Digital Forensic Evidence Analysis in Revealing Defamation on Social Media (Twitter) Using the Static Forensics Method

Reski Badillah¹*, Andi Yulia Muniar², Abd. Rahman³, Febri Hidayat Saputra⁴, Mansyur⁵, Supriadi Sahibu⁶

¹⁻⁴Department of Informatics, Universitas Teknologi Akba Makassar, Makassar, Indonesia
³Educational Research and Evaluation, Universitas Negeri Makassar, Makassar, Indonesia
⁶Computer System, Universitas Handayani Makassar, Indonesia

Abstract

This research addresses the persistent challenge of defamation, notably prevalent on the Twitter platform, where the discovery of digital evidence is hampered by robust privacy protections. The study aims to investigate and identify digital evidence in defamation cases on Twitter, focusing on optimizing the evidence discovery process. Employing static forensics to prevent data alterations during acquisition from devices associated with defamation, the research successfully uncovered various digital evidence, including text from deleted comments, usernames, emails, and deleted image files linked to defamation. Out of the initial 28 reported data instances, 22 pieces of evidence were identified, resulting in an impressive 79% accuracy rate. The investigative procedures align with the chain of custody, ensuring the reliability of the collected evidence. This study not only contributes valuable insights into digital evidence discovery in online defamation cases but also highlights the efficacy of static forensics as a method. These findings provide a foundation for robust digital forensic practices, crucial for addressing challenges posed by online defamation on social media platforms.

Keywords: Digital Evidence, Static Forensics, Defamation, Chain of Custody.

Received: 15 August 2023 Revised: 19 October 2023 Accepted: 6 December 2023

Introduction

In the era of information technology development, social media is becoming increasingly popular and continues to evolve (Bimo, 2017; Yusuf, 2023). Social media is an online service that enables communication through the internet (Noorikhsan et al., 2023). One of the popular platforms is Twitter, where users can express their thoughts through images, videos, or text (Guraba, 2021). Unfortunately, the spread of fake news continues to rise, and many people easily believe in it. In facing this issue, digital forensics becomes an effective method to uncover the truth of digital evidence and identify perpetrators in cases of defamation on social media (DSLA, 2023).

There are several methods and applications used in digital forensics to reveal the truth of digital evidence (Priyono et al., 2022). A study (Anggraini et al., 2022) used the static forensics method to uncover the truth of digital evidence on the TikTok application, using the Magnet AXIOM application and achieving an accuracy rate of 77%. In the research by (Ardiningtias et al., 2021), also using static forensics in a Digital investigation case on Facebook Messenger, the MOBILedit Forensics Express application was employed with an accuracy rate of 85.71%. And in the study by (Utami et al., 2021), the live forensic method was used to prove cases of electronic transaction fraud on WhatsApp web using the FTK Imager and Browser History Viewer applications, with an accuracy rate of 45.6% out of a total of 46 messages and 39% out of a total of 46 timestamps on messages (Leonardo & Indrayani, 2021).

The static forensics method has advantages in uncovering information from digital evidence that is inactive or deleted after the incident, while the live forensic method can obtain evidence directly from a running system quickly (Nurhairani & Riadi, 2019), (Sulianta, 2013). However, static forensics cannot identify real-time changes on devices, whereas live forensics is limited to active systems and cannot trace or obtain evidence if the system is not in an active state (Haryadi, 2022).

*Corresponding author.
E-mail address: reskibadillah@gmail.com (Reski Badillah)
Based on previous research, the conclusion that can be drawn is that not all available or investigated evidence in previous studies was successfully obtained (Abdi, 2021), (Pedapudi & Vadlamani, 2023). Although digital forensics continues to evolve, there are still limitations in collecting evidence from specific devices or platforms, such as strong privacy protection, including the issue of defamation that is currently prevalent (Casey et al., 2022). This research aims to investigate the collection and preservation of activity traces or digital evidence related to defamation on Twitter (Efendi et al., 2020), (Yudhana et al., 2019). The main focus is to measure the accuracy of discovering digital evidence related to the prevalent issue of defamation (Kusbiyanto, 2022). The static forensics method is used to identify and collect relevant evidence from Twitter accounts (Casey, 2009). Analysis tools such as Belkasoft Evidence Center X, Belkasoft Remote Acquisition, and Paraben’s E3 Universal are applied in this study, with the hope of further uncovering cases of defamation on Twitter (Mukti, 2017).

Method

The primary goal of static forensics is to identify, collect, and analyze digital evidence that can be used in legal investigations (Rafique & Khan, 2013). This method can assist in uncovering criminal activities (Sachdeva et al., 2020), cybercrimes, or other illegal activities that occur through digital devices or media (Salamh et al., 2021). In practice, static forensics involves analyzing various types of files and data within image forensics, such as text files, image files, audio files, video files, log files, system files, and metadata (Yaacoub et al., 2021).

