conteúdo malicioso.

Uma das principais razões para que o ataque em conluio produtor-consumidor seja bem-sucedido é o fato de que os pacotes de interesse e de dados usados no ataque são legítimos para a rede e, portanto, não são detectados por mecanismos de verificação de assinaturas. O pacote com o conteúdo malicioso possui uma assinatura válida, carrega a chave do publicador e, assim, passa no teste de verificação de integridade e autenticidade. Logo, não é identificado como malicioso e nem descartado.

Outro objetivo do ataque em conluio é reduzir a eficiência da PIT, ao enviar pedidos de interesses para diferentes conteúdos maliciosos disponibilizados por produtores maliciosos a uma alta taxa. Dessa forma, é possível burlar o mecanismo de *push-back* proposto por Gasti *et al.* [6]. Esse mecanismo é eficiente contra a inundação de pacotes de interesse porque consegue identificar prefixos de nomes de conteúdo que frequentemente estão pendentes na PIT, uma vez que o conteúdo solicitado é inexistente.

Porém, se o consumidor e produtor estiverem agindo em conluio, os pacotes de interesse terão uma entrada na PIT de um nó somente até o conteúdo malicioso, que existe, retornar. Portanto, o mecanismo *push-back* não terá sucesso ao tentar identificar o ataque, pois os pacotes de interesse receberão uma resposta legítima e suas entradas serão removidas da PIT. Nesse caso, o ataque em conluio não provoca o esgotamento de recursos de armazenamento da PIT em um nó. O objetivo do ataque é gerar uma grande quantidade de pacotes de interesses, fazendo com que um nó tenha que manipular muitas solicitações de conteúdo maliciosas em detrimento a

interesses legítimos, o que pode levar a negação de serviço nesse nó.

V. CENÁRIO DE AVALIAÇÃO

A topologia da rede usada na avaliação do impacto do ataque em conluio é composta por 32 nós dispostos em forma de árvore, como mostra a Fig. 2. Os 24 nós folha são consumidores. O número de consumidores legítimos (CL) é fixo em todas as configurações e igual a 16. O número de consumidores maliciosos (CA) varia de 0 a 8. A posição dos CLs e dos CAs é definida aleatoriamente em cada rodada de simulação. O produtor legítimo (PL) é sempre o nó raiz. O produtor malicioso (PA) é o nó filho do nó raiz. Os demais 6 nós que compõem a topologia são os roteadores da rede (RTR). Os enlaces que interconectam os nós possuem taxa de transmissão de 100 Mb/s e atraso de 1 ms.

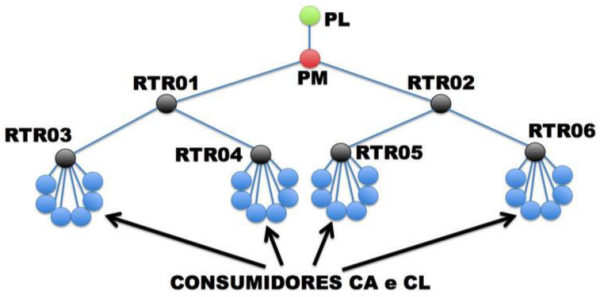


Figura 2. A topologia da rede usada nas simulações.

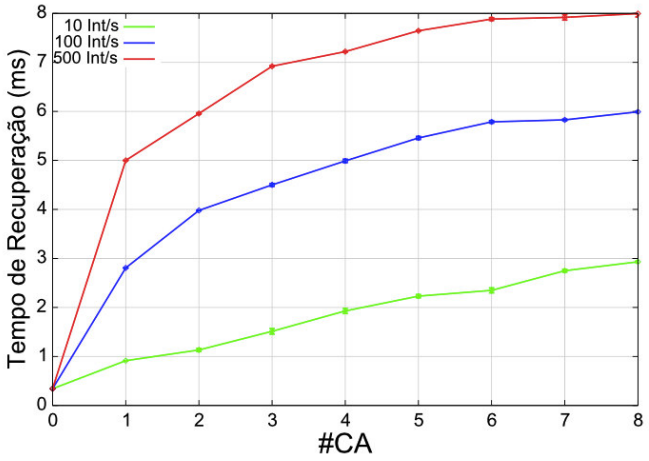
Os conteúdos são solicitados da seguinte forma. Os consumidores maliciosos enviam interesses para 12 conteúdos que são disponibilizados pelo único produtor malicioso a taxas de 10, 100 e 500 interesses por segundo. Cada conteúdo malicioso possui 100 pedaços (*chunks*) e prefixos de nome diferentes. Os pedaços de conteúdo são solicitados de duas formas diferentes: o consumo segundo a popularidade do conteúdo malicioso, seguindo uma distribuição *Zipf* com parâmetro α = 0,7[2], e o consumo sequencial, dito CBR (*Constant Bit Rate*), no qual um consumidor envia pacotes de interesse ordenados pelo nome do conteúdo e de forma cíclica. Os consumidores legítimos sempre enviam 10 interesses/s para outros 12 conteúdos disponibilizados pelo produtor legítimo. Cada conteúdo malicioso possui 100 pedaços (*chunks*) e prefixos de nome diferentes. Os pedaços são solicitados seguindo uma distribuição *Zipf* com parâmetro α = 0,7.

O *cache* dos consumidores legítimos e dos roteadores tem capacidade para armazenar até 1000 pedaços de conteúdo e cada pedaço possui 1024 bytes. Os consumidores maliciosos não possuem *cache* para potencializar o ataque, isto é, sempre enviam interesses independentemente se já receberam o conteúdo anteriormente ou não. A PIT tem tamanho ilimitado para que seja possível avaliar apenas o efeito do aumento da ocupação maliciosa no *cache* dos nós. A política de substituição de *cache* é a *Least Recently Used* (LRU).

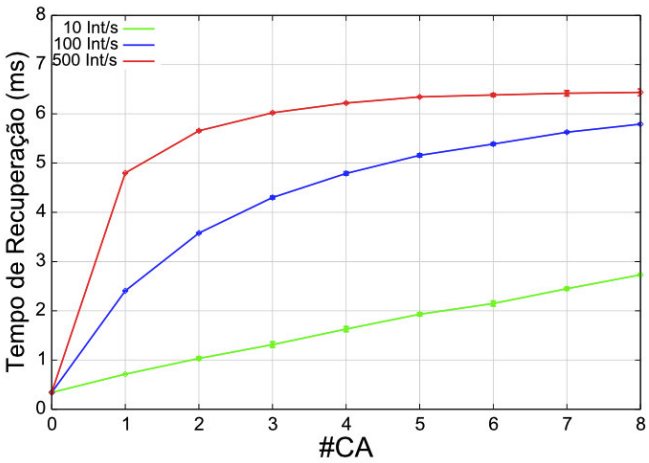
O módulo ndnSIM do simulador NS-3 é usado na avaliação. Para cada configuração, são realizadas 50 rodadas de simulação, cada uma com duração de 180 s. Para os pontos dos gráficos obtidos, são calculados intervalos de confiança representados por barras verticais para um nível de confiabilidade de 95%.

VI. RESULTADOS

Os resultados apresentados têm como objetivo avaliar o impacto do ataque de negação de serviço por conluio produtor-consumidor no desempenho da CCN. As métricas de desempenho são o tempo médio de recuperação de conteúdos legítimos, a ocupação maliciosa média do *cache* dos roteadores, a taxa média de erros de *cache* dos conteúdos legítimos e o percentual de conteúdos legítimos recuperados do produtor.



(a) CBR.

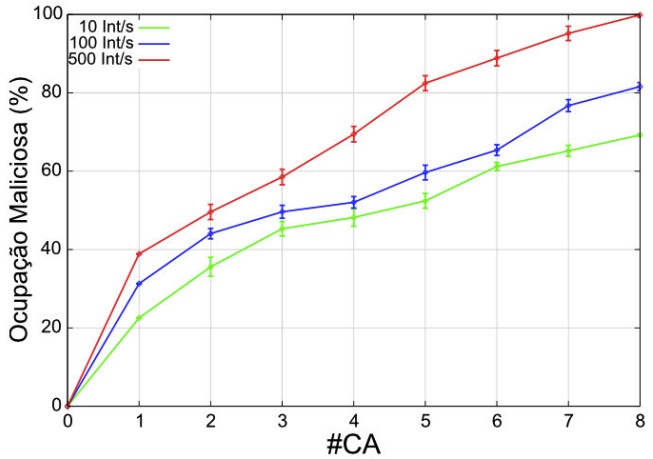


(b) Zipf.

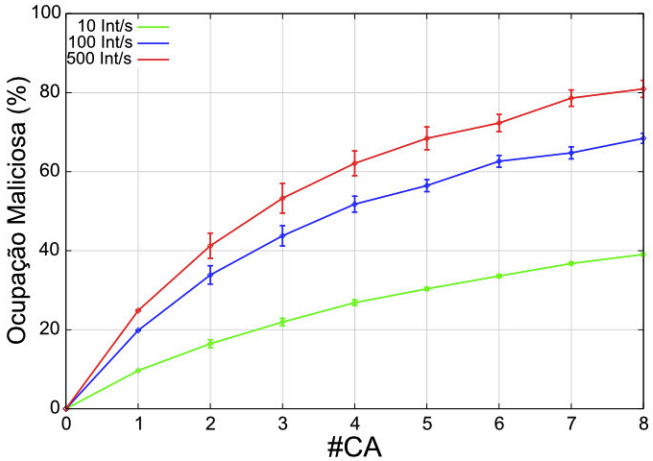
Figura 3. O tempo de recuperação de conteúdos legítimos.

A Fig. 3 mostra o comportamento do tempo médio de recuperação de conteúdos legítimos em função do número de consumidores maliciosos. Nas duas configurações, quando consumidores maliciosos solicitam conteúdos segundo o padrão CBR (Fig. 3(a)) ou quando solicitam conteúdos segundo a distribuição *Zipf* (Fig. 3(b)), o comportamento observado é o mesmo: quanto mais consumidores maliciosos, maior o tempo médio de recuperação de conteúdos. Da mesma forma, quanto maior a taxa de interesses maliciosos, maior o tempo médio de recuperação de conteúdos legítimos. Para a

configuração da Fig. 3(a), por exemplo, quando somente consumidores legítimos solicitam conteúdos, o tempo médio de recuperação de conteúdos legítimos é igual a 0,34 ms. Por outro lado, quando 4 consumidores maliciosos solicitam conteúdos esse tempo é igual a 1,92 ms e 7,21 ms, quando enviam 10 e 500 interesses/s, respectivamente. Quando há 8 consumidores maliciosos, o tempo médio de recuperação de conteúdos legítimos é igual a 2,93 ms e 7,99 ms para as taxas de 10 e 500 interesses/s, respectivamente. Isso mostra que o tempo de recuperação aumentou 23,5 vezes no pior caso para as configurações avaliadas. É importante ressaltar que como os consumidores legítimos possuem *cache* e como eles sempre consomem de acordo com a popularidade (*Zipf*), pode-se observar que sem ataque o consumo é, muitas vezes, feito do próprio *cache* do nó, o que resulta em um tempo de recuperação inferior a 1 ms.



