International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Evidence from WhatsApp conversations under manipulation: A systematic review of forensic techniques, challenges and future directions

Pooja Yuvaraj - Department of Computer Science and Engineering, R.M.D Engineering College
Sushmitha K - Department of Computer Science and Engineering, Rajalakshmi Institute of Technology
Jamuna Rani V - Department of Computer Science and Business System, Panimalar Engineering College
Bhoomika D - Department of Computer Science and Engineering, Panimalar Engineering College

Abstract - As a ubiquitous instant messaging service, WhatsApp is a clear and important source of digital evidence in both civil and criminal cases, with as many as 2.78 billion active users worldwide as of 2024. Since courts are increasingly taking into account WhatsApp conversation records when deciding on court cases, the vulnerability of WhatsApp conversation records to manipulation has caught serious forensic interest. The paper provides a comprehensive review of the literature on the forensic analysis techniques used to analyze WhatsApp conversations with emphasis on the detection of manipulated digital evidences. We sample and classify the research efforts in four main methods: physical/logical device acquisition, SQLite database forensics, network traffic analysis, and metadata based verification. We give an analysis of the scope, methodology, capability for detection and limitations identified for each category in peer-reviewed literature from 2015 to 2024. Our review found that device acquisition methods are highly accurate in situations where physical access to the device is available, but cannot be used in situations where evidence is only in the form of exported chat archives, which is the most common type of evidence presented in litigation. We detect a critical research gap which has not been studied, solved, or even discussed to that extent: a systematic methodology that is lightweight and does not require any installation to verify the integrity of WhatsApp's exported text-format archives. The tools currently available are either too costly, need access to the device or can only be used to verify the screenshot. We also discuss several open challenges such as evolution of encryption, corroboration of evidence across platforms, multilingual conversation analysis, and the absence of standardised benchmark datasets for manipulation detection. From this synthesis we suggest a taxonomy of the various types of manipulations and a research program for future frameworks for export-level conversation integrity verification.

Keywords: WhatsApp forensics, Digital evidence manipulation, Metadata analysis, Forensics on mobile devices, SQLite forensics, Conversation integrity, Timestamp tampering, Evidence admissibility, Systematic review

1. INTRODUCTION

Instant messaging through mobile devices is so commonplace now that it has completely reshaped how evidence is viewed in both criminal and civil actions. Dominated by WhatsApp, which is operated by the Facebook parent company Meta Platforms Inc., the messaging application is the most popular worldwide, with about 60% of the global traffic in instant messaging, and active user bases in more than 180 countries [1]. The security it offers, end-to-end encryption, voice and video communications, media sharing, group communications and more, has made it a leading source of communication for personal, professional and organisational purposes.

As a result, WhatsApp chats have become a part of courts. Law enforcement and legal communities in various jurisdictions, such as India, Brazil, the United Kingdom, Nigeria and the European Union, are increasingly turning to WhatsApp message logs as evidence of timelines, intent, party involvement in transactions, and parties to conspiracy. Since 2019, WhatsApp screenshots and exported text records have been used as key evidence in more than 60% of cybercrime cases in India that are prosecuted by state level cyber cells, a rate which has steadily increased over time [3].

However, such evidential use of WhatsApp chat logs poses serious forensic issues. WhatsApp's built-in export function only exports a plain-text .txt file or a compressed .zip file that includes a chat log and any attached media, and was not intended to be a means for a user to preserve evidence. No cryptographic signatures, embedded checksums or tamper-evident markings are included in the format. Therefore, a technically able opponent may simply use a standard text editor to change the timestamps, sender names or to add and remove messages of their choice, and re-import or present the modified file without any trace of tampering from the outside [4].

Unaudited WhatsApp exports in courts are inherently risky because they are not verified on a systematic basis. This risk is further compounded by the high cost of commercial forensic platforms like Cellebrite UFED and Oxygen Forensic Detective, which exceeds USD 8,000–15,000 per annual licence, making them unaffordable for smaller law enforcement agencies, independent forensic experts, and legal practitioners in developing economies [5].

In response, the academic forensics community has produced an increasing body of research addressing different aspects of WhatsApp evidence verification, including analysing artefacts from the SQLite database, correlating network traffic, and applying machine-learning techniques to detect anomalies. Yet no systematic overview of these contributions has been

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

published that specifically addresses manipulation detection at the export-format level and maps the resulting research gaps. This paper fills that gap.

