ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/324227692

Correlation Analysis Of Forensic Metadata For Digital Evidence

Article *in* International Journal of Computer Science and Information Security, · April 2018

CITATION
1

READS
957

3 authors, including:



Yudi Prayudi
Universitas Islam Indonesia
120 PUBLICATIONS **415** CITATIONS

SEE PROFILE



Bambang Sugiantoro
Universitas Islam Negeri Sunan Kalijaga Yogyakarta
78 PUBLICATIONS **59** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Thesis View project

Project Riset Forensika Digital View project

All content following this page was uploaded by Yudi Prayudi on 05 April 2018.

The user has requested enhancement of the downloaded file.

Corelation Analysis Of Forensic Metadata For Digital Evidence

Zaenudin
Departement of Informatics
Universitas Islam Indonesia,
Yogyakarta, Indonesia
STMIK Mataram,
Indonesia
15917124@students.uii.ac.id

Bambang Sugiantoro
Departement of Information System
UIN Sunan Kalijaga
Yogyakarta, Indonesia
bambang.sugiantoro@uin-suka.ac.id

Yudi Prayudi
Departement of Informatics
Universitas Islam Indonesia
Yogyakarta, Indonesia
prayudi@uii.ac.id

***Abstract*— Metadata is the information that is embedded in a file whose contents are the explanation of the file. In the handling of the main evidence with a metadata-based approach is still a lot of manually in search for correlation related files to uncover various cases of computer crime. However, when correlated files are in separate locations (folders) and the number of files will certainly be a formidable challenge for forensic investigators in analyzing the evidence. In this study, we will build a prototype analysis using a metadata-based approach to analyze the correlation of the main proof file with the associated file or deemed relevant in the context of the investigation automatically based on the metadata parameters of Author, Size, File Type and Date. In this research, the related analysis read the characteristics of metadata file that is file type Jpg, Docx, Pdf, Mp3 and Mp4 and analysis of digital evidence correlation by using specified parameters, so it can multiply the findings of evidence and facilitate analysis of digital evidence. In this research, the result of correlation analysis of digital evidence found that using parameter of Author, Size, File Type and Date found less correlated file while using parameter without Size and File Type found more correlated file because of various extension and file size.**

Keywords: ***Metadata, Forensic, Correlation, Digital, Evidence***

I. Introduction

As the heterogeneity of digital evidence in investigation continues to evolve with technological advances, we are faced with newer digital devices, more artifacts and a variety of file formats, these developments bring benefits, while at the same time providing new opportunities for crime in information technology [1]. In many cases, there is a digital evidence that can assist the officer in uncovering a criminal case. One of them through information about the contents of a data or file called file metadata.

Metadata is information that is embedded in a file in the form of annotation of the file. Metadata contains information about the contents of a data used for the purpose of file or data management that later in a database [2]. Metadata is often called "information about information" or "data about data" [2].

So far, investigators of forensic analysis in the handling of major evidence with a metadata-based approach are still manually in search of correlation of related files. However, when correlated files are in separate locations (folders) and the number of files will certainly be a formidable challenge for forensic investigators in analyzing such digital evidence [1].

Metadata-based researches have been conducted, among others, by [3] linking data with other information, the user accessing it, the file directory where it was stored, the last time it was copied, and so on. Subsequent research Conducting analysis to verify metadata associated with images and track using GPS features [7].

To facilitate the process of correlation analysis, In his research build an AssocGEN analysis system using metadata to determine the association between user file artifacts, logs, and disposal of network packets and identify metadata to classify and determine correlations between artifacts and related artifact groups [5]. Forensic metadata has been done by previous research but by building different tools and parameters. Research with metadata-based forensics has been done by [4]. In his research, a forensic metadata system is used to read metadata characteristics in general and look for metadata correlation files with one parameter: file owner, file size, file date and file type. According to [5]. By using forensic metadata tools will greatly facilitate investigators in analyzing the correlation of digital evidence.