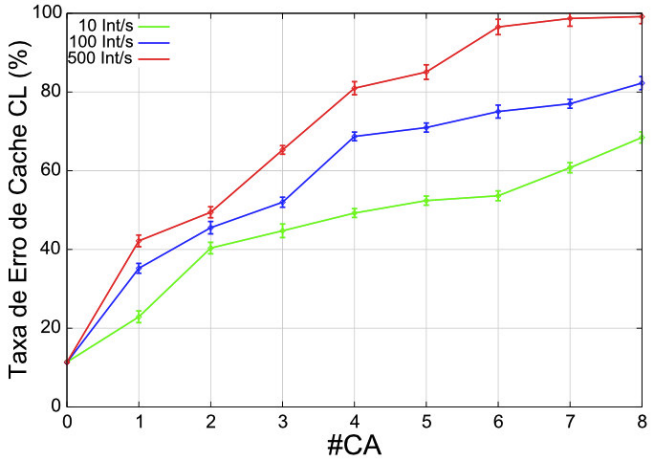
(a) CBR.



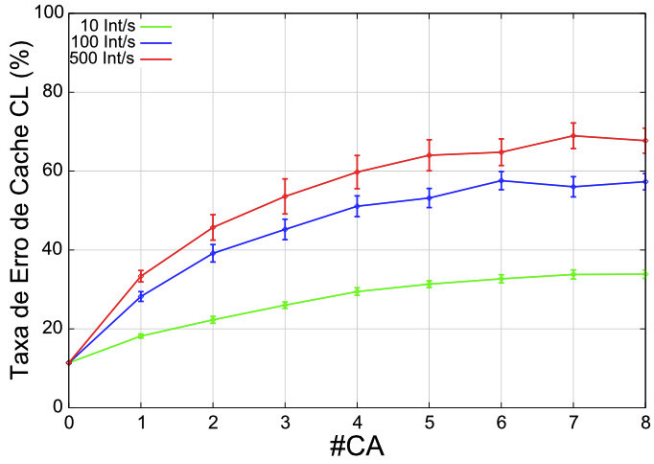
(b) Zipf.

Figura 4. O percentual de ocupação do *cache* dos roteadores por conteúdos maliciosos.

O aumento do tempo de recuperação nas duas configurações é explicado pelo aumento da ocupação maliciosa no *cache* dos nós intermediários e, consequentemente, do aumento da taxa de erros de *cache*, como mostram as Fig. 4 e 5, respectivamente. Quanto maior a ocupação maliciosa, maior a probabilidade do conteúdo solicitado não estar armazenado em *cache*. Para a configuração da Fig. 3(a) e curva para a taxa de 100 interesses/s, por exemplo, nota-se que com 8 consumidores maliciosos, o tempo de recuperação é da ordem de 6 ms. Como o atraso de cada enlace é de 1 ms, conclui-se que os conteúdos legítimos são recuperados mais frequentemente de nós que estão a mais saltos do consumidor do que os nós de borda. Isso indica que os roteadores de borda estão com uma alta ocupação de conteúdos maliciosos em seu *cache*. Nessa situação, os conteúdos maliciosos ocupam em média 80% do espaço total do *cache* dos roteadores, como mostra a Fig. 4(a). Para a configuração na qual os consumidores maliciosos solicitam conteúdos com base na sua popularidade (Fig. 3(b)), o tempo médio de recuperação também aumenta, mas esse aumento é menor do que o observado para a configuração baseada no consumo CBR (Fig. 3(a)).



(a) CBR.

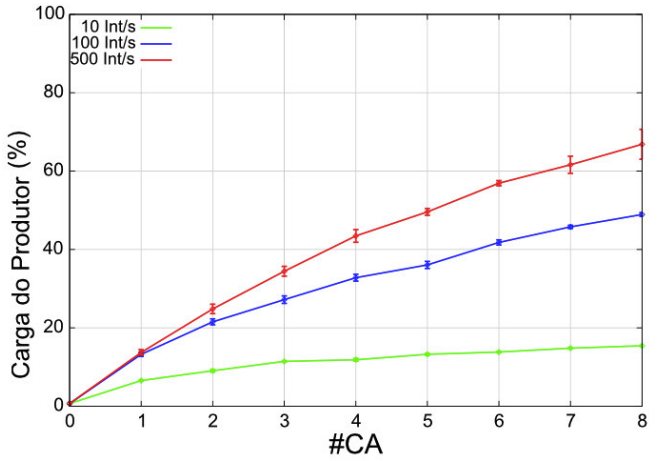


(b) Zipf.

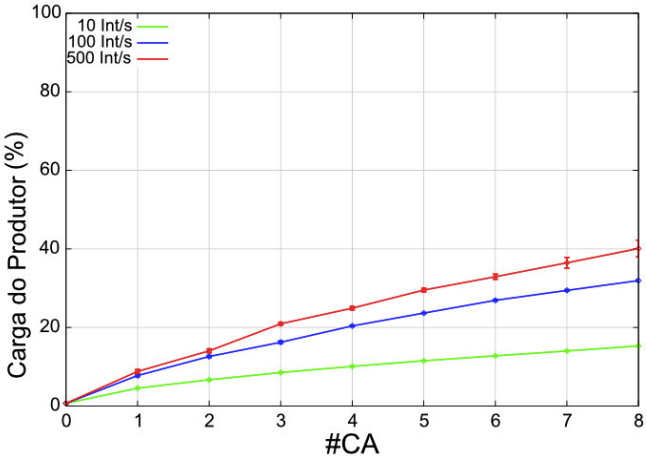
Figura 5. A taxa de erros de *cache* para os conteúdos legítimos.

Isso porque os consumidores maliciosos têm maior probabilidade de encontrarem conteúdos solicitados anteriormente em nós que estão a um ou dois saltos de distância e seus interesses não chegam ao produtor malicioso. Assim, a ocupação maliciosa na rede é menor (Fig. 4(b)), em especial nos nós próximos ao produtor. Assim, tem-se uma taxa de erros de *cache* menor (Fig. 5(b)) e um tempo de recuperação menor para consumidores legítimos.

A Fig. 6 mostra o percentual de conteúdos legítimos recuperados do produtor em função do número de consumidores maliciosos e da taxa de envio de interesses por esses nós. Esses resultados corroboram que o ataque em conluio reduz e eficiência do emprego do *cache* pela CCN. A Fig. 6(a) mostra que se não há ataque cerca de 0,5% dos conteúdos solicitados são recuperados diretamente do produtor. Nesse caso, cada conteúdo legítimo é recuperado do produtor no máximo duas vezes, até que seja armazenado pelos nós RTR1 e RTR2 (Fig. 2). Porém, basta se ter 4 consumidores maliciosos operando a taxa de 10 interesses/s para que esse valor aumente para cerca de 12%. No pior caso, os consumidores legítimos estão recuperando cerca de 67% dos conteúdos legítimos diretamente do produtor.



(a) CBR.



(b) Zipf.

Figura 6. Percentual de carga do produtor legítimo.

Um resultado interessante é que uma ocupação maliciosa de quase 100% quando há 8 consumidores maliciosos enviando 500 interesses/s (Fig. 4(a)) não resulta em 100% de conteúdos recuperados do produtor (Fig. 6(a)). Tal fato é explicado pelo uso de *cache* pelos próprios consumidores. Assim, é possível recuperar o conteúdo do próprio *cache*, sem ter que encaminhar pacotes de interesse para outros nós. No entanto, quando sob ataque, os consumidores legítimos ainda têm solicitações de conteúdo encaminhadas até o produtor, mesmo que eles façam uso de *cache* e solicitem conteúdos de acordo com a popularidade. Portanto, isso comprova que o serviço é negado em virtude da ocupação maliciosa dos *caches* dos roteadores.

Outra observação interessante extraída dos resultados é que a distribuição de consumidores maliciosos é mais efetiva do que o aumento da taxa agregada de envio de interesses maliciosos. Por exemplo, na Figura 3(a), é possível observar que o tempo médio de recuperação de conteúdos legítimos é da ordem de 5 ms quando 4 consumidores maliciosos enviam 100 interesses/s cada um (taxa agregada de 400 interesses/s) ou quando um consumidor malicioso envia sozinho 500 interesses/s. Esse fato é explicado pela ocupação maliciosa dos *caches* ser mais efetiva quando o ataque é distribuído. Para o mesmo exemplo anterior, a Fig. 4(a) mostra que a ocupação maliciosa quando há 4 atacantes enviando 100 interesses/s é da ordem de 50%. Quando há um atacante apenas enviando 500 interesses/s ela é de 40%. Esse fato se repete em outros pontos dos gráficos, considerando também o consumo baseado na popularidade. Por exemplo, na Fig. 3(b), quando 8 consumidores maliciosos enviam 10 interesses/s cada um (taxa agregada de 80 interesses/s) o tempo médio de recuperação de conteúdos legítimos é da ordem de 3 ms. Se um consumidor malicioso envia sozinho 100 interesses/s, esse tempo é menor do que 2,5 ms.

É importante ressaltar que em todos os experimentos realizados, os consumidores legítimos recuperaram todos os conteúdos solicitados. Isso pode ser explicado pelo fato da PIT não ter seu tamanho limitado e por ter sido usado um temporizador para remoção de entradas desta tabela da ordem de 4 s. Nas configurações usadas, esse tempo é muito maior do que o tempo necessário para um pacote de interesse legítimo ser encaminhado até o produtor e o pacote de dados ser encaminhado pelo caminho reverso até o consumidor. No pior caso, como mostra a Fig. 3(a), esse tempo é de aproximadamente 8 ms. Além disso, o produtor nunca remove do seu *cache* o conteúdo produzido por ele próprio. Portanto, nenhum dos nós da rede remove uma entrada da PIT nas configurações usadas antes do consumidor legítimo receber o conteúdo solicitado, mesmo que para isso, os pacotes de interesse tenham que ser encaminhados até o produtor.

