Optimal Approaches for Securing Log Files for Forensic Analysis: A Survey

Sweta Singh
United Institute of Technology
Naini, Allahabad

Prashant Shukla
United Institute of Technology
Naini, Allahabad

ABSTRACT
The log file of any association may include sensitive data which must be protected properly for suitable working of that organization. Maintaining security of such log records is one of the important tasks. Also, over a long period of time maintaining authenticity of such log data is very important. However, deploying such a system for security of log records is a big task for any company and also it needs additional cost. There are many techniques have been proposed so far to secure log records. This paper presents a brief survey of optimal approaches for securing log files for forensic analysis. These techniques are reviewed considering its pros and cons.

General Terms
Security, Compression

Keywords
Log files, Forensic, Privacy, Confidentiality.

1. INTRODUCTION
All internet surfing activities performed by each user in an organization are record in the form of log files. Log files be track of all user behavior with various attributes like date, time web site etc. Also, log files are used to resolve different problems, to identify those users who violating the policies or performing malicious activities. It is also useful to identify intruder. Log file is the most desirable target for attacks [9]. The purpose behind this is that attacker would like to leave no mark outs of the activities executed at the time of attack. Hence, the target of attack is normally log files. Apart from this log files also contained information about private transactions performed in any organization. This susceptible data must be protected [6]. Also log file data can be used for unlawful access to the system [9]. From these scenarios, it is clear that security of log files is one of the most important tasks of an organization. The large no. of information is available on Web and database is still rapidly increasing.

For every website, normally thousands of users will be access a site. The administrator of a system has an access to the server log.

When web users interact with a site, data recording their performance is stored in server logs. Log files may contain very useful information characterizing the users' experience in the website.

Since in a normal size site log files amount to several megabytes a day, there is a requirement of technique and tools to help take benefit of their matter. Usually log files are in the form of text files that can range from 1kb to 100mb, which depends on the congestion at a particular website.

In determining the amount of traffic a site receives during a particular period of time, to understand what exactly the log files are counting and tracking.

Computer forensics is one of the most concerns in IT. Computer forensics is similar to the forensic, Where Police acts like the forensics to clean a crime scene for evidence of what happened, to whom it happened, and who did the crime.

In the case of computer, the crime scene is the machine that was hacked, the injured party is the entity to which the computer belongs, and the hacker is the unlawful. The proof in the case of computer forensics is the track left by the hacker; all are recorded in the log files.

For computer forensics to be effective, it must be precise and truthful log files.

Log files are available for securing a network against intrusion. Log files also record network traffic, a note of the IP addresses is access to your network, on which IP port access was attempted, also date and time , whether or not the attempt was successful, etc.

If used to correct log files can be very helpful to maintain network security and integrity. However, for log files provide to determine security login must be activated and the files must be checked time to time. Log files will provide protection against beginner hackers.

More experienced hackers are aware of log files and to hide their actions from the administrator by either deleting the log file altogether or replace the log file with another duplicate log file showing normal network activity.

In the primary instance, the network administrator will know that an intrusion was detected, but will have no clue as to the determine of the attacker or how the attacker entered the system.

In the second instance, the network manager will have no hints, and will have to switch on other techniques for detecting intrusion.

1.1 Structure of Log files
The log files consists of 19 attributes such as Date, Time, Client IP, AuthUser, ServerName, ServerIP, PortServer, MethodRequest, URI-Stem,query , Protocols Status, Time Taken, Bytes Sent, Bytes Received, Protocol Version, Host, User Agent, Cookies. One of the main problems encounter when dealing with the log files is the amount of data needs to be preprocessed.

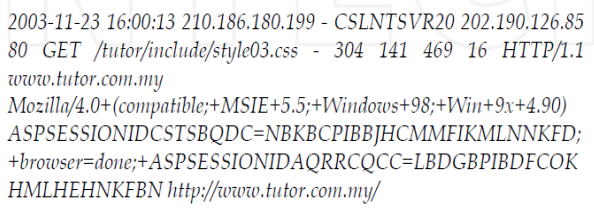

Figure 1. Example of log file

Before we will discuss various aspects of this topic. We start with describing the desirable properties that are needed for securing log files. Various properties are analyzed in this section.