So in this study will build a prototype to understand and read the characteristics of metadata in general and detail the specific metadata and identify, analyze the metadata correlation to group related files or relationships that are considered relevant in the context of investigation automatically based on metadata parameters that is Author, Size , File Type and Date. By using some and all parameters that have been determined, so as to multiply the findings of evidence and facilitate analysis of digital evidence. With this research is expected to contribute to forensic analysts in analyzing the correlation of digital evidence with a metadata-based approach.

II. Literature Review

Several previously conducted studies related to forensic metadata serve as a reference in the writing of this research, among others;

In his research build an AssocGEN analysis system using metadata to determine the association between user file artifacts, logs, and disposal of network packets and identify metadata to group and specify correlations between artifacts and related artifact groups [5].

Other studies use various formats and metadata types to validate different types of documents and files that have a

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

number of formats and metadata types, which can be used to find properties of a file, document or activity of a network. In addition, metadata is widely used in any condition, where metadata can provide a variety of evidence between a group of people, as some do not know the type of information stored in their documents [6].

In his research aims to forensic examination of metadata that is linking data with other information, users who access it, file directory where the storage, last copied, and so forth. In a case, Metadata can produce indirect evidence to support evidence [3]. Next research Perform analysis to verify metadata associated with images and track using GPS features based on GPS Height, Latitude GPS, GPS Longitude and GPS position using Geo tagging feature) [7].

Subsequent research analyzed the BitCurator project to develop an extensible strategy for converting and combining digital forensic metadata into the archive metadata scheme and focusing on metadata generated by the open-source Digital Forensic (DFXML) tool [8]. Related research creates a metadata application for reading file metadata in general and can find files based on file correlation with one of the parameters of the file metadata [4].

From the above literature studies, in this study, will build a prototype for understanding and reading metadata characteristics in general and specific metadata detail and identifying, analyzing metadata correlations for grouping related files or relationships deemed relevant in the context of investigation automatically based on metadata parameters ie Author, Size, File Type and Date. By using some and all parameters that have been determined, so as to multiply the findings of evidence and facilitate analysis of digital evidence. With this research is expected to contribute to forensic analysts in analyzing the correlation of digital evidence with a metadata-based approach.

III. BASIC THEORY

A. Tools

The tools used to build forensic metadata are netbeans. Netbeans is a Java-based Integrated Development Environment (IDE) application from Sun Microsystems that runs on swing. Swing is a Java technology for desktop application development that can run on various platforms such as windows, linux, Mac OS X and Solaris. An IDE is a programming scope that is integrated into a software application that provides a Graphic User Interface (GUI), an editor or text code, a compiler and a debugger [9]

B. Classification of Digital Evidence

In the investigation of the evidence is very important for the sustainability of the case being investigated, because with the evidence that will be analyzed to reveal the motives and perpetrators of the crime. Investigators are expected to understand the types of evidence so that at the time of investigation they recognize the priority of priority evidence. There are several similar terms, namely electronic evidence, digital evidence and evidence findings.

Electronic evidence is physical and visually recognizable (computer, hand phone, camera, CD, hard drive, Tablet, CCTV etc.). While digital evidence is evidence that is extracted or recovered from electronic evidence (file, email, sms, image, video, logs, text). Digital Proof of Evidence is a proof taken from electronic evidence conducted analysis of the evidence, type of digital evidence, among others, Email / Email Address, Web History / Cookies, Image File, logical file, Deleted File, Lost File, Slack files, File Logs, Encrypted Files, Steganography files, Office files, Audio Files, Video Files, User ID and Password, Short Message Service (SMS), Multimedia Message Service (MMS), Call Logs.

Findings of evidence is a digital evidence more meaningful as the output analysis obtained by investigators who directly lead to the reconstruction of the case being faced. In this case, digital evidence is information directly related to the data required by the investigator in the investigation process [10].

