Problems of Forensic Sciences 2019, vol. 120, 335–361

ISSN 1230-7483

THE CHARACTERISTICS OF POPULAR AUDIO RECORDING APPLICATIONS INSTALLED ON SMARTPHONES WITH AN ANDROID OPERATING SYSTEM IN RELATION TO FORENSIC AUDIO ANALYSES

Marcin MICHAŁEK

Institute of Forensic Research, Kraków, Poland

Abstract

The number of evidence recordings made with devices such as smartphones has significantly increased in recent years. Dedicated applications are used for making recordings, which in Poland are installed mainly on the Android operating system. During the research, the properties of free audio recording applications that are most often downloaded from the Internet were examined. Ten different applications were selected and installed on modern smartphones, followed by making test recordings for further analysis. Despite the fact that the selected applications are rather simple, they all have considerable potential. Half of them have an option to edit their recordings. They also enable recording in various conditions and audio formats, such as: 3GP, AAC, AMR, M4A (MP4 with audio), MP3 and WAV, without or with signal compression. It was established that the data structure of these kinds of audio files contains a lot of valuable information that is useful for authenticity analyses, especially in the case of the MPEG multimedia container. This information helps to determine the device and the application used, the exact recording parameters and time when it was made and even whether it has been edited. Using the MATLAB computing environment, our own software was developed for the automatic analysis of the audio files structure in order to search for relevant metadata. The paper also describes the methodology of proceeding with the evidence smartphones and recordings, and, moreover, a way of safely copying data from their memory for further analysis. The obtained results have substantial value for examining the authenticity of audio recordings as part of forensic expert opinions. The acquired recordings also expand the constantly developed database of audio files, which is a useful tool for analyzing the authenticity of recordings.

Keywords

Smartphone; Android recording application; Digital recording; Authenticity analysis; File structure.

Received 19 August 2020; accepted 28 September 2020

1. Introduction

Use of small battery-operated portable multi-media players has become widespread since the time when digital technology combined with the miniaturisation of devices enabled their construction. Popular MP3 or MP4 players, due to their small size and functionality, have dominated the digital multimedia equipment market. In addition to their ability to play audio or video, or open image files, they are often equipped with an audio recording function (Ho, Li, 2015; Savage, Vogel, 2014; Michałek, 2014). Devices dedicated to audio recording constitute a slightly different group; they are also portable and battery-powered, and are often referred to as digital voice recorders. Only a few years ago, more than three-quarters of digital audio recordings subjected to forensic audio testing were recorded using the above-mentioned devices (Michałek,

2016). The product market still offers portable media players and voice recorders (dictaphones), but currently evidence recordings from mobile devices usually originate from other sources.

The dynamic development of wireless mobile phone networks and services offered has, in effect, forced manufacturers to introduce phones onto the market that will make it possible to make use of the full potential of such networks. Smartphones, i.e. devices which combine the function of a mobile phone with that of a portable computer, have gained the greatest popularity, effectively replacing classical mobile phones. Telephone conversations and SMS (text messaging) currently constitute only a part of the capabilities of the average smartphone (Xia, Hsu, Liu, Liu, Ding, Zha, 2013). Usually, among the dozens of applications that have been pre-installed on a smartphone's operating system, at least one is strictly devoted to recording sound. In the event of absence of such an application or the need to use another one, Internet services offer many recording applications for the given selected smartphone model and its operating system, and allow you to independently obtain and install one tailored to your preferences. The usefulness, great adaptability to your own preferences and the general availability of these telephones have made them a basic device in everyday life with many applications (Xia et al., 2013). Recent expert opinions in the field of audio forensics suggest that there has been a significant change over the last few years and that currently, among all mobile devices, smartphones are now the primary source of evidence audio recordings, with a negligible share from portable media players and dictaphones. This has also resulted in a change in the audio saving formats adapted to the operating systems and capabilities of these devices.

Currently, around 2.7 billion people worldwide use 3 billion smartphones. Despite the dynamic situation on the market, for several years now the most popular models of smartphones, both globally and in Poland, have been products of the Korean company Samsung, and subsequent positions in the rankings are occupied by Xiaomi, Huawei or Apple (IDC, 2019, Telepolis, 2020). The current models work mainly under the control of Android or iOS operating systems, which have created a kind of duopoly and serve about 98% of the global smartphone market. Of all the smartphones sold, approximately 85% of devices globally, and 90% in Poland are fitted with the Android operating system (Orange, 2019). The research on smartphone devices presented later in the paper indicates that these statistics are also confirmed in expert practice. Dedicated applications installed by the manufacturer or user are used to record sound in such devices. Experience in the field of making test recordings with the evidence phones studied so far indicates that these are uncomplicated applications with basic functions enabling recording and simple editing. The popular website Google Play (previous name: Android Market) offers about 2.8 million applications, 95% of which are free of charge and can be installed and used in Android systems (AppBrain, 2020). In this service, searching by keywords, for example, "audio (sound) recorder" or "dictaphone", you can find dozens of free applications designed to record audio recordings, which you can easily install on your smartphone on your own. Applications that are pre-installed on new phones (for example: Samsung applications) can also be found on Internet websites.

In this work we have reviewed and selected the most popular applications designed for audio recording working under the Android operating system that can be installed using Internet resources. These applications were installed on a Samsung smartphone, their recording and editing capabilities were established, and then test recordings were made. The principles of procedure during the making of test recordings with the use of such devices have been described. Next, analyses of these applications, and of the parameters of the recordings made with them as well as the properties of audio files and metadata were carried out. Additionally, algorithms and a computer program working in the MATLAB computing environment were developed, designed as a tool for automatic analysis of the structures of saved audio files containing test recordings. The obtained results and observations were evaluated in order to determine the possibility of their use in tests of the authenticity of digital recordings that had mainly been recorded with modern mobile phones.

The main aim of the author was to analyse the properties of generally available applications with audio recording capability and to evaluate whether the obtained results could be applied in the work of audio forensics experts, without any intention to promote or criticise any particular brand, operating system or software.

2. Analysis of evidence and audio test recordings made with smartphones

Evidence digital recordings are not usually recorded with high quality professional recorders, but with commonplace equipment (Michałek, 2014). Audio recordings which are currently being examined in the

Section of Speech and Audio Analysis of the Institute of Forensic Research have relatively often been recorded with the use of the new generation of mobile phones that are popular today, i.e. smartphones. Other types of evidence devices containing (in their memories) recordings that have been commissioned for analyses are also submitted to the Section. The technical condition of each piece of evidence is assessed individually, as is the possibility of securing data from the memory. If a phone and its installed memory card (if there is one) are submitted for testing, then the contents of both are secured during the examination. If possible, an exact copy of the contents of the entire memory or its accessible part in the form of a so-called image is made. Such a copy is the most useful form of collection for further forensic audio analysis (Kajstura, Michałek, Trawińska, 2017). The image of the memory is treated as a reflection of the content of the medium (data carrier) together with the preservation of the file system structure, the properties of the files and their time signatures, which are of great importance for authenticity testing. Creating an image also makes it possible to determine whether there are any deleted files, including audio files, in a memory copy made in this way, and also makes it possible to attempt to recover them (Casey, 2010; Kajstura et al., 2017; Kowalski, Radziszewski, 2017; Willassen, 2005). For the analysed area of the source memory and its image, checksums should be calculated and compared with each other in order to verify the correctness of the copy. If it is not possible to make an image of the memory content, then the data is secured by making a logical copy of copyable files and directories from the analysed data carrier. It is also possible to explore the content of the memory of the evidence carrier, which allows the number and properties of the accessible files to be determined, which has particular significance when it is not possible to carry out the imaging process (Casey, Turnbull, 2011; Willassen, 2005). Securing the memory content of a device or carrier submitted for examination – and the way in which this is performed – is an important element determining the subsequent forensic audio analysis, in particular the analysis of the authenticity of recordings. When a phone is submitted for examination together with a SIM card, the SIM card is also secured as material evidence, but it is not examined for audio forensic purposes, as SIM card memory does not store multimedia files. A write-blocker type device is used for safe exploration and copying of the contents of a smartphone memory, preventing accidental modification of data (Kowalski, Radziszewski, 2017; Michałek, 2018).

On the basis of the smartphones analysed during routine expert opinion activities, it can be concluded that the vast majority of them were Samsung devices, as well as individual models of HTC and Manta brands – all working under the control of the Android operating system. One iPhone 3G device with an iOS operating system was also examined. Each of these devices had one or occasionally two audio recording applications installed. It was noted that these were simple to use applications allowing for basic activities during the recording: starting, pausing and stopping the recording. Some of them, mostly installed on Samsung phones, enabled simple editing of the saved recordings, consisting mainly in deletion of fragments of them and saving changes in source files or creating new ones. All applications analysed so far allowed renaming of files containing audio recordings in the course of or after saving them to memory. It was noted, however, that usually the names of files with evidence recordings are default names for installed applications, and correspond to the names of files containing test recordings made by evidence devices during their testing.

Analysis of collected data originating from the memory of smartphones showed that the audio recordings contained there were mainly saved in M4A and 3GP file formats. Audio data of these types of files together with the accompanying metadata were placed in a so-called MPEG-4 multimedia container, linked with ISO standards (ISO/IEC, 2003; ISO/IEC, 2010; ISO/IEC, 2015). M4A and 3GP formats and the MPEG-4 container are very common in modern mobile devices (3GPP, 2010; Gloe, Fischer, Kirchner, 2014; Ho, Li, 2015; Michałek, 2018). Files of this type are quite easily converted to uncompressed PCM WAVE format or even played in real time by standard software, which constitutes the basis for further forensic audio analysis (Kajstura et al., 2017; Michałek, 2014).

The recognition of recordings as evidence material and the scope of tests defined by judicial bodies makes it necessary to analyse their authenticity (Grigoras, 2005; Kajstura, Trawińska, Hebenstreit, 2005; Koenig, 1990; Korycki, 2011, 2016; Michałek, 2009). The definition of the authenticity of a recording announced by the Audio Engineering Society (AES) indicates the need to assess the recording's continuity and originality, and the method used to make the recording (AES, 1996). Carrying out test recordings is an important part of examinations of both the device itself and the authenticity of the evidence digital recording. They make it possible to determine whether the evidence device could have been used to record the evidence recording and to conduct a comparative analysis of

the parameters and structure of the file containing test recordings (Cooper, 2006; Kajstura et al., 2017; Koenig, Lacey, 2012; Michałek, 2016). They also make it possible to determine whether the device enables the recording of powerline hum and to reveal other possible characteristic features, such as a DC offset (Grigoras, Cooper, Michałek, 2009; Kajstura et al., 2005; Korycki, 2011).

The aspects discussed above also apply to recordings made by smartphones; however, it should be borne in mind that the emergence of new devices recording evidence recordings has forced the development of new ways of preparing and making test recordings. Smartphones are complex multifunction digital devices with their own operating system and with a high probability that every other model will have a different audio recording application installed. If the commissioning party has not submitted the device which was supposed to have been used to make the recording that has been submitted for analysis (most often, s/he has only provided a copy of the recording saved on a different data carrier) and the official request contains questions related to authenticity analysis, s/he should be requested to deliver such a device for testing.

In the process of preparing test recordings on the smartphone, its system settings, including the time clock, and installed applications should be verified. For audio recording applications, the following are established first: the current recording settings, the selected memory and storage folder path, the list of files containing recordings, and possibly other accompanying files, as well as the range of presented information about files. Next, one should determine all possible application settings, such as recording formats and parameters, and file naming, as well as the possible scope of their changes. Before recording "valid" test recordings, which will be subject to later comparative analysis, it is good practice to make a recording or several recordings (at least one for each format) in order to determine which options are available during and immediately after the recording. Experience shows that on smartphones some of the application options are only activated in the course of recording. This also allows determination of how the recordings are saved, i.e. automatically or with user interaction, their real names and readable parameters. Not all applications present the format of recordings or file extensions in a standard way, describing parameters with sometimes enigmatic names in the form of "good quality" or "small files". Familiarisation with the capabilities of a given application allows development of an optimal plan for making test recordings, containing a description of their execution, the sequence of recordings and sequences of reproduced test signals and the application of all possible settings and functions. In practice, test recordings are not stored in the memory built into the evidence smartphone so as not to interfere with its content, but on installed removable microSD memory cards. Cards with a reasonably good saving speed, for example class 10, should be used to in order to ensure that the data stream is saved uninterruptedly on the data carrier. For security purposes, the stage of preparing and making test recordings takes place in a screened chamber without access to the mobile phone network. All tests on the phone are carried out without an installed SIM card or evidence memory cards, and without changing the system settings, folder and file contents. After the forensic audio tests are completed, the memory status of the smartphone is routinely verified, taking into account the files containing the recordings, both from the level of the recording application and the file manager.

