Digital Forensic Analysis for Enhancing Information Security

Ojeniyi Joseph Adebayo, Idris Suleiman & Abdulmalik Yunusa Ade
Department of Cyber Security Science
Federal University of Technology
Minna, Nigeria
Ojeniyija@futminna.edu.ng
Adenijiadedapobolaji@yahoo.com, Sidris27@gmail.com

Ganiyu, S. O, and Alabi, I. O.
Department of Information and Media Technology
Federal University of Technology
Minna, Nigeria
Shefiu.ganiyu@futminna.edu.ng
Isiaq.alabi@futminna.edu.ng

***Abstract*— Digital Forensics is an area of Forensics Science that uses the application of scientific method toward crime investigation. The thwarting of forensic evidence is known as anti-forensics, the aim of which is ambiguous in the sense that it could be bad or good. The aim of this project is to simulate digital crimes scenario and carry out forensic and anti-forensic analysis to enhance security. This project uses several forensics and anti-forensic tools and techniques to carry out this work. The data analyzed were gotten from result of the simulation. The results reveal that although it might be difficult to investigate digital crime but with the help of sophisticated forensic tools/anti-forensics tools it can be accomplished.**

***Index Terms*— Digital forensic, anti-digital forensic, image acquisition, image integrity, privacy.**

I. INTRODUCTION

Forensic Science deals with evidence presentation using scientific processes; while Digital forensic is a branch of Forensic Science that deals with careful extraction or mining of digital evidence that has probative value within a predefined scope in such a way that it can be admissible in the court of law without doubt or question being raised about its integrity.

Due to human inherent nature to invade and circumvent justice they make digital evidence acquisition difficult and almost impossible using some counter measure known as anti-digital forensics (or just anti-forensic). Anti-forensics is the means of thwarting forensic processes thereby making forensic processes difficult, impossible or by delaying it and frustrating the forensic investigators and forensic tools. The rest of the paper is organized as follows: Section 2 presents related work, and two case studies are defined and experimentally investigated all through Sections 3 to 6. Section 7 concludes the paper.

II. RELATED WORKS

A. Enhanced Information through Forensics

Given that there are anti-forensic tools that can obfuscate, minimize or eliminate attack footprints, forensic analysis becomes harder [1]. In their work, Changwei, Anoop, and Duminda [1] aimed to use attack graphs in forensic examinations. The methodology they used included anti-forensic capabilities into attack graphs, so that the missing evidence can be explained by using longer attack paths that erase potential evidence. At the end of their work they were able to show how attack graphs could be used to help forensics investigators narrow down potential attack scenarios, along with evidence left by attackers. Observable limitation of their work from the anti-forensic techniques/tool vulnerability tools they used is the TrueCrypt. Most attackers no longer use TrueCrypt because of the presence it leaves as systems trace on the boot loader.

Balogun and Shao [2] in their work examined what data encoding add to information security and then spotted out its influences on the digital forensics of disk drives. The purpose was to converse the obtainable methods and tools, in digital forensics, to find solutions to the problems posed by encryption. They used TrueCrypt as case study for the encryption solution to illustrate their ideas. They further talked about some features of TrueCrypt software that provides users with plausible deniability and non-repudiation abilities. This makes digital forensics examinations of encrypted disk drives stiffer and less actualizable. The limitation in their work is the TrueCrypt boot loader traces.

Benjamin [3] researched on "Modeling and refinement of forensic data acquisition Specifications", his aim was to "define a model of a special type of digital forensics tools, known as data acquisition tools", the intention of his work is to give a formal description against which implementations of data collecting procedures can be analyzed. The approach he used was the formal refinement language Event-B (Event-B of the data acquisition functionality of digital forensics tools). Event-B is an extension of Abrial's B method Abrial (1996) for modeling distributed systems.

B. Enhanced Information Security through Anti-Forensics

Changwei, Anoop and Duminda [1] worked on a Model using evidence from security events for network attack analysis. They aim at "how to use the information obtained from security events to construct an attack scenario and build an evidence graph." And the objectives of their works were, "To achieve the accuracy and completeness of the evidence

graph, to correlate evidence by reasoning the causality, and use an anti-forensics database and a corresponding attack graph to find the missing evidence", the methodology they employed were: prolog inductive reasoning, abductive reasoning, global reasoning and mapping the evidence to a logical attack graph to construct an evidence graph for network forensics analysis. And they arrived at having a proposed network forensics model, which extends a Prolog logic based system, MulVAL, to automate the causality correlation between evidence collected from security events in an enterprise network.