VII. CONCLUSÃO

Este trabalho avaliou o ataque de negação de serviço em conluio produtor-consumidor para a arquitetura CCN. Esse ataque visa aumentar o tempo de recuperação de conteúdos aumentando a ocupação do *cache* dos nós intermediários com conteúdos maliciosos. Além disso, o ataque em conluio produtor-consumidor, não é identificado pelo mecanismo padrão de verificação de assinaturas da CCN porque os pacotes maliciosos carregam uma assinatura digital válida.

Diferentes configurações foram usadas nas simulações, variando-se o número de consumidores maliciosos, a política de consumo desses consumidores e taxa de pacotes de interesse maliciosos. Os resultados mostram que o ataque em conluio é efetivo, o que compromete o emprego do *cache* pela CCN. No pior caso, o tempo de recuperação aumentou 23,5

vezes para as configurações avaliadas. Esse aumento se deve a uma ocupação maliciosa média de 99% e, consequentemente, a uma taxa de erro de *cache* de 99%. Com isso, os consumidores legítimos recuperando 67% dos conteúdos solicitados diretamente do produtor. Mostra-se também que a distribuição de consumidores maliciosos é mais efetiva do que o aumento da taxa agregada de envio de interesses maliciosos. Os trabalhos futuros incluem a avaliação de outras políticas de substituição de *cache* e do emprego do mecanismo de *push-back* proposto por Gasti *et al.* O objetivo é verificar se tal mecanismo é eficiente como contramedida ao ataque em conluio. Caso não seja, o próximo passo é propor uma contramedida. Sobre os cenários, pretende-se empregar topologias reais e com maior escala nas simulações.

AGRADECIMENTOS

Este trabalho é apoiado por Dinter UFF/UFAC, CNPq, CAPES, FAPERJ, Proppi/UFF, TBE/ANEEL e CELESC/ANEEL.

REFERÊNCIAS

[1] A. Afanasyev, P. Mahadevan, I. Moiseenko, E. Uzun, e L. Zhang, "Interest flooding attack and countermeasures in named data networking," in IFIP Networking, May 2013, pp. 1–9.

[2] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: Evidence and implications," in IEEE Conference on Computer Communications - INFOCOM, Mar. 1999, pp. 126–134.

[3] G. M. Brito, P. B. Velloso e I. M. Moraes, "Redes orientadas a conteúdo: Um novo paradigma para a Internet." Em Minicursos do Simpósio Brasileiro de Redes de Computadores - SBRC, Abr. 2012 pp 211–264.

[4] G. M. Brito, P. B. Velloso, and I. M. Moraes, Information-Centric Networks, A New Paradigm for the Internet, 1st ed., ser. FOCUS - Networks and Telecommunications Series. Wiley-ISTE, 2013.

[5] S. Choi, K. Kim, S. Kim, and B. Roh, "Threat of DoS by interest flooding attack in content-centric networking," in Information Networking International Conference, Jan. 2013, pp. 315–319.

[6] P. Gasti, G. Tsudik, E. Uzun, and L. Zhang, "DoS and DDoS in named-data networking," in International Conference on Computer Communications and Networks - ICCCN, Aug. 2013, pp. 1–7.

[7] F. Q. Guimarães, I. C. G. Ribeiro, A. A. de Rocha e C. V. N. Albuquerque. "Nem tanto nem tão pouco: Existe um timeout Ótimo para PIT CCN na mitigação de ataques DoS," Em Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais - SBSeg, Out. 2013.

[8] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and R. Braynard, "Networking named content," in International Conference on emerging Networking EXperiments and Technologies - CoNEXT, Dec. 2009, pp. 1–12.

[9] Y. Kim, U. Kim, and I. Yeoml, "The impact of large flows in content centric networks," in IEEE International Conference on Network Protocols - ICNP, Oct. 2013, pp. 1–2.

[10] I. C. G. Ribeiro, A. A. de A. Rocha, C. V. N. Albuquerque, and F. Q. Guimarães, "On the possibility of mitigating content pollution in content-centric networking," in Conference on Local Computer Networks (LCN), Sep. 2014, pp. 498–501.

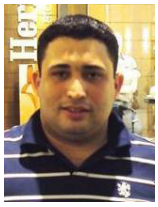
[11] I. C. G. Ribeiro, A. A. de A. Rocha, C. V. N. Albuquerque, and F. Q. Guimarães, "CCNcheck: um mecanismo de mitigação para poluição de conteúdos em redes centradas em conteúdo," Em Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais - SBSeg. Out. 2013.

[12] I. C. G. Ribeiro, F. Q. Guimarães, J. F. Kazienko, A. A. Rocha, P. B. Velloso, I. M. Moraes e C. V. N. Albuquerque, "Segurança em redes centradas em conteúdo: Vulnerabilidades, ataques e contramedidas." Em Minicurso do Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais - SBSeg. Out. 2012, pp 101-150.

[13] D. Smetters and V. Jacobson, "Securing network content," Xerox Palo Alto Research Center - PARC, Tech. Rep. TR-2009-1, 2009.

[14] M. Xie, I. Widjaja, and H. Wang, "Enhancing cache robustness for content-centric networking," in IEEE Conference on Computer Communications - INFOCOM, Mar. 2012, pp. 2426–2434.

[15] L. Zhang, D. Estrin, J. Burke, V. Jacobson, J. Thornton, D. K. Smetters, B. Zhang, G. Tsudik, K. Claffy, D. Krioukov, D. Massey, C. Papadopoulos, T. Abdelzaher, L. Wang, P. Crowley, and E. Yeh, "Named Data Networking (NDN) project," Xerox Palo Alto Research Center - PARC, Tech. Rep. NDN-0001, 2010.



André Luiz Nasserala Pires possui Mestrado pela Universidade Federal Fluminense (2010) e faz Doutorado em Computação na mesma instituição, é graduado em Bacharelado em Sistemas de Informação pela Universidade Federal do Acre (2002). Atualmente é professor da Universidade Federal do Acre. Tem experiência na área de Ciência da Computação, com ênfase em redes de computadores, atuando principalmente no seguinte tema: Segurança em Redes de Computadores.

Igor Monteiro Moraes é Professor Adjunto do Departamento de Ciência da Computação do Instituto de Computação da Universidade Federal Fluminense (UFF) desde 2010. Igor recebeu o título *cum laude* de Engenheiro Eletrônico e de Computação em 2003 e os títulos de Mestre e Doutor em Engenharia Elétrica pela Universidade Federal do Rio de Janeiro (UFRJ), respectivamente, em 2006 e 2009. Seus temas de interesse são as redes orientadas a conteúdo, as redes oportunistas, as arquiteturas para a Internet do Futuro, os sistemas par-a-par, as redes sem fio e a segurança em redes de computadores. Igor é membro da SBC.

SpamBands: a Methodology to Identify Sources of Spam Acting in Concert

E. Fazzion, P. H. B. Las-Casas, O. Fonseca, D. Guedes, W. Meira Jr, C. Hoepers, K. Steding-Jessen and M. H. P. Chaves

***Abstract*— In 2012, estimates indicated that 68.8% of all e-mail traffic was spam, what suggests this is still a relevant problem. Recently, some works have focused on the analysis of spam's traffic inside the network, analyzing the protocols used and the AS which originate the traffic. However, those works usually do not consider the relationships between the machines used to send spam. Such an analysis could reveal how different machines may be used by a single spammer to spread his messages, helping us to understand their behavior. To that end, this work proposes a methodology to cluster the machines used by spammers based on the concept of spam campaigns. The groups identified were characterized to identify different aspects of the spam dissemination process, which suggest different orchestration strategies being used.**

***Keywords*— SpamBands, Spam traffic, Spam orchestration.**

I. Introdução

Há muitos conceitos sobre o que é *spam*, porém todos têm uma base comum: um *spam* é uma mensagem de email de caráter não individual e não solicitada, que é disseminada em larga escala pela rede. As motivações daqueles que realizam essa prática, os *spammers*, são diversas, sendo as mais comuns a venda de produtos, a disseminação de *malware* e ataques de *phishing* [1]. Segundo a companhia Pingdom, cerca de 144 bilhões de mensagens de email foram enviados por dia, em 2012, sendo 68,8% delas *spam* [2]. Isso mostra que recursos para enviar e armazenar 99 bilhões de mensagens, por dia, foram desperdiçados, o que leva a sérios prejuízos financeiros, como revelado em outros trabalhos [3]. Além disto, existe um prejuízo social, onde mensagens legítimas são perdidas por má classificação de filtros de *spam* ou por excesso de tráfego ocasionado por grandes volumes de *spam*[4].

Existem diversas facetas consideradas no combate ao *spam*. Muitos estudos buscam entender o problema do ponto de vista do destinatário e auxiliar na construção de filtros eficazes que descartem mensagens indesejáveis. Outros fazem a análise do comportamento do *spammer* na rede, para entender como o *spam* é disseminado, de onde ele se origina e como ele atravessa a rede sem que os transmissores sejam facilmente identificados. O objetivo, nesse caso, é identificar comportamentos na rede que permitam bloquear as mensagens antes que elas atravessem a rede e consumam recursos para sua filtragem e possível armazenamento [5].

Em ambos os casos, fica visível que o combate ao *spam* requer o entendimento de um sistema complexo de ofuscação usado pelo *spammer* em sua atividade. Esse sistema exige uma complexa orquestração de atores e recursos, cuja existência é reconhecida mas que normalmente é invisível para o profissional que se dedica a esse combate. Para se manter oculto, o *spammer* busca disfarçar sua localização na rede, seja enviando suas mensagens a partir de múltiplas origens, como máquinas infectadas que se organizam em *botnets*, ou usando servidores especializados que podem por sua vez se aproveitar de máquinas mal-configuradas na rede para se ocultar dos destinatários. Além disso, *spammers* também utilizam programas de transmissão que geram diversas mensagens diferentes como versões de um mesmo conteúdo básico, a fim de tentar ludibriar os filtros baseados em conteúdo [4]. Nesse processo, tem importância o conceito de *campanhas de spam*, que são grupos de mensagens que possuem um mesmo objetivo, mas que foram alteradas por métodos de ofuscação para tentar ludibriar filtros [6].

