Information protection when working with log files of the operating system

K Yu Zhigalov¹, M N Orobey² and R I Aliev³

¹ Institute of Control Sciences of Russian Academy of Sciences
² Irkutsk National Research Technical University, 83, Lermontov St., Irkutsk, 664074, Russia
³ Chechen State University, 17a, Dudaev Boulevard, Grozny, 364015, Russia

E-mail: max_1376@mail.ru

Abstract. There are many definitions of the concept of event log files (log-file). In general, a log file means a file containing system information about the functioning of a particular site, server, and operating system. It is a computerized log file of records of events that have occurred over a period of time. Display of all actions initiated by the user in the operating system on the server or site is carried out in the logs. Generally, one record in the log file corresponds to one event. The record is made immediately after the start of the action and is supplemented after its completion. The given records are registered in the log either by the program, in which the user works, or they are sent by another special program, which main purpose is to compile information about the events that have occurred.

1. Introduction

There are many definitions of the concept of event log files (log-file). In general, a log file means a file containing system information about the functioning of a particular site, server, and operating system. It is a computerized log file of records of events that have occurred over a period of time. Display of all actions initiated by the user in the operating system on the server or site is carried out in the logs [1, 2].

Generally, one record in the log file corresponds to one event. The record is made immediately after the start of the action and is supplemented after its completion. The given records are registered in the log either by the program, in which the user works, or they are sent by another special program, which main purpose is to compile information about the events that have occurred [3-5].

The content of the records added to the log can be various. The format of the event record is determined either by the program or by the author working in it. The log file format is established by two types:

• machine-oriented;
• adapted for human reading.

The primary advantage of logs is their focus on security and incident investigation. In such cases, it is required to separate the logs from the very system in which they have functioned. The bottom line is that the criminal is able to destroy the traces of the interaction with the system when penetrating it. In other words, the criminal’s task will be to change the records in logs. “External” logging will not provide an opportunity to do this [6, 7].
2. Results and Discussion

Every action in the information system does not remain unnoticed. On the contrary, each action is recorded in the log directly or indirectly, and sometimes the record is made simultaneously in several logs. The logs can be located anywhere in the computer system, and only a specialist is well positioned to know it.

In order to obtain any information by examining the logs, one should:

- determine how computers and their programs interact;
- define a mechanism for logging events;
- get log files for the specified period of time;
- examine the content of logs in various ways.

Therefore, the study of log file records can play a vital role in the investigation of computer incidents. For instance, web browsing involves the following systems, which may contain information about a given event:

- internet browser;
- user firewall;
- antivirus software;
- operating system;
- DNS server before requesting the user’s browser webpage;
- all routers applied by the computer’s user to access the web page or DNS server;
- billing systems, where routers send their statistics;
- means of protection located in front of the DNS server or web server (antivirus, firewall, intrusion detection system);
- web server;
- CGI scripts run by the web server;
- web servers of third-party applications that the user views on the web page (advertising banners and counters, the operation of which is ensured by third-party providers);
- proxy server;
- equipment taken by the user to access the Internet (ATS, Wi-Fi, GPRS);
- equipment of the System for operative investigative activities both from the user and from the web server side.

Thus, a reasonably simple and particular action creates interrelated records in two to three dozen systems.

3. Log file classification

There is no classification of log files in the scientific literature. The reason is the fact that the logs, regardless the information they contain, do not differ from each other. Work with system logs is carried out in the same way as with mail server logs. It is required to distinguish log files by the external source of information, and not by their internal structure and characteristics. The following logs are highlighted:

- web servers;
- mail servers;
- operating systems (Windows, UNIX, Linux, IOS).

4. Web server logs

The indicated type of logs is one of the basic sources of information about user operations on the site. What information do these logs contain? The dataset is designated by the web server type and the set parameters. They often include information about:

- visitor’s IP address;
- request time, including the time zone;
- customer HTTP request field;
web server response code (status code);
web server response size (excluding HTTP header);
errors when accessing web pages;
errors when starting CGI programs.

The issue of falsifying logs without having access to the web server is relevant when examining web server logs. In other words, can the log files be trusted?

If a criminal does not have access to the web server, then he can only change the HTTP request field, which is entirely formed on the side of the visitor.

Besides, it is permissible to trust the IP address information. It should be borne in mind that the fixed IP may belong to a sox server or proxy server, or another intermediary.