Johannes and Michael [4] researched on a work titled "Anti-forensic resilient memory acquisition", their objectives were:

i. To examine a number of simple anti-forensic techniques and test a representative sample of current commercial and free memory acquisition tools.
ii. To find out if current tools are resilient to very simple anti-forensic measures.
iii. To present a novel memory acquisition technique.
iv. To then evaluate this technique's further vulnerability to subversion by considering more advanced anti-forensic attacks. The method they used was based on direct page table manipulation and PCI hardware introspection, without relying on operating system facilities.

The limitation of their work is its reliance on the operating system on finding the page tables in the first place. All addresses in CR3 and Page Tables are physical addresses.

Przemyslaw and Elias [5] carried out research on "Computer Anti-forensics Methods and Their Impact on Computer Forensic Investigation". The aim of the research was to test whether current known counter-forensics technology can efficiently interfere with computer forensics processes, and their major objectives were to explore the anti-forensics problem in various stages of computer forensic investigation from both a theoretical and practical point of view and to identify the most known computer anti-forensic techniques and test them practically against computer forensic software. They proved that not all counter-forensics techniques are efficient when compared against forensics software. Their major limitation was that the research was not carried out about each individual technique against a range of different forensic tools. Also, the various anti-forensics techniques was not evaluated against packages specifically designed for detection of those techniques in order to develop a much clearer opinion as to whether it is possible to beat counter forensics.

C. *Enhanced Information Security through Counter-Anti-Forensics*

Marco, Alessandro, Alessandro, and Mauro [6] carried out a work on "Countering Anti-Forensics by Means of Data Fusion"; with the aim of analyzing the possibility offered by the adoption of a data fusion framework in a Counter-Anti-Forensic (CAF) scenario". The methodology they employed was a theoretical framework, based on Dempster-Shafer Theory of Evidence. Their objectives were:

i. To synergically merge information provided by Image Forensics (IF) tools and Counter Anti-forensics (CAF) tools.
ii. To reveal traces introduced by anti-forensic algorithms.
iii. To account for the non-trivial relationships between IF and CAF techniques.
iv. To evaluate the proposed method within a representative forensic task, that is splicing detection in JPEG images, with the forger trying to conceal traces.

The limitation of their work was that they did not take into account the following facts:

i. IF tools may be searching for mutually exclusive traces, so some combinations of tool outputs could be excluded.
ii. For a given footprint, IF and CAF algorithms are expected to be in contradiction, so if both kinds of tools detect their footprint this should at least raise some doubts about the correctness of the outputs.
iii. Detecting some kinds of anti-forensic processing does not necessarily imply that the image is a fake.
iv. They were able to arrive at investigating the use of data fusion as a tool for countering Anti-forensics.

III. METHODOLOGY

A. *Introduction*

Before any forensic investigation can be carried out, there must have been a case at hand that needs evidence - especially evidence(s) relating to electronic media. Computer forensic examiners or investigators need to beware of possible circumvention techniques that computer criminals employ to defeat digital forensic approach known as Anti-forensic; but the aim of this work is to make the forensic examiner to be pro-active in investigations and to use reliable techniques. It will be un-wrapping some things the investigator has to put under consideration before, during and after forensic examination.

In this paper it is assumed that an actual crime that involves and greatly relies on digital evidence has actually occurred, some of the problem or cases this paper would be providing solution to would be done through forensic anti-forensics analysis. This would be better understood by analyzing and providing solution to the following fictitious cases assumed.

B. *Assumed Case*

In Federal University of Technology Minna, there was a policy that nobody should for any reason copy any of the institution's file without authorization, but a disgruntled employee who has being asked to resign was caught copying some relevant files into his personal laptop from one of his colleague's company computer whom he had quarrel with recently. When his colleague caught him, he denied it in the presence of others and when his system was searched by the rest colleagues they found nothing incriminating or any

unauthorized file in his possession pertaining to FUT Minna, and now a forensic investigator has being brought to the scene to help out, with the analysis.

Virtual environment would be used to simulate this crime, and the target offender's machine would be assumed to be running Window 7, any windows operating system would work just fine as some of the tools used has also been tested with Window 8 and Window XP they all work fine, most of the tools used to carry out the analysis are operating system independent – meaning Linux, Solaris or Apple IOS are not exceptional, the sharp difference between them boils downs to the their differences in file systems, network configuration and their memory management. The methodology used in this work will be broken down into: Volatile data analysis and Non-volatile/persistent data analysis.

