Malware Detection in Cloud Computing

Safaa Salam Hatem
College of Science,
University of Al-Qadisiyah,
Al-Qadisiyah, Iraq

Dr. Maged H. wafy
Information Technology department,
Faculty of Computers and
Information, Helwan University,
Cairo, Egypt

Dr. Mahmoud M. El-Khouly
Information Technology department,
Faculty of Computers and
Information, Helwan University,
Cairo, Egypt

***Abstract*—Antivirus software is one of the most widely used tools for detecting and stopping malicious and unwanted files. However, the long term effect of traditional host based antivirus is questionable. Antivirus software fails to detect many modern threats and its increasing complexity has resulted in vulnerabilities that are being exploited by malware. This paper advocates a new model for malware detection on end hosts based on providing antivirus as an in-cloud network service. This model enables identification of malicious and unwanted software by multiple detection engines Respectively,This approach provides several important benefits including better detection of malicious software, enhanced forensics capabilities and improved deployability. Malware detection in cloud computing includes a lightweight, cross-Storge host agent and a network service. In this paper Combines detection techniques, static signatures analyze and Dynamic analysis detection. Using this mechanism we find that cloud- malware detection provides 35% better detection coverage against recent threats compared to a single antivirus engine and a 98% detection rate across the cloud environment.**

Keywords— Malware; Security; Cloud Computing

I. INTRODUCTION

Detecting malicious software is a complex problem. The vast, ever-increasing ecosystem of malicious software and tools presents a daunting challenge for network operators and IT administrators. Antivirus software is one of the most widely used tools for detecting and stopping malicious and unwanted software. However, the elevating sophistication of modern malicious software means that it is increased challenging for any single vendor to develop signatures for every new threat. Indeed, a recent Microsoft survey found more than 45,000 new variants of backdoors, Trojans, and bots during the second half of 2006 [1].

In this paper, we suggest a new model for the detection functionality currently performed by host-based antivirus software. This paper is characterized by two key changes.

- *Malware detection as a network service*: First, the detection capabilities currently provided by host-based antivirus software can be more efficiently and effectively provided as an in-cloud network service. Instead of running complex analysis software on every end host, we suggest that each end host runs a lightweight process to detect new files, send them to a network service for analysis, and then permit access or quarantine them based on a report returned by the network service.
- *Multi-detection techniques*: Second, the identification of malicious and unwanted software should be determined by multiple, Different detection engines Respectively. Suggest that malware detection systems should leverage the detection capabilities of multiple, Collection detection engines to more effectively determine malicious and unwanted files.

In the future, we will see an increase in the dependence of cloud computing as consumers increasingly move to mobile platforms for their computing needs. Cloud technologies have become possible by tuberculation in order to share physical server resources between multiple virtual machines (VMs). The advantages of this approach include an increase in the number of clients that can be served for every physical server and the ability to provide software as a service (SaaS).

In this paper, previous work on malware detection had been presented, both conventional and in the presence of cloud as storage in order to determine the best approach for detection in the cloud [2]. We also argue the benefits of multiple detection throughout the cloud and present a new approach to coordinate detection across the cloud.

Section II provides background and related work the research area, specifically: cloud technologies, security system in the cloud, malware detection and detection in the cloud. Section III, we explain our Proposed System. Section IV we show Remarks of our system. Finally, section V Conclusions the points raised in this paper and provide some ideas for future work.

II. BACKGROUND

A. Cloud Computing

With the Internet's ubiquity in modern living, many argue that some level of cloud computing is now a common occurrence. This research heavily focuses on cloud computing technology, and thus requires a formal definition of cloud computing. Cloud computing cannot be easily defined. There are many definitions, which share the same common denominator: the Internet. Cloud computing is a way to use the Internet in the daily life of a single machine or single room, using all the tools installed on computers [Figure 1]. It is also the ability to use shared computing resources with local servers handling applications. With cloud computing users do not worry about the location and the storage of their data. They just start using the services anywhere and at any time. The main driver of this technology is Virtualization (Hypervisor) and virtual appliance [3]