Valid recording of test recordings with the use of a smartphone is a time-consuming task and is preceded by appropriate preparation, but it is important in the process of analysing the device, the application and the saved recordings as well as in the analysis of the authenticity of evidence recordings.

3. Material for testing

3.1. Audio recording device and applications

A popular Samsung smartphone, the Galaxy J3 Dual SIM 2017, constituted the hardware used in the research. At the time of purchase, it had a pre-installed Android version 8.0.0 operating system as well as an application for recording audio: Samsung Voice Recorder version 21.0.22.166.

In order to carry out the intended tests with the use of the Google Play service, descriptions of available applications for making audio recordings were analysed. Next, on the basis of a combined assessment encompassing: number of installations of the given application (popularity), high rating by users and reviewers, as well as capabilities and functionality, eight free applications were selected. The names of these applications are as follows (the version of the application and the developer are given in brackets): Voice recorder (2.34, Splend Apps), Easy Voice Recorder (2.6.1, Digipom), Voice Recorder (5 (36.0), recorder&smart apps), Smart Recorder (1.9.6, Smart Mob), Voice Recorder (3.08, Appliqato Software), Voice Recorder (1.5.6, thinksimple app), EZ Voice Recorder (1.5.11, Top 1), Tape-a-Talk Voice Recorder (2.0.7,

Markus Drösser). For your information, please note that, according to the Google Play service, the number of downloads ranges from about 1 million for the application Tape-a-Talk Voice Recorder to over 1 billion for Samsung Voice Recorder. All the mentioned applications were installed on the Samsung Galaxy J3 Smartphone described above, and after making test recordings with these applications, the Android operating system on this device was upgraded to version 9. After making recordings using the Samsung Voice Recorder version 21.0.22.166, it was upgraded to the newer version 21.1.06.11 and further test recordings were made using it.

In total, properties of 10 applications designated for recording audio recordings in the Android system versions 8 and 9 were analysed. This analysis was performed with the use of:

– information read directly from the menu of the running application,
– the file manager installed on the Android system of the smartphone,
– Windows 7 Ultimate file explorer.

3.2. Audio recordings

With the help of the above mentioned 10 applications installed on the Android system, 142 audio test recordings were made in 6 different formats. These recordings were made using all possible recording settings and functions available during the recording, such as: start and stop recording, pause function, activation of recording by volume level, overwriting with a newer recording, indexing, automatic saving or with user interaction. After the test recordings were made, they were edited in the following ways with the help of applications which had such a capability: deleting a fragment of the recording, copying and pasting a fragment of the recording, overwriting an already saved recording with a newer recording, or repairing the file header. These recordings contain as wide a range of signals as possible, such as: white noise, low, medium and high frequency tones, speech and intentional interference, both impulse and continuous. To safely explore the contents of the memory containing recordings, and to make copies of the recordings, a write-blocker was used to prevent possible modification of data on these types of carriers.

For the purposes of further research, copies of files were made from a device that recorded onto the hard disk of a computer. An analysis of the properties of the collected files containing the test recordings was performed using the following tools:

– programs designed for the analysis of multimedia files, i.e., ffprobe version git-2019-11-13-4e0860e and MediaInfo version 19.09, Windows 7 Ultimate file explorer, the file manager installed on the Android system of the smartphone, as well as information presented by applications making the recordings;
– hexadecimal editors 010 Editor version 6.0.2 and HxD version 1.7.7.0, designed to read and visualize the data structure of the saved files;
– a computer program developed for the needs of this research.

A computer program operating in the MATLAB environment was developed for analysing the data contained in the files with the test recordings. This program enables automatic structure exploration and metadata detection in audio files of any format; however, it was designed primarily for analysis of files saved in MPEG-1 – MPEG-4 and RIFF standards. Files of this kind constitute almost all of the collected research material and may contain more extensive metadata. The operation of the created program can be presented as follows:

1) import of an audio file selected for analysis into a workspace; the import algorithm recognizes the file type on the basis of the file extension and proposes an analysis method that is suitable for the given format, with the user also being able to determine the file type and choose the method by themself,
2) defining of the analysis area, i.e. the whole file or the part of the file containing the metadata; the analysing program may skip the encoded audio data, for example, the whole *mdat* box in MPEG files, and the whole chunk *data* in RIFF files; the file area with encoded media is usually the most extensive part of the file, where there are usually no metadata,
3) exploring a designated area byte by byte and searching for all character strings that may constitute metadata; the detection and decision algorithm is based on the four-character code (FourCC) sequences implemented in the program for boxes and chunks (mainly for MPEG and RIFF) as well as other possible markers, identifiers and tags for all analysed formats; to ensure correct operation, the algorithm also takes into account upper and lower case letters, location in the file, size and interval between consecutive recognized metadata; in the case of detection of a hitherto unknown four-character code sequence, the program also presents it in the results, in order to complete the description and for the user's knowledge; the concepts mentioned here will be discussed in greater detail later in this paper,

4) saving of the metadata structure of the analysed file to a spreadsheet for the purpose of possible further comparative analysis with other files,
5) optional saving to a spreadsheet of metadata indicated by the user, i.e. name, size, location in the file and content.

In the course of the analysis of the collected audio files, the created program was optimized on the basis of a feedback loop with the aim of using it for the analysis of various formats and possible data structures. Apart from the obtained results of the analyses, available research papers and technical documentation were also referred to during development of the program (3GPP, 2010; 3GPP, 2011; ISO/IEC, 1993; ISO/IEC, 2015; Microsoft Corp., 1994; Microsoft Corp., 2010). This program constitute a tool supporting authenticity examinations of audio recordings, which are carried out as part of expert opinion work at the Institute.

4. Results

4.1. Analysis of features of installed applications

The aim of this part of the research was to determine what the possibilities are of recording audio recordings, saving and possibly editing them with the use of applications installed on the Samsung smartphone.

Table 1 presents basic features of the 10 applications with which test recordings were made: the format of saving of the sound, available functions during the recording, the possibility of editing the recordings by a given application and the application options. These applications allow you to save audio recordings in 6 different formats: 3GP, AAC, AMR, M4A (MP4 with audio), MP3 and WAV. An analysis of the parameters of the files containing test recordings, described in detail in a further part of the paper, allowed establishment of the fact that the examined applications most frequently use an MPEG-4 type multimedia container based on ISO standards. It contains a relatively large amount of metadata in relation to other multimedia containers, which is a positive feature from the point of view of analysis of the authenticity of recordings.

In addition to the basic functions of starting and stopping the recording, all applications are equipped with a pause function. This is important information, because activation of this function – intentionally or accidentally – causes discontinuity of the recording. Analysis of the continuity of a recording is one of the basic examinations in the course of analysing its authenticity (Kajstura et al., 2017; Korycki, 2016). In recordings made with older models of phones, which were equipped with function buttons, the use of pause often resulted in the recording of characteristic noises. The use of touchscreens on smartphones significantly reduces the recording of such noises, so discontinuities of this kind are sometimes difficult to reveal (Kajstura et al., 2017). Interestingly, some of the tested applications only allow this function for selected saving formats (items: 2, 3 and 9 in Table 1). Additionally, in the Smart Recorder application, an automatic pause function is available. When this function is turned on,

Table 1
The basic features of installed applications used to make audio test recordings

	Application	Format	Recording options	Audio edit	Audio edit options
1	***Samsung Voice Recorder***	M4A	START, STOP, PAUSE	yes	trim, overwriting
2	***Voice Recorder (Splend Apps)***	WAV, M4A, 3GP, MP3	START, STOP, PAUSE (WAV only)	no	none
3	***Easy Voice Recorder (Digipom)***	WAV, M4A, 3GP	START, STOP, PAUSE (WAV, M4A)	no	none
4	***Voice Recorder (recorder&smart apps)***	WAV, MP3	START, STOP, PAUSE	yes	trim, copy&paste
5	***Smart Recorder (Smart Mob)***	WAV	START, STOP, PAUSE, AUTO PAUSE	no	none
6	***Voice Recorder (Appliqato Software)***	AAC, WAV, AMR, M4A	START, STOP, PAUSE	no	none
7	***Voice Recorder (thinksimple app)***	MP3, WAV	START, STOP, PAUSE	no	none
8	***EZ Voice Recorder (Top 1)***	MP3, M4A	START, STOP, PAUSE	yes	trim
9	***Tape-a-Talk Voice Recorder***	WAV, AAC	START, STOP, PAUSE (WAV only)	yes (WAV only)	trim, append recording
10	***Samsung Voice Recorder (upgrade)***	M4A	START, STOP, PAUSE	yes	trim, overwriting

the recording is automatically paused when the value of sound intensity drops below a set threshold, and it is activated when the sound intensity increases above this threshold.

Half of the analysed applications enable editing of recordings made by them (items: 1, 4 and 8–10 in Table 1), while the Tape-a-Talk Voice Recorder application only does for WAV format. In all these applications, after saving the recording, it is possible to remove the initial and final fragment or a fragment within the recording (trim option). The Samsung Voice Recorder application in both versions allows you to set the initial recording point and overwrite the recording with a new one in a saved or recorded and paused recording. The Voice Recorder application (recorder&smart apps) allows you to copy and paste a fragment of the recording to another place, while Tape-a-Talk Voice Recorder allows you to add a further part of the recording starting from the end of the already saved recording.

Among the significant additional options, in the case of the Samsung Voice Recorder – the possibility of setting checkpoints (so-called bookmarks) by the user together with a description during and after saving of the recording should be mentioned. In turn, the Tape-a-Talk Voice Recorder application has an option to repair the header for WAV files, consisting in automatic entering of the parameters of the file given by the user into the metadata: sampling frequency, bit quantization and number of channels.

Table 2 shows the results of the analysis of features of installed applications, such as assigned extensions and default names for files containing source recordings and recordings after editing, as well as saving options for recordings. Such analysis turns out to be important, since default names of source files in the case of 6 applications (items: 1, 3, 6, 7, 9 and 10) are always different from the default names assigned by other applications. On the other hand, for the 4 remaining applications (items: 2, 4, 5 and 8), the default names of their files are different only for selected formats or with the option to add a time signature to the file name. An important feature of the application is the inclusion in the file names of information about the date and time, both as a fixed element (items: 4, 7 and 8) and as an option set by the user (items 5 and 9). Time signatures in the names of such files are linked with the start of recording, but they depend on the setting of the built-in clock of the smartphone. Most applications, i.e. with the exception of items 4, 7 and 8, number their recordings incrementally. For some applications, there are also some characteristic features in the discussed area. For the Samsung Voice Recorder, the file names stored on the microSD card receive an additional string of *_sd* characters. The Voice Recorder application (Appliqato Software) adds a unique string of alphanumeric characters to the file name, while Easy Voice Recorder for the M4A format allows you to save a file with an m4a or mp4 extension selected by the user.

Edited recordings can be saved either in a new file or in the same (source) file (items: 1, 9 and 10), or only in a new file (items 4 and 8). The common feature of the edited files and those saved as source files is their unchanged name. In turn, all applications add numbering indicating subsequent editing to files saved as new (for example *-1* and *-2* for item 1) or strings of characters indicating execution of editing (for example *_CUT1* for item 8).

Table 2 also shows how a recording is saved to a file after finishing, for specific applications, which can be done in two different ways. In the first case, saving takes place after approval of the name proposed by the application or after entering your own. In this case, the application waits for the user to interact, which may affect the time signature of the file modification. In the second case, the application saves to a file with a default name and automatically after a recording is finished.

The menus of all analysed applications allow the renaming of files containing already saved recordings. However, as previously indicated, usually the names of files containing evidence recordings are the default names for the recording application used. Therefore, the analysis of file names for audio recording applications is justified, as it may support the identification of a specific application, which can be seen both in the examples presented above and in the literature (Kajstura et al., 2017; Michałek, 2016).

4.2. Analysis of the properties of audio recordings and files

In this part of the paper, the properties of the audio recordings and files that were saved using the 10 installed applications will be presented.