Before conducting forensics operation there are lots of things the forensics investigator has to put in mind in other to have valid evidence such as the operation mode of anti-forensics. The knowledge of anti-forensics to an examiner would make the examiner to know when his on track or not to an extent as this would program the mind of the examiner to be pro-active hence the essence of this work.

An anti-forensic researcher examines the forensic tools and their approaches and tries to identify its weakness by thwarting these forensics approaches and devising various techniques and tools. For a forensic examiner to be successful in his job, he must be good in handling most of the forensic tool effectively most especially the free/open source tools, it is a nice idea for him/her to also understand the Anti-forensic approaches and their various relative tools and sometimes weakness in anti-forensic approaches and in tools; for example – when a TrueCrypt is used leaves a trace (its boot loader) even in most cases due to human carelessness some fragment or traces of evidence/artifacts or anti-forensic operation might be detectable by the examiner.

Case Analysis:

From the above model of the case, it can be deduced that if the suspect copies the file to his computer, he may decide to delete it temporary after using it, using shift + delete key to permanently delete the file. Forensic has proven that what most people think they have permanently deleted either with the normal use of shift + delete or deleting files from recycling bin does not actually delete but only raise flag to the file system that the space is available for files to be written on to mark as an unallocated space, within the period the file system spends in doing some house-keeping routing, if certain forensic tools are used it can restore back the file that was taught to have been permanently deleted. Some criminals are aware of this and hence some time find hidden place that the file system and operating system cannot access such as Host protected areas with other lots of areas that even some sophisticated forensic tools would not be able to access. As computer forensic investigators, the investigators must be aware of this areas, this process of data hidden is one of the many anti-forensic processes. Forensic examiner has to be aware of anti-forensic processes. Every anti-forensic effort or approach is being built toward countering phases of forensic analysis either to delay the forensic processes or to even make it impossible.

C. Tools Employed

a. Installation of Virtual Machine(VMware workstation 10.0 as the virtual environment)
b. Installation of Guest Operating systems on the VMware's
   i. Kali Linux (test board)
   ii. Window 7
   iii. Forensic Investigation Tools (FTK)
   iv. AccessData FTK Imager
   v. DumpIt
   vi. WinHex, OSForensics (OSFclone, OSFmount, OSFPassMask)
   vii. Autopsy (Forensic suit),
   viii. SIFT (SAN Investigatory Forensic Tool)
   ix. SecreteLayer
   x. Time Stomp
   xi. Slacker.exe
   xii. FOCA

Guymager.

ALGORITHM AND MODEL DEVELOPMENT

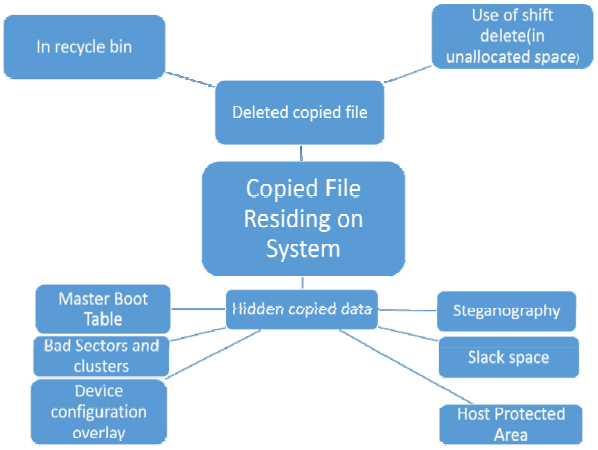

Figure 1: Case model.

Pseudo Code

Seized System;
while Copied file resides on system
The data is hidden;
if data is hidden
check:
- *Master Boot Table,*
- *Bad Sectors/clusters,*
- *Device Configuration overlay,*
- *Steganography,*
- *Slack Space, OR*

- *Host Protected Area;*

else

The data has been deleted;

if data has been deleted

check: Recycle bin OR Unallocated space;

IV. EXPERIMENTAL WORK

After the text edit has been completed, the paper is ready The experimental framework comprises the use of open forensic and anti-forensic tools and techniques to demonstrate how digital forensic analysis could be achieved using one of the numerous open source Forensic and Anti-forensic Tools, it involves the installation and setting up of test bed, the test bed or lab used here is on virtual machine which help to utilized the efficiency but yet less expensive forensic operation. VMware workstation 10 was used and window 7 ultimate was installed as guest operating system (OS) on the virtual machine, other forensic tools (FTs) were installed on the host OS while the Anti-Forensic tools (AFT) was ran on the guest OS. The figure below shows the different interface of the experiment

