Audio Forensics on Smartphone with Digital Forensics Research Workshop (DFRWS) Method

Sunardi[1], Imam Riadi[2], Rusydi Umar[3], and Muhammad Fauzan Gustafi[4*]
[1]Department of Electrical Engineering, Universitas Ahmad Dahlan
[2]Department of Information System, Universitas Ahmad Dahlan
[3–4]Department of Informatics, Universitas Ahmad Dahlan
Yogyakarta 55166, Indonesia
Email: [1]sunardi@mti.uad.ac.id, [2]imam.riadi@is.uad.ac.id, [3]rusydi@mti.uad.ac.id,
[4]muhammad1807048015@webmail.uad.ac.id

***Abstract*—Audio is one of the digital items that can reveal a happened case. However, audio evidence can also be manipulated and changed to hide information. Forensics audio is a technique to identify the sound's owner from the audio using pitch, formant, and spectrogram parameters. The conducted research examines the similarity of the original sound with the manipulated voice to determine the owner of the sound. It analyzes the level of similarity or identical sound using spectrogram analysis with the Digital Forensics Research Workshop (DFRWS) Method. The research objects are original and manipulated files. Both files are in MP3 format, which is encoded to WAV format. Then, the live forensics method is used by picking up the data on a smartphone. Several applications are also used. The results show that the research successfully gets digital evidence on a smartphone with the Oxygen Forensic application. It extracts digital evidence in the form of two audio files and two video files. Then, by the hashing process, the four obtained files are proven to be authentic. Around 90% of the data are identical to the original voice recording. Only 10% of the data are not identical.**

***Index Terms*—Audio Forensics, Smartphone, Digital Forensics Research Workshop (DFRWS)**

I. INTRODUCTION

THE era of advanced technology makes the digital world familiar for people to send messages in the form of text, audio, and video. The ease of application on a smartphone opens up opportunities for someone to commit a crime. The crime involving technology is called cybercrime. For example, it can be a hoax, photo forgery, illegal transaction, and bullying [1].

Illegal transactions using voice and video messages are also digital crimes. Audio can be in the form of voice messages, the sound in the video, sound recordings, and recordings of the wiretapping. Audio files such as WAV, AU, and MP3 can be manipulated to protect the secret messages [2, 3]. Voice and video messages have many drawbacks, which can be manipulated using the software on a computer or smartphone application that can hide the voice owner's identity [4]. The form of manipulation of audio can be done by changing the timbre and pitch.

Presentation of digital evidence is used for authentication and correlation of the cases. The investigator carries out the stage to protect evidence and minimize the damage during the investigations so that the evidence is still original. Hence, the investigator can receive digital evidence of sound recordings using smartphones as recording devices [5]. In a court, the evidence is needed to resolve the case. If the audio evidence is unknown (the voice owner cannot be identified), it cannot be used.

In general, digital forensics aims to analyze the suitability or authenticity of the multimedia content with the original content. Analysis of audio, video, and image on forensics usually not to find evidence but to test the suitability or authenticity of the content of the evidence with the original content [6]. Various methods of computer and digital forensics are developed to deal with various types of cybercrimes. In addition, digital forensics methods are also developed to deal with the increasing number of cyber-attacks which have now become a global trend [7].

To handle data manipulation, it needs forensics methods. One of the methods is the Digital Forensics Research Workshop (DFRWS). It uses scientific methods with a basis for the maintenance, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence origi-

Received: Oct. 01, 2020; received in revised form: Feb. 01, 2021; accepted: Feb. 03, 2021; available online: March 16, 2021.
*Corresponding Author

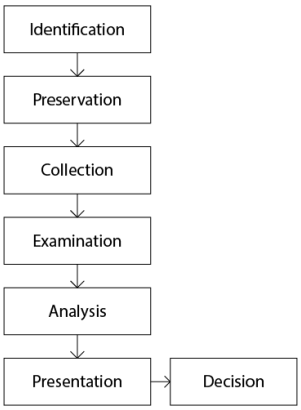

Fig. 1. Stages of the Digital Forensics Research Workshop (DFRWS).

nating from digital sources. It aims to facilitate or continue the reconstruction of events containing crimes. It also helps to anticipate unauthorized actions that are proven to interfere with planned operations [8, 9]. It is one of the forensics methods with complete stages in carrying out the forensics process and is widely used by forensics investigators. In general, the stages of the investigation process of digital evidence using DFRWS in both computers and smartphones have six main stages: identification, preservation, collection, examination, analysis, and presentation. The tools of Oxygen Forensics Suite 2014 and Praat are used in this study.