1.2 Various Properties

The various properties to safe log files are provided below:

1) Correctness**:** The log data stored in log files must be correct. It means that the stored data must be similar with new generated data.

2) Tamper Resistance: Only log file administrator can edit the valid log entries[3]. And if any manipulation is done, it should be noticed.

3) Verification: Every log entry must have sufficient information to check its reliability and authenticity to make sure that the log entries are unchanged.

4) Confidentiality: As an attacker can take vulnerable information from log files, therefore log records should be stored in such a way so that they should not be traceable to anyone in the network.

5) Privacy: Log records should not be detected by the attacker during its transfer and its storage [5].

1.3 Threat Model

In this section we discuss the threats available in present situation while to secure log files [4][7]. Different types of attacks that we need to defend against are given below.

1. Integrity of Log Records during transfer: The attacker can obtain illegal entry to communication medium. And he can not only have access to the data but also he can change the data which is move to the log server for safe storage. \
2. Authenticity of Log Record creator: The attacker may impersonate to be valid network consumer and begin to transmit log records from someone else's identity [8].
3. Confidentiality of Log files: The attacker may try to associate the log records over traffic to find the information about perceptive transactions of an organization.

2. LITERATURE REVIEW

In this section we will discuss the various techniques for securing the log files and their analysis. In this we mainly focus on progress in secure logging technique.

The BSD Syslog Protocol [2] and Reliable Syslog[3] proposed a technique which is used as de facto standard for logging protocols. This technique uses UDP Protocol for release of log records to log server. This technique is the first benchmark to define for classification protocols used in every technique. The working of this is faster due to use of UDP protocol. As UDP protocol is used there is no guarantee of consistent release of the message to log server. There is no guard for log records throughout communication and when log records are present at both client or server side. Syslog is addition to same in which TCP protocol is used for consistent delivery of log records. Bellare [1] was pretend for maintained security by secret key. This secret key is used as initial point to generate hash-chain. The hash-chain is generated strong single-way function. Secret key is generated for every entry of log records. This practice was proposed to keep pre cooperation log data from post addition, removal, modification and recording. Online trustworthy server is requisite in this technique to

Maintain secret keys. If this reliable server is attacked or compromised then log record security is broken.

U. Flegel [4] proposed a technique, in this technique the log records are first analysed by pseudonymizer before they are sent to server. Pseudonymizer filter the specific area of log records and substitutes them by some pseudonymized cost. Every record is individually to provide more safety measures to log data. This technique substitutes some specific log fields by pseudonymized value to create the log data undetectable. The better security provided as compared to previous techniques. This method is not confirmed for correctness of the log records. Once log records are substituted by some values they cannot be retrieved back. This technique does not protect against the attack in which attacker is constantly monitor the data over network. D.Ma [6] suggested an approach, where other than trusted server and user reasonably confidential verifier is use. The moderators verify and audit the data at client and then it check the hash chain of some verification from server. At the end if both outcome are equal then only record is measured as accurate and unchanged so that it can be stored at log server. This technique improves security for log records due to use of moderately trusted verifier. If the confidential server generates hash chains is attack, the log security is gone. If intruder is attacked or cooperation then attacker can alter the log record entry easily.

J.Kelsey [7] presented a technique, This technique is another extension of syslog. At the last of every log record this method adds two blocks-signature block and certificate chunk. The signature block have the digital signature designed based on the aspects of the log files. So this signature is unique for each log record. The certify block contains the record from trusted entity for each log record. This technique provides additional features like authentication, message integrity, and replay attack resistance. In addition to that it has capability to order message and describe the lost messages. If the signature block is misplaced after verification of the record the any alteration made to log records cannot be detected. Syslog-sign is not provide confidentiality to log records during communication and also at end points. Indrajit Ray [9] proposed a technique. This technique uses the client for processing the log files from some log generators. Then this data is sent to the server. The server is accountable for safe storage space and protection service to arriving log data. The log checks to another important aspect which arise on logging server data. Upload of each record, upload mark is attached to it, which is further used for retrieval of data.