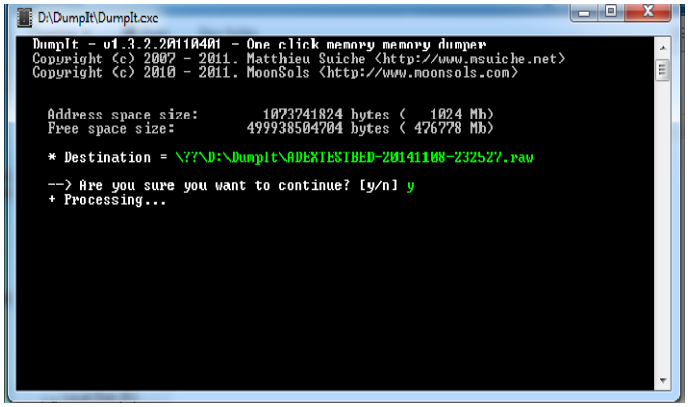

Fig 2: DumpIt capturing RAM image.

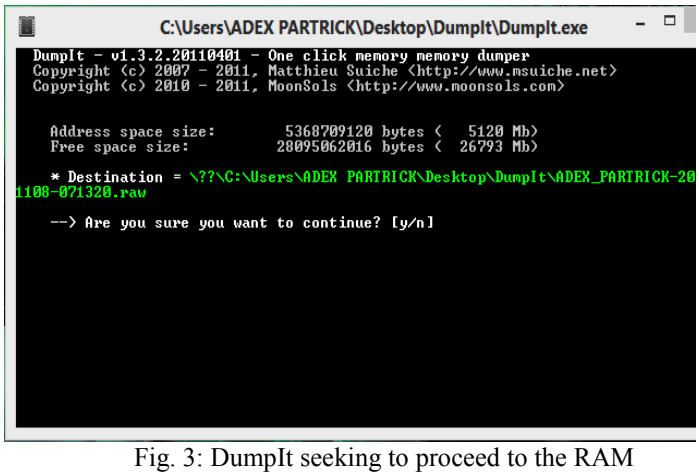

Fig. 3: DumpIt seeking to proceed to the RAM imaging.

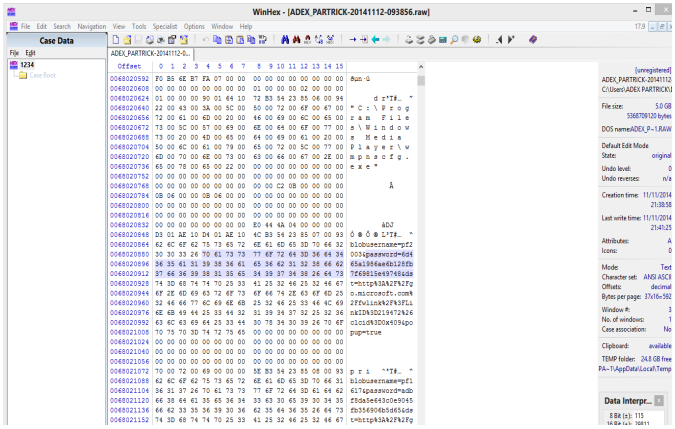

Fig. 4: WinHex returning various password hashes.

V. RESULTS AND DISCUSSION

A. *Volatile Data Analysis*

The following forensic software were used and their results and impact are discussed below.

DumpIt

DumpIt is a volatile data imaging tools that has been used with any WinHex as a Hex Editor tool in other to capture the volatile data by imaging the RAM as shown in Fig. 2. The raw memory dump is generated in the current directory, only a confirmation question is prompted before starting as shown in Fig. 3. This raw memory dump generated was now used as item evidence in the WinHex. The easiest way to disrupt volatile data collection is by putting off the system, considering the fact that once power goes of the system, any stored data on the volatile memory is lost. Once the attacker can switch of the system before the arrival of the forensic examiner, the success of the forensic examination will be drastically reduced hence anti-forensic.