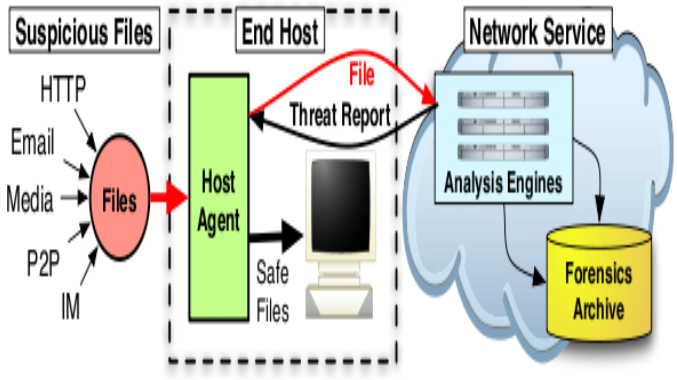


Fig. 1. The Flow of the Process of the cloud computing systems

Cloud computing offers different service models that allow customers to choose the appropriate service model that fits their environment needs, Cloud service models are software as a service (SaaS), Platform as a service (PaaS), and Infrastructure as a service (IaaS) [4] [5]:

- Software-as-a-service (SaaS): The consumer uses the provider's applications, which are hosted in the cloud. For example, Salesforce.com CRM Application.
- Platform-as-a-service (PaaS): Consumers deploy their own applications into the cloud infrastructure. Programming languages and application development tools used must be supported by the provider. For example, Google Apps.
- Infrastructure-as-a-service (IaaS): Consumers are able to provide storage, network, processing, and other resources, and deploy and operate arbitrary software, ranging from applications to operating systems.

B. Related Work

As a matter of fact "cloud computing" concepts date backward to the 1950s, when large-scale mainframes were made available to schools and corporations, In addition, the on-demand computing concept of the cloud model went back to the time-sharing era in the 1960s [7]. Therefore, many of the cloud computing security issues are arguably quite similar to the ones that were introduced during the Internet expansion era. However, Malware detection in a Cloud Computing service was explicitly introduced in [8], what we now commonly refer to as cloud computing is the result of an evolution of the widespread adoption of Virtualization, service-oriented architecture, autonomic, and utility computing. Details such as the location of infrastructure or component

Devices are unknowns to most end-users, who no longer need to be thorough, understand or control the technology infrastructure that supports their computing activities. There are several previous studies related to this research dealing with all of cloud computing and its structure as well as detection systems used for each of the Static analysis, detection: Signature Optimizing Pattern Matching and Dynamic analysis detection: Heuristic, Can be summarized as follows:

C. In Cloud Computing

Oberheide [9] proposed in his thesis "N-Version Antivirus in the Network Cloud" a new model for antivirus deployment by providing antivirus functionality Asa network service using N-version protection. This novel paradigm provides significant advantages over traditional host-based an antivirus, including better detection of malicious software, enhanced forensic's capabilities, improved deployable and manageable retrospective detection. Use a production implementation and real-world deployment of the Cloud AV platform. In addition, Schmidt, et. Al [10], presented an approach for combined malware detection and kernel rootkit prevention in virtualized cloud computing environments, and all running binaries in virtual instance are intercepted and submitted to one or more analysis engines. Besides a complete check against a signature database, lives introspection of all system calls is performed to detect yet unknown exploits or malware.

Malware detection has been an important issue in computing since the late '80s. Since then the predominant method of malware detection has been to scan a computer system for infection by matching malware signatures to files on the computer. Although detection of known samples is extremely reliable, signature based detection only works for malware that has been obtained, analyzed and a suitable signature identified. Murad et al. [11], showed that signature based detection can be thwarted by analyzing the malware instructions and identifying the instructions that comprise the signature.

Li [6] decreased the signature mapping cost by optimizing signature library, taking advantage of common conduct characteristics of viruses such as self-replicate and seasoning, and proposed optimization policy against this scalability issue with the help of data mining. Moreover, he decreased the number of unnecessary signature matching and raises efficiency of that comparison procedure by rearrangement within a signature library. In Heuristic detection, Treadwell [12] suggested analyzing the obfuscation pattern before unpacking, providing a chance to prevent malware from further execution. In this paper, we propose a heuristic detection approach that targets obfuscated windows binary files being loaded into memory.

