



INTEGRATED PROACTIVE FORENSICS MODEL IN NETWORK INFORMATION SECURITY

Gojko Grubor, Ivan Barać

Singidunum University, Serbia

Abstract:

In many cases, web application security cannot provide the required level of security. Proactive collection of network data from all of the network layers in real time and their forensic analysis can help to uncover information about the internal or external attacks and to prevent potential damages. The best way is to combine application and system monitoring and perform centralized traffic monitoring to correlate events. The data collected in such manner can be used to detect traffic anomalies and improve network intrusion detection. Tracing traffic at multiple levels could potentially provide more information about the intrusion features. Analysis of these centralized log data has become an important research area in proactive network security. Any attacks should be detected as soon as possible by monitoring system, to take appropriate corrective measures in timely manner. In this paper deferent types of network events and data sources are described and their integration into centralized log management infrastructure in proactive forensic architecture is researched. The authors of this paper proposed an integrated proactive digital forensic (IPDF) model for internal and external attacks and its contribution to overall network security in context of high – volume network traffic, big data and virtualized cloud computing environment.

Key words:

logging,
monitoring system,
proactive network forensics,
integrated security event
management,
intrusion-detection.

INTRODUCTION

Proactive network forensics is becoming unavoidable in network information security. Two major changes have caused its development. *The first*, costs of high-volume data storage on a network device are affordable [*Related Work*

In the field of DF examination of big (high volume) data [28], there are a few authors' works. The authors in [29] summarized some relevant works who contributed to the subjects such as big data digital forensic investigation, proactive digital forensic and forensic in virtualized environment. Roussev et al. (2004) proposed distributed digital forensics, tool for big data analysis, using Foren-

sics Tool Kit (FTK), and increasing in performance from hours to a few minutes. Golden et al. (2005) presented an open source, high performance file carver, *Scalpel* that increased carving speed for factor of 4. Roussev et al. (2009) emphasized DF investigations on a cloud computing platform. Carrier et al. in [5] adapted the iterative *z* algorithm to speed up the process of imaging, searching and analyzing in DF, detecting outliers via MAC (modified, accessed, and created) times in set of spatial features in order to automate DF analysis and detect infected files. Phillip G. Bradford and Ning H. in their work [25] presented positional PDF architecture to discover insider attacks using monitoring system for following user's behavior in local network.



The authors of this paper proposed an integrated PDF model (PDFI) for internal and external attacks and its contribution to overall network security in context of high – volume network traffic, big data and virtualized environment, such as cloud computing system.

OVERVIEW ON REACTIVE AND PROACTIVE DIGITAL FORENSICS

Digital forensics (DF) can be defined as the set of methods, tools and techniques used to collect preserve and analyses digital data collected from any type of digital media, involved in an incident with the purpose of extracting valid evidence for the court of law. As a response to an incident or computer crime, DF investigation is essentially a reactive process and it is called reactive DF investigation (RDFI) (Fig.1.)

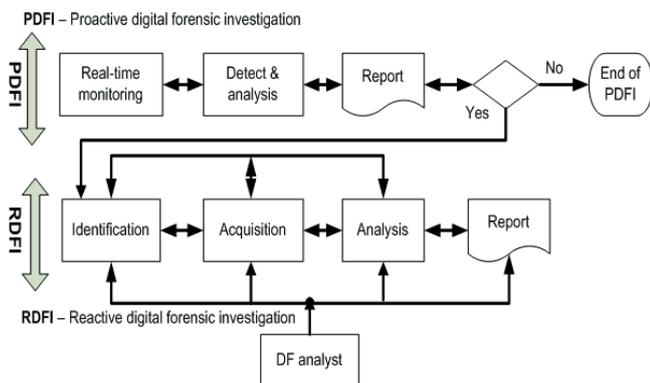


Fig.1. Proactive and Reactive Digital Forensic Framework

Although RDFI or post-mortem forensics is very effective, it is limited, especially in anti-forensic incidents, volatile data and event reconstruction. To overcome this limitation, the Proactive Digital Forensic Investigation (PDFI) [25] is required. The PDFI processes should have the capacity to proactively collect and preserve data, detect suspicious events, analyze extracted evidence and report an incident as it occurs [15, 22, 25]. Therefore, by being proactive, DF is prepared for incidents [16]. Although there is little work done on PDFI, it has many advantages over RDFI such as reducing the effect of anti-forensic methods, providing more accurate and reliable digital evidence in real time, and saving time and money in carrying out DF investigation. The five fundamental principles of computer forensics that could be applied for network RDFI and PDFI are presented in Table 1[9].