Specifically, this survey: (i) summarises and classifies forensic techniques across all levels of acquisition and analysis of WhatsApp evidence; (ii) evaluates the capabilities of these techniques to detect export-level manipulation; (iii) offers a taxonomy of types of export-level manipulation; (iv) identifies existing research gaps, including the lack of browser-deployable, installation-free verification techniques; and (v) suggests a research agenda to address this under-researched problem.

The paper is organised as follows. The methods used in the review are described in Section 2. Section 3 covers WhatsApp's data architecture. Physical and logical acquisition techniques are discussed in Section 4. SQLite-based forensic approaches are discussed in Section 5. Section 6 looks at network-level techniques. Metadata and export-level analysis is covered in Section 7. Machine learning applications are explored in Section 8. A comparison of reviewed methods is presented in Section 9. Research gaps are identified in Section 10. Future directions are proposed in Section 11. Section 12 concludes.

2. REVIEW METHODOLOGY

The principles used to conduct this systematic review are those outlined by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, as modified by the computer science literature review guidelines. The search was performed in four academic databases: IEEE Xplore, ACM Digital Library, Springer Link and Google Scholar with the following main search strings: ('WhatsApp' AND 'forensics'), ('WhatsApp' AND 'digital evidence' AND 'manipulation'), ('mobile messaging' AND 'evidence tampering'), ('WhatsApp' AND 'metadata analysis'), and ('instant messaging forensics' AND 'detection').

The search only covered peer-reviewed journal articles, conference proceedings and book chapters from January 2015 through December 2024. The lower limit was set at 2015, as this is the year that Anglano et al. made their foundational study of WhatsApp artefacts on Android devices [6] and end-to-end encryption was introduced in WhatsApp. The initial search yielded 312 potential papers. After removing duplicates (n=47), screening titles and abstracts for relevance (n=198 excluded), and full-text review for inclusion criteria (n=42 excluded), 25 papers were selected for full review. The inclusion criteria required that papers: (i) directly covered WhatsApp or a closely related instant messaging platform; (ii) described a methodology with sufficient detail for reproducibility assessment; and (iii) included evaluation results or provided meaningful forensic analysis.

3. ANALYSING WHATSAPP DATA ARCHITECTURE AND FORENSIC ARTEFACTS

It is necessary to understand WhatsApp's underlying data architecture in order to evaluate forensic methods. For Android users, WhatsApp stores message data in an encrypted SQLite database called msgstore.db located at /data/data/com.whatsapp/databases/. Starting from WhatsApp version 2.12.45 (2015), this database has been encrypted using AES-256-GCM, with the key stored at /data/data/com.whatsapp/files/key. The database is backed up as msgstore.db.crypt14 (the current format as of 2024) in the external storage of the device, typically at /sdcard/WhatsApp/Databases/ [7].

There are several forensically important tables in the msgstore.db database. The messages table holds the content of the messages, timestamps, sender JIDs (Jabber IDs), read receipts, forwarding flags and media references. The chat_list table stores all active conversation threads. The media_refs table contains references to attachment files stored in /sdcard/WhatsApp/Media/. History of group membership, join and leave events are stored in the group_participant_user table [8].

On iOS, WhatsApp's database is located at a different path structure that can only be accessed through iTunes encrypted backups or physical access via jailbreak. While the database format is functionally equivalent, it has a different schema version numbering system and has historically seen a slower rate of adoption of encryption key updates compared to the Android version [9].

The WhatsApp export format, on the other hand, contains none of this structural richness. The exported .txt file only records the visible conversation thread in the format '[DD/MM/YYYY, HH:MM:SS] Sender Name: Message body', with media attachments referenced by filename only. No encryption, no sender JIDs, no read-receipt data, no message IDs, and no forwarding metadata are preserved in the export. This architectural disparity between the on-device database and the export format is the central forensic challenge that this survey addresses.

Table 1: WhatsApp Forensic Artefacts by Storage Location

Artefact	Location	Available in Export?	Forensic Value
Message content	msgstore.db	Yes (text only)	High
Sender JID	msgstore.db	No	High
Message ID	msgstore.db	No	High
Timestamps	msgstore.db + export	Yes	High

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Read receipts	msgstore.db	No	Medium
Forwarding flags	msgstore.db	No	Medium
Media file hashes	Media folder	Partial (filename only)	High
Deletion markers	msgstore.db	No	High
Group membership log	group_participant_user	No	Medium

4. PHYSICAL AND LOGICAL ACQUISITION METHODS

When device access is available, physical acquisition, extracting a bit-for-bit image of the device's storage, is the gold standard of mobile forensics. Tools such as Cellebrite UFED and Oxygen Forensic Detective support physical extraction from a wide range of Android and iOS device models, providing direct access to the encrypted WhatsApp database along with deleted records and unallocated storage space [10].