III. Proposed System and Results

This paper proposes a malware detection system to be built on cloud environment,

Initially, we will divide the system architecture into two main sections according to the mechanism of action of each part. First Section, relating cloud computing and the second section, explains the two detection techniques that used. A cloud computing, we use cloud as software as a service (SAAS) which is a new service and an information delivery model that utilizes existing technologies [14]. The proposal of this work is to find the optimal solutions to the problems of anti-viruses and improve performance and find possible alternatives for a better working environment without problems with high efficiency and flexibility. In this system, a traditional detection technique as per static signatures and dynamic detection technology has been used. Then, safer

system methods and modern to rival existing anti-virus has been selected, for this a hybrid system of two detection methods has been created:

1) Static analysis: Signature Optimizing Pattern Matching.

2) Dynamic analysis: Heuristic.

Both of them will be explained below.

a) Signature Optimizing Pattern Matching: This method is used depends on the signature, which storage already in the database. For this purpose, we used a string matching algorithm, comparison variants of which arise in finding similar DNA or protein sequences.

It is important to use this method to our system Because of there are several new viruses detected; therefore, it becomes necessary to add their signatures to a library. To this end, a failure comparisons increase, this would negatively affect the efficiency of the signature matching procedure. Based on the virus characteristic of self-replicating and seasoning, this system proposed optimizing policy focus on Signature library; one common feature of virus is that it will scan targeted files and inject the malicious code into the normal files. So lots of replicas coexist within one system. So when any virus is detected by signature match, this virus signature is temporarily stored, so the other replicas do not need to match against the other large amount of signatures in the actual signature library [Figure 2].

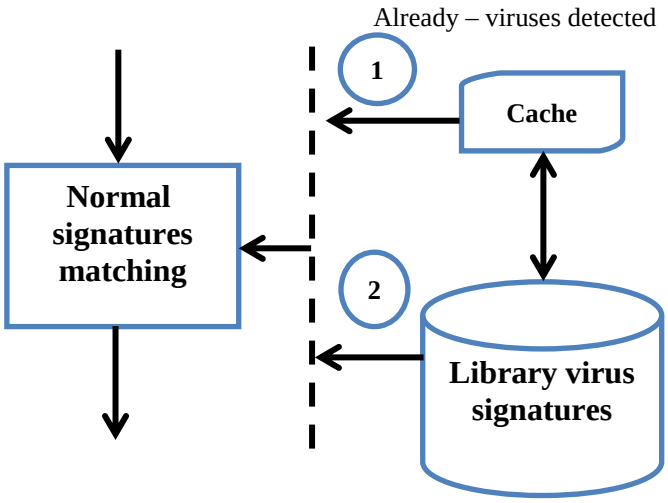


Fig. 2. Shows the process for Optimizing Pattern Matching of Library signatures

So this pre-comparison with already-detected viruses will reduce the signature matching times. [6]

b) Heuristic Detection: antivirus software often used one or several techniques proactively detect malware. This method is dependent on analyzing suspicious file's characteristics and behavior to determine whether it is indeed a malware, Heuristic analyzer (or simply, a heuristic), i.e. A technology in virus detection, which cannot be detected by antivirus databases. It allows detecting objects, which are suspected as being infected by unknown or new modification of known viruses. Files which are found by the heuristic analyzer are considered to be probably infected. [7]

In addition, an analyzer usually begins by scanning the code for a suspicious attributes (commands) characteristic of malicious programs. This method is called static analysis. For example, many malicious programs search for executable programs, open the files found and modify them.

A heuristic examines an application's code and increases its "auspiciousness counter" for that application if it encounters a suspicious command. If the value of the counter after examining the entire code of the application exceeds a predefined threshold, the object is considered to be probably infected.

Moreover ,We connect these processes and initial database in a cloud computing environment to be lighter weight , speed processing and performance; we used a file transfer protocol (FTP) For this purpose to connect between the database system in the cloud and the internal processes of the detection system .

We used the technique of detected in real time (RTP) to detect any suspicious attack on real time for working, In addition, sending notifications to the user in the end -host if there an attack or suspecting files,

Thus the user is using the action required for Eliminate or fix. When finding cases of suspected, unknown virus Signature automatically added to our database system. Figure 3, shows the simple outline of Mechanisms in the proposed system.

