International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 06, Issue 03, September 2025
e-ISSN : 2745-9659
https://ijcis.net/index.php/ijcis/index

Uncovering WhatsApp Fraud Modus Operandi through Digital Artifact Analysis and Cyber Kill Chain Mapping

1st Erika Ramadhani
[1]*Department of Informatics*
[1]*Universitas Islam Indonesia*
[1]Yogyakarta, Indonesia
[1]erika@uii.ac.id

***Abstract*—*WhatsApp fraud has emerged as a significant cybercrime threat, exploiting the platform's wide user base through social engineering and malware-based attacks. This study investigates a WhatsApp fraud case by analyzing digital artifacts to uncover the perpetrator's modus operandi and provide structured guidance for law enforcement. Using the Digital Forensics for Incident Response (D4I) Framework in conjunction with Cyber Kill Chain (CKC) mapping, five key artifacts were identified and evaluated quantitatively based on their strength of evidence (v) and reliability (r). The results show that the malicious APK and source code containing a Telegram bot token constitute primary evidence with the highest probative value, while the Manifest.xml file and hidden background application serve as supporting evidence, and contextual indicators such as sender information provide limited legal weight. These findings highlight the importance of differentiating artifacts by evidentiary significance and demonstrate the value of the proposed scoring methodology. The study has limitations, as it is based on a simulated case and relies partly on expert judgment in scoring criteria. Future research should apply the approach to other platforms and fraud scenarios, and explore automation to enhance objectivity and scalability. Beyond its academic contributions, the study offers a structured rubric for prioritizing evidence and emphasizes the need for standardized evaluation frameworks in digital forensic policy and practice, ultimately strengthening the legal robustness and societal trust in digital investigations.*

Keywords : *Digital Forensics, WhatsApp Fraud, Digital Artifact Analysis, Cyber Kill Chain, D4I Framework, Evidence Prioritization, Forensic Reliability, Digital Forensic Policy*

I. INTRODUCTION.

WhatsApp fraud has escalated into one of the most pervasive forms of cybercrime in recent years. According to the *Global Anti-Scam Alliance* (2023), over 70% of internet users worldwide have been exposed to scams through messaging applications, with WhatsApp consistently ranking as one of the most targeted platforms. In Indonesia alone, the Ministry of Communication and Information reported that digital fraud complaints surged by more than 60% between 2021 and 2023, with WhatsApp scams forming a significant proportion of reported cases. These scams range from financial theft through social engineering to identity fraud, often resulting in both economic losses and long-term psychological trauma for victims. Such trends highlight the urgent need for systematic forensic methods that can support law enforcement and restore public trust in digital platforms [1]

Despite the severity of this phenomenon, law enforcement investigations face persistent challenges in collecting, validating, and prioritizing digital evidence. Artifacts such as chat logs, file metadata, and hidden application traces often go underutilized due to the absence of structured frameworks for analysis. This gap hinders prosecutors from establishing strong legal arguments and undermines the admissibility of digital evidence in court [2]. Therefore, the central research question addressed in this study is: How can digital artifacts from WhatsApp scam cases be identified, analyzed, and validated in a way that strengthens their probative value and supports law enforcement processes?

To address this gap, this study applies the Digital Forensics for Incident Response (D4I) Framework, integrated with the Cyber Kill Chain (CKC) model. The D4I Framework offers a systematic, evidence-driven process for discovering, documenting, analyzing, and interpreting digital artifacts [3], while the CKC model provides a structured lens for mapping the attacker's modus operandi across distinct phases of the attack. By linking these two approaches, this research not only identifies which artifacts hold the highest evidentiary value but also demonstrates how they can be contextualized within a broader attack lifecycle.

The contributions of this study are threefold. First, it identifies and evaluates key artifacts from WhatsApp scam cases, distinguishing between primary, supporting, and contextual evidence based on their probative strength and reliability. Second, it demonstrates the applicability of the D4I Framework in real-world fraud investigations, thereby addressing a critical research gap in forensic methodology. Finally, it provides a transparent and replicable rubric for law enforcement to assess digital evidence admissibility, ensuring both technical rigor and legal soundness. By explicitly linking the D4I Framework to the challenges faced in practice, this study underscores why it is the most suitable approach for enhancing the reliability and effectiveness of digital forensic investigations in messaging-based fraud.