WinHex

The WinHex, was used to analyze the image of the raw memory dump generated by the DumpIt Tool, and as the image was hashed with a Message Digest 5 (MD5), SHA1 and SHA126 algorithm and it was compared before and after forensic operation to make sure its integrity was intact. When the content was searched for passwords WinHex returned various password hashes as show in Fig. 4. After the analysis of the image with WinHex it was able to get various passwords, and file logs such as the type of files and their extensions that was running and the resource they were using. And all this information was documented. AccessData FTK was also used to hash and compare hashes in relative to WinHex and exactly the same was generated, and when AccessData FTK Imager was used to do the raw dump memory analysis the same result was gotten.

B. Persistent Data Analysis

Some part of the analysis was carried out on virtual machine and the other was carried out on the host computer, the reason is that, the analysis could be carried across various forensic laboratories that ordinarily would have involved several machine or impossible with just a single machine. Since the machine under investigation was installed on the virtual machine, the method used in getting evidence from the guest Operating System installed on the virtual machine was a little bit different. It involved imaging the virtual machine Hard Disk Drive (HDD) instead of the host machine or the physical machine.

Analysis of the Volatile Image Gotten from DumpIt Volatile Data Imager of the Seized System

After using WinHex to analyze the volatile image generated by DumpIt, we got the hex representation of the volatile data, since the hex data representation was not readable or understood by human, it was then converted into a messy plain text of which careful observation revealed some sensitive information such the hashed password was found. The hashes were imported into OSForensics tool rainbow table to reveal the plain text. The problem from here was that the username that corresponds to the various passwords was still not known.

Analysis of the Image Gotten from AccessData FTK imager of the Seized System

When the image was analyzed with AccessData FTK imager files incriminating was found in the bad sector and unallocated space. The file found on the unallocated space are possibly those file he deleted using Shit + Delete with the hope that he has actually deleted it from the recycle bin of which only sophisticated forensic tools can retrieve such as AccessData FTK imager.

The file found in the bad sectors were even more incriminating, these files must have been stored there for the sole aim of anti-forensic because merely searching the system will never reveal the presence of files on bad sectors, many forensic software would not even be able to search there for hidden files, when most forensic software comes across bad sectors during scanning, they skip it and continue with the rest of the sectors/clusters. From the result gotten from analyzing his system RAM with DumpIt, one incredible observation is the presence of suspicious programs found, which was Slacker.exe, Slacker.exe is a command line anti-forensic tools used to hide files in slack space and also TimeStomp was found – TimeStomp is a metasploit tools used for manipulating MACB. Table 1 is a fast representation of the meaning of MACB broken down by type of file system:

Table 1 Show the Meaning of MACB on Various File Systems.

File System	M	A	C	B
FAT	Written	Accessed	Changed	-
NTFS	Modified	Accessed	MFT Modified	Created
Ext 2/3	Modified	Accessed	Changed	-
Ext 4	Modified	Accessed	Changed	Created
UFS	Modified	Accessed	Changed	-

Analysis of the Image Gotten from OSFclone Imaging/Cloning Tool

All the passwords discovered in WinHex were not in plain text, but hashes. OSForensics has rainbow table and dictionary list as plug-in to the forensics tools, consequently, the hash value was computed and compared against known hashes on the rainbow table and it was successfully decrypted by finding a text that march the hash. To prevent the investigator from using the common password cracking techniques such as:

i. Dictionary
ii. Rainbow table
iii. And password guessing

A criminal may decide to use random string of data that cannot be found in the dictionary to avoid password guessing and dictionary attack, if the length of this string is considerably long it might just be impossible for even the rainbow table to crack the password, though some desperate criminal might even take the pain of memorizing hash value of a random string for password. For example if the word “forensic” is used as password and MD5 was used to hash it, it would be “ad4642428b76i25b16c6dae5c84a9c”, depending on importance of the file the criminal can memorize that long string of data, the beauty of this is that MD5 is a one hash function, so if the user enters this long string as password depending on the encryption algorithm used the value is going to change to something else. The implication of this is that the password the forensic examiner would be receiving will be confusing to him and even the almighty rainbow table would be worthless in this kind of situation.

Analysis of the Image Gotten from Guymager Imaging Tool

The image that was gotten from Guymager imaging tool was loaded into Autopsy for case analysis. In the analysis of non-volatile data, the original image of the extracted system should be forensically cloned or duplicated; because once the original image is contaminated, the evidence definitely would be doubtful or questionable. Because of the advancement in anti-forensic operation, like the use of logic zip or logic bomb, before trying to crack some of the files, precaution must be taken due to the fact that there are some encryption algorithms that can have more than two encryption keys and decryption keys. One of the key would be meant to alter to the bit of the file into another file entirely, some might even delete the file used to access some of the criminal’s private document. Cases of such have been reported recently.