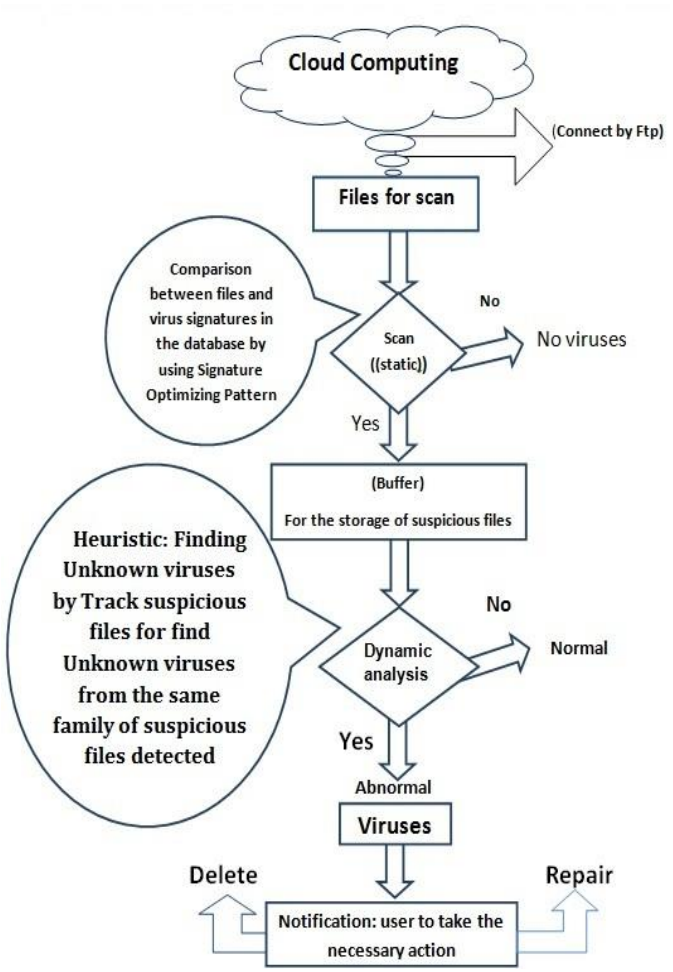


Fig. 3. Simple outline of processes used in proposed architecture

For experimental work, we used hosting services "000webhost.com" as cloud service for uploading the database and execute our system on it, show that in Figure 4

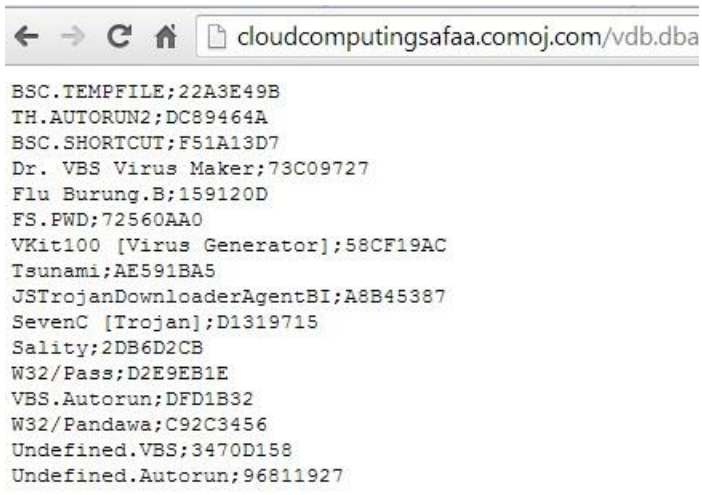


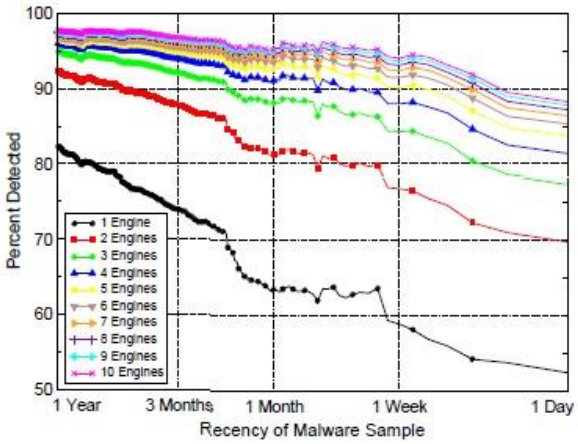
Fig. 4. Snapshot for virus signatures in our cloud environment