Recording parameters

As noted, 142 audio recordings were made with sound saved in 6 different formats using the applications selected for testing. Here are the formats in order of their frequency of occurrence: WAV – in 7 applications, M4A – 6, MP3 – 4, followed by AAC and 3GP – 2, and AMR – 1. Table 3 presents the names of the predefined audio quality settings and their param-

eters for specific applications and possible recording formats: sampling frequency in [kHz] and bit rate in [kbps]. As you can see, the menus of some of the applications allow you to select defined parameter settings for various formats without the possibility of changing them, which are described by names indicating the recording quality. The remaining applications allow the user to freely select the settings to the extent possible. For 2 formats, i.e. 3GP and AAC (items 3 and 9), it was not possible to change any recording parameters, but only to record a recording with one defined setting. On the other hand, for application (item) 2 and for the 3GP format, the menu indicated the possibility of changing the parameters, but the saved files were always characterized by the parameters given in Table 3.

Table 4 shows the codecs used for saving sound for the different formats which have been implemented in the tested applications. In turn, Table 5 presents the multimedia container or data structure used to save the encoded audio data and the metadata describing it.

On the basis of the obtained results, which are presented in Tables 4 and 5, interesting characteristic features can be observed for individual applications. Applications (items) 2 and 3 use AMR codecs in 3GP files and place the data in an MPEG-4 container. Voice Recorder (Splend Apps) uses a rarely applied broadband variant of the AMR codec, i.e. AMR-WB, and – in sound files with an mp3 extension and labelled as MP3 format – it uses AAC-LC coding and places data in an MPEG-4 container. Voice Recorder (Appliqato Software) is the only one to use a "conventional" AMR file structure with an AMR-NB (Narrowband) codec. Several applications (items: 4, 7 and 8) in MP3 file format use MPEG-1, MPEG-2 and MPEG-2.5 standards for different quality of recordings, i.e. MPEG-1 for the best, and MPEG-2.5 for the worst.

Table 2
Default file names, extensions and saving options for analysed audio recording applications

Application	File extension	Example of default file name: original recording	Example of default file name: after editing	Saving options	Saving edited file as
1	m4a	Voice 001.m4a (device memory)		wait for user	new or source
		Voice 031_sd.m4a (microSD)		wait for user	
			Voice 031_sd-1.m4a	automatic	
			Voice 031_sd-2.m4a		
2	wav, m4a, 3gp, mp3	Recording_1.wav	X	automatic	X
3	wav, m4a, mp4, 3gp	My recording 1.wav	X	automatic	X
4	wav, mp3	2019_07_19_12_02_00.wav	2019_07_19_12_02_00_1.wav	wait for user	new
			2019_07_19_12_02_00_2.wav		
5	wav	Recording_2.wav	X	automatic	X
		20190719-1626_Recording_6.wav			
6	aac, wav, amr, m4a	Recording 1 (1271241a-2ea7-4b58-b03c-6138b52fd362).aac	X	wait for user or automatic (option)	X
7	mp3, wav	audio_111400_220819.mp3	X	wait for user	X
8	mp3, m4a	2019_08_23_18_34_01.mp3	2019_08_23_18_34_01_CUT1.mp3	automatic	new
			2019_08_23_18_34_01_CUT2.mp3		
9	wav, aac	Recording_2.aac	editing not available	automatic	X
		Recording-19-12-27-14-16-00.wav	Recording-19-12-27-14-16-00cut1.wav	automatic	new or source
			Recording-19-12-27-14-16-00cut2.wav		
10	m4a	Voice 038.m4a (device memory)		wait for user	new or source
		Voice 045_sd.m4a (microSD)		wait for user	
			Voice 045_sd-1.m4a	automatic	
			Voice 045_sd-2.m4a		

Table 3
Audio quality settings for tested applications. The data are presented in the following order: setting name, sampling frequency [kHz] and bit rate [kbps]

Appli-cation	Formats and settings					
	M4A	3GP	AMR	WAV	MP3	AAC
1	HIGH 48/256 MEDIUM 44.1/128 LOW 44.1/64	X	X	X	X	X
2	GOOD 8-44.1/32-320	SMALL FILES 16/24	X	HIGH 8-44.1/128-706	GOOD 8-44.1/32-320	X
3	HIGH 44.1/96 MEDIUM 16/32 LOW 8/32 or user settings: 8-48kHz extension: m4a or mp4	8/12.2 (no settings)	X	HIGH 44.1/706 MEDIUM 16/256 LOW 8/128 or user settings: 8-48kHz	X	X
4	X	X	X	11.025-44.1/176-706	11.025-44.1/64-128	X
5	X	X	X	8-44.1/128-706	X	X
6	8-44.1/48-128	X	8/12.2	8-44.1/128-706	X	8-44.1/48-128
7	X	X	X	8-48kHz/128-768	8-48/64-320	X
8	LOW 11.025/64 NORMAL 22.05/128 HIGH 44.1/256	X	X	X	LOW 11.025/64 NORMAL 22.05/128 HIGH 44.1/256	X
9	X	X	X	8-22.05/128-352 (free version)	X	44.1/12.2 (no settings)
10	HIGH 48/256 MEDIUM 44.1/128 LOW 44.1/64	X	X	X	X	X

Table 4
Audio codecs implemented in installed applications

Application	Audio codec					
	M4A	3GP	AMR	WAV	MP3	AAC
1	AAC-LC	X	X	X	X	X
2	AAC-LC	AMR-WB	X	WAVE PCM	AAC-LC	X
3	AAC-LC	AMR-NB	X	WAVE PCM	X	X
4	X	X	X	WAVE PCM	LAME 3.99.5, CBR	X
5	X	X	X	WAVE PCM	X	X
6	AAC-LC	X	AMR-NB	WAVE PCM	X	AAC-LC v.2
7	X	X	X	WAVE PCM	LAME 3.100, CBR	X
8	AAC-LC	X	X	X	LAME 3.100, CBR	X
9	X	X	X	WAVE PCM	X	AAC-HE
10	AAC-LC	X	X	X	X	X

The obtained results allow us to state that among the tested applications, the MPEG-4 type container is the most commonly used one; it is very universal and designed to hold various types of multimedia together with metadata. Similarly to the case of the test recordings, it is currently the most common container in which evidence audio recordings are saved when using smartphones.

Time signatures

The analysis also encompassed information on the recording time of source test recordings and recordings that had been edited with the installed applications. The influence of the way of recording of these recordings on their time signatures was also analysed.

During the making of each test recording, the time of its starting and stopping was noted from the clock of the recording device. In this paper, an analysis of default names of recorded audio files has been described above. It was established that 6 tested applications make it possible to add a string of characters, indicating the date and the time, to file names. Time signatures of files containing these recordings were read with the help of the menu of the tested applications, the Android operating system and Windows Explorer directly from the memory where the recordings were saved.

Table 6 shows the names of example files containing audio recordings together with the date and time of the start and end of the recording, and the time signatures of modifications read by the above mentioned three methods. The results indicate that the strings of characters with the date and time contained in the file names correspond to the date and time of the start of the recording. On the other hand, the time signatures of the modifications read with the help of the Android system and Windows Explorer are associated with the end of the making of recordings and saving them to file. Similarly, the menu of most of the analysed applications enables reading of the time signature of the modification of the files, with the exception of Smart Recorder (Smart Mob) and EZ Voice Recorder (Top 1), which give the time signature of creation.

The Android file manager did not allow you to read the time signature of the creation of the saved audio files; in turn, Windows indicated the time signature of creation as being the same as the time signature of modification, i.e. not corresponding to the real time.

The recordings indicated in Table 6 were recorded as continuous ones and were saved directly after finishing with names proposed by the applications. It was established that 5 of the tested applications indicated in Table 2, after finishing the recording, wait for interaction from the user before saving the recording to file. In order to determine whether this time affects the time signatures of the files, additional continuous test recordings were made, during which, from the end of the recording to the moment of confirmation and saving, the user waited for the smartphone clock to show the next minute. It was established that the time of modification of files saved in this way corresponded to the time of their finishing in the case of applications (items) 4, 6 and 7, i.e. Voice Recorder (recorder&smart apps), Voice Recorder (Appliqato Software) and Voice Recorder (thinksimple app). However, for the Samsung Voice Recorder in both tested versions (items 1 and 10), the time of modification of the file

Table 5
Multimedia container or file structure used by tested applications

Application	Multimedia container or file structure					
	M4A	3GP	AMR	WAV	MP3	AAC
1	MPEG-4	X	X	X	X	X
2	MPEG-4	MPEG-4	X	RIFF	MPEG-4	X
3	MPEG-4	MPEG-4	X	RIFF	X	X
4	X	X	X	RIFF	MPEG-1, -2, -2.5	X
5	X	X	X	RIFF	X	X
6	MPEG-4	X	AMR	RIFF	X	MPEG-2 ADTS
7	X	X	X	RIFF	MPEG-1, -2, -2.5	X
8	MPEG-4	X	X	X	MPEG-1, -2, -2.5	X
9	X	X	X	RIFF	X	MPEG-4
10	MPEG-4	X	X	X	X	X

corresponded to the moment of confirmation of the name and saving, which can be done by the user after a varying amount of time.

The influence of editing recordings in applications that have such an option on time signatures of modifications was also analysed. It was established that all files after the performed edition described in Table 1 were characterised by a changed, i.e. later, time of modification, but for applications (items) 4, 8 and 9, i.e., Voice Recorder (recorder&smart apps), EZ Voice Recorder (Top 1) and Tape-a-Talk Voice Recorder, the new time signature of modification corresponded to the moment of saving of the recording after changes. What is very interesting is that the time signature of modification for files edited in the Samsung Voice Recorder in both versions already changed at the moment of making the modification of the recording, even before confirming changes, and the fact of saving the file did not affect the time signature of the modification.

As can be seen, using the three methods given in Table 6 allows you to read the time signature of a modification rounded to the nearest minute. It should also be added that the file names mentioned in the table with the time of start of recording and time signatures of their modification indicated depend on the settings of the clock built into the smartphone. Usually the Android menu allows you to change the settings of this clock and then the time information indicated above may differ from the real time. However, in the default configuration, smartphones have an option of automatically updating the time provided by the network operator, which is significant for authenticity examinations, as then the time set in the device corresponds to the real time. As already mentioned, usually the names of evidence audio files from the memory of smartphones correspond to default names; however, the menu of all hitherto analysed applications, both in evidence devices and in phones studied in this research work, allowed you to rename files containing recordings. The described methods of reading information about the time of recording and its possible editing on the basis of names of files and time signatures of creation or modification are helpful in the process of analysis of the authenticity of recordings, but they should always be verified by other methods, such as the method using powerline hum or analysis of file metadata.

Analysis of metadata

Visualisation and analysis of the structure of files and information recorded in metadata are currently

Table 6
Default file names and time signatures of the audio test recordings

Appl.	Default file name	Beginning of the recording	End of the recording and save	Modification time (application)	Modification time (Android)	Modification time (Windows)	Duration
1	Voice 025_sd.m4a	2019-07-11 11:48:00	2019-07-11 11:52:00	2019-07-11 11:52	2019-07-11 12:52	2019-07-11 12:52	03:59
2	Recording_1.wav	2019-07-17 15:44:00	2019-07-17 15:48:00	2019-07-17 15:48	2019-07-17 15:48	2019-07-17 15:48	04:00
3	2019-07-18 15-39-00My recording 5.m4a	2019-07-18 15:39:00	2019-07-18 15:44:00	2019-07-18 15:44	2019-07-18 15:44	2019-07-18 15:44	04:59
4	2019_07_19_11_11_00.wav	2019-07-19 11:11:00	2019-07-19 11:16:00	2019-07-19 11:16	2019-07-19 11:16	2019-07-19 11:16	04:59
5	20190719-1626_Recording_6.wav	2019-07-19 16:26:00	2019-07-19 16:27:00	Created only: 2019-07-19 16:26:00	2019-07-19 16:27	2019-07-19 16:27	01:00
6	Recording 1.aac	2019-08-20 12:52:00	2019-07-20 12:57:00	2019-07-20	2019-07-20 12:57	2019-07-20 12:57	04:59
7	audio_114200_220819.mp3	2019-08-22 11:42:00	2019-08-22 11:45:00	2019-08-22 11:45	2019-08-22 11:45	2019-08-22 11:45	02:59
8	2019_08_23_18_34_01.mp3	2019-08-23 18:34:01	2019-08-23 18:36:00	Created only: 2019-08-23 18:34	2019-08-23 18:36	2019-08-23 18:36	01:59
9	Recording-19-12-27-13-54-00.wav	2019-12-27 13:54:00	2019-12-27 13:57:00	2019-12-27 13:57	2019-12-27 13:57	2019-12-27 13:57	03:00
10	Voice 039_sd.m4a	2019-12-27 14:48:00	2019-12-27 14:50:00	2019-12-27 14:50	2019-12-27 14:50	2019-12-27 14:50	02:00

among the basic methods of examining the authenticity of digital recordings (Ho, Li, 2015; Kajstura et al., 2017; Korycki, 2016; SWGDE, 2018). Metadata, i.e. additional information describing data with encoded sound, may be very important and include recording parameters, time signatures, data relating to the device and operating system and many others depending on the format and multimedia container (Gloe et al., 2014; Koenig, Lacey, 2012; Korycki, 2016; Michałek, 2018). In real cases, making test recordings with the studied device enables a comparative analysis of the structure of the test recording files with the structure of the evidence file. The analysis is carried out using software enabling visualization of files in hexadecimal and ASCII code form.