Table 1 Fundamental principles of computer forensics

Principle	Description
1	Consider the entire system (e.g. the user spaces, file system)
2	There is no trust either in user or in policy
3	Analyze the cause and effects of events.
4	Understanding context and interpreting meaning of an event.
5	All actions and results must be done by forensic analyst.

Functional model of network proactive digital forensic investigation

There are two type of network traffic monitoring systems [26, 12]: (1) *Catch-it-as-you-can* and (2) *Stop-look-and-listen system*. In the first approach all packets passes through one determined traffic point where it is recorded and saved for later on analysis. This approach is not appropriate for PDA monitoring system, as it requires large amount of memory and batch type of analysis. In second approach each packet is analyzed in real time and only certain type of predefined information are stored for later on analysis. The type of stored information could be suspicious and malicious data. This approach is appropriate for PDF monitoring system. However it requires much faster processor to respond incoming data traffic.

In case of PDFI the entire history of the system must be preserved and sometime the analysis and report the results should be perform in real-time. Proactive forensics depends on strong network monitoring system that makes main part of PDFI infrastructure. It must be designed to perform monitoring of internal user activities and to collect potential forensic evidence of the insider and outsider threats [6]. There are few weaknesses in current NIDS based monitoring system such as: false positive and false negative, detection of non-critical events and low level and slow deviations [29] in user behavior. Most effective proactive forensic system can be designed with NIDS and/or IPS devices as triggers for appropriate forensic tool. These triggers should be forensically relevant data collected and generated by NIDPS in monitoring system. Depending on NIDPS model security vector usually has three states; *suspicious, normal and anomalous* [1, 18]. Collected data are aggregated into one of the three states. To select the relevant security features general access procedure can be used [29]. The results of PDF and monitoring system are presented in Fig.2.

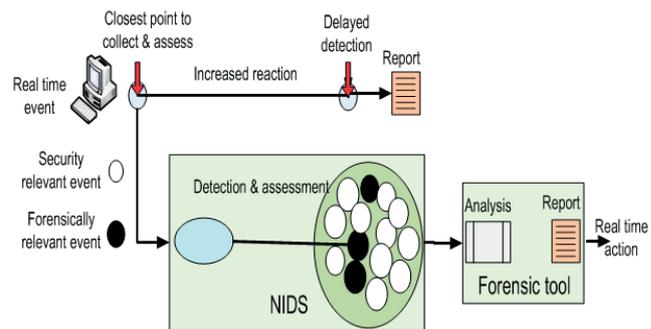


Fig 2. Functional model of PDFI real-time monitoring system

An effective monitoring system is needed for effective and proactive incident management. In PDF system a permanent monitoring provides a real-time verification of the network security system. As the security layered infrastructure provides best protection of network information asset, most appropriate is a layered architecture of IDS in implementing PDF monitoring system. The authors in [1, 29] proposed typical NIDS design containing the three layers: top, middle and bottom. The model is proposed for network proactive protection from internal attacks.



Even though internal attacks are still prevailing, malicious programs and direct attacks over Internet can't be underestimated. The authors of this paper, using the idea from [29], proposed the three layered IDS architecture for external and internal threats (Fig.3). Traffic from the Internet is filtered in border firewall. Network IDS (NIDS) in this location registers attacks from Internet that braking through border firewalls. The NIDS registers attacks on the web, FTP, exchange and all other servers located in DMZ (perimeter network), too. It indicates problems with security policy, firewall configuration or its malfunctioning. The top layer of this model quickly registers malicious attacks from black-listed web sites and unauthorized internal user processes by malicious sites name and users' processes names. The middle layer utilizes a role based access control (RBAC) rule generated by the GA module to capture the internal unauthorized processes associated with particular user role. It is supposed that some malicious codes can pass through this layer. The bottom layer performs statistical analysis over the remaining users' processes for any "low-and-slow" deviations from the referenced process patterns associated with user and group of users' roles [29]. This layer can detect potential malware using signature and heuristic methods.