Increased smartphone users cause the increase in crimes involving Android smartphones. Therefore mobile forensics is needed to solve this problem [10–12]. Mobile forensics is a live forensics process [13]. It is the science of recovering digital evidence from mobile devices using data that are compatible with forensics [14]. The data taken from a cell phone can be used as evidence. This evidence can be used as a basis when investigating a case by law enforcement agencies. There is some evidence that can be extracted from mobile phones, such as contact numbers, call logs, messages, audio files, emails, Internet history, and other evidence related to the investigated case [15]. Logical or physical methods can extract the data. The purpose of logical extraction is to get data from the system by directly interacting with the device using tools or software specifically for mobile device forensics.

Moreover, audio forensics is a field of science that analyzes audio, sound, or recorded evidence. Voice recordings contain data and information similar to frequency characteristics used to determine the identity [16, 17]. In audio forensics, listening and analyzing often use visual and sound of spectrograms. It uses applications to investigate and construct evidence in court. However, it often leads to wrong conclusions due to data loss when transferring data or editing it electronically [18]. The analyzed audio can be through the parameters of pitch, formant, and spectrogram to show identities [19]. With the audio forensics method that has been developed, analysis can reduce the possibility of errors in the conclusions [20].

Spectrogram analysis aims to assess the identical common patterns and specific patterns on each analyzed word's formant. It looks at the energy levels in each formant. If the pronunciation of certain words from the manipulated recording and the original recording does not show a definite change, the pronunciation of the sentence has the same spectrogram [21–23].

Previous research on audio forensics is not carried out using forensics steps that have been determined by forensics associations such as the National Institute of Standards and Technology (NIST), Association of Chief Police Officers (ACPO), and DFRWS. Previous research uses high pitch and low pitch audio samples as the used audio [18]. The novelty in the research is using the proven forensics steps, namely DFRWS. It compares the original audio with the manipulated audio with the feature of converting humans' voices into robotic voices. The research aims to find the best way to determine the identity of the owner of the voice that has changed.

II. Research Method

A. Research Material

The research object is two audio files. There are original and manipulated files. The original audio file is 15-seconds long with a file size of 125 KB. Meanwhile, the manipulated audio file is 14-seconds long with a file size of 111 KB. Both files are in MP3 format, which is encoded WAV format later. Then, the live forensics method is used to recover the data on a smartphone.

B. Research Flow

The research uses a digital forensics method created by DFRWS. The DFRWS method helps obtain evidence and a centralized mechanism for recording the information. The data are collected in several stages, as shown in Fig. 1 [24]. First, the identification stage determines the needs for the investigation and searches for the evidence. Second, the preservation or maintenance stage requests the digital evidence to ensure the authenticity of the evidence and refutes sabotage claims. Third, the collection stage identifies certain parts of digital evidence and data sources. Fourth, the examination stage determines the data filtering on

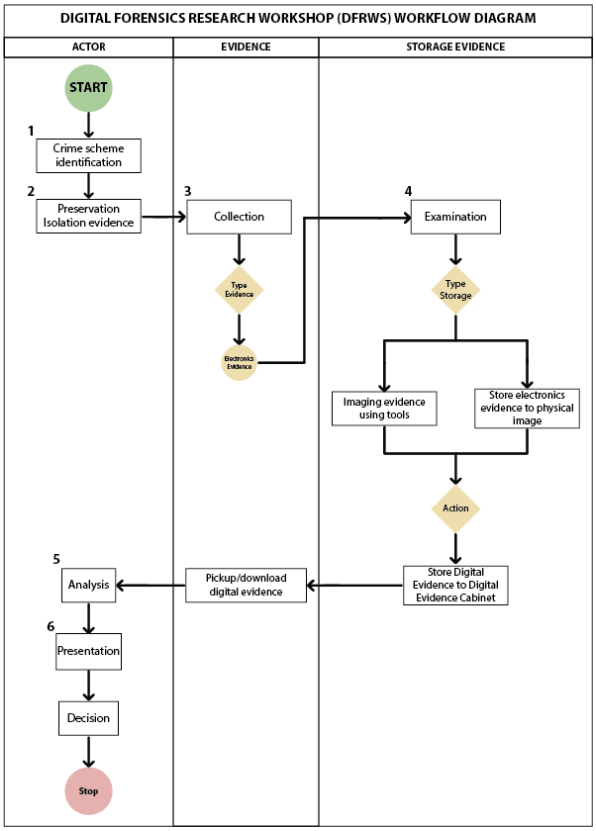

Fig. 2. Digital Forensics Research Workshop (DFRWS) workflow diagram.

certain parts of the data source by changing the shape of the data. However, it does not change the contents of the data to maintain authenticity. Last, the analysis stage determines where, by whom, how, and why the data are produced.