On the basis of analysis of all the test recordings, it was established that most of the data were located on files saved in an MPEG-4 container. It was further established that it is also the most frequently used by the analysed applications, which is indicated in Table 5. Traces unambiguously indicating editing of the audio recording using installed applications were revealed in the structure of this type of test files.

The internal structure of test audio files in the MPEG-4 container is based on the guidelines indicated in the ISO/IEC 14496 set of standards (ISO/IEC, 2003; ISO/IEC, 2010; ISO/IEC, 2015). Within this structure, you can distinguish so-called boxes, i.e. separated and autonomous parts of the file, containing fields with metadata (information) or coded multimedia data, forming an ordered and hierarchical file structure. Data in boxes are saved in the *big-endian* convention, i.e. the most significant byte appears first.

Analysis of the structure of all test audio files stored in the MPEG-4 container showed that each of them contains 3 main boxes: File Type (identifier in structure: *ftyp*) with information about compatible specifications to play the file, Media Data (*mdat*) with coded multimedia data and Movie Box (*moov*) with metadata. Analysis of the File Type box for test files indicates that both versions of the Samsung Voice Recorder application make use of the basic specification 3GPP version 4, while the remaining applications saving files in MPEG-4 use the MP4 specification version 2. Additional metadata were not revealed in any test file within the Media Data box, except from its header. In the MPEG-4 container the most important object for analysis is a box called Movie Box, which contains metadata for coded media within numerous internal boxes – those boxes which are significant for this study will be discussed later. According to ISO recommendations, the Movie Box should be located at the end of the whole file (ISO/IEC, 2015).

In the case of all installed applications, test files in M4A, MP3 and AAC format in Movie Box contain 28 internal boxes. In M4A files in both versions of the Samsung Voice Recorder, there are an additional 4 boxes informing about: bookmarks placed by the user, their number, location and possible description. This application also adds an important box labelled with the identifier *vrdt* with a sequence of ASCII characters *com.sec.android.app.voicenote.common.util.VoiceRecorder* informing that the file has been recorded with the Samsung Voice Recorder, designed for the Android operating system. Test files saved in 3GP format contain 29 internal boxes in the Movie Box object. All analysed test files in the MPEG-4 container contain a *User Data* (*udta*) box, within which are located *SDLN*, *smrd* and *smta* boxes indicating that the file was saved using a Samsung device. Another significant object in all analysed MPEG-4 files is the *meta* box, in which *com.android.version* entries with values *8.0.0* or *9* indicate an Android operating system and the version. Examples of such metadata for recordings made with Samsung Voice Recorder applications (items 1 and 10 in Table 1) installed in both mentioned systems are presented in Fig. 1.

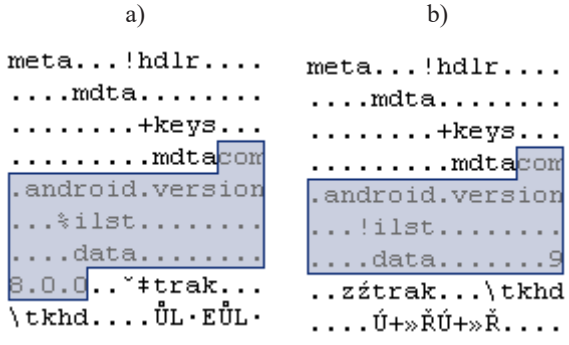

Fig. 1. Metadata within the *meta* box in MPEG-4 audio files relating to the smartphone operating system and the version: Android 8.0.0 (part a) and Android 9 (part b) for recordings made using the Samsung Voice Recorder application.

An *stsd* box was also noted in MPEG-4 files – it contained information on the codec and parameters of sound saving: for M4A, MP3 and AAC formats they are in an internal *esds* box, whilst for the 3GP format in a *sawb* or *samr* box (which corresponds to AMR-WB for application (item) 2 or AMR-NB for application (item) 3) as well as in a *damr* box. Information decoded from the *stsd* object allowed verification of the codecs and parameters of recording read using the menu of the applications and tools mentioned in section 3.2 of this work.

In the *Movie Box* object in internal *mvhd*, *trak\tkhd* and *mdia\mdhd* boxes, metadata contain a total of 6 entries allowing decoding of time signatures of creation and modification of a presentation, track and media rounded to the nearest 1 second for a recording saved in an MPEG-4 container. It was established that for each of the recordings that were not subjected to editing, the time signatures of creation decoded in this way are the same as the time signatures of modification and are linked with saving of a recording to file. They also correspond to the pre-defined system time signatures of modification, taking into account the difference in UTC time applied in metadata from system time and possible rounding to the nearest 1 minute. What is significant for authenticity testing is that time signatures in metadata in an MPEG-4 file do not change after copying the file to another data carrier. However, they depend on the smartphone clock setting at the moment a recording is saved. Metadata in *mvhd*, *trak* and *mdia* boxes also contain the duration for the audio recording rounded to the nearest 1 ms. Additionally, all saved test files in an MPEG-4 container in an *hdlr* box possess entries indicating only the content of the soundtrack.

The tested applications which save their recordings in WAV format use a PCM codec without signal compression. Recordings of this format are placed in a RIFF container and contain data blocks called chunks saved in the *little-endian* convention. Analysis of the structure of the RIFF test files showed that they all contain basic metadata for this type of file in canonical form: a RIFF identifier informing about the type of container, a WAV string, *fmt* chunk with recording parameters, and *data* chunk with the coded recording. Amongst applications saving recordings in this container, one that stands out is Smart Recorder (Smart Mob), which places an additional chunk *list* with internal objects *INFO* and *INAM*, which contain strings of characters with the name of the file. What is significant is that the name of a file saved in this way in metadata does not change after a possible name change in the file system.

The information about the applied version of the LAME codec is stored in structures of files in MP3 format (items: 4, 7 and 8 in Table 1). Only the Voice Recorder application (thinksimple app) places metadata with a TAG identifier, which contain a string of 4 characters with the year of recording, in the structure of such files.

As was established, among the installed applications, only the Voice Recorder (Appliqato Software) saves recordings in AMR and AAC formats in a "conventional" way. Analysis of the structures of test files in these formats did not show the presence of any individual metadata, but only standard and basic information about recording parameters that are necessary to play a recording.

The influence of the editing of recordings on metadata

As part of the research, the structures of files with recordings after editing – taking into account all possible options indicated in Table 1 – were analysed. Among applications having the option of editing, three of them allow for the editing of M4A files (items: 1, 8 and 10 in Table 1) and two applications each allow for editing of MP3 files (items 4 and 8) and WAV files (items 4 and 9).

Comparative analysis of files with source recordings in relation to files with edited recordings for WAV and MP3 formats, which were studied in this work, did not show changes in the structure of this type of files nor in the content of their metadata.

Significant changes in edited files were, however, noted for recordings in M4A format placed in an MPEG-4 container, which have been summarized in Table 7.

As can be seen, editing a recording with Samsung Voice Recorder in both versions and EZ Voice Recorder (Top 1) results in reorganization of the structure of M4A files: the metadata content is modified, fragments or even whole boxes are deleted, and their location or size are changed. The metadata modifications described in Table 7 were noted both in files after edition saved as new files, and in files after edition saved as source files.

In *mvhd*, *tkhd* and *mdhd* internal boxes, there are metadata allowing you to decode the time of creation and modification of the presentation, track and media saved in the track. Fields with time signatures contain values indicating the number of seconds since midnight 1 January 1904 in UTC time format to the moment of saving or modification of the recording. All test files saved in this research work contain tracks with recorded sound.

The effect of editing of a recording on metadata in an M4A file that allow the decoding of time signatures was analysed. The content of fields labelled as *creation_time* and *modification_time* in *mvhd*, *tkhd* and *mdhd* boxes was read using hexadecimal editors (ISO/IEC, 2015). In this way 6 values were read, which were decoded to UTC time and then to Central European time.

The EZ Voice Recorder application (Top 1), after edition of a recording, only allows it to be saved

to a new file (Table 2). For this application, analysis of the mentioned 6 fields from *mvhd*, *tkhd* and *mdhd* boxes showed that all decoded time signatures are the same and correspond to the moment of saving of the recording post edition to a new file.

In turn, for both versions of the Samsung Voice Recorder, after edition of a recording, there is a possibility to save the changes in the source file or in a new one, which has an influence on time signatures in metadata. It was established that after editing a recording and saving it:

– to a new file: time signatures of creation and modification of a presentation as well as time signatures of modification of a track and media correspond to the moment of edition of the recording (for example: deleting a fragment); in turn, time signatures of the creation of a track and media correspond to the time of saving of the source recording; it was also established that further modifications of the file cause a change in the time signature of creation and modification of the presentation as well as modification for the track and media; however, the time signatures of creation of the track and media remain without changes,
– in the source file: time signatures of creation and modification of the presentation and modification of the track and media correspond to the moment of edition of the recording, whilst the time signature of creation for the track and media are linked with saving of the recording to the file before its edition.

On the basis of the results presented above, it can be stated that editing a recording using an installed application causes changes in the structure of M4A files and modification of significant metadata. These changes, however, help in the evaluation of the authenticity of such a recording.

The research carried out on the applications and test recordings together with the obtained results constitute a source of information that can be used in the analysis of the authenticity of evidence recordings. The application of dedicated programs and system tools as well as the designed program in the MATLAB environment allowed analysis of parameters of files with recordings, their time signatures and metadata. The saved audio files were subjected to global comparative analysis and on this basis it was established that each of the installed applications places at least one distinctive feature within its files. Table 8 shows the properties of test files with an indication of those features that are distinctive for the given application in relation to others.

In each case the basic material for forensic audio analysis is a recording saved in a file, and the revealed distinctive features and their comparison with test recordings can be helpful in the identification of an application, software or device used to record the recording.

5. Conclusions

Summarising the performed research, one may say that it has enabled a broadening of knowledge on the most frequently used applications for making audio recordings that are installed on the Android system. Analysis of selected applications allowed familiari-

Table 7
Changes within metadata in M4A audio files after recording editing

Application	Metadata in boxes		
	Before editing	After Editing	Contents and meaning
1 and 10	meta	missing	operating system and version
	stts	resized (2B less)	time indexes for samples
	stsz	replaced with stsc	sample sizes
	stsc	replaced with stsz	groups (chunks) of media samples
	stco	resized (various lengths)	chunks offset table
	hdlr	missing string: SoundHandle	media type of the track
8	ftyp; specifications: mp42, isom	specifications: M4A, mp42, isom	file type (specification)
	mdat	replaced with moov	encoded media data
	moov	replaced with mdat	metadata for recording
	udta and sub-boxes	missing	typical boxes for Samsung device
	meta	missing	operating system and version

zation with their capabilities and the making of test recordings with them. The results of analysis showed that these applications, although free, allow for a relatively broad spectrum of possibilities of audio recording: in 6 different audio formats, with predefined or customised parameters, in high quality uncompressed files or with sound compression, as well as with the possibility of including additional information in their names. Half of the analysed applications enable editing consisting in deletion of a fragment of the recording, copying and pasting it in another place or overwriting it with a new recording. Using system tools and dedicated programs as well as the algorithm developed in the MATLAB environment, properties of the collected recordings and file structures were analysed. On this basis, it was ascertained that it was possible to establish parameters of the test recordings, their time signatures and metadata content. The results of the conducted analyses unambiguously point to the possibility of making use of them in authenticity testing of evidence audio recordings.