The method used in this research is static forensics (Waseem et al., 2021), which consists of several stages:

a. Identification
   In this identification stage, preparing equipment for the investigation process is based on the created case scenario. This includes determining what is needed to find digital evidence.

b. Data Collection
   Collecting data or gathering evidence, in the form of the target investigative smartphone. The evidence is the Samsung Galaxy J7 smartphone, as seen in Table 1.

Table 1. Smartphone Specifications

<table>
<thead>
<tr>
<th>Specifications</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>SM-J700F</td>
</tr>
<tr>
<td>Android Version</td>
<td>6.0.1</td>
</tr>
<tr>
<td>IMEI 1</td>
<td>352846072333925</td>
</tr>
<tr>
<td>IMEI 2</td>
<td>352847072333923</td>
</tr>
<tr>
<td>SIM Card</td>
<td>Yes</td>
</tr>
</tbody>
</table>

c. Analysis
   Copying data from the target device using applications like Belkasoft Evidence Center X (Delija et al., 2022), Belkasoft Remote Acquisition, and Paraben’s E3 Universal (Perumal & others, 2022). The data acquisition process from the smartphone or target device to the laptop is done using a data cable. The acquisition process takes a considerable amount of time depending on the storage capacity of the smartphone. After completing the data acquisition process, an examination of metadata information is conducted, searching for evidence such as image files, text messages, comments, account information, or Twitter IDs related to the initial data. The data search process also takes time, depending on the number of files acquired and the amount of data or evidence that needs to be found.

d. Reporting
   This stage involves displaying the data resulting from the analysis, followed by data matching or accuracy calculation using the percentage formula between the initial data (victim) and the data resulting from the analysis (perpetrator). The percentage calculation results determine the accuracy level of discovering digital evidence between the initial data and the analysis results. After accuracy calculation, the suspect is identified, and then a report on the investigation results of the defamation case on the Twitter application is presented. The investigation report takes the form of a chain of custody report.
The design method of Digital Forensic Evidence Analysis in Revealing Defamation on Social Media (Twitter) Using the Static Forensics Method is as follows:

1. Case Scenario Design

This research creates a scenario in which the perpetrator posts negative content about the victim using a fake account. The purpose of running this scenario is to facilitate the investigation of social media Twitter defamation cases. The scenario unfolds as follows, figure 1.

Initially, the perpetrator creates a fake account on the Twitter application and uses it to comment on one of the victim's posts. Subsequently, the perpetrator uses the fake account to send messages to the victim. After not receiving a response from the victim, the perpetrator sends messages containing negative or demeaning words. The perpetrator also resumes commenting on the victim's posts with defamatory remarks (Javed et al., 2022). The victim, responding to these messages, tries to identify the perpetrator's identity. Next, the perpetrator posts false information about the victim. After making posts, and comments, and sending negative messages, the perpetrator promptly deletes the text messages and comments. Feeling disturbed by the negative messages and comments, the victim reports the suspected perpetrator to the authorities for investigation. In this study, it is assumed that the perpetrator is a coworker who initially showed interest in the victim.
However, after being rejected, the perpetrator did not accept the rejection, leading to the decision to engage in defamation against the victim.

2. Investigation Flow

The investigation process in Figure 2, begins with data collection or preparing evidence for examination. This involves copying or retrieving data from the target device using applications such as Belkasoft Evidence Center X, Belkasoft Remote Acquisition, and Paraben’s E3 Universal. After acquiring or retrieving data, the next step is to examine the data, searching for evidence such as image files, text messages, comments, posts, account information, or Twitter IDs related to the initial data (Khweiled & Jazzar, 2021). Following the examination of relevant data or evidence, the next step is to match the evidence between the initial data and the results of the analysis by calculating the accuracy level to determine the validity of the evidence. Subsequently, identifying who needs further action, the investigation results are presented in a report handed over to the appropriate authorities.
Results and Discussion

Result

Through the application of the static forensics method, this study successfully identified key defamatory elements embedded within social media posts on Twitter. The method proved instrumental in extracting and analyzing textual content, images, and associated metadata, offering a comprehensive understanding of the context and nature of defamatory material. The results of the study from Digital Forensic Evidence Analysis in Revealing Defamation on Social Media (Twitter) Using the Static Forensics Method are as follows:

1. Case Scenario
   In this research, utilizing the static forensics method, a scenario is outlined within a Twitter scene, involving the victim's Twitter account named @DillaLyla and the perpetrator's Twitter account named @Velakhu. Initially, the perpetrator merely commented on one of the victim's posts, greeting the victim, and also sent a message to which the victim did not respond to. Subsequently, the perpetrator began commenting on the victim's posts with unpleasant words. Moreover, on the perpetrator's Twitter account, derogatory comments accompanied by the victim's photo were posted. Feeling disturbed by the perpetrator, the victim reported the case to the authorities, citing violations of the ITE Law Article 27 Paragraph (4) on defamation and/or character assassination through electronic media. Upon learning of the report, the perpetrator then deleted chats, and posts, and changed the Twitter ID and name that were previously associated with the victim.