C. Spectrogram

A spectrogram is a form of visualization of each formant value. The value is equipped with an energy level or known as the formant bandwidth. It is very suitable for voice identification cases of sound falsification using the pitch shift technique to eliminate the character of the recording sound owner [22, 23]. Spectrogram analysis is used to identify the general patterns in words and specific patterns that are unique to each formant and the energy level of the analyzed syllables. If the distinctive spectral patterns of the two analyzed audios do not show any significant difference, it can be said that the two sounds are identical [19, 21, 23].

III. Results and Discussion

Figure 2 shows the DFRWS workflow in the research. It starts from a physical investigation of a smartphone that has been designated as evidence. This stage shows the specifications to help investigators to take the necessary data. Then, the authenticity of evidence is maintained to prevent data manipulation by shutting down the existing network on the smartphone according to the procedure by turning on the airplane mode. Next, data collection with the live forensics method is conducted. The data retrieval is done when the smartphone is turned on. It retrieves and selects the required data for investigators. After the data are extracted successfully, check the result of the hash value on the digital evidence. The analysis is done to determine the identity of the voice owner of the recording. The results can be legally proven according to the law. Audio forensics has several methods that need to be done, such as statistical analysis of pitch, formant, and bandwidth (hood ratio, graphical distribution, and spectrogram). Last, it reports the initial to the end result in the presentation stage containing physical evidence, digital evidence, and analysis results. The results can be considered whether the evidence is valid or not and show the most effective method in this case.

TABLE I
SMARTPHONE IDENTIFICATION.

Types	Specification
Brand	Samsung
Series	Galaxy
Model	Young
Model Number	SM-G130H
IMEI	352716071399351/00 352717071399359/00
Operating Sytem	Android
Operating Sytem Version	4.4.2 (KitKat)
Processor	ARM Cortex-A7

A. Identification Stage

According to the case scenario, this research uses a smartphone with digital evidence of two audio files. The research uses the Samsung Galaxy Young smartphone with several specifications. It has Android KitKat as the operating system and ARM Cortex-A7 as the processor. The specification can be seen in Table I.

B. Preservation Stage

The next stage is preservation. It keeps the digital evidence, ensures the authenticity of the evidence, and denies the claim that the evidence has been sabotaged. The process is done with the isolation technique of physical evidence and the creation of backups in the form of cloning or processing the image files of the evidence using smartphone cloning hardware to storage. First, the researchers isolate the smartphone from the communication network. Isolation needs to be done to

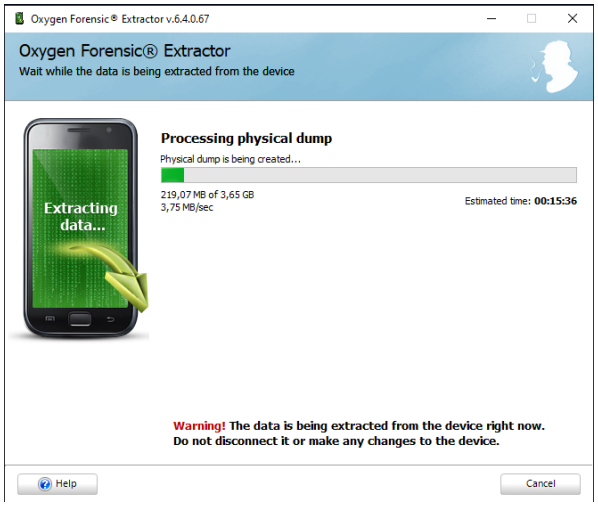

Fig. 3. Smartphone backup process.

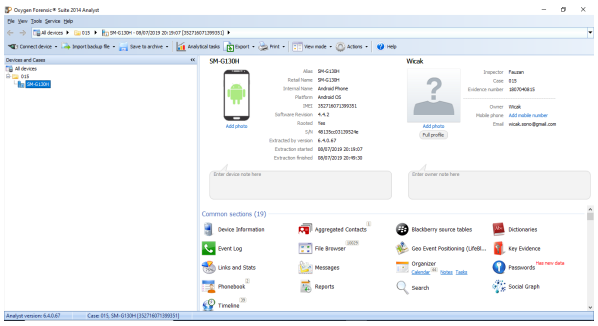

Fig. 4. Result of the backup process.

avoid things that can damage digital evidence or affect data integrity. It is done by using airplane mode in the smartphone.