First of all, the work carried out allowed for collection and analysis of audio recordings recorded with various mobile applications, which will make it possible to perform comparative analyses of their properties with those of evidence recordings submitted for examination in the future. This will also make it possible to supplement the database of audio recordings which is constantly being developed by the author, and which constitutes an additional tool supporting authenticity testing. Secondly, the analysis of metadata of test files saved in the popular MPEG-4 container allows you to obtain much important information about the recording parameters, time signatures, the device used, and the recording application. It also enables indication of the differences between source recordings and edited recordings. Additionally, each of the installed applications is characterised by at least one feature that distinguishes its files from the rest, which may be helpful in identifying the specific application used to save the given examined audio file. The results obtained are of significant importance, since questions concerning authenticity of recordings are frequently included in the official request of institutions commissioning forensic audio expert opinions.

Table 8
Distinctive features of the analysed audio files recorded using installed applications

Application	Distinctive feature					
	Default file name	File extension	Container	Format	Codec	Metadata
Samsung Voice Recorder and Samsung Voice Recorder (upgrade)	Yes					Yes (additional boxes)
Voice Recorder (Splend Apps)	Yes/No(WAV)		MPEG-4 (MP3)		AMR-WB (3GP) AAC-LC (MP3)	Yes (additional boxes)
Easy Voice Recorder (Digipom)	Yes	mp4 (M4A/MP4)			AMR-NB (3GP)	Yes (AMR library)
Voice Recorder (recorder&smart apps)	Yes(WAV)/ No(MP3)				LAME3.99.5 (MP3)	
Smart Recorder (Smart Mob)	No/ Yes(with time signature)					Yes (additional chunks; WAV)
Voice Recorder (Appliqato Software)	Yes	amr (AMR)	MPEG-2 ADTS (AAC)	AMR-NB		Yes (file structure)
Voice Recorder (thinksimple app)	Yes					Yes (TAG; MP3)
EZ Voice Recorder (Top 1)	Yes(M4A)/ No(MP3)					
Tape-a-Talk Voice Recorder	Yes		MPEG-4 (AAC)		AAC-HE (AAC)	Yes (box sizes; AAC)

References

1. 3rd Generation Partnership Project (3GPP). Technical Specification Group Services and System Aspects. (2010). Technical Specification 26.244 V 9.2.0: Transparent end-to-end packet switched streaming service (PSS). April 9, 2020 from: http://www.3gpp.org.
2. 3rd Generation Partnership Project (3GPP). Technical Specification Group Services and System Aspects. (2011). Technical Specification 26.090 V 10.1.0: Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions. April 9, 2020 from: http://www.3gpp.org.
3. AppBrain. (2020). Android and Google Play statistics. Retrieved April, 10, 2020 from: https://www.appbrain.com/stats.
4. Audio Engineering Society (AES). (1996). AES *recommended practice for forensic* purposes – Managing recorded audio materials intended for examination, AES27–1996. Audio Engineering Society Inc. Retrieved April 9, 2020 from: http://www.aes.org.
5. Casey, E. (2010). *Handbook of Digital Forensics and Investigation*. Elsevier Academic Press.
6. Casey, E., Turnbull, B. (2011). Digital evidence on mobile devices. (In) E. Casey (Eds.), *Digital evidence and computer crime, 3rd Edition* (pp. 603–606). Elsevier Academic Press.
7. Cooper, A. J. (2006). Detection of copies of digital audio recordings for forensic purposes, PhD thesis. United Kingdom: Faculty of Technology, Department of Information and Communication Technology, The Open University.
8. Gloe, T., Fisher, A., Kirchner, M. (2014). Forensic analysis of video file formats. *Digital Investigation*, *11*, 68–76.
9. Grigoras, C. (2005). Digital audio recording analysis: The electric network frequency (ENF) criterion. *International Journal of Speech, Language and the Law*, *12*, 64–76.
10. Grigoras, C., Cooper, A., Michałek, M. (2009). Forensic Speech and Audio Analysis Working Group – Best practice guidelines for ENF analysis in forensic authentication of digital evidence; REF. CODE: FSAAWG-BPM-ENF-001. Retrieved October 13, 2017 from: http://www.enfsi.eu.
11. Ho, A. T. S., Li, S. (2015). Handbook of digital forensics of multimedia data and devices. United States of America: John Wiley & Sons.
12. International Data Corporation, IDC. (2019). Smartphone market share. Retrieved April 10, 2020 from: https://www.idc.com.
13. ISO/IEC 11172-3:1993: Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 3.
14. ISO/IEC 14496-14:2003. Information technology – Coding of audio-visual objects – Part 14: MP4 File Format.
15. ISO/IEC 14496-1:2010. Information technology – Coding of audio-visual objects – Part 1: Systems.
16. ISO/IEC 14496-12:2015. Information technology – Coding of audio-visual objects – Part 12: ISO base media file format.
17. Kajstura, M., Michałek, M., Trawińska, A. (2017). Ekspertyza fonoskopijna. (In) M. Kała, D. Wilk, J. Wójcikiewicz (Eds.), *Ekspertyza sądowa. Zagadnienia wybrane* (pp. 674–726). Warszawa: Wolters Kluwer.
18. Kajstura, M., Trawińska, A., Hebenstreit, J. (2005). Application of the electrical network frequency (ENF) criterion. A case of a digital recording. *Forensic Science International, 155*(2–3), 165–171.
19. Koenig, B. E. (1990). Authentication of forensic audio recordings. *Journal of Audio Engineering Society*, *38*(1), 3–33.
20. Koenig, B. E., Lacey. D. S. (2012). Authenticity analyses of the header data encoded WMA files from small Olympus audio recorders. *Journal of Audio Engineering Society*, *60*(4), 255–265.
21. Korycki, R. (2011). Analiza autentyczności cyfrowych nagrań fonicznych. *Problemy Kryminalistyki*, *271*(1), 5–21.
22. Korycki, R. (2016). *Methods of investigating the authenticity of recording of telephone conversations*. *Problems of Forensic Sciences*, *107*, 515–536.
23. Korycki, R. (2016). Problematyka badania autentyczności cyfrowych nagrań fonicznych. *Prokuratura i Prawo*, *12*, 138–157.
24. Kowalski, B., Radziszewski, R. (2017). Ekspertyza informatyczna. (In) M. Kała, D. Wilk, J. Wójcikiewicz (Eds.), *Ekspertyza sądowa. Zagadnienia wybrane* (pp. 634–673). Warszawa: Wolters Kluwer.
25. Michałek, M. (2009). The application of powerline hum in digital recording authenticity analysis. *Problems of Forensic Sciences*, *80*, 355–364.
26. Michałek, M. (2014). Database and parameters of digital audio test recordings. 16th ENFSI Forensic Speech And Audio Analysis Working Group Meeting, Wiesbaden, Germany.
27. Michałek, M. (2016). Test audio recordings and their use in authenticity examinations; Database of properties of digital audio recorders and recordings. *Problems of Forensic Sciences*, *105*, 355–369.
28. Michałek, M. (2018). Metadata in audio files compliant with ISO/IEC 14496-12 and their characteristics as well as the evaluation of usability in the investigation of the authenticity of recordings. *Problems of Forensic Sciences*, *115*, 241–261.
29. Microsoft Corporation. (1994). Microsoft Multimedia standards update, New Multimedia Data Types and Data Techniques, Revision 3.0.
30. Microsoft Corporation. (2010). Advanced Systems Format (ASF) Specification, Revision 01.20.05.

31. Orange.pl. (2019). iOS vs. Android – porównanie systemów operacyjnych smartfonów od Apple'a i Google'a. Retrieved April 10, 2020 from: https://www.orange.pl.
32. Savage, T. M., Vogel, K. E. (2014). *An introduction to digital multimedia*, *Second Edition*. United States of America: Jones and Barlett Learning.
33. Scientific Working Group on Digital Evidence, SWGDE. (2018). Best practices for digital audio authentication. Retrieved November 30, 2018 from: https://www.swgde.org.
34. Telepolis. (2020). Rynek smartfonów w Polsce i na świecie w 1Q2020 według Canalys. Samsung wciąż na czele. Retrieved May 8, 2020 from: https://www.telepolis.pl.
35. Xia, F., Hsu, Ch. H., Liu, X., Liu, H., Ding, F., Zha, W. (2013). The power of smartphones. Retrieved April 9, 2020 from: https://www.researchgate.net.
36. Willassen, S. (2005). Forensic analysis of mobile phone internal memory. (In) M. Pollit, S. Shenoi (Eds.), *Advances in Digital Forensics, IFIP International Conference on Digital Forensics, National Center for Forensic Science* (pp. 191–204). Orlando: Springer.

Corresponding author
Dr. Marcin Michałek
Institute of Forensic Research
ul. Westerplatte 9
PL 31-033 Kraków
e-mail: mmichalek@ies.krakow.pl

Problems of Forensic Sciences 2019, vol. 120, 335–361

CHARAKTERYSTYKA POPULARNYCH APLIKACJI DO REJESTRACJI DŹWIĘKU INSTALOWANYCH W SMARTFONACH Z SYSTEMEM OPERACYJNYM ANDROID W ODNIESIENIU DO BADAŃ FONOSKOPIJNYCH

1. Wstęp

Od czasu gdy technika cyfrowa połączona z miniaturyzacją urządzeń pozwoliła na konstruowanie niewielkich rozmiarów przenośnych odtwarzaczy multimedialnych zasilanych bateryjnie, stały się one bardzo często wykorzystywane. Popularne empetrójki czy empeczwórki ze względu na swoje małe gabaryty oraz funkcjonalność opanowały rynek multimedialnego sprzętu cyfrowego. Oprócz możliwości odtwarzania dźwięku, wideo lub plików graficznych wyposażone są one często w funkcję rejestracji nagrań dźwiękowych (Ho, Li, 2015; Savage, Vogel, 2014; Michałek, 2014). Nieco odmienną grupę stanowią urządzenia przeznaczone do rejestracji dźwięku, również przenośne i zasilane bateryjnie, określane często jako dyktafony cyfrowe. Jeszcze kilka lat temu ponad trzy czwarte cyfrowych nagrań dźwiękowych poddawanych badaniom fonoskopijnym rejestrowane było z wykorzystaniem wymienionych wyżej urządzeń (Michałek, 2016). Rynek sprzedaży w dalszym ciągu oferuje przenośne odtwarzacze multimedialne i dyktafony, jednak obecne źródła dowodowych nagrań z urządzeń mobilnych są już najczęściej inne.

Dynamiczny rozwój bezprzewodowych sieci telefonii komórkowej i oferowanych usług wręcz wymusił na producentach wprowadzenie na rynek aparatów telefonicznych, które umożliwią wykorzystanie pełnych możliwości takich sieci. Największą popularność zdobyły tzw. smartfony, czyli urządzenia łączące funkcję telefonu komórkowego i przenośnego komputera, wypierając skutecznie klasyczne telefony komórkowe. Funkcja rozmów telefonicznych czy komunikacji esemesowej stanowi aktualnie tylko część możliwości przeciętnego smartfonu (Xia, Hsu, Liu, Liu, Ding, Zha, 2013). Zwykle wśród dziesiątek aplikacji fabrycznie zainstalowanych w jego systemie operacyjnym co najmniej jedna służy *stricte* do rejestracji dźwięku. W przypadku braku takiej aplikacji lub potrzeby korzystania z innej serwisy internetowe oferują wiele tego rodzaju programów dla wybranego modelu smartfonu i jego systemu operacyjnego oraz umożliwiają samodzielne pozyskanie i zainstalowanie aplikacji dopasowanej do preferencji użytkownika. Użyteczność, duże możliwości dostosowywania do własnych upodobań i potrzeb oraz ogólna dostępność tych aparatów sprawiły, że aktualnie jest to podstawowe urządzenie w życiu codziennym o wielu zastosowaniach (Xia i in., 2013). Działalność opiniodawcza w zakresie fonoskopii pozwala na stwierdzenie, że na przestrzeni ostatnich kilku lat nastąpiła znaczna zmiana i obecnie spośród wszystkich urządzeń mobilnych to smartfony stanowią podstawowe źródło dowodowych nagrań dźwiękowych przy znikomym udziale przenośnych odtwarzaczy i dyktafonów. Konsekwencją tego jest również zmiana formatów zapisu dźwięku dostosowanych do systemów operacyjnych i możliwości tych urządzeń.