Oberheide ET. Al. [9] showed the detection rate with N-Version of Protection in the Network Cloud. N-version protection is closely related to our approach, a paradigm in which multiple implementations of critical software are written by independent parties to increase the reliability of software by reducing the probability of concurrent failures [15]. Traditionally, N-version programming has been applied to systems requiring high availability such as distributed file systems [16]. N-version programming has also been applied to the security realm to detect implementation faults in web services that may be exploited by an attacker [17].

Moreover, A handful of online services has recently been constructed that implement N-version detection techniques. For example, there are online web services for malware submission and analysis [18, 19, and 20]. However, these services are designed for the occasional manual upload of a virus sample, rather than the automated and real-time protection of end hosts, which results in vastly different architectural decisions and performance characteristics.

Engines	3 Months	1 Month	1 Week
1	73.9%	63.1%	59.6%
2	87.7%	81.0%	77.6%
3	92.0%	87.8%	84.8%
4	93.8%	90.9%	88.4%
5	94.8%	92.4%	90.5%
6	95.4%	93.4%	91.8%
7	95.9%	94.0%	92.8%
8	96.2%	94.5%	93.5%
9	96.5%	94.8%	94.0%
10	96.7%	95.0%	94.4%

(a)



(b)

Fig. 5. Detection rate for ten popular antivirus products as a function of the age of the malware samples [9]

In figure 5 (a) demonstrates how the use of multiple heterogeneous engines allows cloud to significantly improve the aggregate detection rate. While in figure 5 (b) shows the detection rate over malware samples ranging from one day old to one year old. The graph shows how using ten engines can increase the detection rate for the entire year-long AML dataset as high as 98%.

The proposed system uses multiple engines detection, and then uploads this process and database to cloud storage by using FTP (File Transfer Protocol), for dealing easily with large files and huge databases of viruses.

Consequently, our results Increase rates in detection rates up to 98% with an increase in the speed of detection and time spent and easy dealing with large files back up.

As Li [6] improved the detection by using signature matching optimization policy for anti-virus programs, we also, combine heuristic and signature matching optimization for detecting known viruses and unknown viruses not for known signature only, this increased detection rate as well as the development of system databases And exploit Time and effort in the detection of signatures in the full library.

IV. CONTRIBUTIONS AND REMARKS

Contributions

Our approach of moving the detection of malicious software into the cloud is aligned with a strong trend toward moving services from end host and monolithic servers into the cloud. For example, in-network email [21, 22] and HTTP [23] filtering systems are already popular and are used to provide an additional layer of security for enterprise networks. In addition, there have been several attempts to provide cloud services as overlay networks [24]; the main relevant contributions of our approach are the following:

The first main contribution of this thesis is the design and development of malware detection system in cloud, a system for providing anti-virus scanning for desktop computers. Cloud – malware detection (CMD) combines a set of pre-existing, third-party scanning services and offload the scanning of les from the host computer to these services. The evaluation of CMD found that performance of the system was highly dependent on the le system activity while the system was active, but that there were specific instances where the system performed well. The findings from this thesis can help to address the performance concerns involved in cloud-based malware scanning. This could result in a system that would be capable of performing nearly transparent anti-malware protection from the cloud.

The second contribution of this thesis was an extension of the desktop version of N-version protection. The system was designed and developed for the Window operating system, and the evaluation of the system showed favorable performance, suggesting that cloud-based anti-malware scanning may be a very good fit for providing a level of security to computer devices.

Finally, our thesis includes a comprehensive examination and summary of the current body of academic research pertaining to cloud-based security for both desktop computers and mobile devices, as well as research regarding low-impact anti-malware techniques which might also be suitable for Detect malware in cloud computing.

A. Why cloud computing applied to C.M.D?

1) Reduction of costs – unlike on-site hosting the price of deploying applications in the cloud can be less due to lower hardware costs from a more effective use of physical resources.

2) Universal access - cloud computing can allow remotely located employees to access applications and work via the Internet.