C. Collection Stage

The third stage is the collection. It is the process of collecting the identification of a particular part of the digital evidence and data source. The retrieval process of digital evidence in smartphones has a high risk. In case of fatal errors, data and digital evidence can be lost or corrupted so that it is not readable. Therefore, it is necessary to do the preservation stage first (backup). This process is also called logical acquisition. The used tool to perform the backup process is the Oxygen Forensic which has a good backup system in a smartphone. Figure 3 shows the backup process with Oxygen Forensic. The result of this backup process is an image with varied document sizes. It depends on the amount of data on the smartphone. It is shown in Fig. 4.

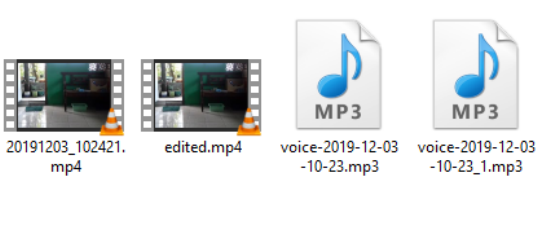

Fig. 5. Extraction Results.

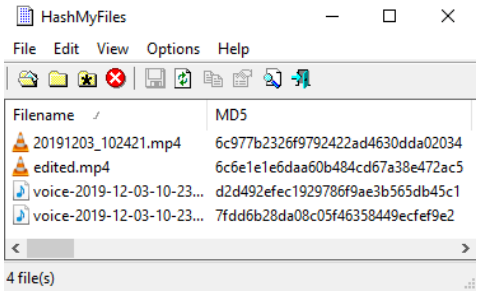

Fig. 6. Hashing Process.

D. Examination Stage

The extracted files are in the form of two video files and two audio files. From these results, the original video file is 20191203_102421.mp4. Meanwhile, the video that has been manipulated is edited.mp4. Then, the original audio file is voice-2019-12-03-10-23.mp3, and the manipulated audio file is voice-2019-12-03-10-23_1.mp3. The extraction results can be seen in Fig. 5.

E. Hashing Process

The extraction results that have been stored are continued to the hashing process. It determines the hash value in the files, as shown in Fig. 6. It shows the result of the hash process using HashMyFile. The hashing results in the original video file (6c977b2326f9792422ad4630dda02034) and the original audio file (d2d492efec1929786f9ae3b565db45c1). Meanwhile, the manipulated video file is 6c6e1e1e6daa60b484cd67a38e472ac5, and the manipulated audio file is 7fdd6b28da08c05f46358449ecfef9e2. The hash results can conclude that both the original and the fake files are different. The next step is that the video will be converted to audio in the form of WAV format. Then, the files are analyzed.

F. Analysis Stage

The data are processed by cutting it per syllable and taking the pitch value of the syllable. The analyzed

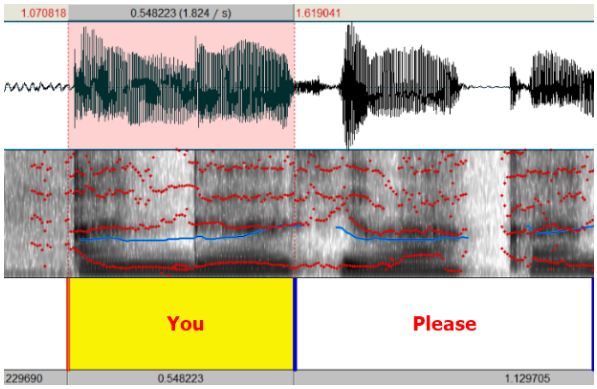

Fig. 7. Dividing audio process.

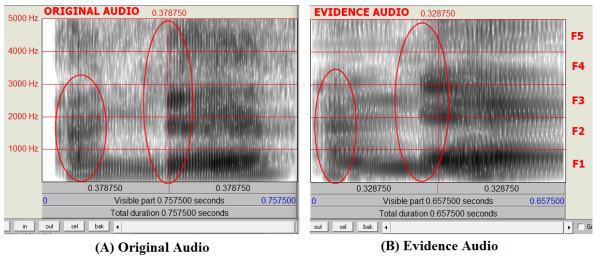

Fig. 8. Spectrogram analysis.

audio contains the words, "You, please make a transfer of twenty million after that the ordered goods will be sent three days later after transfer". Separating each word is done twenty times on each obtained audio file. The Praat application is used to divide the audio for each word. The example is in Fig. 7.