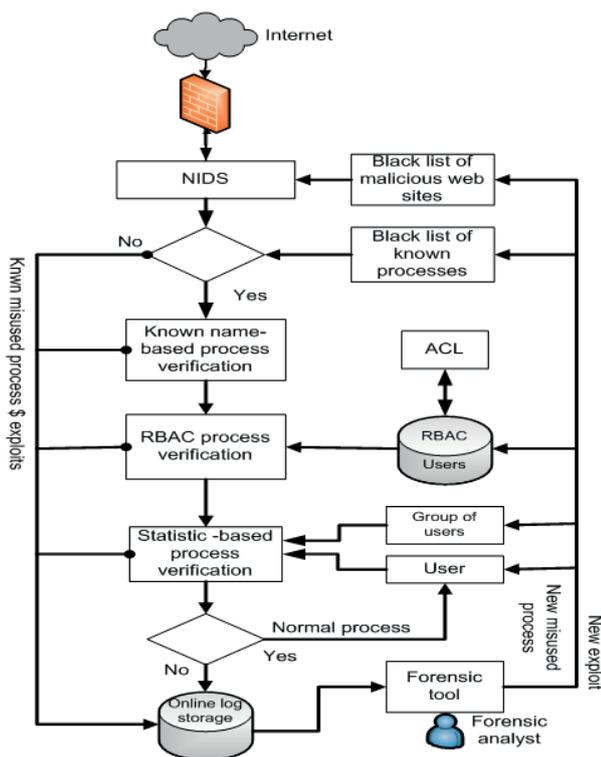


Fig.3. Integrated PDFI model for internal and external attacks investigation

Suspicious users' or malicious processes from the three layers are logged securely at a separate log storage providing input in offline forensic tool. Forensic analyst, as a member of security team [30, 32] follow monitoring system alarms and security log storage on daily basis. Any misused or malicious process or code, discovered in forensic analysis process, automatically is sent to update black list of known processes and known malicious web-sites.

Proactively collecting and categorizing network data

In Fig.3 information of network security mechanisms are used in security practice for data collection from network devices and monitoring system. Obviously firewalls, AV programs and security log files are used most often. Many people use NIDPS systems, too. However, the most interesting issues on this chart is position of the server and client honey pots that they have used before but stopped due to additional workload involved in setting them up [User profiling according to his/her most often performed actions.

- ◆ Attacker profiling according to activities performed to unauthorized access.
- ◆ Signature analysis or „typing signature” analysis is a complex stylometric problem and can't be fully reliable as unique factor for attacker's identification.
- ◆ Attack signature that uses e.g. attacker's favorite type of vulnerabilities to perform an attack such as: security hole, misconfiguration, buffer overflow, SQL injection or cross-site scripting [13]. Due to continuous new smart attacks created almost every day, any IDS systems must be regularly updated.

Network forensic data location and sources

In PDF architecture design is necessary to know where forensically relevant data are located. The key sources of forensic evidences on network and Internet are well known (Table 2) [1].

Table 2 Key forensic data sources on network and Internet

DF evidence source	Type of evidence source
Attacker's computer	Log file, working files, ambient data (slack and non allocated space of HD)
Corrupted computer	Log file, working files, ambient data
Firewalls	Log file
Network device	Log files, buffers, memories
ISP (Internet Service Provider)	Client's traffic logged data that are mandatory retained (1 year)
Victim's computer	Log file, working files, ambient data change of configuration, remained malicious files (Trojans, viruses, root-kit), hash value changed files, store stolen files, web traces of the attack, unknown extension files, etc.