C. Metadata Concepts

Metadata can be interpreted as "data (spatial) data", containing information about data characteristics and plays an important role in data exchange mechanisms. Through metadata information expected data users can interpret the data in the same way, when users see directly spatial data. The metadata document contains information that describes the characteristics of the data, especially the content, quality, condition, and manner of obtaining it. Metadata is used to perform pertinent spatial data documentation about who, what, when, where, and how spatial data is prepared.

There are several types of metadata files such as Descriptive Metadata is Data that can identify the source of information so that it can be used to facilitate the process of discovery and selection. Coverage included in this data is the author, title, year of publication, subject or keyword headers and other information that the process of filling is the same as the traditional catalog. Administrative Metadata is Data that can not only identify the source of information but also how it is managed. The scope of this data is the same as the descriptive data only with the data maker, the time of manufacture, the file type, other technical data. In addition, this data also contains information about access rights, intellectual property rights, storage and preservation of information resources. Structural Metadata is Data that can make between the related data can be related to each other. More explicitly, this metadata is used to determine the relationship between physical files and pages, pages and chapters and chapters with books as the final product [11].

D. Test Flow Metadata Forensic Systems

In forensic metadata research for the analysis of evidence, correlation includes several stages of testing is the testing phase to read the characteristics of metadata and testing to perform metadata correlation.

a) Metadata File Characteristic Reading Flow

Here is described in detail the steps of use of this application in viewing the characteristics of the metadata file in Figure 2 flowchart below:

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

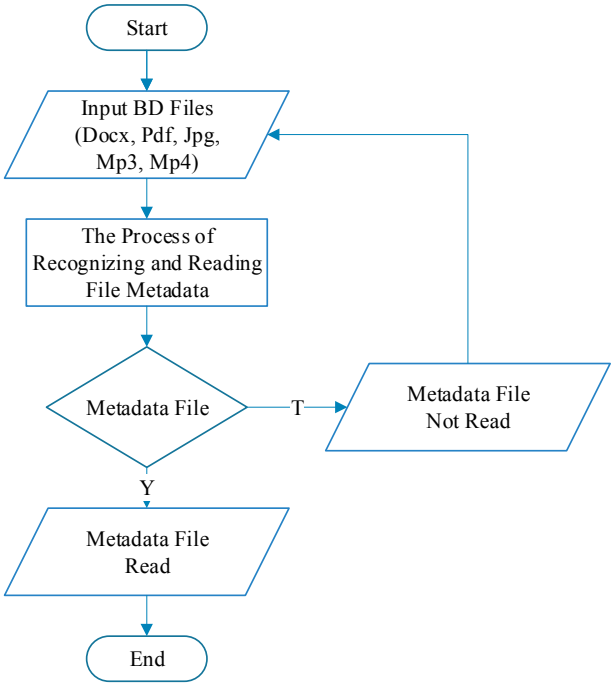


Figure 2. Flowchart Reading Characteristics of Metadata File

Explanation of the testing process to read the characteristics of metadata file using forensic metadata system that is built first to start or forensic metadata system is run, then input file digital evidence that will read metadata, the process of multiplying and reading the metadata file, there are conditions where the metadata file cannot read will return to the input file object evidence, then metadata that can be read metadata will be directly displayed metadata last program in closing or finished.

b) Metadata File Correlation Testing Flow

Here is described in detail the steps of the use of this application program to perform the correlation of the file in figure 3. flowchart below:

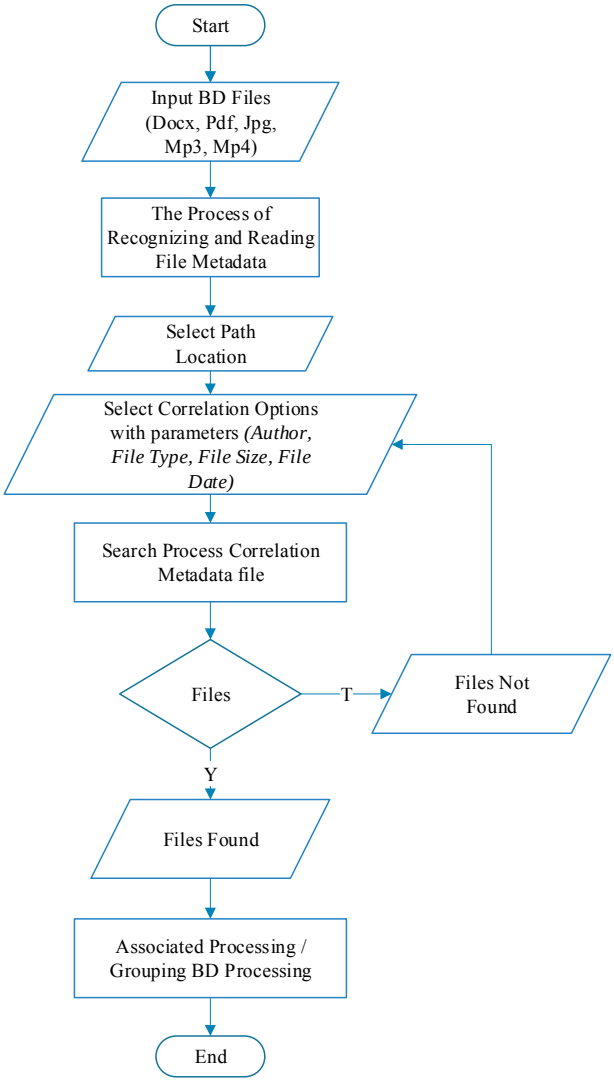


Figure 3. Flowchart Process Testing System / Tools Correlation Metadata file

First start the forensic metadata system, then do input the main evidence file to read metadata, then the process of understanding and reading the metadata file, then select the location of the correlation path and then select the correlation option with parameters, than the system will find the metadata correlation based on parameter selection, if the file is not found it will return to the correlation option but the correlation file found then will proceed to the analysis process and the last system is completed.

IV. RESEARCH METHODS

The method used in forensic metadata research for this correlation analysis of digital evidence can be seen in Figure 1 below:

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

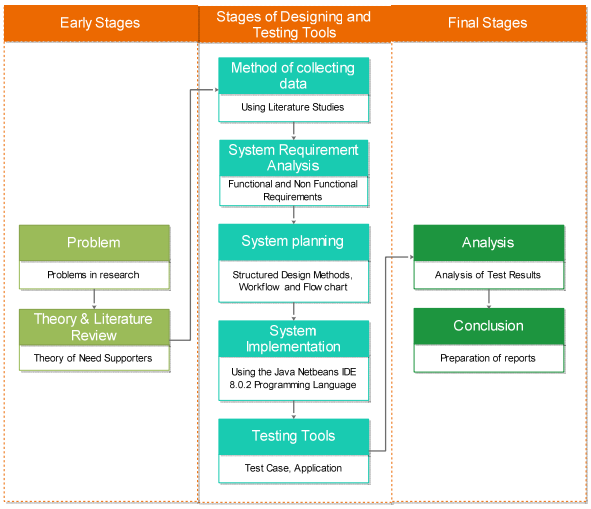


Figure 1. The Proposed Methodology

research methodology that will be built outline is divided into three stages, namely the first stage consists of problem identification and literature review, second stage or stage design and testing tools consist of data collection methods, system requirements analysis, system design, system implementation and testing tools, analysis of test results and the final stage of the completion stage of the conclusion contains the preparation of research reports.

V. ANALYSIS AND RESULT

In this study, the prototype has been built from the implementation until the results of analysis and discussion. Test prototypes built with some predefined files and to analyze the metadata correlation with specified parameters.