II. LITERATURE REVIEW

Research on fraud in instant messaging platforms has evolved along both social and technical dimensions. From the social perspective, Lee et al. (2023) demonstrated that psychological factors and low user self-efficacy significantly increase susceptibility to phishing attempts on messaging apps [4]. Such findings explain why WhatsApp, with its massive user base, remains a prime target for social engineering attacks. However, these studies often stop short of linking behavioral vulnerabilities to forensic challenges, leaving a gap in how user behavior translates into digital artifacts that can be used in investigations.

On the technical side, several works have analyzed WhatsApp artifacts as potential sources of evidence. Meng et al. (2022) explored IndexedDB in WhatsApp Web as a

forensic data source [5], while Kim et al. (2025) examined artifacts in the Web and UWP versions of WhatsApp [6]. Son et al. (2022) extended this line of research by demonstrating decryption methods in other encrypted instant messaging platforms such as Signal and Threema [8]. Together, these works establish that even with end-to-end encryption, residual digital traces can be extracted and analyzed. Yet, most of these studies focus on identifying data sources without offering systematic frameworks for evaluating the evidentiary strength and legal admissibility of such artifacts.

Malware exploitation has emerged as another critical theme in WhatsApp-related fraud. Schmutz et al. (2024) analyzed hook-type Android malware that runs covertly in the background [8], while Faruki et al. (2023) surveyed malware evasion techniques, highlighting the growing sophistication of attacks [9]. Palma et al. (2024) further introduced explainable machine learning for Android malware detection [10]. These contributions illustrate the technical complexity of APK-based fraud, such as the "Wedding Invitation" scam analyzed in this study. Nevertheless, prior research rarely connects these malware-focused findings to the broader lifecycle of fraud attacks or considers how they map onto established cybercrime frameworks such as the Cyber Kill Chain.

Finally, the literature also touches on communication channels and legal implications. Al lelah et al. (2023) revealed how attackers increasingly exploit legitimate cloud services—such as the Telegram Bot API—as command-and-control (C2) infrastructures [11]. From a legal standpoint, Heath et al. (2023) emphasized that forensic soundness and chain of custody are essential for ensuring that digital evidence from ephemeral messaging applications is admissible in court [12]. While these studies underscore important technical and legal considerations, they often treat them in isolation, with limited effort to integrate social, technical, and legal aspects into a unified investigative approach.

Taken together, existing studies provide valuable insights into the social engineering tactics, technical artifacts, and legal challenges of instant messaging fraud. However, three key gaps remain. First, prior work has primarily summarized artifacts or malware without critically evaluating their relative probative strength or reliability as legal evidence. Second, little effort has been made to bridge social/behavioral insights with technical forensic analysis, resulting in fragmented understandings of fraud cases. Third, there is a lack of integrated frameworks that connect digital artifact analysis with structured models of cybercrime progression.

This research addresses these gaps by applying the Digital Forensics for Incident Response (D4I) Framework in conjunction with Cyber Kill Chain (CKC) mapping. By doing so, it not only identifies key digital artifacts but also evaluates their evidentiary value, integrates social engineering perspectives with technical findings, and demonstrates a replicable methodology that supports both academic inquiry and law enforcement practice.

III. RESEARCH METHODS

This study adopts a case study approach focusing on fraud attacks conducted through the WhatsApp application using the *Wedding Invitation.apk* file, a malicious Android package depicted in Figure 1. This case is representative of prevalent fraud schemes in Indonesia over the past two years, where perpetrators employ social engineering to convince victims to install disguised malware. Once installed, the application requests excessive permissions and operates covertly in the background, enabling the attacker to intercept sensitive data such as one-time passwords (OTPs).