Obecnie na świecie około 2,7 miliarda osób użytkuje 3 miliardy smartfonów. Pomimo dynamicznej sytuacji na rynku od kilku lat najpopularniejszymi modelami smartfonów, zarówno globalnie, jak i w Polsce, są produkty koreańskiej marki Samsung, a na dalszych pozycjach znajdują się produkty Xiaomi, Huawei czy Apple (IDC, 2019, Telepolis, 2020). Aktualne modele pracują głównie pod kontrolą systemów operacyjnych Android albo iOS, które stworzyły swego rodzaju duopol i obsługują około 98% globalnego rynku smartfonów. Spośród wszystkich sprzedawanych smartfonów w system operacyjny Android wyposażonych jest, w przybliżeniu, 85% urządzeń w ujęciu globalnym, a 90% w Polsce (Orange, 2019). Z przedstawionego w dalszej części pracy opisu badań nad urządzeniami typu smartfony wynika, że statystyki te znajdują również potwierdzenie w praktyce eksperckiej. Do nagrywania dźwięku w tego rodzaju urządzeniach służą dedykowane aplikacje zainstalowane przez producenta lub użytkownika. Doświadczenie w zakresie rejestracji nagrań testowych badanymi do tej pory dowodowymi telefonami wskazuje, że są to nieskomplikowane aplikacje z podstawowymi funkcjami umożliwiającymi rejestrację nagrania i prostą edycję. Popularny serwis internetowy *Google Play* (poprzednia nazwa: *Android Market*) oferuje około 2,8 miliona aplikacji, z czego 95% bezpłatnych z możliwością ich zainstalowania i użytkowania w systemach Android (AppBrain, 2020). W serwisie tym wyszukując po słowach kluczowych przykładowo, „dyktafon" czy „rejestrator dźwięku", odnaleźć można dziesiątki bezpłatnych aplikacji przeznaczonych do zapisu nagrań dźwiękowych, które można w prosty sposób samodzielnie zainstalować w smartfonie. Zasoby internetowe zawierają również takie aplikacje, które instalowane są fabrycznie w nowych aparatach (przykładowo: aplikacje Samsunga).

W niniejszej pracy dokonano przeglądu i selekcji najpopularniejszych aplikacji przeznaczonych do rejestracji dźwięku pracujących pod kontrolą systemu Android i możliwych do zainstalowania z wykorzystaniem zaso-

bów sieci Internet. Aplikacje te zainstalowano w smartfonie marki Samsung, ustalono ich możliwości rejestracji i edycji, a następnie utrwalono nagrania testowe. Opisano również zasady postępowania w trakcie wykonywania nagrań testowych z wykorzystaniem tego rodzaju urządzeń. W dalszej kolejności przeprowadzono analizę tych aplikacji, jak również parametrów wykonanych nimi nagrań i właściwości plików dźwiękowych oraz metadanych. Dodatkowo opracowano algorytmy i program komputerowy pracujący w środowisku obliczeniowym MATLAB, przeznaczony jako narzędzie do automatycznej analizy struktur zapisanych plików dźwiękowych z nagraniami testowymi. Uzyskane wyniki i obserwacje poddano ocenie w celu ustalenia możliwości ich wykorzystania w badaniach autentyczności nagrań cyfrowych wykonanych przede wszystkim nowoczesnymi telefonami komórkowymi.

Głównym celem autora było zbadanie właściwości ogólnie dostępnych aplikacji pozwalających na rejestrację dźwięku oraz ocena zastosowania otrzymanych wyników w pracy eksperta z zakresu fonoskopii bez zamiaru promowania czy krytyki jakiejkolwiek marki, systemu operacyjnego lub oprogramowania.

2. Problematyka analizy dowodowych i testowych nagrań dźwiękowych rejestrowanych smartfonami

Dowodowe nagrania cyfrowe utrwalane są przeważnie nie za pomocą wysokiej jakości profesjonalnych rejestratorów, ale urządzeń powszechnego użytku (Michałek, 2014). Nagrania dźwiękowe, które poddawane są obecnie badaniom w Pracowni Analizy Mowy i Nagrań Instytutu Ekspertyz Sądowych, relatywnie często zarejestrowano z wykorzystaniem popularnych dziś aparatów telefonicznych nowej generacji, tj. smartfonów. Do analizy przekazywane są także tego typu dowodowe urządzenia, w pamięci których znajdują się nagrania mające zostać poddane zleconym badaniom. Indywidualnie dla każdego dowodu dokonuje się oceny jego stanu technicznego oraz możliwości zabezpieczenia danych z pamięci. W sytuacji gdy do badań dostarczono telefon oraz ewentualnie zainstalowaną w nim kartę pamięci, zawartość obu zabezpieczana jest podczas analizy informatycznej. Jeżeli jest to możliwe, wykonuje się wówczas wierną kopię zawartości całej pamięci albo jej dostępnej części w postaci tzw. obrazu. Tak wykonana kopia to najbardziej użyteczna forma zabezpieczenia danych do dalszej analizy fonoskopijnej (Kajstura, Michałek, Trawińska, 2017). Obraz pamięci traktowany jest jako odzwierciedlenie zawartości nośnika wraz z zachowaniem układu struktury plików, ich właściwości i sygnatur czasowych, co ma duże znaczenie dla badań autentyczności. Wykonanie obrazu pozwala również ustalić, czy w tak wykonanej kopii pamięci znajdują się ewentualne usunięte pliki, w tym dźwiękowe, oraz podjęcie próby ich odzyskania (Casey, 2010; Kajstura i in., 2017; Kowalski, Radziszewski, 2017; Willassen, 2005). Dla analizowanego obszaru pamięci źródłowej oraz jej obrazu należy wyliczyć sumy kontrolne i porównać je ze sobą w celu weryfikacji poprawności kopii. Jeżeli nie można wykonać obrazu zawartości pamięci, dane zabezpiecza się poprzez sporządzenie logicznej kopii możliwych do skopiowania plików i katalogów z badanego nośnika. Możliwa jest również eksploracja zawartości pamięci dowodowego nośnika, która pozwala na ustalenie liczby i właściwości dostępnych plików, co ma szczególne znaczenie, gdy nie jest możliwe przeprowadzenie procesu obrazowania (Casey, Turnbull, 2011; Willassen, 2005). Zabezpieczenie zawartości pamięci przekazanego do badań urządzenia lub nośnika oraz sposób jego wykonania jest istotnym elementem determinującym późniejszą analizę fonoskopijną, w szczególności analizę autentyczności nagrań. W przypadku dostarczenia do badań aparatu telefonicznego wraz z kartą SIM zostaje ona zabezpieczona jako dowód rzeczowy, lecz na potrzeby fonoskopijne nie jest poddawana badaniom, gdyż ich pamięć nie przechowuje plików multimedialnych. Do bezpiecznej eksploracji i wykonania kopii zawartości pamięci smartfonów wykorzystuje się urządzenie typu *write-blocker*, uniemożliwiające przypadkową modyfikację danych (Kowalski, Radziszewski, 2017; Michałek, 2018).

Na podstawie przeanalizowanych dotychczas smartfonów w ramach działalności opiniodawczej można stwierdzić, że w znacznej większości były to urządzenia marki Samsung oraz pojedyncze modele marek HTC i Manta – wszystkie pracujące pod kontrolą systemu operacyjnego Android. Przebadano również jedno urządzenie iPhone 3G z systemem operacyjnym iOS. Każde z tych urządzeń miało zainstalowaną jedną lub sporadycznie dwie aplikacje przeznaczone do rejestracji dźwięku. Zauważono, że były to nieskomplikowane w obsłudze aplikacje pozwalające na podstawowe czynności podczas rejestracji: rozpoczęcie, wstrzymanie i zakończenie nagrywania. Niektóre z nich, przeważnie zainstalowane w telefonach Samsung, umożliwiały prostą edycję zapisanych nagrań polegającą głównie na usunięciu ich fragmentów i zapisaniu zmian w plikach źródłowych lub utworzenie nowych. Wszystkie analizowane dotąd aplikacje umożliwiały zmianę nazw plików z nagraniami dźwiękowymi w trakcie albo po ich zapisaniu w pamięci. Zauważono jednakże, że zwykle nazwy plików z dowodowymi nagraniami są nazwami domyślnymi dla zainstalowanych aplikacji i odpowiadają nazwom plików z nagraniami testowymi wykonywanymi dowodowymi urządzeniami w trakcie ich badań.

Analiza zabezpieczonych danych pochodzących z pamięci smartfonów wykazała, że znajdujące się tam nagrania dźwiękowe odznaczały się zapisem głównie

w formatach M4A i 3GP. Dane dźwiękowe tego rodzaju plików wraz ze współtowarzyszącymi metadanymi umieszczane były w tzw. kontenerze multimedialnym typu MPEG-4, związanym ze standardami ISO (ISO/IEC, 2003; ISO/IEC, 2010; ISO/IEC, 2015). Formaty M4A i 3GP oraz kontener MPEG-4 są bardzo często spotykane w nowoczesnych urządzeniach mobilnych (3GPP, 2010; Gloe, Fischer, Kirchner, 2014; Ho, Li, 2015; Michałek, 2018). Pliki tego typu są w dość łatwy sposób konwertowane do postaci nieskompresowanej PCM WAVE lub nawet odtwarzane w czasie rzeczywistym przez standardowo wykorzystywane oprogramowania, co stanowi podstawę do dalszej analizy fonoskopijnej (Kajstura i in., 2017; Michałek, 2014).

Uznanie nagrań jako materiału dowodowego oraz zakres badań określanych przez organa procesowe sprawia, że konieczna jest analiza ich autentyczności (Grigoras, 2005; Kajstura, Trawińska, Hebenstreit, 2005; Koenig, 1990; Korycki, 2011, 2016; Michałek, 2009). Definicja autentyczności nagrania ogłoszona przez *Audio Engineering Society* (AES) wskazuje na potrzebę oceny jego ciągłości i oryginalności oraz związanej z tym metody rejestracji (AES, 1996). Ważną częścią w przebiegu badań zarówno samego urządzenia, jak i autentyczności cyfrowego nagrania dowodowego są nagrania testowe. Pozwalają one na stwierdzenie, czy dowodowe urządzenie mogło być użyte do utrwalenia dowodowego nagrania oraz na przeprowadzenie analizy porównawczej parametrów i struktury pliku z nagraniami testowymi (Cooper, 2006; Kajstura i in., 2017; Koenig, Lacey, 2012; Michałek, 2016). Pozwalają również ustalić, czy urządzenie umożliwia rejestrację przydźwięku sieciowego oraz ujawnić ewentualne inne charakterystyczne cechy, jak na przykład składową stałą (Grigoras, Cooper, Michałek, 2009; Kajstura i in., 2005; Korycki, 2011).

Omówione wyżej aspekty dotyczą również nagrań rejestrowanych przez smartfony, przy czym pojawienie się nowych urządzeń rejestrujących nagrania dowodowe wymusiło opracowanie nowych sposobów przygotowania i realizacji nagrań testowych. Smartfony to skomplikowane cyfrowe urządzenia wielofunkcyjne z własnym systemem operacyjnym i wysokim prawdopodobieństwem, że każdy inny model będzie miał zainstalowaną odmienną aplikację do rejestracji dźwięku. Jeżeli zleceniodawca nie dostarczył urządzenia, które miało zostać użyte do rejestracji nagrania przekazanego do analizy (najczęściej przekazał jedynie kopię nagrania zapisaną na innym nośniku), a postanowienie zawiera kwestie związane z analizą autentyczności, należy zwrócić się do niego z wnioskiem o dostarczenie do badań tego urządzenia.