A. Results Read File Characteristics File Method

The main evidence file that will read metadata first in browse after the program will process until identified metadata then will appear metadata in general table, checksum and detail as in table 1 below:

Table 1. The result of reading metadata image file TTD.jpg

No	Kind of Metadata	Value
1	Location file	E:\Bahan-Bahan\TTD.jpg
2	Name File	TTD.jpg
3	Type File	Jpg
4	Author	Zen Alkarami
5	Computer	DESKTOP-HJQGNJT
6	Owner	46 DESKTOP-HJQGNJT\Zen

B. Results of File Metadata Correlation Analysis

The result of correlation analysis of metadata file based on parameter ie; Author, Size, File Type and Date. By testing files with extension Jpg, Docx, Pdf, Mp3, and Mp4. In one folder As follows:

a) Correlation Results with Author, Size, File Type and Date Parameters

The result of metadata analysis of correlated file is TTD.jpg file which metadata Author "Zen Alkarami", File Size "327946 byte", file type "Jpg" and with date in file TTD.jpg i.e. "January 24, 2018", conducted file- files are located in the materials folder with the option "equals", then found 2 files that its Author "Zen Alkarami", File size "327946 bytes", Extension file "Jpg" and the date is the same as "January 24, 2018" from metadata the date of the existing TTD.jpg file in that location. The following can be seen in the implementation view in Figure 4 and the results of the analysis from table 4 below:

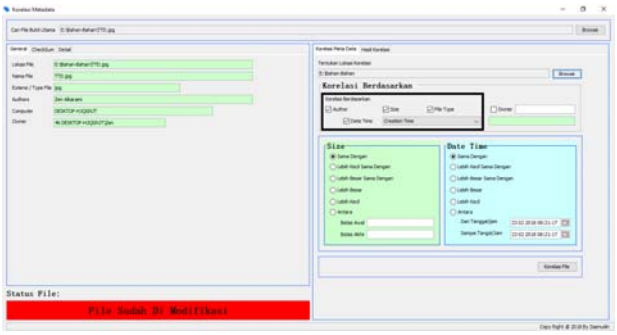


Figure 4. Display of Correlation Implementation with Author, Size, File Type and Date Parameters

Table 4. Correlation Results Based on Author, Size, File Type and Date Parameters

Nama File	Size	Date Creation	Date Modification	Path
gambar.jpg	327946	2018-01-24 04:13:54	2018-01-25 10:51:09	E:\Bahan-Bahan\gambar.jpg
TTD.jpg	327946	2018-01-24 04:13:52	2018-01-24 04:13:54	E:\Bahan-Bahan\TTD.jpg

b) Correlation Results Without Parameters Size and File Type

Results Correlation Analysis Without Parameters Size and File Type in question is to search for various types of files and sizes so obtained correlation results that vary or more with the evidence file TTD.jpg. Then got 6 file result of analysis which metadata Author its "Zen Alkarami", date "24-Januari-2018" with file type in the form of "Mp3, Pdf, Jpg and Docx" and file size different Here can be seen view implementation at Figure 5 and the results in table 5 below:

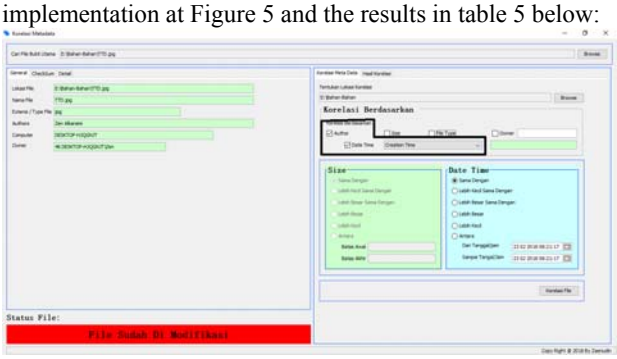


Figure 5. Show Correlation Implementation Without Parameter Size and File Type

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table 5. Results Correlation Without Parameters Size and File Type