Other fields belong to the web server and cannot be changed by a criminal.

5. Examining mail server logs and email headers

The user, creating messages, applies a special program called an email client. Then it is addressed to the sender to the mail server, and further to the recipient’s email server. Once received, the message is placed in the customer’s mailbox. In order to pick up a message, the user should take advantage of an email client program.

The message is generally stored in both the sender’s client and the recipient’s client. A copy of the message is not saved on the server, but a record in the log file is registered (about receiving and sending). Moreover, the message has a special “Received” routing header.

The sender and the recipient can use the web interface of the mail server instead of email customer programs, their functions will be the same. The sender and the recipient will communicate with web interface through a browser.

We identify the following traces in the process of transferring messages from the sender to the recipient:

- copy of the message in the sender’s computer;
- record in the logs belonging to the email server;
- copy of the message with added headers in the recipient’s computer.

The main tracks have been listed above. Besides, other traces confirming the fact of transmission and reception of a message are singled out:

- additional traces in the sender’s computer (in antivirus, browser, etc.);
- traces in the provider logs, through which the connection is made between the customer and the sender’s email server;
- traces in the logs of e-mail server programs (antivirus, anti-spam) in order to transfer the message;
- supplementary traces in a computer owned by the recipient.

If the web interface of the mail server is used instead of the email customer programs, then other traces appearing through web browsing will be added to the above mentioned.

Recently, most mail servers have been using the technology of automatic sign of messages sent by means of a DKIM digital signature. The signature is intended to protect information about the sender from the spammer falsification. In this case, not only the “From” field is signed, but also other fields, and frequently the text of the message. It is significant that the given digital signature can be brought into practice as evidence in the case. It confirms the letter is from the server that has signed it at the appropriate time. If one carry out the identification of users on the server, then it is possible to find out the sender identity. Moreover, the point of the letter’s detection and its extraction way do not matter for establishing the above information. Compliance with the law is a priority.
6. System logs
The security model of the information system includes three constituents:

• 1. authentication;
• 2. authorization;
• 3. audit.

7. Conclusion
Log files are the content of the audit constituent. The registration of all events related to the system security is the audit essence. It is worth noting that logging will not prevent a criminal from gaining access to the system, however, it will allow him to be identified, exposed and found more likely. Moreover, logging helps to reveal system vulnerabilities. The amount of information, when investigating a computer crime, depends on the degree of the audit completeness. An expert is capable of finding out a lot of required information using the recorded data.

References
[1] Kondratiev V V, Karlina A I, Guseva E A, Konstantinova M V and Gorovoy V O 2018 Structure of Enriched Ultradisperse Wastes of Silicon Production and Concretes Modified by them IOP Conference Series: Materials Science and Engineering 463(4) 042064
[2] Sysoev I A, Kondrat’ev V V, Zimina T I and Karlina A I 2018 Simulation of the Energy States of Electrolyzers with Roasted Anodes at Elevated Currents Metallurgist 61(11–12) 943–949
[3] Gura D A, Gura A Y, Volkova T A and, Lipilin D A 2020 Recreational space as a factor of sustainable development of coastal geosystems of the Krasnodar region IOP Conference Series: Earth and Environmental Science 579(1) 012007 doi:10.1088/1755-1315/579/1/012007
[4] Buryanina N, Korolyuk Y, Koryakina M, Lesnykh E and Suslov K 2019 Algorithm of Current Protection Based on Three Instantaneous-Value Samples Proceedings of IEEE PES Innovative Smart Grid Technologies Europe 890542
[5] Suslov K V, Solonina N N and Smirnov A S 2011 Smart Grid: A new way of receiving primary information on electric power system state IEEE PES Innovative Smart Grid Technologies Conference Europe 6162654
[6] Sivtsov A V, Yolkin K S, Kashlev I Mб Karlina A I 2020 Processes in the Charge and Hearth Zones of Furnace Working Spaces and Problems in Controlling the Batch Dosing Mode during the Smelting of Industrial Silicon and High-Silicon Ferroalloys. Metallurgist 64(5-6) 396–403
[7] Ershov V A, Karlina A I, Karlina Y I 2021 Process Control System for Producing a Silica-Based Modifying Additive. Metallurgist 65(3-4) 446–453