Many external sources offer forensic evidence for an incident in forms of IPs and URLs addresses, malicious URLs addresses of DNS-a, type of malware, botnets or C&C servers and malicious scanning. In general, methods of data acquisition for incident identification are quite different in surveyed organization (Fig.4.) [11]

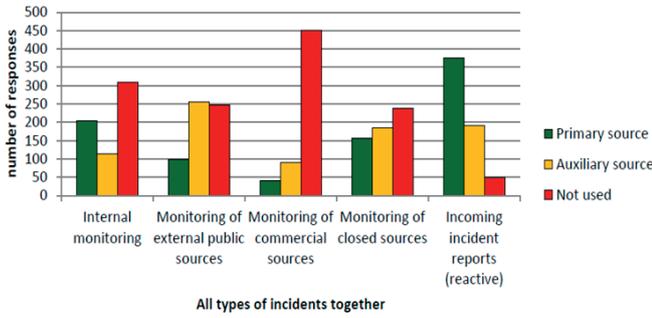


Fig.4. Computer incident data providing [11]

Forensically pertinent data are collected into log files started from border routers and firewalls, via web servers and all other network active devices. The types of collected data should be designed according to PDF system purpose – detection of internal, external or both threats [10, 31 and 32]. The sources of collected data are shown in Table 1.

Table 3: Network forensic data sources

Network device	Proactively collected forensic data
Host	Storage device images, RAM contents, and any static evidence located within agents reach, can be transmitted over the network to the forensic server. It can be network server, such as e-mail, file, print, and DB servers.
Router	The type of information on a router is related to traffic logs that may contain routing process errors, router status, such as the interfaces, or even suspicious activity, depending on logging configuration.
Firewall	It keeps detailed logs of network activity such as regular network traffic, recognized attacks, rejected packets, allowed applications, and sources of suspicious activity (e.g. IPs), keeping track of which protocols or services tried to break in.
Switch	Data in the content addressable memory (CAM), where the mapping of a MAC address to a specific port is located, and information about the Virtual Local Area Network (VLAN).
NIDS	It can log the following (not all inclusive list): Port scans; Traffic coming in on uncommon ports or protocols; Recognized threats, such as worms or viruses; Anonymous FTP attempts or other services; Originating IP addresses of attacks; Bandwidth usage, etc.
NIPS	It blocks or shuts down perceived threats on the network. NIPS can log the same events as an NIDS does, but it analysis data on the network in real-time and scanning them for threats.
Network printer	It usually logs in print jobs with the associated metadata. Someone uses Linux or other NIX operating system, so an agent can be put on a printer to capture its data.
Network copier	It keeps logs and can handle agent installed as network printer does.
Wireless access point-WAP	It logs everything similarly to a normal router, with the addition of wireless-specific information, authentication, SSIDs and other incoming connections.
Storing logged events	The amount of stored data is limited by network transport and storage capacity. The solution is in use of agents to analyze network data in real-time and save only suspect data for later on analysis. The agents can alert on and forward suspicious activity, and reduces the storage load on the system.

Network device	Proactively collected forensic data
Storage area network (SAN)	It is a separate network consisting of devices dedicated to data storage. Its implementation is sometimes complex because it can rival the size of the entire network.
Network attached storage (NAS)	It connects to the network with file level protocols such as NFS or Samba. NAS system can be stripped down to be a server dedicated to storage access, like a normal file server but with even less general-purpose functionality.
Direct attached storage (DAS)	As a non-networked storage connected to the server, it extends the server storage capability by attaching another computer that is solely dedicated to storage. It is extremely fast as it has no network structure to contend with, but suffer from not being able to share storage space with other servers except with its connected host.

Aspects of network collected data forensic analysis

1) Forensic analysis of time stamps

Correlating the time stamps from all network devices is the first step taken in any network investigation. It is impossible to establish a baseline from which to compare data timestamps if timestamps are not synchronized [4]. Best way to synchronize all devices on a network is to use Network Time Protocol (NTP) and to keep all network components accurate within milliseconds of Coordinated Universal Time (UTC). This accuracy is necessary because network communication relies on accurate timestamps to function correctly. Challenges in correlating network events are numerous.