Este trabalho utiliza uma abordagem que combina aspectos de campanhas com aspectos de comportamento de rede a fim de tentar lançar mais luz sobre esse elemento orquestrador subjacente ao processo de envio de *spam*. Para este fim, utilizamos tanto elementos baseados no conteúdo da mensagem, para permitir a identificação das *campanhas de spam*, quanto elementos do tráfego de rede, para identificar as máquinas originadoras de cada campanha. Com isso, propomos um método capaz de identificar os grupos de máquinas na rede que se encontram em um certo momento sob o controle de um orquestrador oculto, o *spammer*. A esses grupos denominamos *SpamBands*.

Segundo a abordagem adotada neste trabalho, um(a) *SpamBand* é um grupo de máquinas correlacionadas pelo fato de terem enviado mensagens identificadas como pertencentes a um mesmo conjunto de campanhas de *spam*. Utilizando essa estrutura em nossas avaliações, conseguimos mostrar relações importantes como o período de atividade de cada *SpamBand* e a forma como o *spammer* escolhe o protocolo utilizado. Com relação ao período de atividade, mostramos a tendência desses grupos se manterem estáveis ao longo do tempo, podendo

E. Fazzion, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, elverton@dcc.ufmg.br

P. H. B. Las-Casas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, pedro.lascasas@dcc.ufmg.br

O. Fonseca, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, osvaldo.morais@dcc.ufmg.br

D. Guedes, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, dorgival@dcc.ufmg.br

W. Meira Jr, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, meira@dcc.ufmg.br

C. Hoepers, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, cristine@cert.br

K. Steding-Jessen, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, jessen@cert.br

M. H. P. Chaves, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, mhp@cert.br

se estender por diversas campanhas e que a técnica pode identificar, como efeito adicional, possíveis partes de redes *botnets*. Quando consideramos a forma como as mensagens são enviadas, observamos que, *em geral*, *SpamBands* utilizam apenas *proxies* (HTTP ou SOCKS) ou apenas *mail relays* abertos (SMTP) em seus envios, apesar de algumas *SpamBands* apresentarem um comportamento híbrido, utilizando os dois tipos de protocolos.

A definição de *SpamBand* pode facilitar a identificação de *botnets* e outras infra-estruturas de distribuição utilizadas pelos *spammers*. Com isso, ações podem ser desenvolvidas para impedir a ação das máquinas envolvidas, removendo-as da rede ou procedendo à remoção de qualquer *malware* nelas instalado. Além disso, pela identificação dos grupos pode se tornar mais eficaz o uso de *blacklists* no bloqueio ao *spam*: se uma máquina é identificada como fazendo parte de um grupo que contém elementos já incluídos em uma lista negra, essa nova máquina também pode ser automaticamente adicionada àquela lista.

II. Metodologia de identificação de *SpamBands*

O conceito de *SpamBands* foi desenvolvido durante a análise dos dados de *spam* coletados em diversos pontos da Internet, onde percebemos que várias origens surgiam na análise do spam observado em diferentes pontos da rede. Nesta seção detalhamos a metodologia proposta para a identificação das *SpamBands* e um exemplo real de aplicação que ressalta alguns elementos importantes da proposta.

Como mencionado, a base do conceito de *SpamBands* é a premissa de que máquinas que enviam mensagens pertencentes às mesmas campanhas são controladas por um mesmo agente orquestrador, estando, assim, relacionadas a uma mesma origem. A relação entre máquinas e campanhas pode ser modelada como um grafo G, onde as máquinas são vértices e há uma aresta entre duas máquinas se elas enviaram mensagens associadas a uma mesma campanha. A Fig. 1 ilustra a construção desse grafo.

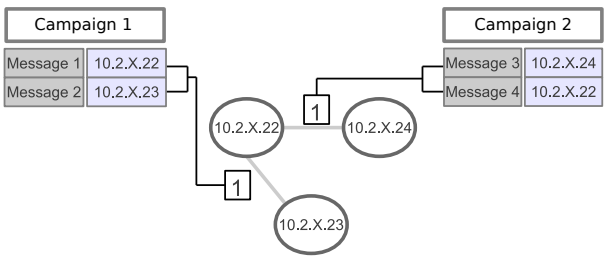


Figura 1. Modelo de grafo para relação entre campanhas e endereços IP.

A partir do grafo G, um *SpamBand* pode ser identificado como um sub-grafo denso (diversas origens que compartilham um mesmo conjunto de campanhas). A identificação desses subgrafos pode ser obtida aplicando-se algoritmos de agrupamento de grafos; entretanto, tais algoritmos tendem a ser bastante complexos e difíceis de calibrar [7]. Com base nas características particulares do problema em questão, adotamos uma estratégia mais simples e interativa, descrita a seguir.

Inicialmente, cada componente conectado de G poderia ser identificado como um *SpamBand*. Entretanto, aspectos práticos exigem que essa definição seja refinada. Por exemplo, quando um endereço IP pode se referir a diferentes máquinas atrás de um mecanismo de NAT: duas máquinas podem estar atuando sob coordenações diferentes, mas serem vistas no resto da rede como um mesmo endereço de origem. Em outros casos, um endereço é visto participando de uma campanha até certo instante do dia e a partir de então passa a participar de outra. Os nós referentes a esses endereços IP aparecem no grafo como nós de ligação entre sub-grafos mais densos, que na prática se referem a *SpamBands* diferentes.

A forma adotada para identificar esses casos e isolar os *SpamBands* envolvidos foi utilizando-se o conceito de *betweenness*, que mede o grau de centralidade de nós em um grafo. Essa métrica quantifica o número de caminhos mínimos entre todos os pares de nós no grafo que passam por um vértice em questão. A premissa é que, se alguns vértices possuem um valor de *betweenness* muito elevado em relação ao que seria esperado para um grafo fortemente conectado, existe uma chance maior desses vértices conectarem dois sub-grafos internamente mais densos. Assim, se removemos esses vértices, acentuamos a separação entre os sub-grafos densos desejados.

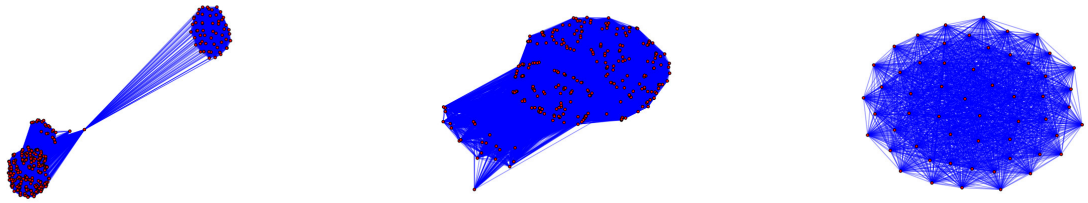
A determinação de *SpamBands* é então apresentada no algoritmo 1, que recebe três parâmetros de entrada: o grafo (**G**), o limiar de *betweenness* mínimo a ser considerado (**limiar_bt**) e o número máximo de endereços IP (vértices) que podem ser removidos para dividir um componente (**limiar_ips**). O primeiro passo determina os componentes conectados de G, que constituem uma primeira aproximação dos *SpamBands*. A seguir, identificamos sub-grafos densos em cada componente conectado removendo nós com *betweenness* acima de **limiar_bt**, respeitando o limite **limiar_ips**, que define o tamanho mínimo de um sub-grafo denso, para evitar a geração de conjuntos muito pequenos. O algoritmo retorna o conjunto S que contém todos os *SpamBands*.

Por exemplo, a Fig. 2(a) mostra um dos componentes conectados com maior número de máquinas observados em um dos dias da nossa análise. Claramente, podemos verificar que há pelo menos dois grupos praticamente disjuntos de nós, unidos por um nó que aparece entre eles. Aplicando o algoritmo 1 naquele componente conectado, isolamos os dois *SpamBands* relativos aos grupos mais densos, mostrado nas figuras 2(b) e 2(c).

Analisamos os *SpamBands* revelados através do componente conectado da Fig. 2(a). O *SpamBand* da Fig. 2(b) está distribuído em quatro ASes (17816, 17623, 4837 e 17430). Por outro lado, apesar do *SpamBand* da Fig. 2(c) estar localizado no mesmo *Country Code*, seu único AS (4134) difere de todos os outros ASes do *SpamBand* da Fig. 2(b), o que traz uma forte diferença e sugere que esses *SpamBands* identificam *botnets* distintas.

III. Coleta de dados

Os dados utilizados na análise foram coletados utilizando-se oito *honeypots* de baixa interatividade instalados em diferentes



(a) *SpamBand* original (b) *SpamBand* 1 (c) *SpamBand* 2

Figura 2. *SpamBands*: componente original e aqueles revelados a partir da aplicação do algoritmo 1.

Algorithm 1: *SpamBands* (Grafo G, Real limiar_bt, Real limiar_ips)

```
S = ∅;
C=G.ComponentesConectados(); ;
for comp em C do
    ips_a_remover = ∅ ;
    for ip em comp do
        if ip.Betweenness() >
          limiar_bt*comp.MaiorBetweenness() then
            ips_a_remover.Adiciona(ip);
        end
    end
    if ips_a_remover.Tamanho() >
      limiar_ips*comp.Numvertices() then
        S += comp;
    end
    else
        S += comp.RemoveVertices(ips_a_remover);
    end
end
retorna S;
```

partes do mundo: Austrália (AU-01), Áustria (AT-01), Brasil (BR-01 e BR-02), Equador (EC-01), Holanda (NL-01), Taiwan (TW-01) e Uruguai (UY-01). A distribuição desses *honeypots* teve por objetivo capturar dados de diferentes pontos da Internet, a fim de obter uma visão mais global do spam que viaja pela rede.

Todos os *honeypots* foram desenvolvidos de modo a simular computadores com *proxies* HTTP e SOCKS e *mail relays* SMTP abertos, que frequentemente são abusados para o envio de spam. Quando uma máquina se conecta à porta 25 de um dos *honeypots*, ela tem a impressão de estar interagindo com um servidor SMTP operando como um *open relay*, que repassa mensagens de correio para outros servidores. Já uma máquina que se conecta a um *honeypot* através dos protocolos HTTP ou SOCKS, é levada a crer que é capaz de estabelecer conexões para outros servidores SMTP na rede. Toda a interação do atacante com o suposto servidor de correio é registrada e as mensagens de spam são armazenadas localmente — nenhuma mensagem de spam é realmente entregue ao seu destino, exceto mensagens classificadas como mensagens de teste, segundo regras pré-definidas. Periodicamente, ao longo de cada dia, todo o spam armazenado nos *honeypots* é copiado para os servidores centrais do projeto.