Autopsy

After the image gotten from the Guymager was feed into Autopsy for analysis it showed all visited URL, deleted folders, directories, the content of the recycling bin, the MACE of each of the files, bookmarks, browsing history, cookies and all removable drives.

From the lists of URLs, it was discovered that he has visited Facebook and Gmail, this means from the passwords gotten from WinHex we can now relate them each to his account respectively, but there is still something missing which is his usernames for the two accounts. For the username to be extracted, looking at Facebook for example requires the following:

Email or Phone: Password:

From the above we can figure out the authentication SQL query to be:

If (email OR Phone AND Password) grant Access;

From the above it is obvious that since we have gotten his Facebook password we only need either the email and password or the phone number and his password to gain access.

For one to have access to Gmail, the credential requires are password and the corresponding email address, from the password gotten from the RAM and the email gotten from FOCA access to the Google account is sure. To get the email and/or phone number FOCA would get the job done. Observably, many a times when a site is visited with browsers, for instance Facebook, if the system is short down or the user opt out from surfing the site (Facebook) without logging out, the next time the browser is fired up, it loads the last logged in session, the magic behind this at the client side is the cookies which is like a token, or a value the server stores on the browser cache for identity remembrance, which this cookies the server can then remember who the user is and serve or load his last browsing session to him, the forensic examiner can use this little piece of information for forensic investigation.

Other traces of the forensic examiner may use is the browsing history, remembered password and bookmarks; if this artifacts are cleared it will harden the forensic investigation. Some dubious criminals that suspects forensic investigation might do the following to cover track:

i. Change file extension and header before deletion
ii. Alter the MACB file attributes
iii. Over slack spaces and unallocated spaces

FOCA

When FOCA was used to query Federal University of Technology, using Google query option in FOCA search options, lots of information was harvested, but the one that was picked was CV of FUT Minna staff and among it was his own inclusive. From his CV, his active emails including Gmail account name, and phone number were found.

Most of this tools used to do this forensic analysis have almost exactly the same features to carry out basic forensic analysis, sadly on the seized system there was stegno files of which none of this tools actually took note of, probably because it was a free/evaluation version.

Further researched work should be carried out using same tools that are licensed with forensic analysis with lots of features. The best thing to do in other to avoid this FOCA from successfully given out information the forensic investigator may find interesting like your email address is to have many email accounts for various purpose.

For example; the email that should be used in Curriculum Vitae should be different from the one used for social media, and also different from the one use for online trading or e-commerce, this same goes for phone numbers. The kind off sensitive information pasted on the internet should be minimal especially the social media as this is a pool or a rich resources for information retrieval

VI. CONCLUSION AND RECOMMENDATIONS

In this paper, widely acceptable forensic methodology such as identifying, collecting, analyzing and reporting have been investigated. Hidden places on the logical and physical structure of the computer where evidence may reside were described and some forensic tools and their application to real life situation presented. Due to human inherent element such as being bias as a result of sentiment cases, measure should be put in place to avoid investigators handling cases that he may pick interest in or that has to do with people that know him directly or indirectly, if such issue arises where investigator has interest in the case, it should be awarded to external professional examiner.

REFERENCES

[1] L.Changwei, S. Anoop, & W. Duminda, (2014). A Model Towards Using Evidence From Security Events. *A Model Towards Using Evidence From Security Events*, vol. *10*, 103-122

[2] M. Balogun & Y. Z. Shao, "Privacy Impacts of Data Encryption on the Efficiency". *International Journal of Advanced Computer Science and Applications,* vol. 4, no. 5, 2013, pp. 36-40

[3] A. Benjamin, "Modelling and refinement of forensic data acquisition", *Digital Investigation,* vol. 11, no. 2, 2014, pp. 90-101

[4] S. Johannes, & C. Michael, "Anti-forensic resilient memory acquisition", *Digital Investigation,* 2014

[5] P. Przemyslaw, and P. Elias, "Computer Anti-forensics Methods and Their Impact on", 2009, Retrieved September 3, 2014 from http://hdl.handle.net/10552/1508.

[6] F. Marco, B. Alessandro, P. Alessandro, and B. Mauro, "Countering Anti-Forensics by Means of Data Fusion", 2014.