2) Logs filters and archivers analysis

Software such as event log analyzers collects, analyzes, reports, and archives SysLog from networked Windows hosts and other networked active devices. These applications generate reports helping in threats monitoring and network forensic analysis. The event analyzer software proactively reduces system downtime, helps system administrators and increases network performances as whole. Some relevant features of typical event log analyzers are centralized event log management, compliance reporting, security analysis, automatic alerting, etc. [19]. Major benefits of an event log analyzer are shown in Table 4 [38].

Table 4: Benefits of an event log analyzer

Num.	Benefit
1	Network visibility and control
2	Manage <i>Windows Event Logs, Unix/Linux SysLogs, W3C Logs and SQL Server Audit Logs</i>
3	Log data collection at a single centralized location
4	Data integration and normalization
5	Derive reports both on compliance and security
6	Pre-configured to address different compliance needs



Num.	Benefit
7	Support for new compliance reports
8	Report scheduling and distribution
9	Real-time alerts
10	Powerful filtering capacity
11	Automatic, flexible and secured log data archive
12	Important events separated from a pile of events
13	Continuous watch
14	Customizable solution to suit any requirements
15	Monitor the performance of network
16	Customized dashboard view for administrative purposes

There were several design goals when implementing logging infrastructure to capture the log data from every device on the network at all times. Logging into one central place where all the information would converge is the best

way to correlate events across multiple devices. Results of our system operation are presented in Fig.5. (Registered attempt port scans detected by the Snort sensor). For this reason a common time source and log format are desirable. It needs to have integrity in case the logs were required for use as evidence. This demands well designed archiving processes and handling procedures. From a usability perspective the log data needs to be easily reported on and manipulated by common database tools. A final requirement is to have the system immediately alert administrators whenever suspicious items appeared in the logs. SysLog is most feature-laden product in this arena [8, 31] and majority of features and functionality is on its native OS UNIX.

Configuring some of the network devices logs

Before we begin configuring all the devices, a little planning should be done. We need to understand a little about how the SysLog protocol itself was structured, to be able to design an efficient plan for using it at our company. Each SysLog message includes a priority value which is made up of two parts. The value is expressed as *facility severity*, where *facility* is a type of category for the message and *severity* is its relative importance. Some of the facilities are assigned for static purposes and others are user definable. The user definable facilities are those that we need to consider.

When configuring our devices to use SysLog we often have the option to determine which facility and sometimes which severity the device would use. This provided us with an opportunity to organize the incoming log data in a way that would make it easier to manipulate both in an immediate (e.g. live response) and an historical (e.g. forensic analysis) context. Configuring all of our devices according to this scheme lend itself to a number of applications. This approach allowed us to easily focus our attention on high-risk devices. Those messages coming from perimeter hosts could be isolated from the rest of the traffic flow for increased scrutiny. Some examples of this approach will be detailed later in the paper, but now will be examined how some specific devices should be configured.

For the strong monitoring system in PDF infrastructure a log server, as the core of the centralized log architecture, is the most important. It enables proactive security features and makes easier digital forensic analysis later on. This collection of data in the log server, documents and specifications provide the first response to the computer incident and make easier forensic analysis in case of computer crime [9].

The screenshot shows the LogAnalyzer web interface. At the top, there are navigation tabs: Search, Show Events, Statistics, Reports, Help, Search in Knowledge Base, Admin Center, and Logoff. A search filter is set to 'severity:0,1,2,3,4'. Below the search bar, a table titled 'Recent syslog messages' is displayed. The table has columns: Date, Facility, Severity, Host, Syslogtag, ProcessID, and Message. The messages are sorted by date, showing several 'AUTH ALERT' messages from 'snort-ips1' and 'MikroTik'. A detailed view for a message with ID '23983' is shown at the bottom right, displaying fields like Date, Facility, Severity, Host, Syslogtag, ProcessID, and Message.

Fig.5: Central log server-logging SNORT alerts



Snort logging to central log server

We have designed a system for centralized logging alerts and notifications generated by multiple Snort sensors in a network using open source software solutions such as: Cent OS ver. 6.3, Snort ver. 2.9.5.6, and Barnyard, BASE, Rsyslog and Adiscon Log Analyzer ver. 3.6.3.