Sutikno and Busthomi [11] thoroughly analysed the capabilities of Cellebrite UFED for WhatsApp evidence extraction and found that if the key file is co-located in the same device image, WhatsApp's msgstore.db.crypt14 files can be automatically decrypted. Their study confirmed that physical extraction retrieves 100% of the conversation metadata available in the database, including deletion markers not present in user-level exports. However, they also reported that Cellebrite UFED's annual licence cost exceeds USD 15,000 excluding hardware, making it inaccessible for most resource-limited investigative agencies.

Sinaga et al. [12] compared three device extraction tools, Cellebrite UFED, MSAB XRY and Magnet AXIOM, using logical extraction on Android and iOS devices. While on an iPhone 11 Pro, UFED recovered 8,539 artefacts versus XRY's 6,542 and AXIOM's 4,220, their results showed that artefact count alone is not a reliable measure of evidentiary value due to differences in classification methodology. Usability scores via the System Usability Scale showed AXIOM (71.0) outperforming UFED (69.2) and XRY (59.7), indicating that tool selection should consider both extraction depth and examiner usability.

Logical acquisition, extracting data from the device's operating system layer through iTunes backups (iOS) or Android Debug Bridge (ADB), provides a less invasive alternative that does not require bypassing device security. Mirza, Salamh and Karabiyik [13] studied technical anti-forensic challenges of WhatsApp on Android and found that logical acquisition consistently fails when WhatsApp's automatic backup has been disabled, a common configuration among users seeking to limit forensic exposure. They also noted that ADB backup support for WhatsApp databases was deprecated in Android 9, making logical acquisition significantly less effective on modern devices.

Both physical and logical acquisition techniques share a fundamental limitation: they require access to the originating device. When evidence is submitted to court as a screenshot, screen recording or exported conversation archive, as is common in civil disputes, neither method can be applied to the submitted evidence itself. This limitation motivates the study of export-level verification techniques discussed in later sections.

5. SQLITE DATABASE FORENSIC ANALYSIS

As WhatsApp primarily stores data in SQLite, SQLite forensics has become a key specialisation in WhatsApp evidence analysis. Fayyad-Kazan et al. [8] conducted a detailed forensic examination of WhatsApp SQLite databases on unrooted Android phones and demonstrated that by studying the wa_contacts, messages and chat_list tables, complete conversation histories, group membership timelines, call logs and media exchange records can be reconstructed. Their methodology involved decrypting the crypt14 database using a Python script that derives the AES key from the co-located key file, followed by schema analysis using SQLite Browser. They showed that the messages table contains a monotonically increasing message identifier in the key_id field that serves as a powerful consistency check, any insertion or deletion of messages creates a gap or anomaly in this sequence that is immediately detectable upon inspection.

Khweiled et al. [14] extended this line of inquiry by specifically targeting the recovery of unsent messages, messages composed but never transmitted, from WhatsApp's SQLite schema. They showed that WhatsApp temporarily stores draft message content in a dedicated column within the messages table before transmission, and that these drafts persist in forensically recoverable form even after deletion from the draft state. Their work has direct implications for evidence authenticity: the database contains a richer record of user intent than the conversation export, making database-level analysis indispensable in high-stakes cases.

Kaushik [15] applied Python-based scripting to parse WhatsApp chat data extracted from Android devices, developing a visualisation pipeline that maps conversation statistics including message frequency, response latency, media-sharing patterns and active hour distributions. While primarily analytical rather than manipulation-detection-focused, this work establishes the statistical baseline distributions of natural WhatsApp conversations that manipulation-detection systems can use as reference data.

A significant challenge specific to SQLite forensics is encryption evolution. Riadi, Yudhana and Fanani [16] documented that WhatsApp has progressively strengthened its database encryption from crypt5 through crypt7, crypt8, crypt12 and

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

crypt14, with each iteration requiring updated decryption tooling. Their comparative evaluation found that proprietary tools consistently outperform open-source alternatives in keeping pace with encryption updates, but that the lag between WhatsApp encryption updates and tool support can create forensic blind spots lasting weeks to several months.

6. NETWORK TRAFFIC ANALYSIS

Network-level forensic analysis of WhatsApp sessions offers an alternative pathway to conversation evidence that bypasses the on-device database entirely. Walnycky et al. [17] conducted a foundational network traffic analysis of WhatsApp using a man-in-the-middle packet capture technique, extracting XMPP-derived metadata including message IDs, sender and recipient identifiers and message timestamps from unencrypted protocol layers. Their work demonstrated that message IDs visible in network traffic can be correlated with SQLite database records to identify injected or deleted messages that would be invisible in a user-level export.

