Header investigation for spam email forensics using framework of national institute of standards and technology

Imam Riadi a,1, *; Rusydi Umar a,2; Mustafa a,3

a Universitas Ahmad Dahlan, Jl. Kapas No.9, Semaki, Kec. Umbulharjo, Kota Yogyakarta, Daerah Istimewa Yogyakarta 55166
1 imam.riadi@is.uad.ac.id; 2 russydi.umar@mti.uad.ac.id; 3 mustafa.ramlannor@gmail.com
* Corresponding author

Abstract
Today's technology makes communication very easy and can be used anywhere, even a distance of hundreds to thousands of kilometres is not a barrier in communicating. One of the tools or media that is widely used is via email. However, there are many disadvantages that may be obtained from the email, one of which is spamming or mail. The purpose of this research is to know the stages of spamming email analysis through header analysis. The method used in this study is the National Institute of Standards and Technology (NIST) and this method uses 4 stages, namely collection, examination, analysis, and reporting. The results of this study are expected to be able to find out the spam sender's email address, the spam email sender's IP address, and other information needed.

Keywords: Digital Forensics; Forensic e-mail; Forensic Tools; NIST

Introduction
Nowadays internet services make it easy for humans to do all their activities anywhere and anytime. This convenience is fortified by the reach of the internet that goes beyond various boundaries that makes the growth of the internet very fast every day [1]. One of the internet services that is widely used and very popular is e-mail which is used in an organizational, corporate or individual environment [2]. With technological developments, e-mail is not only able to send text files, but can also send several files such as audio, video, photo, and other extension files [3]. There are threats that follow by utilizing these features as a medium of crime in the cyber world, because email is the easiest tool to become a medium for sending spam (phishing, scam, malware, computer viruses and mail worms) and malicious programs that are camouflaged and attached to attachments. One of the crimes found involving email was email spamming and email spoofing [4]. Spamming is the sending of unwanted news or advertisements or what is called bulk mail or junk email [5]. While email spoofing is an email that is intentionally faked so that it seems as if it was sent from a legitimate email [6].

With so many crimes happening today, more and more technologies are being developed to check and protect e-mails, including spam e-mail detection [7]. One way to develop this technology is to conduct internet forensic investigations [8]. The results of testing and analysis on the system are designed for useful forensic evidence [9]. In general, there are two types of internet forensic investigations, namely dead forensics and live forensics [10]. Dead forensics is a technique that requires data to be stored permanently in a storage media device, generally a hard disk. Live forensics is an analytical technique that involves running data which is generally stored in Random Access Memory (RAM) or in transit on a network [11]. An important part in digital forensics is the authenticity of digital evidence [12]. Conducting an investigation through the stages of a digital forensics digital examination procedure approach is the correct way to obtain the evidence [13].

The National Institute of Standards and Technology (NIST) is a method used to perform forensic analysis. This method has been widely used as a reference for forensic analysis research. For example, in android-based analysis, Wijaya (2017) used the NIST method to analyze the telegram application on smartphones [14]. Also, Anshori (2018) analyzed the digital evidence for Facebook Messenger which also used the NIST method [15].

This research can later become new knowledge about how forensic investigations deal with crimes in the cyber world, especially in the case of e-mail spam. As in the real world, crimes in the cyber world also require a forensic process, which combines elements of law and computer science. This research can be a first step to solving a complex crime case and can help with previous research. E-mail header investigation is an important aspect of
investigation because e-mail metadata and other information are contained in e-mail headers. Analysis of e-mail headers can reveal the source, destination, e-mail client, sender IP, identification of fake or authentic e-mails, and more. In practice, e-mails have been repeatedly adopted as evidence by legal departments. With the continuous advancement of national legal processes and the continuous improvement of electronic evidence laws, e-mail forensics is indispensable in the detection of computer crime cases.

Method

The analysis process of this research uses the National Institute of Standards and Technology (NIST) method. This method refers to the basic stages in a forensic analysis, namely collection, examination, analysis, and reporting [16] which is shown in Figure 1.

![Figure 1. Stages in the method National Institute of Standards and Technology (NIST)](image)

A. Collection

Collection was the stage of conducting the forensic process to identify sources that are considered potential to be used as evidence, and the steps needed in data collection.

B. Examination

Examination was the stage of processing the data collected forensically, either automatically or manually.

C. Analysis

Analysis was the stage of analyzing the results of the examination using technically and legally justified methods to obtain useful information and answer questions that encourage collection and examination.

D. Reporting

The reporting stage was reporting the results of the analysis which includes a description of the actions taken.

Results and Discussion

The results of this study were earned through evidences on spam emails by opening the header in the email.

A. Collection

This stage was the stage for identifying the header section for digital evidence and conducting data sources. The first step in the forensic process was to identify sources that were considered potential to be used as evidence. Further, we described the steps required in data collection.

![Figure 2. Inboxes in a spam folder](image)
B. Examination

The examination was run to determine data filtering in certain parts of the data source. Data filtering was carried out by changing the shape of the data but we did not make changes to the data content because the authenticity of the data was very important.

In Figure 4 is a display of spam email headers. This email header was used for analysis of emails that entered the spam folder.
C. Analysis

The step taken was to analyze the generated. We analyzed where, how and why the data was generated, and by whom.

**Figure 5. Spam Email Header Details**

![Figure 5: Spam Email Header Details](image)

From the analysis of the spam email headers, we could trace where it came from, but we did not know yet who did it because only the IP Address was recorded (74.6.134.215). IP address trace results show in **Figure 6**.

There were 5 IP address administrators in the world. These five administrators are those who regulate the use of IP addresses on every computer in the world. The five administrator ip addresses are:

1. ARIN (North America and Sub-Saharan Africa) website: www.arin.net.
2. RIPE (Europe and Northern Africa) website: www.ripe.net.
3. APNIC (Asia Pacific) website: www.apnic.net.
4. LACNIC (Southern and Central America and the Caribbean) website: www.lacnic.net.
5. AFRINIC (The African Network Information Centre) website: www.afrinic.net.

**Figure 6. IP address trace results**

![Figure 6: IP address trace results](image)
After analyzing the IP of the spam email sender, the eToolz application notified that the IP (74.6.134.215) belonged to ARIN.

D. Reporting

At these digital forensics reporting stage, from the 3 stages that we run, digital forensic evidence was obtained. At the previous stage, we managed to get digital evidence in the form of an IP address contained in the email header.

Conclusion

Based on the results of the tests conducted using the NIST method, it can be concluded that e-mail header investigation was an important aspect of the investigation because e-mail metadata and other information were contained in the e-mail headers. Analysis of e-mail headers could reveal the source, destination, e-mail client, sender IP, identification of fake or authentic e-mails, and more. IP addresses could be tracked using applications to make it easier to find the sender of the e-mail. Once an IP address was tracked, it was easy to find routes, geographic locations, network providers, and more.

Based on the results of the research, the authors suggest that hopefully the current study can be useful for developing better tools so that it can be more effective and the results obtained are more detailed. More previous studies reviews are needed so that the research results are more accurate and digital evidence can be used as evidence.

References