3) Up-to-date software - a cloud provider will also be able to upgrade software keeping in mind feedback from previous software releases.

4) Choice of applications- This allows flexibility for cloud users to experiment and choose the best option for their needs. Cloud computing also allows a business to use, access and pay only for what they use, with a fast implementation time.

5) Potential to be greener and more economical - the average amount of energy needed for a computational action carried out in the cloud is far less than the median amount for an on-site deployment. This is because different organizations can share the same physical resources securely, leading to more efficient use of the shared resources.

6) Flexibility – cloud computing allows users to switch applications easily and rapidly, using the one that suits them needs best. However, migrating data between applications can be an issue.

The proposed system includes two types of protection built in remote-server protection; make sure that it has a backup system by File Transfer Protocol (FTP); FTP is normally used to transfer files between computers on a network. Cloud FTP enables files to be transferred to Storage Clouds, for transforming data and process to the cloud.

Consequently ,these processes saves latest malware protection in a local cache on your computer then send to cloud server, So that it protects your PC even when you aren't connected to the cloud.

Elsewhere, our system also features a Lightweight, Dispute of the famous anti-virus Products. Figure 6, shows the effect of cloud on size comparison between Cloud malware detection (CMD) and famous antivirus.

Thus, in today's antivirus programs, static analysis is used in combination with dynamic analysis. The idea behind this combined approach is to emulate the execution of an application in a secure virtual environment; the following figure shows the detection rates for viruses of this system and Interface scan.

In view of different detection, methods must be combined to determine whether a file is secure to open, access, or execute. Several variables may impact this process, to be more powerful and more-safe at malware Known and unknown to continuous update of the database of viruses and automatically.

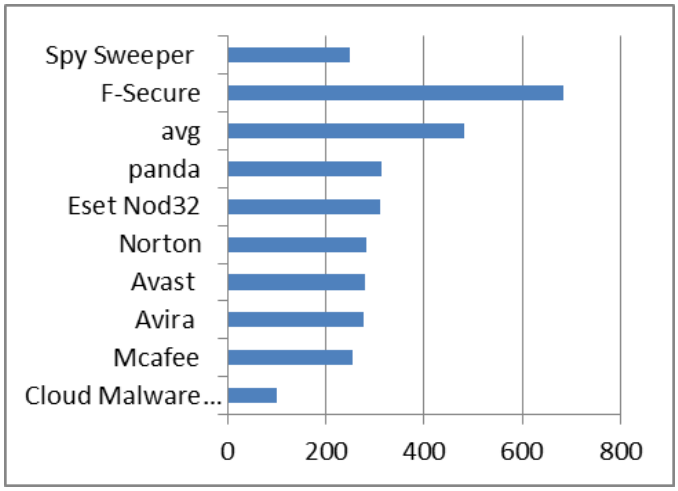


Fig. 6. Size of each installed anti-virus software - has been download updates after installation to being included in the results.

V. Conclusions and Future Work

To conclude, we have proposed a system for combined malware detection systems and cloud computing environments, all running binaries and malware are intercepted by submitting to one or more analysis engines, a complete check against a signature database to detect yet unknown exploits or malware. We will suggest increasing in the dependence of cloud computing as consumers increasingly move to cloud computing platforms for their computing needs. In this paper, we reviewed previous work on malware detection, both conventional and in the presence of storage in order to determine the best approach for detection in the cloud. We also argue the benefits of distributing detection throughout the cloud and present a new approach to coordinate detection across the cloud.

In the proposed system, we have used traditional detection techniques (optimizing pattern) as per static signatures and dynamic detection technology (heuristic). Then, we have chosen for safer system methods as well as speed and modern to rival existing anti-virus.

The proposal of this work is to find the best solutions to the problems of anti-viruses and improve performance and find possible alternatives for a better working environment without problems with high efficiency and flexibility.

We used the optimal traditional methods and modern to detect viruses, for unknown and already detected viruses through the signatures and the Heuristic.

Future work in this field will focus on the development of detection systems based on memory introspection and heuristic or statistical detection, as opposed to signature-based detection.

References

[1] Microsoft, “Microsoft security intelligence report”, [online]:http://www.microsoft.com/technet/security/default.mspx, July December 2006.