W procesie przygotowywania nagrań testowych smartfonu należy zweryfikować jego ustawienia systemowe, w tym zegara czasu, oraz zainstalowane aplikacje. Dla aplikacji rejestrujących dźwięk ustalane są w pierwszej kolejności bieżące ustawienia zapisu, wybrana pamięć i ścieżka folderu przechowywania, lista plików z nagraniami i ewentualnie innych plików towarzyszących oraz zakres prezentowanych informacji o plikach. W dalszej kolejności należy ustalić wszystkie możliwe ustawienia aplikacji, takie jak formaty i parametry zapisu oraz nazewnictwo plików, jak również określić możliwy zakres ich zmian. Przed rejestracją „właściwych" nagrań testowych, które będą objęte późniejszą analizą porównawczą, dobrą praktyką jest wykonanie nagrania lub kilku nagrań (dla każdego formatu przynajmniej jedno) w celu ustalenia, jakie opcje dostępne są w trakcie rejestracji i bezpośrednio po jej zakończeniu. Doświadczenie wskazuje, że w smartfonach niektóre opcje aplikacji uaktywniają się tylko w trakcie rejestracji. Pozwala to również ustalić, w jaki sposób zapisywane są nagrania, tj. automatycznie czy z interakcją użytkownika, ich rzeczywiste nazwy i możliwe do odczytania parametry. Nie wszystkie aplikacje podają w sposób standardowy format zapisu nagrań lub rozszerzenia plików, określając parametry enigmatycznymi niekiedy nazwami w postaci „jakość dobra" albo „małe pliki". Zapoznanie się z możliwościami danej aplikacji pozwala na opracowanie optymalnego planu wykonania nagrań testowych, zawierającego opis ich realizacji, kolejność nagrań i sekwencji odtwarzanych sygnałów testowych oraz zastosowanie wszystkich możliwych ustawień i funkcji. W praktyce nagrania testowe zapisywane są nie w pamięci wbudowanej w dowodowy smartfon, aby nie ingerować w jej zawartość, ale na zainstalowanych wymiennych kartach pamięci microSD. Należy przy tym stosować karty o możliwie dobrej prędkości zapisu, na przykład klasy 10, aby zapewnić nieprzerwany zapis strumienia danych na nośniku. Etap przygotowywania i wykonywania nagrań testowych w celach bezpieczeństwa odbywa się w ekranowanej komorze bez dostępu do sieci telefonii komórkowej. Wszelkie badania aparatu telefonicznego przeprowadzane są bez zainstalowanej karty SIM i ewentualnych dowodowych kart pamięci oraz bez zmiany ustawień systemu, zawartości folderów i plików. Po zakończonych badaniach fonoskopijnych rutynowo dokonuje się zweryfikowania stanu pamięci smartfonu z uwzględnieniem plików zawierających nagrania, zarówno z poziomu aplikacji nagrywającej, jak i menedżera plików.

Właściwe wykonanie nagrań testowych z wykorzystaniem smartfonu to zadanie czasochłonne i poprzedzone odpowiednim przygotowaniem, jednak istotne w procesie analizy urządzenia, aplikacji i zarejestrowanych nagrań oraz badań autentyczności nagrań dowodowych.

3. Materiał do badań

3.1. Urządzenie i aplikacje do rejestracji dźwięku

Za warstwę sprzętową realizacji badań odpowiadał popularny smartfon marki Samsung, model Galaxy J3 Dual SIM 2017. W chwili zakupu posiadał on fabrycznie zainstalowany system operacyjny Android w wersji 8.0.0 oraz aplikację do rejestracji nagrań dźwiękowych *Samsung Voice Recorder* w wersji 21.0.22.166.

W celu wykonania zamierzonych badań z wykorzystaniem serwisu *Google Play* przeanalizowano opis dostępnych aplikacji do rejestracji nagrań dźwiękowych. Następnie na podstawie połączonej oceny wynikającej z: liczby instalacji aplikacji (popularności), wysokiej oceny użytkowników i recenzentów oraz możliwości i funkcjonalności, wykonano selekcję ośmiu bezpłatnych aplikacji. Nazwy tych aplikacji są następujące (w nawiasie podano wersję aplikacji i przez kogo opublikowana): *Voice recorder* (2.34, Splend Apps), *Easy Voice Recorder* (2.6.1, Digipom), *Voice Recorder* (5 (36.0), recorder&smart apps), *Smart Recorder* (1.9.6, Smart Mob), *Voice Recorder* (3.08, Appliqato Software), *Voice Recorder* (1.5.6, thinksimple app), *EZ Voice Recorder* (1.5.11, Top 1), *Tape-a-Talk Voice Recorder* (2.0.7, Markus Drösser). Dla orientacji należy uzupełnić, że informacje zamieszczone w serwisie *Google Play* podają liczbę pobrań od około 1 miliona dla aplikacji *Tape-a-Talk Voice Recorder* do ponad 1 miliarda dla *Samsung Voice Recorder*. Wszystkie wyszczególnione aplikacje zainstalowano w opisanym wyżej smartfonie Samsung Galaxy J3, a po wykonaniu nimi nagrań testowych uaktualniono system operacyjny Android tego urządzenia do wersji 9. Po wykonaniu nagrań aplikacją *Samsung Voice Recorder* w wersji 21.0.22.166 uaktualniono tę aplikację do nowszej wersji 21.1.06.11 i z jej wykorzystaniem zarejestrowano kolejne nagrania testowe.

Łącznie przeanalizowano właściwości 10 aplikacji przeznaczonych do rejestracji nagrań dźwiękowych w systemie Android i w wersjach 8 oraz 9. Analizę tę wykonano za pomocą:

- informacji odczytywanych bezpośrednio z menu uruchomionej aplikacji,
- menedżera plików zainstalowanego w systemie Android w smartfonie,
- eksploratora plików systemu Windows 7 Ultimate.

3.2. Nagrania dźwiękowe

Za pomocą wymienionych wyżej 10 aplikacji zainstalowanych w systemie Android zarejestrowano 142 nagrania testowe z zapisem dźwięku w 6 różnych formatach. Nagrania te wykonano z wykorzystaniem wszystkich możliwych ustawień zapisu oraz funkcji dostępnych w trakcie rejestracji, takich jak: uruchomienie i zakończenie nagrywania, funkcja pauzy, aktywacja nagrywania poziomem głośności, nadpisanie nowszym zapisem, indeksowanie, zapis automatyczny i z interakcją użytkownika. Po wykonaniu nagrań testowych poddano je edycji za pomocą tych aplikacji, które miały taką możliwość, polegającą na: usunięciu fragmentu nagrania, skopiowaniu i wklejeniu fragmentu nagrania, nadpisaniu utrwalonego już nagrania nowszym zapisem lub naprawie nagłówka pliku. Nagrania te zawierają możliwie szeroki zakres dźwięków, takich jak: szum o ciągłym paśmie, tony w zakresie niskich, średnich i wysokich częstotliwości, mowę oraz celowe zakłócenia o charakterze impulsowym i ciągłym. Do bezpiecznej eksploracji zawartości pamięci z nagraniami i wykonania ich kopii wykorzystano urządzenie uniemożliwiające ewentualną modyfikację danych na tego typu nośnikach, tj. *write-blocker*.

W celu realizacji dalszych badań wykonano kopie plików z urządzenia nagrywającego na dysk twardy komputera. Analizę właściwości zgromadzonych plików z nagraniami testowymi wykonano za pomocą następujących narzędzi:

- programów przeznaczonych do analizy plików multimedialnych, tj. ffprobe w wersji git-2019-11-13-4e0860e oraz MediaInfo w wersji 19.09, eksploratora plików systemu Windows 7 Ultimate, menedżera plików zainstalowanego w systemie Android w smartfonie oraz informacji prezentowanych przez aplikacje rejestrujące nagrania;
- edytorów szesnastkowych 010 Editor w wersji 6.0.2 i HxD w wersji 1.7.7.0, przeznaczonych do odczytu i wizualizacji struktury danych utrwalonych plików;
- programu komputerowego opracowanego na potrzeby niniejszych badań.

Do analizy danych zawartych w plikach z nagraniami testowymi opracowano program komputerowy działający w środowisku obliczeniowym MATLAB. Program ten umożliwia automatyczną eksplorację struktury i detekcję metadanych w plikach dźwiękowych o dowolnym formacie, jednak zaprojektowano go przede wszystkim do analizy plików zapisanych w standardach MPEG-1 – MPEG-4 oraz RIFF. Pliki tego rodzaju stanowią niemal całość zgromadzonego materiału do badań i mogą zawierać bardziej rozbudowane metadane. Działanie utworzonego programu można przedstawić następująco:

1) import do przestrzeni roboczej pliku dźwiękowego wybranego do analizy; algorytm importu na podstawie rozszerzenia pliku rozpoznaje jego typ i proponuje metodę analizy odpowiednią dla danego formatu, przy czym użytkownik ma również możliwość samodzielnego określenia typu pliku i wyboru metody,
2) zdefiniowanie obszaru analizy, tzn. całość pliku albo jego część zawierająca metadane; program analizujący może pominąć zakodowane dane dźwiękowe, przykładowo w plikach MPEG cały boks *mdat*,

a w plikach RIFF cały chunk *data*; obszar pliku z zakodowanymi mediami to najczęściej najobszerniejsza jego część, w której zwykle nie występują żadne metadane,
3) eksploracja wskazanego obszaru bajt po bajcie i wyszukanie wszystkich ciągów znaków mogących stanowić metadane; algorytm detekcji oraz decyzyjny opiera się na zaimplementowanych w programie sekwencjach *four-character code* (FourCC) dla boksów i chunków (głównie dla MPEG i RIFF), jak również pozostałych możliwych znaczników, identyfikatorów i tagów dla wszystkich analizowanych formatów; dla zapewnienia poprawności działania algorytm uwzględnia również wielkie i małe litery, położenie w pliku, rozmiar i odstęp pomiędzy kolejnymi rozpoznanymi metadanymi; w przypadku detekcji nieznanej dotąd sekwencji FourCC program przedstawia ją również w wynikach, w celu uzupełnienia opisu i do wiedzy użytkownika; rozwinięcie przytoczonych tutaj pojęć zostanie przedstawione w dalszej części pracy,
4) zapis struktury metadanych analizowanego pliku do arkusza kalkulacyjnego celem ewentualnej dalszej analizy porównawczej z innymi plikami,
5) opcjonalny zapis do arkusza metadanych wskazanych przez użytkownika, tj. nazwy, rozmiaru, położenia w pliku i zawartości.

W trakcie wykonywania badań zgromadzonych plików dźwiękowych tworzony program poddawano optymalizacji na zasadzie pętli sprzężenia zwrotnego w celu jego wykorzystania do analizy różnorodnych formatów i możliwych struktur danych. Oprócz uzyskanych wyników badań, podczas opracowywania programu posłużono się również dostępnymi opracowaniami i dokumentacją techniczną (3GPP, 2010; 3GPP, 2011; ISO/IEC, 1993; ISO/IEC, 2015; Microsoft Corp., 1994; Microsoft Corp., 2010). Oprogramowanie to stanowi narzędzie wspomagające badania autentyczności nagrań dźwiękowych, jakie wykonywane są w ramach działalności opiniodawczej Instytutu.

4. Wyniki

4.1. Analiza właściwości zainstalowanych aplikacji

Celem niniejszej części badań było ustalenie, jakie są możliwości rejestracji nagrań dźwiękowych, ich zapisu oraz ewentualnej edycji z wykorzystaniem aplikacji zainstalowanych w smartfonie Samsung.

W tabeli 1 przedstawiono podstawowe właściwości 10 aplikacji, którymi wykonano nagrania testowe: format zapisu dźwięku, dostępne funkcje w trakcie rejestracji, możliwość edycji nagrań przez daną aplikację oraz jej opcje. Aplikacje te pozwalają na zapis nagrań dźwiękowych w 6 różnych formatach: 3GP, AAC, AMR, M4A (MP4 z audio), MP3 i WAV. Analiza parametrów plików z nagraniami testowymi, opisana szczegółowo w dalszej części pracy, pozwoliła ustalić, że badane aplikacje najczęściej wykorzystują kontener multimedialny typu MPEG-4 oparty na standardach ISO. Zawiera on relatywnie dużo metadanych w odniesieniu do pozostałych kontenerów multimedialnych, co jest cechą pozytywną dla analizy autentyczności nagrań.