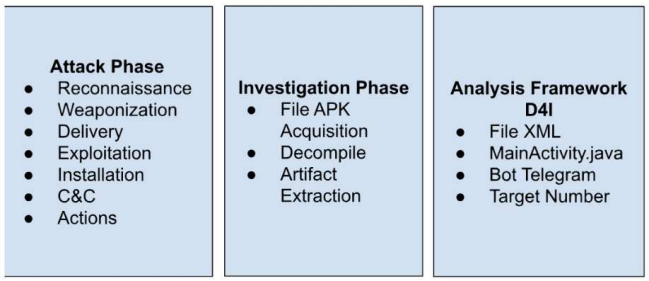

Figure 1. Diagram simulation of Whatsapp Scam

3.1 Data Collection and Validation

The primary data source for this study consists of digital artifacts generated from a controlled simulation of a WhatsApp scam attack. These include chat messages, APK files, metadata, and extracted source code. To ensure reproducibility, the data collection process adhered strictly to the Discovery and Acquisition stages of the D4I Framework. Each artifact was acquired using write-blocking mechanisms to preserve integrity, while hashing (MD5 and SHA-256) was applied to confirm that no alterations occurred during acquisition. Validation was achieved by repeating the extraction process on multiple devices and cross-verifying the consistency of the recovered artifacts.

3.2 Tool selection and Justification

The technical analysis was conducted using a combination of Apktool and Show Java for decompiling the APK file, Autopsy for artifact examination, and FTK Imager for creating forensic disk images. These tools were selected based on their proven reliability, availability, and acceptance in forensic practice. Apktool and Show Java were chosen because they allow detailed inspection of AndroidManifest.xml and embedded Java code, which are critical for detecting excessive permissions and hidden malicious logic. Autopsy was preferred over alternatives such as EnCase due to its open-source accessibility and extensibility, making it suitable for academic and law enforcement environments. FTK Imager was selected for its robustness in generating forensically sound disk images while maintaining the integrity of the original data.

3.3 Ethical and Legal Considerations

All experiments were conducted using simulated data and malware samples in an isolated environment to avoid risk to real users or devices. No actual victim data was collected. Ethical compliance was ensured by anonymizing identifiers and restricting the analysis to synthetic or test accounts. From a legal standpoint, the methodology adheres to the principles of forensic soundness, including maintaining a clear chain of custody for all artifacts and ensuring that evidence acquisition

methods align with standards that support admissibility in court.

3.4 Integration of D4I framework and CKC Mapping

The forensic investigation followed the four stages of the D4I Framework—Discovery, Documentation, Dynamics, and Interpretation—and mapped each artifact to the corresponding phase of the Cyber Kill Chain (CKC). For example, the attacker's phone number was linked to *Reconnaissance*, the malicious APK to *Weaponization*, and the hidden background process to *Installation*. This dual approach ensures that both the technical behavior of the malware and the attacker's modus operandi are systematically reconstructed. Figure 2 illustrates the methodological flow, showing how each stage of the investigation aligns with the D4I Framework and CKC phases

3.5 Quantitative Evaluation of Evidence

In addition to descriptive artifact analysis, this study employed a quantitative rubric to assess each artifact along two dimensions: (1) Strength of Evidence (*v*): representing the probative value of an artifact in proving the attacker's actions within the CKC phases; (2) Reliability (*r*): representing the trustworthiness of an artifact as legal evidence, considering acquisition method, integrity, and chain of custody. Each dimension was further divided into six sub-criteria:

(a) For *v*: relevance, specificity, connection to CKC phase, causal proximity, corroboration by other artifacts, and evidentiary clarity.
(b) For *r*: integrity, authenticity, acquisition method, chain of custody, reproducibility, and independence from external bias.

Scores were assigned on a scale of 1–5 for each sub-criterion and then normalized to a 0–1 range. Weighted averages were calculated to produce final *v* and *r* scores for each artifact. This structured approach ensures transparency, reproducibility, and objectivity in distinguishing between primary evidence, supporting evidence, and contextual evidence.

Figure 2 illustrates the overall research methodology, structured according to the D4I Framework and Cyber Kill Chain (CKC) mapping. The process begins with Discovery, where suspicious artifacts such as the *Wedding Invitation.apk* are identified and extracted. This is followed by Documentation, in which each artifact is systematically recorded with screenshots, metadata, and tabular mapping to CKC phases to preserve forensic integrity. The third stage, Dynamics, analyzes the interaction and progression of artifacts across the CKC phases, showing how the attack evolves from reconnaissance to data exfiltration. Finally, **Interpretation** answers the 5W1H questions (Who, What, When, Where, Why, and How), linking the artifacts back to the attacker's modus operandi. Together, the diagram emphasizes that the methodology is both systematic and reproducible, integrating forensic acquisition with structured cyberattack analysis.