Nama File	Size	Date Creation	Date Modification	Path
audio.mp3	327946	2018-01-24 04:13:54	2018-01-25 07:03:23	E:\Bahan-Bahan\audio.mp3
Daftar TTD.pdf	6507	2018-01-24 04:17:18	2018-01-24 04:17:17	E:\Bahan-Bahan\Daftar TTD.pdf
format.pdf	327946	2018-01-24 04:13:54	2018-01-25 07:03:23	E:\Bahan-Bahan\format.pdf
Gambar.jpg	327946	2018-01-24 04:13:54	2018-01-25 10:51:09	E:\Bahan-Bahan\Gambar.jpg
Surat Pernyataan.docx	12490	2018-01-24 04:17:00	2018-01-24 04:16:59	E:\Bahan-Bahan\Surat Pernyataan.docx
TTD.jpg	327946	2018-01-24 04:13:52	2018-01-24 04:13:54	E:\Bahan-Bahan\TTD.jpg

VI. CONCLUSION

Based on the results obtained in the discussion, the forensic metadata research for the correlation analysis of digital evidence can be deduced as follows. Built-in forensic metadata can read all file types specifically on the computer both in general and in detail including the tested file as sample. Based on the test to read the characteristics of metadata can be understood in general that is divided into three main parts; General Metadata ie File location, File name, File type / Extension file, Outhors Owner and Computer. Metadata Checksum is MD5 and SHA-256 Value. Metadata detail is cration time, last access time, last modified time, directory, other, regular file symbolic link, size, Make, Model, Orientation, X Resolution, Y Resolution, Resolution Unit, Software, Date / Time, Positioning, Exposure Time, F-Number, Exposure Program and so on. The method used to find metadata and metadata correlation characteristics is by forensic metadata tools. Tools used are the work of the researchers themselves. Based on the test of metadata correlation analysis with parameter of Author, Size, File Type, and Date then found fewer file compare to without parameter size and file type hence found file with various extension and file size.

VII. FUTURE WORK

The suggestions that need to be developed for further research are as follows. In the next research need to be done correlation analysis not only with parameter of metadata. Further development and research needs to be added multi local or multi drive option to browse the main evidence file.

REFERENCES

[1] S. Raghavan and S. V. Raghavan, 2014. “AssocGEN: Engine for analyzing metadata based associations in digital evidence,” Int. Work. Syst. Approaches Digit. Forensics Eng., SADFE,

[2] J.Riley, 2017 Understanding Metadata: What Is Metadata, and What is it for?.

[3] A. Spore, 2016.“Report Information from ProQuest,” no. June,

[4] Subli, Sugiantoro & Prayudi, 2017. “ Forensic Metadata to support the investigation process of the "scientific journal DASI

[5] S. Raghavan and S. V Raghavan, 2013. “A study of forensic & analysis tools,” 2013 8th Int. Work. Syst. Approaches to Digit. Forensics Eng., pp. 1–5,

[6] F. Alanazi and A. Jones, “The Value of Metadata in Digital Forensics,” Proc. - 2015 Eur. Intell. Secur. Informatics Conf. EISIC 2015, vol. 8, no. 2011, p. 182,

[7] P. R. Kumar, C. Srikanth, and K. L. Sailaja, 2016. “Location Identification of the Individual based on Image Metadata,” Procedia Comput. Sci., vol. 85, no. Cms, pp. 451–454, 2016.

[8] L. Drive, M. Hall, C. Hill, K. Woods, A. Chassanoff, and C. a Lee, 2013. “Managing and Transforming Digital Forensics Metadata for Digital Collections,” 10th Int. Conf. Preserv. Digit. Objects, no. November, pp. 203–208,

[9] R. Sharma and S. Koshy, 2011. “Promoting Open Source Technology in Education : NetBeans : The Perfect Open Source IDE,” vol. 4333, pp. 571–575,

[10] Y. Prayudi, 2014 “Problema Dan Solusi Digital Chain Of Custody Dalam Proses Investigasi,”April,

[11] U. Salama, V. Varadharajan, M. Hitchens, and DUMMY, 2012. “Metadata Based Forensic Analysis of Digital Information in the Web,” Annu. Symp. Inf. Assur. Secur. Knowl. Manag., pp. 9–15,

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

View publication stats