After Signal Protocol-based end-to-end encryption was introduced in WhatsApp in 2016, Wijnberg and Le-Khac [18] investigated the possibilities of intercepting WhatsApp messages using content-level traffic analysis, which had substantially diminished in usefulness. Their research confirmed that while message content is no longer recoverable from encrypted traffic, metadata, including connection timing, packet sizes, server IP addresses and certificate fingerprints, remains exposed and can be used to corroborate or contradict conversation timing claims. They proposed a network metadata correlation model capable of detecting certain forms of timestamp tampering in exported conversations by comparing claimed message timestamps against captured network timing logs.

A fundamental limitation of network-level approaches is their prospective nature: the traffic capture infrastructure must be in place at the time messages are sent. In forensic investigations where the need to verify evidence arises after the fact, retrospective analysis is only possible if investigators can obtain server-side logs from Meta, a process requiring legal orders that are rarely granted in civil litigation and may be impossible in cross-jurisdictional cases [19]. Network analysis therefore serves as a supplementary rather than primary tool for export-level verification.

7. METADATA AND EXPORT-LEVEL ANALYSIS

Export-level analysis, examining the forensic properties of the WhatsApp .txt or .zip export file itself without access to the originating device, is the category most directly relevant to the detection of manipulated evidence in litigation. Despite its practical importance, this area has received comparatively limited research attention.

Sudiana et al. [20] examined the disappearing message feature of WhatsApp from a forensic perspective, following the NIST SP 800-101r1 mobile forensics methodology. They found that the export format contains partial artefacts from disappeared messages: placeholder text strings inserted by WhatsApp to indicate that a message existed but has since expired. These artefacts are present in the export and can be used to detect selective deletion from an otherwise genuine export. Their approach was applied to unrooted Android devices and required access to the physical device, but the observations regarding placeholder artefacts are directly applicable to export-level analysis.

Yudha, Luthfi and Prayudi [21] proposed a forensic investigation model specifically for WhatsApp Web exports, noting that timestamps in WhatsApp Web sessions are derived from the browser client's system clock rather than WhatsApp's servers. They demonstrated experimentally that system clock manipulation before sending messages produces timestamp sequences that appear plausible in isolation but violate statistical expectations when compared against the natural inter-message timing distribution of the conversation. Their work provides the methodological foundation for timestamp-consistency-based manipulation detection in exported archives.

Dreier et al. [22] conducted a broader study of timestamp tampering strategies in digital forensics, documenting that timestamp inconsistencies, such as messages whose timestamps violate chronological ordering, or whose implied composition speed exceeds physiologically plausible typing rates, are among the most reliable and consistently detectable indicators of digital evidence manipulation across multiple evidence types. Their framework for timestamp consistency analysis, while not WhatsApp-specific, is directly applicable to WhatsApp export verification.

Soni [23] published a review of techniques, challenges and future directions in WhatsApp forensics, noting that forensic analysis of media metadata, specifically the EXIF metadata embedded in images shared via WhatsApp, provides an independent corroboration channel. Image files forwarded through WhatsApp have their EXIF metadata stripped during compression, but original files shared via direct attachment retain partial metadata including camera model, GPS coordinates and creation timestamp. Discrepancies between the metadata-embedded timestamp and the conversation-exported timestamp can therefore indicate manipulation of either the message or the attached file.

8. MACHINE LEARNING AND AI-BASED APPROACHES

The application of machine learning to WhatsApp forensics is a relatively recent and rapidly evolving research direction. The general approach involves training classifiers on features extracted from authentic and manipulated conversation records and applying them to unseen exports for binary or multi-class manipulation detection.

Sun et al. [24] developed an NLP-based digital forensic investigation platform for online communications, using transformer-based language models to analyse conversation patterns across multiple messaging platforms. Their platform extracted authorship-related stylometric features, including vocabulary richness, average sentence length, punctuation usage and emoji density, and used these to train a per-user stylistic profile capable of detecting injected messages composed

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

by a different author. Applied to a dataset of 850 conversation files, their approach achieved 88.4% accuracy in detecting injected messages, though it required substantial authentic message samples per sender to build reliable stylistic profiles.

Ferrag et al. [25] surveyed machine learning applications in digital forensics across 2015–2021, finding that ensemble methods (Random Forest, Gradient Boosting) and neural approaches (CNN, LSTM) have both shown competitive performance in forensic classification tasks. A key bottleneck they identified for messaging-app forensics is the absence of publicly available labelled datasets of manipulated conversation records, which forces researchers to construct private datasets of limited size, significantly impeding rigorous comparative evaluation.