An open source NIDPS, Snort, is developed by Sourcefire and designed to scan combined signature, protocol, and anomaly. It is the most deployed NIDPS technology worldwide [34]. Its output system, Barnyard, creates a special binary output format called *unified*. Barnyard 2 can read this file, and resend its data to a database, and stores them when the database temporarily cannot accept connections [34]. Based on the code from the Analysis Console for Intrusion Databases (ACID) project, BASE (Basic Analysis and Security Engine), provides a web-based application to query and analyze the alerts created by SNORT IDS system [35]. Rsyslog, open source software utility, based on UNIX and Unix-like platforms, forwards log messages in an IP network [36]. Adiscon Log Analyzer is a web interface to syslog and other network event data. The database can be populated by Monitor Ware Agent, Win Syslog or Event Reporter on the Windows side and by rsyslog on the Unix/Linux side [37].

The system operates in the following manner: SNORT generates alerts or notification and transferred it to Barnyard. Barnyard parses the received messages and store it in SNORT events database (BASE uses this database). Same message sends to local rsyslog on local IDS/IPS sensor, which is forwarded to rsyslog on central log server. Log Analyzer provides easy browsing and analysis of real time network events and reporting services.

CONCLUSION

Deployment of proactive digital forensic (PDF) infrastructure seems to be inevitable in order to assure required level of network security in complex network environment. The first step in implementing PDF into the network is to provide collecting pertinent security and forensic data from all active networked devices in real time. Most useful tool is layered IDS monitoring system for proactive detection and reaction to internal and external network attacks. The event log analyzer software is becoming most important tool for proactive forensic deployment into network security system. The next step in proactive forensic infrastructure deployment should be to standardize event logs from the entire active log devices and to collect them into centralized log server. This log server provides creation of a proactive security system and a proactive forensic architecture. It makes easier digital forensic analysis following computer incident or crime.

The next step in proactive forensic infrastructure deployment should be to standardize event logs from the entire active log devices and to collect them into centralized log server. This log server provides creation of a proactive security system and a proactive forensic architecture. The collection of data, documents and specifications of the key

components such as the data dictionary, syntax specifications, and event taxonomies provide the first response to computer incident [10]. It makes easier digital forensic analysis following computer incident or crime [8].

In this paper authors proposed the theoretical functional models of the strong monitoring and PDF systems for real - time response to both internal and external attacks.

REFERENCES

- [1] Acces Data white paper, The importance of Integrating Host and Network Forensics, www.acessdata.com, 2013.
- [2] Alec Yasinsac, Yanet Manzano, Policies to Enhance Computer and Network Forensics, Proceedings of the 2001 IEEE, Workshop on Information Assurance and Security, United States Military Academy, West Point, NY, 5-6 June, 2001.
- [3] Bayuk, J., Cyber Security Policy Guidebook, ISBN-10: 1118027809, Edition: 1 April 24, www.amazon.com, 2012.
- [4] Boyd Ch., Forster P., Time and date issues in forensic computing a case study, www.elsevier.com, 2004.
- [5] Brian D. Carrier and E.H. Spaford, Automated digital evidence target definition using outlier analysis and existing evidence, Proceedings of the 2005, Digital Forensic Research Workshop (DFRWS). Citeseer, 2005.
- [6] Bruce J. Nikkel, The Role of Digital Forensics within a Corporate Organization, IBSA Conference, Vienna, May 2006.
- [7] Dan Rathbun, Using SysLog in a Microsoft & Cisco Environment, SANS Institute Reading Room, June 27, 2003.
- [8] David E. Learner editor, Electronic crime scene investigation, Nova Science Publishers, Inc., Library of congress cataloging-in-publication data, ISBN: 978-1-60876-493-8 (E-Book) New York, 2009.
- [9] Douglas Schweitzer, Incident Response: Computer Forensics Toolkit, Wiley Publishing Inc. 2003.
- [10] NCJRS, USA, Electronic Crime Scene Investigation: A Guide for First Responders, www.ncjrs.org, 2001.
- [11] ENISA, Proactive Detection of Network Security Incidents, 2011.
- [12] Garfinkel, S. Network Forensics: Tapping the Internet, <http://www.oreillynet.com/pub/a/network/2002/04/26/net-tap.html>
- [13] Gottlieb, J., Key challenges in proactive threat management, CEO of Sensage, Help Net Securita News, 29 August 2012.
- [14] Grance T., Kent K., Kim B., Computer Security Incident Handling Guide, NIST SP 800-61, January 2004.
- [15] Gregory Leibolt, The Complex World of Corporate Cyber Forensics Investigations, Springer's Forensic Laboratory Science Series, 2011.
- [16] Icove D., Segar K., VonStorch W., Computer Crime, A Crimefighter's Handbook, O'Reilly & Associates, 2006.
- [17] ISF, The standard of Good Practice for Information Security, www.isf.com, 2006.
- [18] Kaufman J. Robert, Intrusion Detection and Incident Response, IS 3523 course, UTSA Spring, 2012, <http://faculty.business.utsa.edu/rkaufman/IDLsn4.ppt>, (accessed: 24.03.2012).