2. Identification
   To carry out the investigative process to find digital evidence, careful preparation of all necessary equipment is essential, including computers, forensic software, and required cables throughout the investigation.

<table>
<thead>
<tr>
<th>Tool's Name</th>
<th>Description</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laptop</td>
<td>Lenovo ThinkPad X260, RAM 8.00 GB, Windows 10 Pro</td>
<td>Hardware</td>
</tr>
<tr>
<td>Smartphone</td>
<td>Samsung Galaxy J7</td>
<td>Hardware</td>
</tr>
<tr>
<td>USB Cable</td>
<td>USB Micro</td>
<td>Smartphone Connector</td>
</tr>
<tr>
<td>Belkasoft Evidence Center X</td>
<td>V.2.0.132777</td>
<td>Software</td>
</tr>
<tr>
<td>Belkasoft Remote Acquisition</td>
<td>V.1.2.12161</td>
<td>Software</td>
</tr>
<tr>
<td>Paraben's E3 Universal</td>
<td>V. 3.7.16206.20452</td>
<td>Software</td>
</tr>
</tbody>
</table>

3. Collection of Evidence
   After receiving a report related to the violation of the Electronic Information and Transactions Law (ITE Law), the investigative team promptly took action by securing items suspected to be evidence in the defamation case on the Twitter platform. During the investigation at the perpetrator's residence, the investigative team found a Samsung Galaxy J7 smartphone suspected to be used in the commission of the alleged crime.

<table>
<thead>
<tr>
<th>Name of Evidence</th>
<th>Picture</th>
<th>Description</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung Galaxy J7</td>
<td><img src="image" alt="Samsung Galaxy J7" /></td>
<td>Version Number MMB29K.J700FXXU4BVH1 6.0.1 IMEI (1) 352846072333925</td>
<td>Hardware</td>
</tr>
</tbody>
</table>

4. Device-related Case
   The perpetrator in this issue is faced with a case involving the use of their mobile device in acts of defamation and derogatory remarks against the victim. This case is related to applicable laws, and in the submitted report, the victim attached screenshots documenting the criminal actions performed by the perpetrator. This serves as tangible evidence supporting the charges against the perpetrator in the context of this case.
In the image, a compelling narrative unfolds as it captures the perpetrator’s actions. Initially, the wrongdoer downloads a picture from the victim’s account, setting the stage for an act of digital transgression. Subsequently, the perpetrator takes a step further by posting the victim’s photo on a public platform, accompanied by an unpleasant caption. This deliberate attempt to tarnish the victim’s image is intensified by the inclusion of a tagged mention to the victim’s Twitter account. The act of tagging not only amplifies the visibility of the defamatory content but also directly involves the victim, ensuring their notification and potential exposure to the harmful post. This image encapsulates a critical moment in the perpetration of online
Defamation, serving as digital evidence that can be instrumental in legal proceedings and investigations to hold the perpetrator accountable for their actions.

Article 27 paragraph (3) of the ITE Law also regulates defamation. Perpetrators charged under this article may face imprisonment for up to 4 years and/or a maximum fine of IDR 750,000,000.00 (seven hundred fifty million Indonesian Rupiah). Furthermore, in the revision of Law No. 19 of 2016, it is explained that the provision in Article 27 paragraph (3) constitutes a criminal complaint. Spreading information to incite hatred or hostility towards individuals and/or specific groups based on ethnicity, religion, race, and inter-group (SARA) is also prohibited under Article 28 paragraph (2) of the ITE Law. The punishment for perpetrators of hate speech as described in Article 28 paragraph (2) is imprisonment for a maximum of 6 years and/or a fine of up to IDR 1,000,000,000.00 (one billion Indonesian Rupiah).

**Discussion**

The temporal analysis conducted through static forensics emerges as a valuable tool for unraveling the chronological evolution of defamatory content on Twitter. This temporal perspective provides a holistic view of the lifecycle of such content, aiding in the identification of patterns, modifications, and dissemination dynamics over time. This stage, begins with the acquisition of data on the Samsung Galaxy J7 device using tools such as Belkasoft Evidence Center X, Belkasoft Remote Acquisition, and Paraben’s E3 Universal (Lwin et al., 2020).