[2] Dropbox, Inc., dropbox.com webpage, [Online]: https://www.dropbox.Com/ (accessed 13/04/12).

[3] C. Grace. “Understanding intrusion-detection systems" [J], PC Network Advisor, vol. 122, pp. 11-15, 2000.

[4] S. Subashini, V. Kavitha s.l “A survey of security issues in service delivery models of cloud computing.” .Science Direct, Journal of Network and Computer Applications, pp. (1-11) January (2011).

[5] Shirlei Aparecida de Chaves, Rafael Brundo Uriarte and Carlos Becker Westphall “Toward an Architecture for Monitoring Private Clouds." S.l. IEEE December (2011).

[6] Bo Li, Eul Gyu I'm “ A signature matching optimization policy for anti-virus programs” Electronics and Computer Engineering, Hanyang University, Seoul, Korea. © IEEE 2011

[7] Chen, Z. & Yoon, J. “IT auditing to assure a secure cloud computing”, (2010). [Online]: http://doi.ieeecomputersociety.orgezproxy.umuc.edu/10.1109/SERVICES.2010.118.

[8] J. Oberheide, E. Cooke, and F. Jahanian “CloudAV: N-Version Antivirus in the Network Cloud”, In Proceedings of the 17th USENIX Security Symposium (Security'08). San Jose, CA, 2008.

[9] Jon Oberheide, Evan Cooke and Farnam Jahanian “Cloud N-Version Antivirus in the Network Cloud”,Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 48109 (2007).

[10] Matthias Schmidt, Lars Baumg Artner, Pablo Graubner, David Bock and Bernd Freisleben “Malware Detection and Kernel Rootkit Prevention in Cloud Computing Environments." University of Marburg, Germany (2011).

[11] K. Murad, S. Shirazi, Y. Zikria, and I. Nassar, “Evading Virus Detection Using Code Obfuscation” in Future Generation Information Technology, vol. 6485 of Lecture Notes in Computer Science, pp. 394–401, Springer Berlin , Heidelberg, 2010.

[12] Scott Treadwell, Mian Zhou “A Heuristic Approach for Detection of Obfuscated Malware”, Bank of America, 1201 Main St, Dallas, TX 75202, © IEEE 2009.

[13] Carlin, S., & Curran, K. “Cloud computing security”, International Journal of Ambient Computing and Intelligence.

[14] “Heuristic analysis in Kaspersky Internet Security” [Online]: http://support.kaspersky.com , ID: 8936 , 2013 Mar 01 2013

[15] Algirdas Avizienis, “The n-version approach to fault-tolerant software”, IEEE Transactions on Software Engineering, 1985.

[16] Rodrigo Rodrigues, Miguel Castro, and Barbara Liskov. Base, “using abstraction to improve fault tolerance”, In Proceedings of the eighteenth ACM symposium on Operating systems principles, New York, NY, USA, 2001.

[17] Lajos Nagy, Richard Ford, and William Allen, “N-version programming for the detection of zero-day exploits”, In IEEE Topical Conference on Cybersecurity, Daytona Beach, Florida, USA, 2006.

[18] Carsten Willems and Thorsten Holz. Cwsandbox.[Online]: http://www.cwsandbox.org/, 2007.

[19] Hispasec Sistemas. “Virus total”, [Online]: http://virustotal.com, 2004.

[20] Norman Solutions. Norman sandbox whitepaper. http://download.norman.no/whitepapers/whitepaper_Norman_SandBox.pdf, 2003.

[21] Barracuda Networks. “Barracuda spam firewall”,[Online]: http://www.barracudanetworks.com, 2007.

[22] Cloudmark, “Cloudmark authority anti-virus”,[Online] :http://www.cloudmark.com, 2007.

[23] Alexander Moshchuk, Tanya Bragin, Damien Deville, Steven D. Gribble, and Henry M. Levy, “Spyproxy: Execution-based detection of malicious web content”, In Proceedings of the 16th USENIX Security Symposium, August 2007.

[24] Stelios Sidiroglou, Angelos Stavrou, and Angelos D. Keromytis, “Mediated overlay services (moses): Network security as a composable service”, In Proceedings of the IEEE Sarnoff Symposium, Princeton, NJ, USA, 2007.