O período de coleta usado nesta análise foi de 07/10/2013 a 25/10/2013, totalizando 19 dias consecutivos. A Tabela I oferece uma visão geral dos dados coletados.

Cerca de 225 milhões de mensagens foram coletadas, provenientes de endereços IP associados a 93 *country codes* distintos. Apesar do protocolo SOCKS ser o responsável pela maior parte do tráfego, representando 51,8% das mensagens enviadas, o número de endereços IP que utilizam o protocolo SMTP é maior, com 69,4% do total, mesmo enviando um número inferior de mensagens.

A Tabela II mostra o número de IPs, o número de mensagens e o número de ASes observados em cada *honeypots*. É importante notar que existe uma sobreposição de IPs entre *honeypots*, que indica que grupos de disseminação de spam estão atuando em mais de um coletor. Este fato será detalhado posteriormente.

IV. Resultados

Nesta seção, apresentamos os principais resultados obtidos utilizando a técnica descrita na Seção II. Inicialmente, na Subseção IV-A, fazemos um estudo de caso detalhado de forma a mostrar diferentes tipos de *SpamBands* e como estes atuam.

Na Subseção IV-B, damos uma visão geral do comportamento dos *SpamBands* encontrados nos *honeypots* e uma possível orquestração de máquinas. Ainda mais, mostramos, por meio de um exemplo, que existem *SpamBands* atuando em diferentes *honeypots*, reforçando a existência de uma orquestração e a eficácia da técnica.

A Subseção IV-C mostra um resultado imediato, obtido através do estudo dos *SpamBands*, no aprimoramento de *blacklists*. Por último, na Seção IV-D, apresentamos um estudo temporal dos *SpamBands* com resultados interessantes, realçando o quão valiosa a técnica exposta neste artigo pode ser no estudo dos *spammers* nessa dimensão.

A. Estudo de caso

Nesta seção, detalhamos os *SpamBands* descobertos nos dados do exemplo ao final da Seção II. Todos os 7 *SpamBands* podem ser vistos na Tabela III.

TABELA I
VISÃO GERAL DA BASE

	HTTP(%)	SMTP (%)	SOCKS (%)	Total
Mensagens (milhões)	76,25 (33,7)	32,82 (14,5)	116,58 (51,8)	225,66
Endereços IP	11135 (29,3)	26313 (69,4)	4372 (11,5)	37895
Prefixos de rede	40 (1,5)	2218 (87,7)	342 (13,5)	2529
Sistemas Autônomos (AS)	11 (1,6)	591 (89,0)	125 (18,8)	664
***Country Codes* (CC)**	6 (6,4)	92 (98,9)	31 (33,3)	93
Volume de Tráfego (GB)	211,18 (28,6)	160,74 (21,7)	365,97 (49,7)	737,90

TABELA II
MENSAGENS E IPS POR *honeypot*

	AT-01	AU-01	BR-01	BR-02	EC-01	NL-01	TW-01	UY-01
Mensagens (milhões)	25,27	6,51	13,89	38,64	16,57	57,52	53,92	13,33
Endereços IP	10438	19420	26762	11261	25494	11053	11145	10138
ASes	330	330	473	142	274	130	122	327

TABELA III
SpamBands DESCOBERTOS NO *honeypot BR-01* DO EXEMPLO AO FINAL DA SEÇÃO II

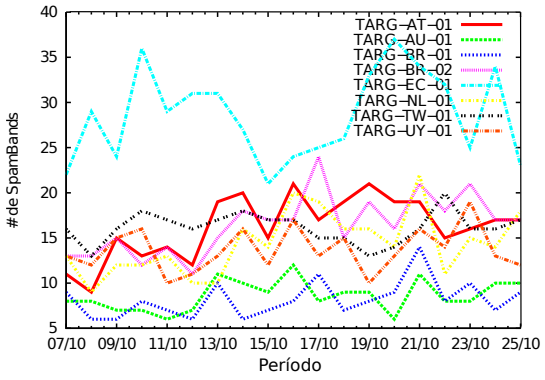
	Msg	IPs	ASes	CC (Top)	SMTP (%)	SOCKS (%)	HTTP (%)	XBL	PBL	Número de horas ativo
***SpamBand* 1**	48.244	971	1	1 (TW)	100	0	0	55	971	24
***SpamBand* 2**	475.971	910	198	52 (CN)	100	0	0	636	597	24
***SpamBand* 3**	2.711	303	4	1 (CN)	100	0	0	224	257	18
***SpamBand* 4**	1.795	56	1	1 (CN)	0	100	0	1	53	23
***SpamBand* 5**	35.389	200	96	26 (BR)	0	100	0	0	56	24
***SpamBand* 6**	28.680	5	1	1 (TW)	0	100	0	0	5	24
***SpamBand* 7**	16.679	3	1	1 (TW)	0	100	0	0	3	24

Entre os *SpamBands* da tabela, podemos identificar três grupos. O primeiro é composto pelos *SpamBands* 6 e 7. Estes *SpamBands* utilizam o procoloco SOCKS e mandam muitas mensagens em relação ao número pequeno de endereços IP que possuem, indicando o uso de servidores dedicados para a disseminação de spam.

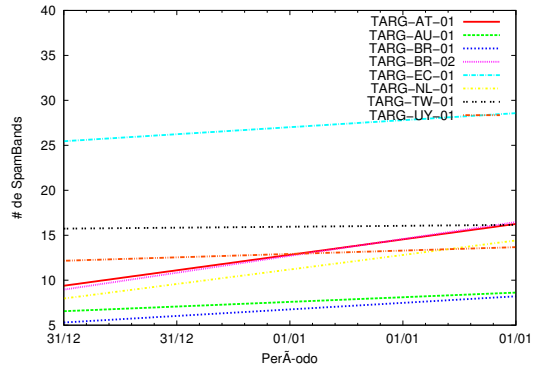
O segundo grupo é formado pelos *SpamBands* 1 e 2, que enviam um número de mensagens muito maior que os demais. Os dois possuem um alto número de endereços IP que estão distribuídos, no *SpamBand* 1, em um AS (3462) do tipo ISP (*Internet Service Provider*) e, no *SpamBand* 2, em 198 ASes. Além disso, grande parte dos endereços IP estão na XBL, o que sugere que estes *SpamBands* podem fazer parte de grandes *botnets* que estão ao redor do mundo e que mandam muito spam.

O terceiro grupo, formado apenas pelo *SpamBand* 5, possui características muito similares aos *SpamBands* 1 e 2, mas o protocolo utilizado é SOCKS. Quatro dos cinco ASes que mais enviaram mensagens neste *SpamBand* são ASes de *hosting*. Isso leva a crer na possibilidade de que algum grupo de disseminação contratou diversos servidores dedicados para enviar suas mensagens.

O *SpamBand* 3 tem características muito similares ao segundo grupo. Entretanto, ele envia um baixo número de mensagens de *spam* e está concentrado em poucos ASes que estão localizados no mesmo *Country Code*. Isso sugere que este *SpamBand* faz parte de uma pequena *botnet*. Já o *SpamBand* 4 traz indícios de que um único serviço de *hosting* está enviando campanhas de spam pela rede.



(a) Número de *SpamBands* por *honeypots* durante o período



(b) Regressão linear do comportamento observado na Fig. 3(a)

Figura 3. Distribuição dos *SpamBands* no período.

B. Visão geral dos SpamBands

A técnica aplicada ao longo de 19 dias gerou um total de 2306 *SpamBands*. A Fig. 3(a) mostra a distribuição desses *SpamBands* ao longo dos dias. Como observado na Tabela II, os dois *honeypots* que mais possuem endereços IP são o *BR-01* e o *EC-01*. Entretanto, observamos no gráfico da

Fig. 3 que esses dois *honeypots* estão em extremos diferentes no gráfico ao longo dos dias, onde o *honeypot EC-01* é o que mais possui *SpamBands* e o *honeypot BR-01*, o que possui menos. Isso sugere que o *honeypot EC-01* é atacado por mais grupos de disseminação de spam do que o *honeypot BR-01*. O restante dos *honeypots* se mantém bem relacionados, mostrando que eles são atacados por um número parecido de grupos de disseminação de spam. A Fig. 3(b) mostra uma regressão linear do número de *SpamBands* por dia por cada *honeypot*. A linha de tendência revela retas com inclinação suave reforçando que a variação observada na Fig. 3(a) tem uma regularidade e representa algum tipo de ofuscação utilizada pelo *spammer*.

Relação entre protocolos

TABELA IV
RELAÇÕES DOS PROTOCOLOS ENTRE *SpamBands*.

	***SpamBands* (%)**
Somente HTTP	12 (0,52)
Somente SMTP	925 (40,10)
Somente SOCKS	891 (38,62)
Somente HTTP e SMTP	1 (0,05)
Somente HTTP e SOCKS	383 (16,60)
Somente SMTP e SOCKS	42 (1,82)
HTTP e SMTP e SOCKS	53 (2,29)

A Tabela IV mostra a distribuição dos protocolos nos *SpamBands*. Através dessa tabela é possível notar uma relação interessante entre HTTP e SOCKS: entre todos os *SpamBands* que utilizam HTTP, 97,10% também utilizam SOCKS. Como ambos protocolos são utilizados para atacar o *honeypot* como *proxy*, isso leva a um forte indício de que o uso desses dois protocolos esteja relacionado com algum tipo de ofuscação, a qual não abordamos mais profundamente neste trabalho.

É possível verificar que muito poucos *SpamBands* utilizam SMTP em conjunto com outro protocolo. Entretanto, existem *SpamBands* que utilizam os protocolos HTTP/SOCKS e SMTP ao mesmo tempo, levando a crer na existência de um grupo de disseminação de spams que utiliza dois ou mais tipos de redes distintas para enviar suas mensagens. Uma possibilidade é o uso tanto de redes *botnets* quanto servidores dedicados para o envio de campanhas de spam. O primeiro tipo de rede tende a utilizar o protocolo SMTP pois o spammer está interessado em apenas repassar suas mensagens, visto que o mesmo já está oculto na rede. No entanto, o uso dos protocolos HTTP e SOCKS, no segundo tipo de rede, indica que o grupo de disseminação de spam utiliza servidores dedicados para o envio de suas mensagens.