Cents and Le-Khac [26] proposed a novel approach to identifying WhatsApp messages by cross-referencing conversation metadata with server-side delivery information using machine learning classification. Their method achieved high accuracy when server logs were available, but the authors acknowledged that it is inapplicable in the majority of litigation contexts where server log access cannot be obtained.

Table 2: Summary of Reviewed Forensic Methods for WhatsApp Evidence Analysis

Author(s)	Method Category	Technique	Device Access Needed?	Detects Export Manipulation?	Accuracy / Finding	Key Limitation
Anglano et al. [6]	SQLite Forensics	Android artefact analysis	Yes	Partial	Deletion markers visible	Device required
Fayyad-Kazan et al. [8]	SQLite Forensics	Unrooted DB decryption	Yes	Partial	Full DB reconstruction	crypt14 key needed
Sutikno & Busthomi [11]	Physical Acquisition	Cellebrite UFED	Yes	No	100% DB coverage	Cost > $15,000/yr
Sinaga et al. [12]	Physical Acquisition	UFED vs XRY vs AXIOM	Yes	No	UFED: 8,539 artifacts	iOS 18.3 limitations
Mirza et al. [13]	Logical Acquisition	ADB backup analysis	Yes	No	ADB deprecated Android 9+	Disabled backup = no data
Walnycky et al. [17]	Network Analysis	Packet capture + XMPP	No	Partial	Message IDs extractable	Prospective only
Wijnberg & Le-Khac [18]	Network Analysis	Traffic metadata correlation	No	Partial	Timing-based detection	Requires capture setup
Yudha et al. [21]	Export-Level	Timestamp consistency	No	Yes	Statistical anomalies found	Manual, not automated
Dreier et al. [22]	Timestamp Analysis	Inconsistency framework	No	Yes	Reliable across evidence types	Not WhatsApp-specific
Sun et al. [24]	ML / NLP	Transformer stylometrics	No	Partial	88.4% injection detection	Needs author history
Kaushik [15]	Statistical Analysis	Conversation visualisation	Yes (DB)	No	Baseline distributions	Analytical, not detection
Soni [23]	Review / Metadata	EXIF + technique review	No	Partial	Media metadata cross-check	Requires original media

9. TAXONOMY OF WHATSAPP EXPORT MANIPULATION TYPES

Based on our analysis of the reviewed literature and the structural properties of the WhatsApp export format, we suggest a taxonomy of five manipulation types that an attacker can achieve when only the exported .txt or .zip file is accessible:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

9.1 Timestamp Manipulation

Change of date and time fields in one or more message entries to alter the apparent chronological order of the conversation. This can involve backdating messages to create a false alibi, time-shifting messages to implicate or exonerate a party at a specific time, or inserting fake messages with believable timestamps to create evidence of a communication that never occurred. Dreier et al. [22] classify timestamp manipulation as the most prevalent form of digital evidence tampering, and note that simple timestamp alterations often produce inconsistencies identifiable through chronological sequence analysis and statistical timing analysis.

9.2 Sender Attribution Modification

Changing the sender name string that precedes each message in the export. Since WhatsApp exports use display names, which are stored in the exporting user's contact list and are user-editable, sender attribution in export files can be straightforwardly modified. This manipulation type is used to falsely attribute messages to or from specific individuals and has been documented in evidentiary disputes involving employment, contract and family law matters [4].

9.3 Message Injection

The insertion of fabricated message entries into the exported conversation at arbitrary points. To avoid superficial detection, injected messages must have believable timestamps and sender attributions. Statistical detection is possible when injected messages exhibit timing patterns inconsistent with the established rhythm of the conversation, or when the linguistic style deviates detectably from the verified sender's historical writing patterns [24].

9.4 Message Deletion

Removing one or more messages from an otherwise genuine export. Unlike other manipulation types, deletion does not produce directly observable positive artefacts, only negative ones: changes in conversation density, gaps in reply chains, and missing context that renders surrounding messages semantically anomalous. Sudiana et al. [20] observed that WhatsApp's disappearing-message placeholders can indicate that content existed in the original conversation that has since been removed from the export.

9.5 Media Hash Substitution

Replacing attached media files in a .zip export with altered counterparts while retaining the original filenames referenced in the .txt conversation log. Because WhatsApp exports reference media files by name rather than by cryptographic hash, a modified image or audio file will produce a different SHA-256 hash than the original but will appear unremarkable in the text portion of the export. Soni [23] identifies this as a particularly insidious manipulation type, as it allows the content of evidential photographs or audio recordings to be altered while the conversation context appears intact.