This technique is new and the generally safe way for securing log files. The aim to maintain reliability and confidentiality is assigned to cloud architecture of the organization.

2.1 OTHER TECHNIQUES

2.1.1 Make Log Files Append Only

Set up log files to write any information to the end of the file only. Not to permit log files again to write or delete information already in the log files. Log files are the only facts of the hacker's events. If a hacker can break into the organization, this protective measure will not permit the hacker to alter the log files.

2.1.2 Set Permissions

Permissions should be configured to reject access to any user who is not the system administrator.

Allow read access to log files will explain a hacker characteristic traffic pattern on the system and which ports are open, exposing possible vulnerabilities in the system. Set the permissions low down, hacker will be able to more easily take-off network traffic, thus thrashing all trace of his presence.

2.1.3 Password Protect Log Files

Set a password on the log files or the directory which contains the log files. If a hacker gets into the system, he will then have to split into the log files.

This method wishes to be accompanying by other techniques to be successful. If a hacker breaks into the system as the root user, password protection will be of little use.

2.1.4 Create Duplicate Log Files

To create the log files written in more than single place. Even if set of log files is compromise, the other place will authorize you to trail the intrusion.

This technique for defending log files should be check the traffic with using a individual logging server.

2.1.5 Hide Log Files

Do not put your log files in a clear spot. Hackers know the location easily of the log files. To keep away from log files, put them in an unexpected mark.

This is a new way of technique that can be used in conjunction with multiple techniques**.**

2.1.6 Use Separate Logging Server

Using a separate logging server can be a very effective way of securing log file. The log files stored on different servers such as DNS, web, or file servers will be copied to this central log server.

This technique adds a significant layer of security to your log files. An attacker would have to enter your server and then enter the security on the log server.

2.1.7 Save Log Files to Write-Once Media

Saving log files to write-once media, such as Compact Disks to protect log files from removal or alteration.

This physical security protects the log files from tamper because an attacker would have to actually impair the CD for logs to be lost.

2.1.8 Encrypted Log Files

Encryption the log files will make them tough for the hacker to read. A hacker will not be able to make significant changes in the log files without the encryption key. It will be easier to delete the log files than to change them.

3. CONCLUSION

In this paper we have discussed several existing state of art techniques for securing log files. For each technique we have provided detailed contributions along with its benefits and limitations. Because of this analysis, a number of future scopes of research are highlighted in this topic.

4. REFERENCES

[1] M. Bellare and B. S. Yee, ―Forward integrity for secure audit logs,‖ Dept.Comput. Sci., Univ. California, San Diego, Tech. Rep., Nov. 1997

[2] C. Lonvick, The BSD Syslog Protocol, Request for Comment RFC 3164,Internet Engineering Task Force, Network Working Group, Aug. 2001.

[3] D. New and M. Rose, Reliable Delivery for Syslog, Request for Comment RFC 3195, Internet Engineering Task Force, Network Working Group, Nov. 2001.

[4] U. Flegel, ―Pseudonymizing unix log file,‖ in Proc. Int. Conf. Infrastruture Security, LNCS 2437. Oct. 2002, pp. 162–179.

[5] J. E. Holt, ―Logcrypt: Forward security and public verification for secure audit logs,‖ in Proc. 4th Australasian Inform. Security Workshop, 2006, pp. 203–211.

[6] D. Ma and G. Tsudik, ―A new approach to secure logging,‖ ACM Trans. Storage, vol. 5, no. 1, pp. 2:1–2:21, Mar. 2009.

[7] J. Kelsey, J. Callas, and A. Clemm, Signed Syslog Messages, Request for Comment RFC 5848, Internet Engineering Task Force, Network Working Group, May 2010.

[8] BalaBit IT Security (2011, Sep.). Syslog-ng— Multiplatform Syslog Server and Logging Daemon [Online]. Available: http://www.balabit.com/network-security/syslog-ng

[9] Indrajit Ray,K.Belyaev,‖Secure Logging As A Service-Delegating log management to the cloud ‖, IEEE Systems Journal,June 2013