Relações entre número de endereços IP, mensagens, CCs e ASes

O gráfico da Fig. 4(a) mostra que apenas 10% dos *SpamBands* com protocolos SOCKS e HTTP têm mais de 100 endereços IP, o que sugere o uso de servidores para o envio. Entretanto, cerca de 37,5% do total de *SpamBands* que possuem o protocolo SMTP têm mais de 100 endereços IP, o que não surpreende, pois redes *botnets*, em geral, são constituídas por um número maior de endereços IP no envio se comparado com HTTP e SOCKS, além de ter como característica o uso do protocolo SMTP. Entretanto, observando a Fig. 4(b) verificamos uma inversão: *SpamBands* HTTP e SOCKS tendem a enviar mais mensagens do que *SpamBands* SMTP. Isso sugere que *SpamBands* SMTP, apesar de serem formados por um grande número de endereços IP, enviam poucas mensagens.

Os gráficos das figuras 4(c) e 4(d) são bastante semelhantes. Aplicando a correlação de Pearson entre o número de *Country codes* e *ASes*, obtemos um coeficiente de 0.95, o que indica que um mesmo *SpamBand* tende a ter comportamento semelhante nos dois gráficos. Dessa forma, a análise para o gráfico 4(c) espelha-se no gráfico 4(d).

O gráfico da Fig. 4(c) sugere que os *SpamBands* que mais estão espalhados pelos países são SMTP, o que mostra uma característica típica de *botnets*. Todavia, cerca de 85% dos *SpamBands* que utilizam o protocolo SMTP contém endereços IP vindos de menos de 10 CCs, o que indica pequenas *botnets*, similar ao *SpamBand* 3 da Tabela III. Por outro lado, todos os *SpamBands* que possuem HTTP e cerca de 90% que possuem SOCKS têm endereços IP de, no máximo, 5 *Country Codes*, indicando grupos de disseminação que utilizam servidores para o envio de suas mensagens. Entretanto, alguns *SpamBands* que usam o protocolo SOCKS (cerca de 10%) chegam a ter mais de 5 *Country Codes*, indicando um comportamento similar ao *SpamBand* 5 da Tabela III.

Interseção de *SpamBands* entre *honeypots*

Como visto na Tabela II, existe uma recorrência de máquinas entre os *honeypots*. Isso leva a crer que um *SpamBand* pode também participar de outros *honeypots*. Para ilustrar essa reincidência, utilizamos os *SpamBands* do *honeypot BR-01* como referência de comparação com *SpamBands* de outros *honeypots*. A Fig. 5 apresenta essa visão.

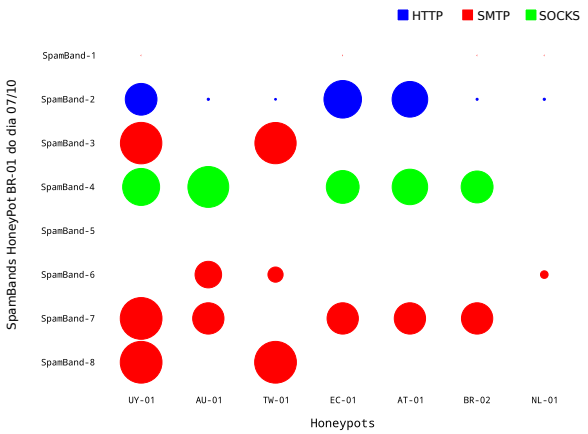
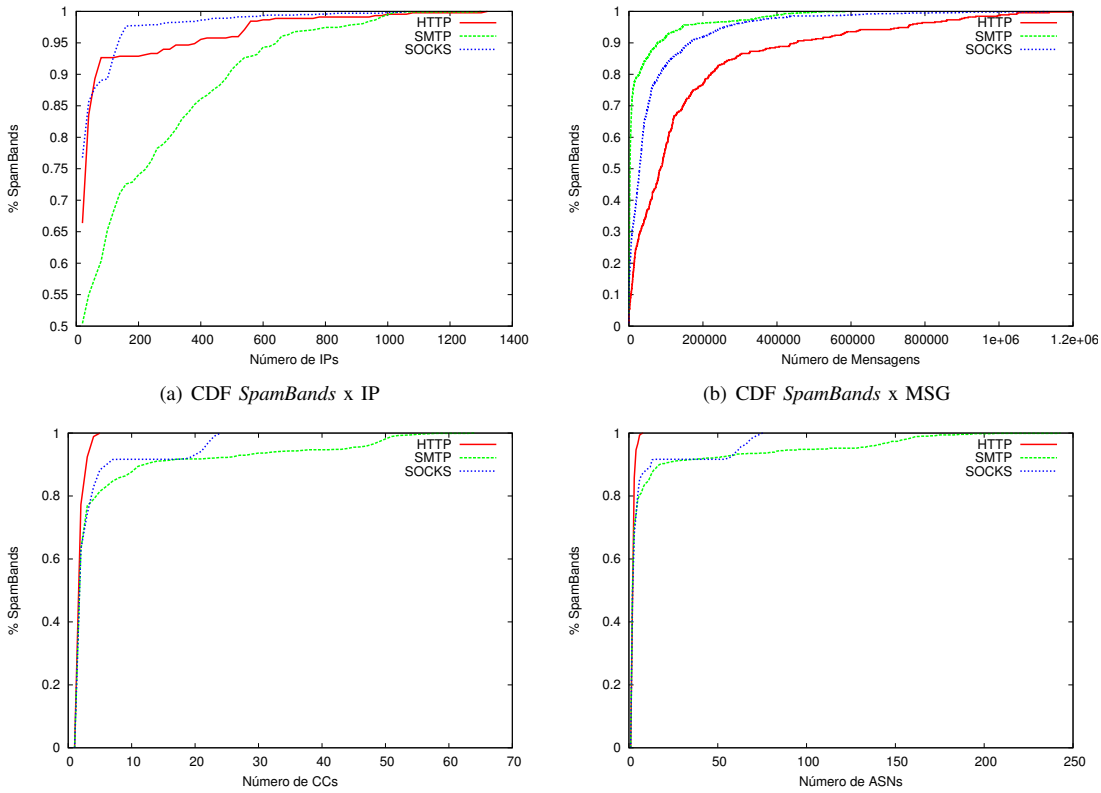


Figura 5. Interseção dos *SpamBands* do *honeypot BR-01* com *SpamBands* de outros *honeypots*.

Analisando a figura, é possível verificar que se máquinas de algum *SpamBand* aparecem em outros *honeypots*, elas



(a) CDF *SpamBands* x IP

(b) CDF *SpamBands* x MSG

(c) CDF *SpamBands* x CC

(d) CDF *SpamBands* x AS

Figura 4. *SpamBands* em relação ao número de endereços IP, mensagens, CCs e ASes.

tendem a estar no mesmo *SpamBand*. O *SpamBand* 5 do *honeypot BR-01*, não possui máquinas em outros *honeypots*, o que indica que esse grupo de máquinas têm visão apenas do *honeypot* usado como referência. Esse fato se assemelha com o *SpamBand* 1, que é o maior em número de máquinas. Entretanto, esse *SpamBand* possui uma única máquina nos *honeypots UY-01, EC-01, BR-02 e NL-01*. Isto leva a crer que o *SpamBand* possui conhecimento dos *honeypots* citados mas, por algum motivo desconhecido, está utilizando apenas o *honeypot BR-01*.

Averiguando os *SpamBands* 2, 3, 4, 6, 7 e 8, vê-se que estes grupos conseguem alcançar outros *honeypots*. Além disso, eles não têm recorrência nos mesmos *honeypots*, o que reforça a hipótese de estes grupos serem independentes. Outro fato importante é que esses *SpamBands* também não utilizam todas as máquinas em todos os *honeypots*. Essa evidência leva a crer que existe algum tipo de distribuição de atividades desses *SpamBands* na rede.

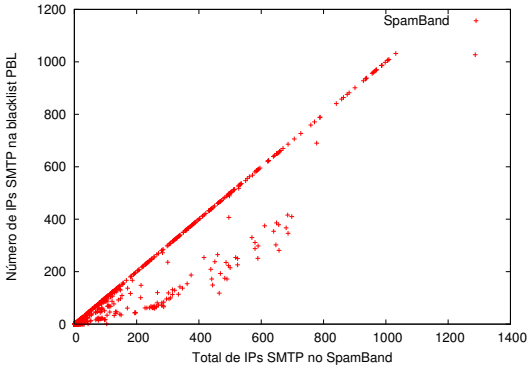
C. *Relação entre SpamBands e* blacklists

A Tabela V fornece a relação entre o número de IPs dos *SpamBands* que estão em *blacklists* e o número de IPs que o *SpamBand* possui em cada protocolo. Observamos uma correlação muito forte no número de IPs na *PBL (Policy Block List)* e o número de IPs do protocolo SMTP nos *SpamBands*. Conforme observado na seção IV-B, 90% dos *SpamBands* possuem somente o protocolo SMTP, o que leva a um forte indício desses *SpamBands* serem partes de *botnets*. Pela Fig. 6(a) observamos que a *PBL* captura grande parte desses IPs que estão nos *SpamBands* e que possivelmente fazem parte de *botnets*. Por outro lado, os protocolos HTTP e SOCKS possuem uma correlação fraca, o que era esperado visto que *SpamBands* desse tipo tendem a enviar suas mensagens de serviços de *hosting*.

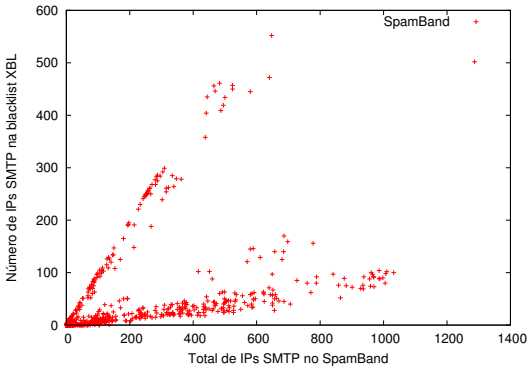
TABELA V
COEFICIENTE DE DETERMINAÇÃO ENTRE PROTOCOLOS DOS *SpamBands* E *Blacklists*.