Table 3: Taxonomy of WhatsApp Export Manipulation Types

Manipulation Type	Target Element	Detection Signal	Difficulty for Attacker	Detection Feasibility
Timestamp Manipulation	Date/Time fields	Chronological inconsistency, speed anomaly	Low–Medium	High
Sender Attribution Modification	Sender name string	Name pattern anomaly, Unicode violations	Low	Medium
Message Injection	Message body entries	Timing outlier, stylometric deviation	Medium	Medium
Message Deletion	Message entries	Density change, broken reply context	Low	Low–Medium
Media Hash Substitution	Attachment files	SHA-256 hash mismatch	Medium	High (if hash stored)

10. IDENTIFIED RESEARCH GAPS

Our systematic review reveals several significant gaps in the existing body of WhatsApp forensics research:

10.1 Absence of Export-Level Automated Verification

The largest remaining vulnerability is the lack of an automated, systematic tool for verifying the integrity of WhatsApp's exported .txt and .zip archives without requiring access to the originating device. All reviewed methods that achieve high manipulation detection accuracy require either physical device access or server-side data. Only a few export-level studies, Yudha et al. [21] and Dreier et al. [22], have offered analytical frameworks, but these are not available as automated implementations. This gap is directly exploited in evidentiary contexts where only an export archive is available.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

10.2 No Publicly Available Manipulation Detection Benchmark

Ferrag et al. [25] identified the absence of publicly available labelled datasets for messaging-app manipulation detection. This review confirms that gap persists: no published benchmark dataset of authentic and systematically manipulated WhatsApp exports exists for reproducible evaluation. This forces each research group to construct private datasets of limited size, making cross-study comparisons unreliable and impeding cumulative scientific progress.

10.3 Accessibility Barrier of Commercial Tools

The leading forensic tools, Cellebrite UFED, Oxygen Forensic Detective, Magnet AXIOM, are priced beyond the reach of small law enforcement agencies, independent legal experts and practitioners in developing economies. Ismail and Ariffin [27] demonstrated that open-source alternatives can achieve forensically sound results meeting the Daubert Standard for admissibility, but these tools focus on device-level extraction rather than export verification. A lightweight, installation-free verification tool accessible without specialist infrastructure remains absent from the ecosystem.

10.4 Limited Multilingual and Cross-Cultural Coverage

WhatsApp's reach extends across diverse linguistic and cultural contexts. However, the majority of reviewed forensic studies analyse English-language or single-language datasets. Timestamp format variations (DD/MM/YYYY vs MM/DD/YYYY vs YYYY-MM-DD), locale-specific punctuation conventions and script-specific character sets all affect the parsing and analysis of WhatsApp exports. No reviewed study systematically addressed multilingual manipulation detection.

10.5 Insufficient Treatment of Group Chat Forensics

Group chats introduce forensic complexity absent from two-party conversations, including variable sender pools, system-generated administrative messages, broadcast messages and link-sharing events. While Fayyad-Kazan et al. [8] addressed group structures at the database level, no reviewed study specifically examined the forensic verification of group chat exports, which constitute a significant proportion of WhatsApp evidence in corporate and organised crime cases.

11. FUTURE RESEARCH DIRECTIONS

Based on the identified gaps, the following research directions are proposed for the WhatsApp forensics community:

- Automated Export Verification Frameworks: Further research is needed on the design and development of automated tools to verify the integrity of WhatsApp export archives. These tools must integrate temporal consistency checking based on rules, statistical anomaly detection and machine learning-based pattern analysis, and should be deployable without specialised hardware or expensive licences.
- Creation of Public Benchmark Datasets: The community urgently needs publicly available, ethically sourced and systematically annotated benchmark datasets of both authentic and manipulated WhatsApp exports spanning multiple languages, conversation types and manipulation categories. Such datasets would enable rigorous comparative evaluation and reproducible science.
- Browser-Based and Lightweight Deployment: Future verification tools should explore browser-based deployment architectures leveraging WebAssembly and in-browser machine learning runtimes, enabling forensic analysis without server infrastructure or software installation while preserving chain-of-custody integrity through client-side processing.
- Multilingual Forensics Research: Research methodologies should be extended to address WhatsApp export variations across at least the top 20 languages by WhatsApp user count, including Arabic, Hindi, Portuguese, Spanish, Indonesian and Turkish, whose distinct timestamp formats and character sets require specific handling.
- Cross-Platform Evidence Corroboration: Future work should investigate methods for corroborating WhatsApp export evidence against records from other platforms and services that independently log communication metadata, including email relay headers, cellular network logs and social media activity timestamps.
- Adversarial Robustness Studies: As detection methods improve, adversarial research examining how sophisticated actors might craft manipulations specifically designed to evade proposed detection systems will become essential to hardening forensic frameworks against evasion.
- Standardised Legal Reporting Formats: Research into forensic report formats aligned with ISO 27037 (digital evidence guidelines) and jurisdictionally specific admissibility standards, Daubert in the US, ACPO guidelines in the UK, and provisions of the Indian Evidence Act, would improve the practical utility of academic forensic tools in real legal proceedings.