Spectrogram analyzes the specific patterns on the formant. If the spectral characteristic pattern on two analyzed sounds does not indicate a distinction, the sounds are identical [22]. Figure 8 shows the audios are identical. The red circle in Fig. 8 shows that the two audios are identical. There is a distinctive pattern occurring at the beginning of the 3000 Hz to 1000 Hz frequency, and the typical pattern appears again at 5000 Hz to 1000 Hz. Then, each formant in the audios is analyzed.

G. Presentation Stage

The final stage is the presentation stage done by redisplaying that information generated from the previous stage once the reserchers obtain the evidence from the analysis process. The Oxygen Forensic application can retrieve data from evidence in a smartphone as seen in Fig. 9. Oxygen Forensic can retrieve 227 data. There are 2 audio files, 2 video files, 5 deleted audio files, 9 deleted video files, 15 image files, 59 deleted

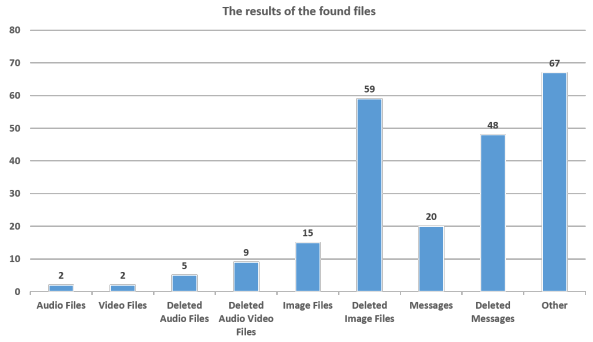

Fig. 9. The results of the found files.

TABLE II
THE RESULT OF THE ANALYSIS.

Syllables	Spectrogram Analysis
You	Matched
Please	Matched
Make a	Matched
Transfer	Matched
Of	Matched
Twenty	Matched
Million	Matched
After	Matched
That	Matched
The	Matched
Ordered	Matched
Goods	Matched
Will	Matched
Be	Matched
Sent	Matched
Three	Unmatched
Days	Matched
Later	Matched
After2	Unmatched
Transfer2	Matched

images, 20 message files, 48 deleted message files, and 67 other files.

The result of spectrogram analysis shows that the sound sample is identical to the evidence (see Table II). The formant frequency waves do not differ much between the two data. The results show that 90% of the data are identical to the original voice recording. Only 10% of the data are not identical. The research has the same results as previous studies.

IV. Conclusion

Audio is one of the digital items that can reveal a happened case. However, audio evidence can also be manipulated and changed to hide information. To handle data manipulation, it needs forensics methods. One of the methods is the DFRWS. It uses scientific methods with a basis for the maintenance, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence originating from digital sources.

Based on the obtained results, the research successfully gets digital evidence on a smartphone with the Oxygen Forensic Suite 2014 application. It extracts digital evidence in the form of two audio files and two video files. Then, by the hashing process, the four obtained files are proven to be authentic. The spectrogram analysis that samples the audio with video evidence is identical. The specific pattern on each formant does not have a significant difference, so that the results are identical to the audio sample. It proves that 90% of the sound of the evidence is identical to the original voice recording. Only 10% of the data are not identical.

There is a suggestion for further research related to audio forensics. Future research can use other forensic methods in the forensics analysis, such as pitch statistical analysis, formant and bandwidth analysis, Mel Frequency Cepstrum Coefficients (MFCC), or the Itakura-Saito Distance method.

ACKNOWLEDGEMENT

The research was supported by a grant from Universitas Ahmad Dahlan (Decree: PUPS-025/SP3/LPPM-UAD/IV/2019). The authors are indebted to the Faculty of Industrial Technology Universitas Ahmad Dahlan, which provided a good research facility and environment.

REFERENCES

[1] A. P. U. Siahaan, "Pelanggaran cybercrime dan kekuatan yurisdiksi di Indonesia," *Jurnal Teknik dan Informatika*, vol. 5, no. 1, pp. 6–9, 2018.

[2] N. Alsaidi, M. Alshareef, A. Alsulami, M. Alsafri, and A. Aljahdali, "Digital steganography in computer forensics," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 18, no. 5, pp. 54–61, 2020.

[3] A. Wicaksono, S. Adinandra, and Y. Prayudi, "Penggabungan metode Itakura Saito Distance dan Backpropagation Neural Network untuk peningkatan akurasi suara pada audio forensik," *JUITA: Jurnal Informatika*, vol. 8, no. 2, pp. 225–233, 2020.