	PBL	**XBL**
HTTP	0.38	0.11
SMTP	0.86	0.55
SOCKS	0.35	0.08

Em relação a *XBL (Spamhaus Exploits List)*, observamos uma correlação moderada, o que não é esperado visto que diversas máquinas que estão em *botnets* estão infectadas por algum tipo de *malware*. Analisando o gráfico da Fig. 6(b) observamos o porquê da relação ter sido moderada: existem dois eixos de tendência. O primeiro é uma relação linear entre o número de IPs na XBL do *SpamBand* e o número de IPs SMTP, que seria esperado: os endereços IP de todos os participantes de uma *botnet* tendem a acabar sendo identificados por *blacklists*. Entretando, o segundo eixo possui uma relação 1:10, o que sugere algum comportamento especial por parte daqueles *SpamBands*. Eles não só conseguem mascarar bem as atividades de suas máquinas na rede do ponto de vista



(a) Distribuição de IPs SMTP por IPs SMTP na PBL nos *SpamBands*



(b) Distribuição de IPs SMTP por IPs SMTP na XBL nos *SpamBands*

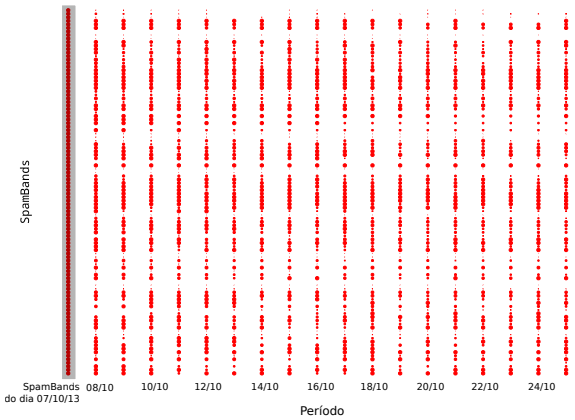
Figura 6. Relação entre o protocolo SMTP e as *blacklists PBL* e *XBL* nos *SpamBands*.

das *blacklists*, mas mantêm uma taxa comum de identificação em tais listas, o que sugere um comportamento planejado. Determinar a razão para tal comportamento, entretanto, exige novas análises e coletas, sendo considerada como trabalho futuro.

D. Relação Temporal

Nesta seção, procuramos entender o comportamento dos *SpamBands* do primeiro dia do período avaliado (07/10/2013) em outros dias. O método utilizado para verificar a continuidade do *SpamBand* é recuperar o *SpamBand* do mesmo *honeypot*, no dia seguinte, que mais possui IPs em comum com o *SpamBand* do dia de referência. Observe que esta técnica permite que novos IPs apareçam no *SpamBand* ao longo dos dias e que iremos discutir mais adiante. A Fig. 7 mostra que existe uma tendência dos *SpamBands* permanecerem ao longo do tempo. O tamanho dos pontos do gráfico indicam quantos IPs permaneceram em relação ao dia de referência.

Pode-se notar pelo gráfico da Fig. 7 que os *SpamBands* mudam constantemente seu tamanho ao longo dos dias. Para uma visão geral do comportamento temporal dos *SpamBands* por protocolo, procuramos observar dois quesitos: a variação do tamanho e a estabilidade dos IPs que participam do *SpamBand* no período avaliado. O primeiro quesito é calculado através do coeficiente de variação e o segundo, dividindo a média



(a) Comportamento dos *SpamBands* do dia 07/10/2013 ao longo dos dias

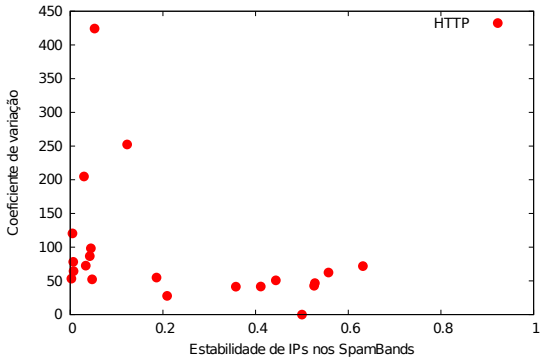
Figura 7. Comportamento geral dos *SpamBands* ao longo dos dias.

de IPs pelo número total de IPs distintos que apareceram no período. A Fig. 8 mostra uma relação global entre protocolos, estabilidade e variação dos *SpamBands*. Pelas figuras 8(a) e 8(c), observamos que os *SpamBands* HTTP e SOCKS tendem a manter seu tamanho e possuir maior estabilidade. Isso reforça, mais uma vez, que os *SpamBands* baseados nesses protocolos utilizam serviços de *hosting* para enviar suas mensagens. Por outro lado, vemos um comportamento diferenciado do protocolo SMTP na Fig. 8(b), que indica que esses *SpamBands* são bem menos estáveis que os dos protocolos HTTP e SOCKS e possuem maiores variações no tamanho, o que indica uma dinamicidade nesses *SpamBands*.

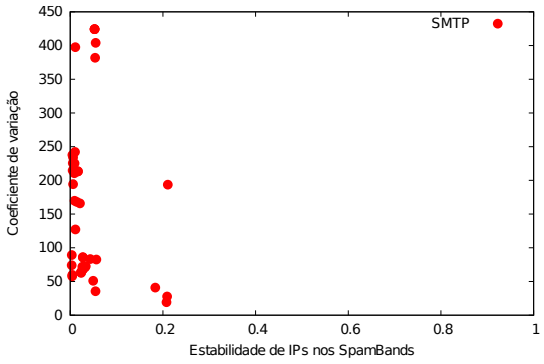
Exemplo de relação temporal entre campanhas e IPs nos *SpamBands*

Para ilustrar o comportamento das campanhas em relação a mudança dos IPs nos *SpamBands*, realizamos o mesmo método aplicado anteriormente para identificar *SpamBands* semelhantes, entre dias, através de IPs. Entretanto, utilizamos campanhas ao invés de IPs. Nos dois gráficos da Fig. 9, mostramos uma relação entre os IPs e campanhas nos *SpamBands*. A Fig. 9 mostra o experimento realizado para o *honeypot BR-01*. Como podemos observar na figura, existe uma relação entre o comportamento dos grupos de IPs e campanhas. O *SpamBand* 6 do *honeypot BR-01* no dia 07/10/2013 desaparece completamente no dia 12/10/2013, como mostra a Fig. 9(a). Todos os 47 IPs deste *SpamBand* são SMTP e fazem parte de apenas um *country code* (CN) e um AS (4134) do tipo DSL (4134), o que indica um *SpamBand* que faz parte de uma pequena *botnet*. Como os IPs do tipo DSL tendem a ser dinâmicos, é possível essas máquinas finais tenham mudado seu endereço IP durante os dias. Entretanto, a possibilidade delas terem saído da *botnet* é maior visto que as campanhas que elas suportavam também desapareceram.

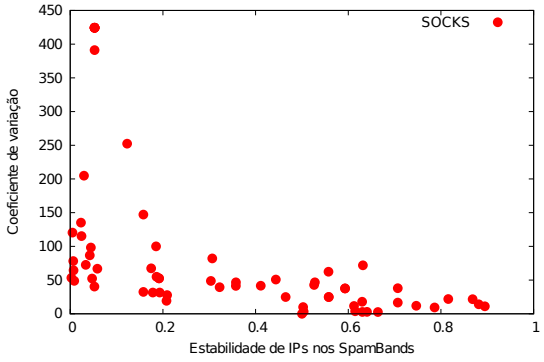
O *SpamBand* 5 retrata um grupo puramente SOCKS, onde IPs estão distribuídos em 23 *country codes* e 72 ASes. Esse grupo é semelhante ao *SpamBand* 5 do estudo de caso da Seção IV-A, que possivelmente contratou diversos serviços



(a) Variação x Estabilidade HTTP



(b) Variação x Estabilidade SMTP

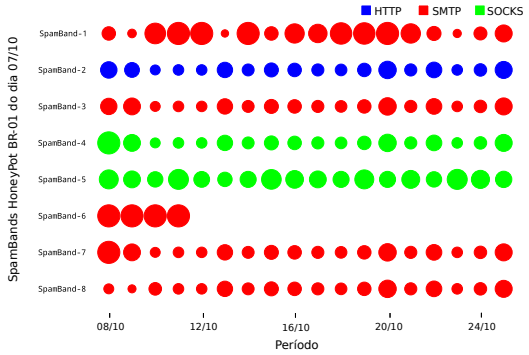


(c) Variação x Estabilidade SOCKS

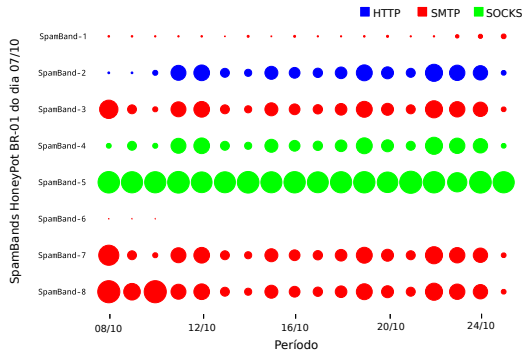
Figura 8. Estabilidade e desvio padrão relativo de IPs nos *SpamBands* , ao longo do dia, por protocolo.

de *hosting* para o envio das campanhas. Observe que o grupo de máquinas permanece similar ao longo do tempo e existe uma periodicidade no grupo de campanhas, mostrando que este grupo de máquinas enviam um mesmo grupo de campanhas alternadas ao longo dos dias avaliados.

Um outro comportamento interessante que podemos observar é sobre os *SpamBands* 2,3,4,7 e 8. No dia 11/10, estes *SpamBands* se unem, formando um único *SpamBand* e, por isso, o comportamento temporal desses 5 *SpamBands* nos dois gráficos da Fig. 9 são iguais. Este fato indica que estes *SpamBands* podem estar oferecendo serviços que são contratados por grupos de disseminação e que, em algum momento, um grupo contratou estes serviços para enviarem as mesmas campanhas.



(a) Comportamento dos IPs dos *SpamBands* do *honeypot BR-01* do dia 07/10 ao longo do período analisado.



(b) Comportamento das campanhas nos *SpamBands* do *honeypot BR-01* do dia 07/10 ao longo do período analisado.

Figura 9. Exemplo de relação entre campanha e IPs nos *SpamBands*.

V. Trabalhos Relacionados

Alguns autores já focaram no comportamento dos *spammers* de formas que tiveram impacto sobre este trabalho.