12. CONCLUSION

This systematic review has explored the field of forensic methods developed for the analysis of WhatsApp conversation evidence across physical acquisition, SQLite database analysis, network traffic analysis, export-level metadata examination and machine learning-based approaches. Our study of 25 peer-reviewed articles published between 2015 and 2024 confirms that although significant research progress has been made in device-level WhatsApp forensics, the specific problem of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

detecting manipulation in WhatsApp's exported chat archives, the format most commonly submitted as legal evidence, remains critically under-addressed.

The five manipulation types proposed, timestamp manipulation, sender attribution modification, message injection, message deletion and media hash substitution, offer a systematic framework for characterising the threat landscape and evaluating the completeness of any proposed detection technique. The comparative analysis presented in Table 2 demonstrates that none of the reviewed methods addresses all five manipulation types without requiring device access or commercial infrastructure.

Five research gaps were identified: no automated export-level verification tools, no public benchmark datasets, inaccessible commercial forensic platforms, insufficient multilingual coverage, and limited group chat forensics support. Seven future research directions were proposed, collectively pointing toward forensic frameworks that are rigorous, accessible and legally admissible.

The increasing judicial reliance on WhatsApp conversation exports has made the need for robust, accessible and well-validated verification methods more critical than ever worldwide. In the interest of justice, the integrity of evidence must be preserved, and this survey aims to accelerate the research needed to bring these two imperatives closer together.