**Acquisition using Belkasoft Remote Acquisition:**

1. After logging into Belkasoft Remote Acquisition, to initiate the acquisition, connect the Samsung Galaxy J7 device using a USB cable. Once connected, click the Acquire button > Mobile > ADB Backup > Click the folder for storing the acquisition results > Start. Then, wait for the acquisition process to complete. The duration of the acquisition process depends on the amount of data on the Samsung Galaxy J7 device.

2. Searching for evidence in Belkasoft Evidence Center X

Open Belkasoft Evidence Center X > Add Data Source > Add Existing > Mobile Image > Then select the acquisition result file from Belkasoft Remote Acquisition.
The digital evidence successfully found includes image files, the perpetrator's accounts @Velakhu and @tershangKha, where the @Velakhu account was changed to @tershangKha. Additionally, there are ongoing tweet posts with the caption, “Assalamualaikum dunia,” created on October 8th. Acquisition using Belkasoft Remote Acquisition Paraben’s E3 Universal:

1. After accessing Paraben’s E3 Universal > Create New Case > Enter the Case name "Digital Evidence" > choose the Case storage location. Connect the Samsung Galaxy J7 device using a USB cable, then click Start Acquisition > Android > Custom Logical Acquisition > Continue > Start Acquisition > Wait for the acquisition process to complete.

2. Searching for digital evidence
   The obtained digital evidence includes deleted text comments and posts, image files, emails, as well as the username of the perpetrator.
One of the deleted posts by the perpetrator had the caption, ‘@DillaLyla “Si kurus kerempeng, merasa paling tersakiti, padahal enang cewek ular,”’ accompanied by a photo of the perpetrator.
Twitter ID, the perpetrator's username @Velakhu, which was changed to @tershangKha.

Table 4. Successful evidence found

<table>
<thead>
<tr>
<th>Type Data</th>
<th>Jumlah</th>
<th>Belkasoft Evidence Center X</th>
<th>Paraben’s E3 Universal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Username / Id</td>
<td>4</td>
<td>2</td>
<td>✓</td>
</tr>
<tr>
<td>Deleted Image Files</td>
<td>2</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Profile Image File</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deleted Comment</td>
<td>10</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Deleted Post</td>
<td>2</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Chat</td>
<td>6</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Ongoing Posts</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>5</td>
<td>21</td>
</tr>
</tbody>
</table>

The results of the acquisition process and the search for digital evidence are attached in the table. The initial data reported by the victim amounted to 27, and the data successfully found through analysis amounted to 21. The percentage is calculated as follows:

\[ \text{Percentage} = \frac{\text{Partial Value}}{\text{Total Value}} \times 100\% \]

\[ = \frac{22}{28} \times 100\% \]

\[ = 0,7857 \times 100\% \]

\[ = 78,57\% \]

The success rate in uncovering digital evidence reaches an impressive 79%. The investigation procedure stipulates that once the accuracy rate surpasses the 70% threshold, legal authorities are mandated to initiate follow-up actions against the perpetrator. Specifically, these actions fall under the defamation case outlined in Article 27 paragraph (3) of the ITE Law. This stringent criterion emphasizes the gravity of online defamation and the importance of meeting a substantial evidentiary standard before legal measures are taken. The procedural framework for conducting this investigation is meticulously governed by the order and comprehensive report detailing the chain of custody investigation results. This approach ensures a systematic and accountable process in handling cases of digital defamation, providing a robust foundation for law enforcement to pursue decisive and just actions by the stipulations of the law.
Conclusions and Suggestions

Conclusions

After analyzing the discovery of digital evidence in the case of defamation on the social media platform Twitter, the following conclusions can be drawn. The research results indicate that out of the initial 28 data (victim's reporting data), 22 data were successfully found through analysis. The investigation procedure aligns with the order and the report of the chain of custody investigation, ensuring the preservation chain of data or evidence in the defamation case. The investigation report revealed successfully found data, including emails, usernames/IDs, deleted comments and posts by the perpetrator, the perpetrator's profile picture, and images deleted along with two posts by the perpetrator. The evidence also showed that the perpetrator changed their username after committing the crime. The success rate of finding digital evidence in this investigation is 79%. Based on the investigation process, if the accuracy rate is above 70%, the authorities will proceed with taking action against the perpetrator under Article 27 paragraph (3) of the ITE Law.

Suggestions

Based on the research on defamation on the Twitter social media platform, the following recommendations are proposed for further development. Utilize the latest and most advanced tools in the data collection process. By using these advanced tools, it is expected that data searches with keywords will become more efficient, saving time, and facilitating the identification of digital evidence without the need to search each acquisition data individually. And consider adding superior tools to find deleted data such as chat records to achieve a better accuracy rate in the discovery of digital evidence.

Acknowledgments: Thank God Almighty for the blessings given in completing this research.

References


https://www.kompasiana.com/muhammad99498/653b18cc110fcee5247396822/tren-penggunaan-internet-dan-media-sosial-dalam-era-konvergensi-media