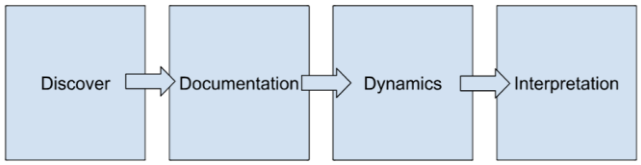

Figure 2. Research methodology

IV. RESULT AND ANALYSIS

The results of the study should be written clearly and concisely. The discussion should describe the importance of the results of the study, not repeat it.

3.1 Artefact Identification

The forensic investigation produced eight key digital artifacts (a_1–a_8) that correspond to different stages of the Cyber Kill Chain (CKC). These relationships can be formalized in the mathematical relation *R*, where each pair (p_i, a_j) indicates that artifact *aj* plays a role in CKC phase *pi*. For example, (p_1, a_1) maps the WhatsApp sender's number to *Reconnaissance*, while (p_2, a_2) associates the malicious APK file with *Weaponization*. This formalization ensures analytical rigor by explicitly linking evidence to the sequential stages of the attack lifecycle.

$$R=\{(p_1,a_1), (p_2,a_2), (p_2,a_8), (p_3,a_3), (p_4,a_4), (p_5,a_5), (p_6,a_6), (p_7,a_7)\} \quad \ldots\ldots\ldots(1)$$

where $p1,\ldots,p7$ are the stages in CKC phase and a_1,...a_8 are the

a_1: victim's phone number
a_2: APK file
a_3: Manifest.xml file
a_4: android permission
a_5: hidden installed application
a_6: telegram boot token
a_7: captured SMS
a_8: file metadata.

3.2 Artefact Analysis

In this study, there are five main artifacts that were analyzed to uncover the modus operandi of fraud perpetrators using WhatsApp. The first artifact is the victim's phone number (a_1) used during the reconnaissance phase, indicating that the perpetrator had conducted initial reconnaissance to determine the target. The second artifact is an APK file named "Wedding Invitation" (a_2), which plays a role in the Weaponization phase, where malicious applications are sent to trap victims. Next, the Manifest.xml file (a_3) containing excessive permission requests was analyzed as part of the Exploitation phase, demonstrating how the attacker exploited Android system vulnerabilities. The fourth artifact, a hidden application running in the background (a_4), represents the Installation phase, proving that the malware was installed silently without the user's knowledge. Finally, the application source code and Telegram bot token (a_5) show both Command and Control (C2) activity and Actions on Targets, as these artifacts serve as a means for the attacker to receive sensitive data such as OTPs directly from the victim's device. These five artifacts sequentially depict the systematic stages

of the attacker in carrying out the attack according to the CKC model.

Based on this assessment methodology, each artifact in this case was then assigned a score for strength of evidence (v) and reliability (r), allowing for a clear view of the relative contribution of each artifact in proving the stages of the attack. The assessment results are shown in the following Table 1.

Table 1. Strength of evidence (v) and reliability (r)

Artefact	CKC Phase	v (strength of evidence)	r (reliability)
a_1	Reconnaissance	0.588	0.488
a_2	Weaponization	0.913	0.775
a_3	Exploitation	0.813	0.688
a_4	Installation	0.838	0.713
a_5	C2/Actions of Objective	0.950	0.713

According to Table 1, each artifact was evaluated along two dimensions: strength of evidence (v) and reliability (r). Table 1 presents the results, showing that the source code containing the Telegram bot token (a_5) scored the highest (v=0.950, r=0.713), followed by the malicious APK (a_2) with strong values (v=0.913, r=0.775). These results indicate that artifacts directly tied to malware functionality and attacker infrastructure have the strongest probative value and reliability for legal proceedings. By contrast, reconnaissance-related artifacts such as the sender's number (a_1) had lower scores (v=0.588, r=0.488), demonstrating that contextual indicators alone cannot establish culpability in court.