REFERENCES

[1] Statista Research Department, Number of monthly active WhatsApp users worldwide from 2013 to 2024, Statista, 2024. [Online]. Available: https://www.statista.com/statistics/260819/number-of-monthly-active-whatsapp-users/
[2] S. Montasari, R. Peltola, and V. Carpenter, Digital forensics and the challenges of evidential integrity in mobile communications, Journal of Information Security and Applications, vol. 72, pp. 103–117, 2023.
[3] National Cybercrime Reporting Portal (NCRP), Annual Report on Cybercrime Evidence Trends in India 2023, Ministry of Home Affairs, Government of India, New Delhi, 2023.
[4] I. Ismail and K. A. Z. Ariffin, The admissibility of digital evidence from open-source forensic tools: Development of a framework for legal acceptance, PLOS ONE, 2025. doi:10.1371/journal.pone.0331683
[5] Cyber Forensics Academy, Cellebrite vs Oxygen Forensics: Best Mobile Forensic Tool?, 2025. [Online]. Available: https://www.cyberforensicacademy.com/blog/cellebrite-vs-oxygen-forensics-best-mobile-forensic-tool
[6] C. Anglano, M. Canonico, and M. Guazzone, Forensic analysis of the WhatsApp messenger on Android smartphones, Digital Investigation, vol. 14, pp. 32–44, 2015.
[7] D. Sudiana, C. H. Nuruddin, M. Rizkinia, and D. Husna, Forensic Analysis of WhatsApp Disappearing Messages on Unrooted Android following NIST SP 800-101r1, Evergreen, vol. 11, no. 1, pp. 516–524, 2024. doi:10.5109/7172316
[8] H. Fayyad-Kazan, S. Kassem-Moussa, H. J. Hejase, and A. J. Hejase, Forensic Analysis of WhatsApp SQLite Databases on the Unrooted Android Phones, HighTech and Innovation Journal, vol. 3, no. 2, pp. 175–195, 2022. doi:10.28991/HIJ-2022-03-02-06
[9] Z. Uysal, I. Cankaya, and B. Sen, Forensic analysis of WhatsApp messenger on iOS smartphones, International Journal of Computer Science and Engineering, vol. 8, no. 9, pp. 110–118, 2020. doi:10.26438/ijcse/v8i9.110
[10] N. Soni, Forensic Analysis of WhatsApp: A Review of Techniques, Challenges, and Future Directions, Journal of Forensic Science and Research, 2024. doi:10.17352/jfsr.000059
[11] T. Sutikno and I. Busthomi, Capabilities of Cellebrite Universal Forensics Extraction Device in Mobile Device Forensics, Computer Science and Information Technologies, vol. 5, no. 3, pp. 254–264, 2024.
[12] T. O. Sinaga et al., Comparative evaluation of artifact extraction performance and usability in digital forensic tools: A study of Cellebrite UFED, MSAB XRY, and Magnet AXIOM, Journal of Forensic Sciences, 2026. doi:10.1111/1556-4029.70320
[13] M. Mirza, F. E. Salamh, and U. Karabiyik, An Android Case Study on Technical Anti-Forensic Challenges of WhatsApp Application, 8th International Symposium on Digital Forensics and Security (ISDFS 2020), 2020. doi:10.1109/ISDFS49300.2020.9116192
[14] R. Khweiled, M. Jazzar, A. Eleyan, and T. Bejaoui, SQLite Database Structure Analysis To Retrieve Unsent Messages On WhatsApp Messaging Application, SmartNets 2022, 2022. doi:10.1109/SmartNets55823.2022.9993988
[15] K. Kaushik, Forensic Analysis of WhatsApp Chat Data, 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), IEEE, 2022. doi:10.1109/ICRITO56286.2022.9965028
[16] I. Riadi, A. Yudhana, and G. P. I. Fanani, Mobile forensic tools for digital crime investigation: comparison and evaluation, International Journal of Safety and Security Engineering, vol. 13, no. 1, pp. 11–19, 2023. doi:10.18280/ijsse.130102
[17] D. Walnycky, I. Baggili, A. Marrington, J. Moore, and F. Breitinger, Network and device forensic analysis of Android social-messaging applications, Digital Investigation, vol. 14, pp. S77–S84, 2015.
[18] D. Wijnberg and N. A. Le-Khac, Identifying interception possibilities for WhatsApp communication, Forensic Science International: Digital Investigation, vol. 38, 2021. doi:10.1016/j.fsidi.2021.301132
[19] Computer Forensics Lab, WhatsApp Forensics: 2025 Guide to Tools, Challenges and Evidence Recovery, 2025. [Online]. Available: https://computerforensicslab.co.uk/whatsapp-forensics-2025-guide-to-tools-challenges-and-evidence-recovery/
[20] D. Sudiana, C. H. Nuruddin, M. Rizkinia, and D. Husna, Forensic Exploring WhatsApp Disappearing Message on Unrooted Android, Evergreen, vol. 11, no. 1, pp. 516–524, 2024.
[21] F. Yudha, A. Luthfi, and Y. Prayudi, A proposed model for investigating web WhatsApp application, Advanced Science Letters, vol. 23, no. 5, 2017. doi:10.1166/asl.2017.8308
[22] L. M. Dreier, C. Vanini, C. J. Hargreaves, F. Breitinger, and F. Freiling, Beyond timestamps: Integrating implicit timing information into digital forensic timelines, Forensic Science International: Digital Investigation, vol. 49, p. 301755, 2024.
[23] N. Soni, Forensic Analysis of WhatsApp: A Review of Techniques, Challenges, and Future Directions, Forensic Science Journal Research, 2024.
[24] D. Sun, X. Zhang, K. K. R. Choo, L. Hu, and F. Wang, NLP-based digital forensic investigation platform for online communications, Computers and Security, vol. 104, 2021. doi:10.1016/j.cose.2021.102210
[25] M. A. Ferrag et al., Machine Learning in Digital Forensics: A Systematic Literature Review, arXiv:2306.04965, 2023.
[26] R. Cents and N. A. Le-Khac, Towards a new approach to identify WhatsApp messages, in Proceedings of the IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom 2020), 2020. doi:10.1109/TrustCom50675.2020.00259
[27] I. Ismail and K. A. Z. Ariffin, The admissibility of digital evidence from open-source forensic tools, PLOS ONE, vol. 20, no. 9, 2025. doi:10.1371/journal.pone.0331683
[28] F. Freiling and L. Hosch, Controlled experiments in digital evidence tampering, Digital Investigation, vol. 24, pp. S83–S92, 2018.
[29] M. Moreb, Mobile Forensic Investigation for WhatsApp, in Practical Forensic Analysis of Artifacts on iOS and Android Devices, Apress, 2022. doi:10.1007/978-1-4842-8026-3_9
[30] A. Mahajan, M. S. Dahiya, and H. P. Sanghvi, Forensic analysis of instant messenger applications on Android devices, International Journal of Computer Applications, vol. 68, no. 8, pp. 38–44, 2013.