[4] H. Wu, Y. Wang, and J. Huang, "Identification of electronic disguised voices," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 489–500, 2014.

[5] V. A. Hadoltikar, V. R. Ratnaparkhe, and R. Kumar, "Optimization of MFCC parameters for mobile phone recognition from audio recordings," in *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. Coimbatore, India: IEEE, June 12–14, 2019, pp. 777–780.

[6] R. Böhme, F. C. Freiling, T. Gloe, and M. Kirchner, "Multimedia forensics is not computer forensics," in *International Workshop on Computational Forensics*. Springer, 2009, pp. 90–103.

[7] A. C. K. Wardana, R. Pedrason, and T. B. Prasetyo, "Implementasi digital forensik Brunei Darussalam dalam membangun keamanan siber," *Peperangan Asimetrik*, vol. 4, no. 1, pp. 1–22, 2018.

[8] I. O. Ademu, C. O. Imafidon, and D. S. Preston, "A new approach of digital forensic model for digital forensic investigation," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 2, no. 12, pp. 175–178, 2011.

[9] I. Zuhriyanto, A. Yudhana, and I. Riadi, "Analisis perbandingan tools forensic pada aplikasi Twitter menggunakan metode Digital Forensics Research Workshop," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 5, pp. 829–836, 2020.

[10] I. Riadi, R. Umar, and A. Firdonsyah, "Forensic tools performance analysis on Android-based Blackberry Messenger using NIST measurements," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 5, pp. 3991–4003, 2018.

[11] A. Al-Sabaawi and E. Foo, "A comparison study of Android mobile forensics for retrieving files system," *International Journal of Computer Science and Security (IJCSS)*, vol. 13, no. 4, pp. 148–166, 2019.

[12] A. Wirara, B. Hardiawan, and M. Salman, "Identifikasi bukti digital pada akuisisi perangkat mobile dari aplikasi pesan instan "WhatsApp"," *Teknoin*, vol. 26, no. 1, pp. 66–74, 2020.

[13] R. Umar, A. Yudhana, and M. N. Faiz, "Experimental analysis of web browser sessions using live forensics method," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 5, pp. 2951–2958, 2018.

[14] W. Jansen and R. Ayers, "Guidelines on cell phone forensics," *NIST Special Publication*, vol. 800, no. 101, pp. 1–104, 2007.

[15] I. Riadi, Sunardi, and P. Widiandana, "Investigasi cyberbullying pada Whatsapp menggunakan Digital Forensics Research Workshop," *RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 730–735, 2020.

[16] R. R. Huizen, N. K. D. A. Jayanti, and D. P. Hostiadi, "Model evaluasi rekaman percakapan di

audio forensik," *Jurnal Sistem dan Informatika (JSI)*, vol. 11, no. 2, pp. 133–140, 2017.

[17] A. G. Boyarov and I. S. Siparov, "Forensic investigation of MP3 audio recordings," *Theory and Practice of Forensic Science*, vol. 14, no. 4, pp. 125–136, 2019.

[18] E. N. Gural and M. Pazarci, "A supporting method to detect manipulated zones in digitally edited audio files," *Medicine Science*, vol. 7, no. 2, pp. 1–4, 2017.

[19] V. R. C. Putri and Sunarno, "Analisis rekaman suara menggunakan teknik audio forensik untuk keperluan barang bukti digital," *Unnes Physics Journal*, vol. 3, no. 1, pp. 50–59, 2014.

[20] S. Camacho, D. M. Ballesteros, and D. Renza, "A cloud-oriented integrity verification system for audio forensics," *Computers & Electrical Engineering*, vol. 73, pp. 259–267, 2019.

[21] A. Subki, B. Sugiantoro, and Y. Prayudi, "Membandingkan tingkat kemiripan rekaman voice changer menggunakan analisis pitch, formant dan spectogram," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 5, no. 1, pp. 17–22, 2018.

[22] M. N. Al-Azhar, *Audio forensic: Theory and analysis*. Pusat Laboratorium Forensik Polri Bidang Fisika dan Komputer Forensik, 2011.

[23] P. Rose, *Forensic speaker identification*. CRC Press, 2002.

[24] A. L. Suryana, R. El Akbar, and N. Widiyasono, "Investigasi email spoofing dengan metode Digital Forensics Research Workshop (DFRWS)," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 2, no. 2, pp. 111–117, 2016.