3.3 Relevance to Law Enforcement

From a legal enforcement perspective, the scoring highlights the differential evidentiary weight of each artifact shown in Table 2. The a_5, with the highest strength and solid reliability, are especially relevant because they establish a direct causal link between the malicious application and the perpetrator's command-and-control channel. Such linkage is probative in court, as it can demonstrate not only the presence of malware but also the perpetrator's active control over the victim's data.

The a_2 also provides strong evidentiary value because it is a tangible artifact that can be verified through hashing, reverse engineering, and reproducibility testing. Its higher reliability score makes it admissible as digital evidence, since integrity and chain of custody can be more easily documented. This artifact, therefore, can serve as a primary exhibit in legal proceedings.

By contrast, the a_3 and the a_4, though technically strong in proving exploitation and stealth, carry slightly lower reliability scores. In legal terms, these artifacts are considered supporting evidence, useful to corroborate the malware's behavior but requiring cross-validation with stronger artifacts such as the APK and source code.

Table 2. Relevance to law enforcement

Artefact	Legal Relevance	Justification
a_1	Contextual Evidence	Indicates reconnaissance but weak probative value and low reliability; cannot stand alone in court.
a_2	Primary Evidence	Malware payload with verifiable hash; directly demonstrates weaponization and attack vector.
a_3	Supporting Evidence	Shows exploitation of Android permissions; corroborates malicious intent but less reliable alone.
a_4	Supporting Evidence	Confirms stealth installation; relevant but needs corroboration with APK and source code.
a_5	Primary Evidence	Direct causal link to attacker's C2 channel; high probative value for proving intent and control.

Finally, the a_1, while useful in reconstructing the reconnaissance phase, is the least probative . It has weaker relevance in court because phone numbers alone do not establish malicious intent or direct perpetrator involvement. Instead, they function primarily as contextual evidence to frame the beginning of the attack.

The results have three main implications for law enforcement and forensic practitioners:

(a) Evidence Prioritization: By quantifying v and r, investigators can prioritize artifacts with the greatest impact in court, ensuring that primary evidence is foregrounded while contextual artifacts are used strategically to frame the attack.

(b) Forensic Soundness: The scoring highlights the importance of acquisition integrity and reproducibility. For instance, the APK's high reliability score reflects that its authenticity can be independently verified through hashing and reverse engineering.

(c) Legal Admissibility: The clear distinction between artifact categories (primary, supporting, contextual) provides prosecutors with a structured rubric for presenting evidence that meets admissibility standards, reducing the risk of dismissal due to weak or unreliable data.

Overall, this scoring system provides law enforcement with a structured method to prioritize digital evidence: artifacts with high v and high r (a_5 and a_2) should be presented as primary evidence, while those with moderate scores (a_3 and a_4) serve to strengthen the narrative, and lower-scoring artifacts (a_1) act as supplementary context. This ensures that legal arguments rest on the most admissible and probative evidence, thereby enhancing the robustness of the prosecution's case.

VI. CONCLUSION

This study demonstrates that digital artifact analysis can systematically reconstruct the modus operandi of WhatsApp fraud, from reconnaissance to actions on objectives, when

mapped to the Cyber Kill Chain (CKC) and evaluated through the D4I Framework. By quantifying artifacts according to strength of evidence (v) and reliability (r), the analysis revealed that the malicious APK (a_2) and source code with Telegram bot token (a_5) constitute primary evidence with the highest probative value, while the Manifest.xml file (a_3) and hidden background application (a_4) function as supporting evidence, and contextual indicators such as sender information (a_1) provide limited legal weight. These findings underscore the importance of differentiating artifacts by evidentiary significance rather than treating all digital traces as equal.

This study is based on a simulated WhatsApp scam case, which, while representative of real-world attacks, may not capture the full complexity of actual investigations involving diverse devices, operating system versions, or cross-platform evidence. Additionally, the quantitative rubric, although systematic, is dependent on expert judgment in scoring sub-criteria, which may introduce some subjectivity.