- [19] Keith J. Jones, Richard Bejtlich, Curtis W. Rose, *Real Digital Forensics, Computer Security and Incident Response*, Addison-Wesley, 2008.
- [20] Matson J.V., *Effective Expert Testimony*, 3rd edition Boca Raton, Press, p.71, 1999.
- [21] Milosavljević, M., Grubor, G., *Computer crime investigation*, Univerzitet Singidunum, 2011.
- [22] Nelson B., Phillips A., Enfinger F., Christopher S., *Guide To Computer Forensics and Investigations*, Second Edition, Published by Course Technology, 25 Thompson Learning, Inc., Printed in Canada, 2006.
- [23] Norman ASA, *Proactive Forensic Toolkit*, RSA Conference 2010, San Francisco, 2 March, 2010.
- [24] Paul Taylor, *Proactive Forensics in the Workplace*, Litigation and Forensics, Data Recovery Services, Inc. www.legal-forensics.com, 2010.
- [25] Phillip G. Bradford, Ning Hu, *A Layered Approach to Insider Threat Detection and Proactive Forensics*, The University of Alabama, Department of Computer Science, Box 870290, Tuscaloosa, AL 35487-0290 pgb@cs.ua.edu, nhu@cs.ua.edu, 2006.
- [26] [1] Rajdeep A, Niyogi, R.C. Joshi, Emmanuel S. Pilli, *Generic Framework for Network Forensics* 2010 International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 11.
- [27] RFC 3227, *Guidelines for Evidence Collection and Archiving*, www.faqs.org/rfcs/rfc3227.html, 2002.
- [28] RSA Conference 2013, *Big Data Capabilities*, Teradata Corporation and the Ponemon Institute, San Francisco, 2013.
- [29] SoltanAlharbi, Belaid Moa, Jens Weber-Jahnke and IssaTraore, *Performance Proactive Digital Forensics*, High Performance Computing Symposium 2012 (HPCS2012) IOP Publishing, *Journal of Physics: Conference Series* 385 (2012) 012003 doi:10.1088/1742-6596/385/1/012003.
- [30] S. Waldbusser, *Remote Network Monitoring Management Information Base – IETF RFC 1757*, Carnegie Mellon University, USA (Last updated 2013-03-02), 2013.
- [31] Terrence Lillard et al., *Digital forensics for network, Internet, and cloud computing*, British Library Cataloguing-in-Publication Data ISBN: 978-1-59749-537-0, Elsevier, 2010.
- [32] Tim Grance, et al., *Guide to Computer and Network Data Analysis: Applying Forensic Techniques to Incident Response*, National Institute of Standards and Technology SP 800-86, 2008.
- [33] Zimmerman, S., *Proactive Computer Forensics*, *Digital Forensics Magazine*, Issue 3, 1st May 2010.
- [34] www.snort.org.
- [35] <http://sourceforge.net/projects/secureideas/>.
- [36] <http://en.wikipedia.org/wiki/Rsyslog>.
- [37] <http://logalyzer.adiscon.com>.
- [38] <http://www.manageengine-sales.co.uk/group28.html>