Guerra et al. apresentam uma técnica que utiliza uma estrutura de mineração de dados denominada FPTree para agrupar mensagens de spam [8]. As mensagens assim agrupadas definem o conceito de campanha de spam, como usadas neste trabalho: uma campanha é um conjunto de mensagens que foram enviadas com um mesmo objetivo mas que foram diferenciadas por algum tipo de ofuscação, com a finalidade de não serem captadas por filtros spam.

Ramachandran et al. mostram que o *spammer* alterna as máquinas usadas para envio, de modo ocultar sua origem e contornar diversos filtros de spam na rede [9]. Esse trabalho sugere que as mensagens de uma mesma campanha podem ser enviadas por diferentes máquinas, o que motiva nosso trabalho para encontrar uma forma de agrupar essas máquinas.

Moreira Moura et al. introduz o conceito de *Bad Neighborhoods*, que são vizinhanças de rede com alta probabilidade de um IP enviar spam [10]. *Fonseca et al.* estende esse conceito e estabelece uma relação direta de vizinhança com Sistemas Autônomos (AS), por esses terem fronteiras bem definidas. Nosso trabalho apresenta uma visão complementar a esses conceitos: ao invés de focarmos diretamente nos locais de

origem do *spam*, estamos procurando entender como diferentes origens (máquinas em diferentes pontos da rede) se relacionam para atender às necessidades do *spammer*, o orquestrador por trás de todo o processo.

Por fim, Zhuang *et al.* associa características de spam a *botnets*, que são um meio de envio de mensagens de spam [11]. Contudo, existe a possibilidade de que vários *spammers* utilizem a mesma *botnet* ou combinações delas, visto que essas redes muitas vezes são alugadas para terceiros [12]. Dessa forma, sem uma identificação das campanhas de spam sendo enviadas, grupos de máquinas utilizadas por diferentes spammers podem ser vistas como uma só entidade, não levando a um bom agrupamento de máquinas.

VI. Conclusão e Trabalhos Futuros

Neste trabalho, buscamos entender melhor o comportamento dos *spammers* correlacionando as máquinas utilizadas para envio através das campanhas de spam enviadas. Para realizar essa análise, propusemos o conceito de *SpamBands*, grupos de máquinas que participam das mesmas campanhas e sugerem a existência de um único orquestrador por trás de seu comportamento e desenvolvemos uma metodologia baseada em grafos para identificar esse grupos. Inicialmente, conectamos todas as máquinas que enviam as mesmas campanhas. Os grupos revelados por esta metodologia passam por um processo de refinamento, de forma a separar subgrafos densos, que revelam os *SpamBands*.

Descobrimos que a grande maioria dos *SpamBands* tendem a utilizar apenas o protocolo SMTP ou os protocolos HTTP/SOCKS, o que faz uma distinção entre grupos que utilizam servidores dedicados e redes *botnets* para o envio de spam. Além disso, mostramos que esse conceito permite revelar grupos que não são inteiramente detectados pela *blacklist XBL*, podendo ajudar a inferir outras máquinas que deveriam pertencer à *blacklist*. Ademais, encontramos *SpamBands* que utilizam os dois tipos de protocolos, levando a crer na existência de grupos de disseminação de spam que utilizam tanto servidores dedicados quanto redes *botnets* para enviar mensagens.

Por fim, realizamos ainda um estudo sobre os *SpamBands* ao longo dos dias avaliados, revelando que eles se repetem ao longo do tempo. Nesta avaliação, descobrimos que *SpamBands* que utilizam os protocolos HTTP/SOCKS tendem a ser mais estáveis em relação ao número de IPs, o que não acontece com *SpamBands* que utilizam o protocolo SMTP. Como trabalho futuro, pretendemos analisar mais profundamente o comportamento dos *SpamBands* ao longo dos dias, buscando a existência de uma interação entre eles, de forma a entender o comportamento temporal.

Agradecimentos

Este trabalho foi parcialmente financiado por NIC.BR, Fapemig, CAPES, CNPq e InWeb.

Referências

[1] D. Crocker, "Challenges in anti-spam efforts," *The Internet Protocol Journal*, vol. 8, no. 4, 2006. [Online]. Available: "http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_8-4/anti-spam_efforts.html"

[2] Royal Pingdom, "The internet 2012 in numbers," Artigo na Web, Visitado em 2014. [Online]. Available: http://royal.pingdom.com/2013/01/16/internet-2012-in-numbers/

[3] J. C. Sipior, B. T. Ward, and P. G. Bonner, "Should spam be on the menu?" *Commun. ACM*, vol. 47, no. 6, pp. 59–63, Jun. 2004. [Online]. Available: http://doi.acm.org/10.1145/990680.990681

[4] G. V. Cormack, "Email spam filtering: A systematic review," *Found. Trends Inf. Retr.*, vol. 1, no. 4, pp. 335–455, Apr. 2008. [Online]. Available: http://dx.doi.org/10.1561/1500000006

[5] P. H. B. Las-Casas, D. Guedes, W. M. Jr., C. Hoepers, K. Steding-Jessen, M. H. P. Chaves, O. Fonseca, E. Fazzion, and R. E. A. Moreira, "Análise do tráfego de spam coletado ao redor do mundo," in *Anais do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*. SBC, 2013.

[6] P. H. C. Guerra, D. Guedes, W. M. Jr., C. Hoepers, and K. Steding-Jessen, "Caracterização de estratégias de disseminação de spams," in *Anais do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*. SBC, 2008.

[7] H. Almeida, D. Guedes, W. Meira, and M. J. Zaki, "Is there a best quality metric for graph clusters?" in *Proceedings of the 2011 European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part I*, Athens, Greece, 2011, pp. 44–59.

[8] P. H. C. Guerra, D. E. V. Pires, D. Guedes, J. Wagner Meira, C. Hoepers, and K. Steding-Jessen, "A campaign-based characterization of spamming strategies," in *Proceedings of the 5th Conference on e-mail and anti-spam (CEAS)*, Mountain View, CA, 2008.

[9] A. Ramachandran and N. Feamster, "Understanding the network-level behavior of spammers," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 291–302, Aug. 2006. [Online]. Available: http://doi.acm.org/10.1145/1151659.1159947

[10] G. C. Moreira Moura, R. Sadre, and A. Pras, "Internet bad neighborhoods: the spam case," in *7th International Conference on Network and Services Management (CNSM 2011), Paris, France*, O. Festor and E. Lupu, Eds. USA: IEEE Communications Society, October 2011, pp. 1–8.

[11] L. Zhuang, J. Dunagan, D. R. Simon, H. J. Wang, I. Osipkov, and J. D. Tygar, "Characterizing botnets from email spam records." in *LEET*, F. Monrose, Ed. USENIX Association, 2008.

[12] D. Raywood, "The botnet market and what you get for your money," *SC Magazine UK*, 2010.

Elverton Fazzion é aluno de mestrado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da computação pela Universidade Federal de Minas Gerais (2014). Seus interesses são na área de redes de computadores, mineração de dados e algoritmos.

Pedro Las-Casas é aluno de doutorado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da Computação pela PUC Minas (2010) e mestrado pela UFMG (2013). Seus interesses são em redes de computadores, processamento massivo de dados e mineração de dados.

Osvaldo Fonseca é aluno de mestrado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da computação pela Universidade Federal de Minas Gerais (2013). Seus interesses são na área de redes de computadores e mineração de dados.

Dorgival Guedes possui graduação e mestrado em Ciências da Computação pela Universidade Federal de Minas Gerais e doutorado pela University of Arizona, Tucson (1999). É professor associado do departamento de Ciência da Computação da UFMG, Brasil. Atuou como professor visitante no International Computer Science Institute (ICSI) e na University of California, Berkeley, em 2011. Suas áreas de pesquisa incluem Redes de Computadores, Sistemas Distribuídos e Sistemas Operacionais, especialmente quando elas são relacionadas com escalabilidade de aplicações distribuídas, incluindo áreas como Computação em Núvem, Big-Data, e Redes Definidas por Software.

Wagner Meira Jr. Bacharel e mestre em Ciência da Computação pela Universidade Federal de Minas Gerais (1990 e 1993) e doutor em Ciência da Computação pela University of Rochester (1997). Já publicou mais de 200 artigos em veículos de comunicação de grande importância e é co-autor do livro *Data Mining and Analysis - Fundamental Concepts and Algorithms* publicado pela Cambridge University Press em 2014. Atualmente é professor titular da Universidade Federal de Minas Gerais. Suas áreas de interesse são sistemas paralelos e distribuídos e mineração de dados, assim como a sua aplicação em redes sociais, comércio eletrônico, recuperação de informação e bioinformática, entre outros.

Cristine Hoepers possui graduação em Ciências da Computação pela Universidade Federal de Santa Catarina (1996) e doutorado em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2008). É Gerente Geral do CERT.br/NIC.br, onde está desde 1999, e atua no desenvolvimento de boas práticas de segurança, na conscientização de usuários e na coordenação do honeyTARG Project, Capítulo do Honeynet Project Mundial. Tem experiência nas áreas de Redes de Computadores, Segurança, Gestão de Incidentes e uso de Honeypots para Análise de Tendências e Detecção de Ataques.

Klaus Steding-Jessen possui graduação em Engenharia da Computação pela UNICAMP (1996) e doutorado em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2008). É Gerente Técnico do CERT.br/NIC.br, onde está desde 1999, atuando nas áreas de infraestrutura, treinamento e análise de tendências, esta última como parte do honeyTARG Project, capítulo do Honeynet Project Mundial. Tem experiência nas áreas de Redes de Computadores, Segurança, Tratamento de Incidentes e uso de Honeypots para Análise de Tendências e Detecção de Ataques.

Marcelo Chaves é bacharel em Ciência da Computação pela Universidade Federal de Ouro Preto (1999) e mestre em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2002). É analista de projetos de segurança senior do CERT.br/NIC.br, onde está desde 2002. Atua na área de infraestrutura, pesquisa e desenvolvimento do CERT.br, além de ser membro e desenvolvedor do honeyTARG Project, capítulo do Honeynet Project Mundial. Suas especialidades incluem tecnologias envolvendo honeypots, análise de incidentes (incluindo fraudes via Internet), metodologias antispam, arquiteturas de segurança, monitoramento de redes e análise de logs.

Website of the National Network of Information Security and Cryptography
http://www.renasic.org .br

ISSN 2358-8963