Further studies should validate and refine this methodology across multiple fraud cases and platforms, such as Telegram, Signal, or Facebook Messenger, to test its generalizability. The integration of machine learning or automated scoring mechanisms could also reduce subjectivity and enhance reproducibility in assessing artifact reliability and evidentiary strength.

Beyond its academic contributions, this study provides law enforcement and forensic practitioners with a structured rubric for prioritizing digital evidence and improving legal admissibility. More broadly, the findings highlight the need for standardized frameworks and scoring systems in digital forensics policy, ensuring that courts, investigators, and policymakers adopt consistent criteria when evaluating digital artifacts in cybercrime cases. By doing so, digital forensic practice can become not only more scientifically rigorous but also more impactful in strengthening public trust in the security and governance of digital ecosystems.

REFERENCES

[1] S. Nishchal, “Forensic Analysis of WhatsApp: A Review of Techniques, Challenges, and Future Directions,” J Forensic Sci Res, vol. 8, no. 1, pp. 019–024, June 2024, doi: 10.29328/journal.jfsr.1001059.

[2] R. Nurdin and E. Ramadhani, “Investigasi Forensika Digital WhatsApp Scam dengan Menggunakan Framework D4I,” jurnal.mdp.ac.id, Mar. 2024, doi: 10.35957/jatisi.v11i1.6616.

[3] A. Dimitriadis, N. Ivezic, B. Kulvatunyou, and I. Mavridis, “D4I - Digital forensics framework for reviewing and investigating cyber attacks,” Array, vol. 5, p. 100015, Mar. 2020, doi: 10.1016/j.array.2019.100015.

[4] Y. Y. Lee, C. L. Gan, and T. W. Liew, “Thwarting Instant Messaging Phishing Attacks: The Role of Self-Efficacy and the Mediating Effect of Attitude towards Online Sharing of Personal Information,” IJERPH, vol. 20, no. 4, p. 3514, Feb. 2023, doi: 10.3390/ijerph20043514.

[5] W. Meng, T. Giannetsos, and C. D. Jensen, “Information and Future Internet Security, Trust and Privacy,” Future Internet, vol. 14, no. 12, p. 372, Dec. 2022, doi: 10.3390/fi14120372.

[6] G. Kim, U. Hur, S. Kang, and J. Kim, “Analyzing the Web and UWP versions of WhatsApp for digital forensics,” Forensic Science International: Digital Investigation, vol. 52, p. 301861, Mar. 2025, doi: 10.1016/j.fsidi.2024.301861.

[7] J. Son, Y. W. Kim, D. B. Oh, and K. Kim, “Forensic analysis of instant messengers: Decrypt Signal, Wickr, and Threema,” Forensic Science International: Digital Investigation, vol. 40, p. 301347, Mar. 2022, doi: 10.1016/j.fsidi.2022.301347.

[8] D. Schmutz, R. Rapp, and B. Fehrensen, “Forensic analysis of hook Android malware,” Forensic Science International: Digital Investigation, vol. 49, p. 301769, June 2024, doi: 10.1016/j.fsidi.2024.301769.

[9] P. Faruki, R. Bhan, V. Jain, S. Bhatia, N. El Madhoun, and R. Pamula, “A Survey and Evaluation of Android-Based Malware Evasion Techniques and Detection Frameworks,” Information, vol. 14, no. 7, p. 374, June 2023, doi: 10.3390/info14070374.

[10] C. Palma, A. Ferreira, and M. Figueiredo, “Explainable Machine Learning for Malware Detection on Android Applications,” Information, vol. 15, no. 1, p. 25, Jan. 2024, doi: 10.3390/info15010025.

[11] T. Al lelah, G. Theodorakopoulos, P. Reinecke, A. Javed, and E. Anthi, “Abuse of Cloud-Based and Public Legitimate Services as Command-and-Control (C&C) Infrastructure: A Systematic Literature Review,” JCP, vol. 3, no. 3, pp. 558–590, Sept. 2023, doi: 10.3390/jcp3030027.

[12] H. Heath, Á. MacDermott, and A. Akinbi, “Forensic analysis of ephemeral messaging applications: Disappearing messages or evidential data?,” Forensic Science International: Digital Investigation, vol. 46, p. 301585, Sept. 2023, doi: 10.1016/j.fsidi.2023.301585.