Oprócz podstawowych funkcji uruchomienia i zakończenia nagrywania wszystkie aplikacje wyposażone są w funkcję wstrzymywania nagrywania, tj. pauzy. To istotne informacje, ponieważ aktywacja tej funkcji – w sposób celowy czy przypadkowy – powoduje nieciągłość zapisu nagrania. Analiza ciągłości nagrania jest jednym z podstawowych badań w trakcie analizy jego autentyczności (Kajstura i in., 2017; Korycki, 2016). W nagraniach wykonanych starszymi modelami telefonów, które wyposażone zostały w przyciski funkcyjne, zastosowanie pauzy powodowało często rejestrację charakterystycznych odgłosów. Wykorzystanie w smartfonach ekranów dotykowych znacznie ogranicza rejestrację takich odgłosów, dlatego nieciągłości tego rodzaju bywają trudne do ujawnienia (Kajstura i in., 2017). Co ciekawe, niektóre z badanych aplikacji pozwalają na tę funkcję jedynie dla wybranych formatów zapisu (pozycje: 2, 3 i 9 w tabeli 1). Dodatkowo w aplikacji *Smart Recorder* dostępna jest funkcja automatycznej pauzy. Jej włączenie powoduje automatyczne wstrzymanie nagrywania, gdy wartość natężenia dźwięku spadnie poniżej ustalonego progu, oraz jego aktywację dźwiękiem o natężeniu powyżej tego progu.

Połowa analizowanych aplikacji umożliwia wykonanie edycji zarejestrowanych przez siebie nagrań (pozycje: 1, 4 oraz 8–10 w tabeli 1), przy czym aplikacja *Tape-a-Talk Voice Recorder* jedynie dla formatu WAV. We wszystkich tych aplikacjach po zapisie nagrania możliwe jest usunięcie początkowego i końcowego fragmentu lub też fragmentu w obrębie nagrania (opcja *trim*). Aplikacja *Samsung Voice Recorder* w obu wersjach umożliwia w zapisanym albo rejestrowanym i wstrzymanym nagraniu ustawienie początkowego punktu rejestracji i nadpisanie nagrania nowym zapisem. Aplikacja *Voice Recorder* (recorder&smart apps) pozwala na skopiowanie i wklejenie fragmentu nagrania w inne jego miejsce, natomiast *Tape-a-Talk Voice Recorder* – dogranie dalszej części zapisu, począwszy od końca zapisanego już nagrania.

Z istotnych opcji dodatkowych wymienić należy dla aplikacji *Samsung Voice Recorder* możliwość ustawiania przez użytkownika punktów kontrolnych (tzw. zakładek) wraz z ich opisem w trakcie i po zapisie nagrania. Z kolei aplikacja *Tape-a-Talk Voice Recorder* posiada opcję naprawy nagłówka dla plików WAV, polegającą na automatycznym wpisaniu do metadanych parametrów pliku podanych przez użytkownika: częstotliwości próbkowania, kwantyzacji bitowej i liczby kanałów.

W tabeli 2 przedstawiono wyniki analizy takich właściwości zainstalowanych aplikacji, jak nadawane rozszerzenia i domyślne nazwy dla plików z nagraniami źródłowymi i po edycji oraz opcje zapisu nagrań. Analiza taka okazuje się istotna, gdyż domyślne nazwy plików źródłowych w przypadku 6 aplikacji (pozycje: 1, 3, 6, 7, 9 i 10) różnią się zawsze od nazw pozostałych aplikacji. Z kolei dla 4 pozostałych aplikacji (pozycje: 2, 4, 5 i 8) nazwy domyślne ich plików są różne tylko dla wybranych formatów albo z opcją dodania sygnatury czasu do nazwy pliku. Istotną właściwością aplikacji jest umieszczanie w nazwach plików informacji o dacie i godzinie, zarówno jako element stały (pozycje: 4, 7 i 8), jak i jako opcja ustawiana przez użytkownika (pozycje 5 i 9). Sygnatury czasu w nazwach tego rodzaju plików związane są z rozpoczęciem rejestracji nagrania, jednak są one zależne od ustawienia wbudowanego zegara smartfonu. Większość aplikacji, tj. z wyjątkiem wskazanych w pozycji 4, 7 i 8, numeruje swoje nagrania rosnąco. Dla niektórych aplikacji istnieją też w omawianym zakresie pewne cechy charakterystyczne. Dla aplikacji *Samsung Voice Recorder* nazwy plików zapisane na karcie microSD otrzymują dodatkowy ciąg znaków *_sd*. Aplikacja *Voice Recorder* (Appliqato Software) dodaje do nazwy pliku niepowtarzalny ciąg znaków alfanumerycznych, natomiast *Easy Voice Recorder* dla formatu M4A umożliwia zapis pliku z wybranym przez użytkownika rozszerzeniem *m4a* albo *mp4*.

Nagrania poddane edycji można zapisać w nowym lub w tym samym (źródłowym) pliku (pozycje: 1, 9 i 10), albo wyłącznie w nowym pliku (pozycje 4 i 8). Cechą wspólną dla plików poddanych edycji i zapisanych jako źródłowe jest ich niezmieniona nazwa. Z kolei wszystkie aplikacje do plików zapisanych jako nowe dodają numerację oznaczającą kolejną edycję (przykładowo *-1* i *-2* dla aplikacji z pozycji 1) lub ciągi znaków wskazujące na jej wykonanie (na przykład *_CUT1* dla aplikacji z pozycji 8).

W tabeli 2 przedstawiono również, jak odbywa się zapis nagrania do pliku po jego zakończeniu dla konkretnych aplikacji, co może nastąpić na dwa różne sposoby. W pierwszym przypadku zapis następuje po zatwierdzeniu proponowanej przez aplikację nazwy lub po wprowadzeniu własnej. Wówczas aplikacja oczekuje na interakcję ze strony użytkownika, co może mieć wpływ na sygnaturę czasu modyfikacji pliku. W drugim przypadku aplikacja zapisuje do pliku z nazwą domyślną i automatycznie po zakończeniu nagrywania.

Menu wszystkich analizowanych aplikacji umożliwia zmianę nazwy plików z zapisanymi już nagraniami. Jednak, jak zaznaczono wcześniej, zwykle nazwy plików z nagraniami dowodowymi są nazwami domyślnymi dla użytej aplikacji nagrywającej. W związku z tym analiza nazw plików dla aplikacji do rejestracji nagrań dźwiękowych jest uzasadniona, gdyż może wspomagać identyfikację konkretnej aplikacji, co stwierdzono zarówno na przedstawionych wyżej przykładach, jak i wskazano w literaturze (Kajstura i in., 2017; Michałek, 2016).

4.2. Analiza właściwości nagrań i plików dźwiękowych

W dalszej części pracy przedstawione zostaną właściwości nagrań i plików dźwiękowych utrwalonych za pomocą zainstalowanych 10 aplikacji.

Parametry zapisu

Jak zaznaczono, za pomocą wyselekcjonowanych do badań aplikacji zarejestrowano 142 nagrania z zapisem dźwięku w 6 różnych formatach. Odpowiednio do częstości ich występowania są to następujące formaty: WAV – w 7 aplikacjach, M4A – 6, MP3 – 4, następnie AAC i 3GP – 2 oraz AMR – 1. W tabeli 3 dla konkretnych aplikacji i możliwych formatów zapisu podano nazwy predefiniowanych ustawień jakości dźwięku oraz ich parametry: częstotliwość próbkowania w [kHz] oraz przepływność bitową w [kbps]. Jak widać, menu części aplikacji umożliwia wybór zdefiniowanych ustawień parametrów dla różnych formatów bez możliwości ich zmiany, które określone są nazwami wskazującymi na jakość zapisu. Pozostałe aplikacje umożliwiają swobodny wybór ustawień przez użytkownika w możliwym zakresie. Dla dwóch formatów, tj. 3GP i AAC (pozycje 3 i 9), nie było możliwości zmiany jakichkolwiek parametrów zapisu, a jedynie rejestracja z jednym zdefiniowanym ustawieniem. Z kolei dla aplikacji z pozycji 2 i dla formatu 3GP menu wskazywało na możliwość zmiany parametrów, jednak zapisane pliki odznaczały się zawsze takimi parametrami, które podano w tabeli 3.

W tabeli 4 przedstawiono kodeki wykorzystywane do zapisu dźwięku dla poszczególnych formatów, które zaimplementowano w testowanych aplikacjach. Z kolei w tabeli 5 zaprezentowano zastosowany kontener multimedialny lub strukturę danych w celu zapisu zakodowanych danych dźwiękowych i opisujących je metadanych.

Na podstawie uzyskanych wyników, które przedstawiono w tabelach 4 i 5, dla poszczególnych aplikacji można zaobserwować interesujące cechy charakterystyczne. Aplikacje z pozycji 2 i 3 stosują kodeki AMR w plikach 3GP i umieszczają dane w kontenerze MPEG-4. *Voice Recorder* (Splend Apps) wykorzystuje rzadko stosowany kodek AMR w odmianie szerokopasmowej, tj. AMR-WB, a w plikach dźwiękowych z rozszerzeniem *mp3* i oznaczonych jako format MP3 stosuje kodowanie AAC-LC i umieszcza dane w kontenerze MPEG-4. Aplikacja *Voice Recorder* (Appliqato Software) jako jedyna stosuje „klasyczną" strukturę pliku AMR z kodekiem AMR-NB (*Narrowband*). Kilka aplikacji (pozycje: 4, 7 i 8) w plikach formatu MP3 wykorzystują standardy

MPEG-1, MPEG-2 i MPEG-2.5 dla różnej jakości nagrań, tj. MPEG-1 dla najlepszej, a MPEG-2.5 dla najgorszej.

Otrzymane wyniki pozwalają stwierdzić, że wśród testowanych aplikacji najczęściej stosowany jest kontener typu MPEG-4, bardzo uniwersalny i przeznaczony do umieszczania różnego rodzaju multimediów wraz z metadanymi. Podobnie jak w przypadku nagrań testowych jest on aktualnie najczęściej spotykanym kontenerem, w którym zapisywane są dowodowe nagrania dźwiękowe z wykorzystaniem smartfonów.

Sygnatury czasowe

Analizie poddano również informacje dotyczące czasu rejestracji nagrań testowych źródłowych oraz poddanych edycji za pomocą zainstalowanych aplikacji. Zbadano także wpływ sposobu zapisu tych nagrań na ich sygnatury czasowe.

W trakcie wykonywania każdego z nagrań testowych odnotowywano z zegara urządzenia rejestrującego czas ich rozpoczęcia oraz zakończenia. W niniejszej pracy opisano powyżej analizę domyślnych nazw zarejestrowanych plików dźwiękowych. Ustalono, że 6 testowanych aplikacji umożliwia dodanie do nazw plików ciągu znaków wskazujących na datę i godzinę. Sygnatury czasowe plików z tymi nagraniami odczytywane były za pomocą menu testowanych aplikacji, systemu operacyjnego Android oraz Eksploratora systemu Windows bezpośrednio z pamięci, gdzie zapisane zostały nagrania.

W tabeli 6 przedstawiono nazwy przykładowych plików z nagraniami dźwiękowymi wraz z datą i godziną początku i końca rejestracji oraz sygnatury czasowe modyfikacji odczytane podanymi wyżej trzema sposobami. Wyniki wskazują, że ciągi znaków z datą i godziną zawarte w nazwach plików odpowiadają dacie i godzinie rozpoczęcia ich rejestracji, tzn. uruchomienia nagrywania. Z kolei sygnatury czasowe modyfikacji odczytane za pomocą systemu Android i Eksploratora Windows związane są z zakończeniem rejestracji nagrań i zapisem ich do pliku. Podobnie, menu większości analizowanych aplikacji umożliwia odczyt sygnatury czasowej modyfikacji plików, za wyjątkiem *Smart Recorder* (Smart Mob) i *EZ Voice Recorder* (Top 1), które podają sygnaturę utworzenia.

Menedżer plików systemu Android nie umożliwiał odczytu sygnatury czasowej utworzenia zarejestrowanych plików dźwiękowych, z kolei system Windows wskazywał na sygnaturę czasową utworzenia taką samą, jak sygnatura modyfikacji, czyli nieodpowiadającą czasowi rzeczywistemu.

Nagrania wskazane w tabeli 6 zarejestrowano jako ciągłe i zapisano je bezpośrednio po zakończeniu z nazwami proponowanymi przez aplikacje. Jak ustalono, 5 spośród testowanych aplikacji wskazanych w tabeli 2 po zakończeniu nagrywania oczekuje na interakcję użytkownika przed zapisem nagrania do pliku. W celu ustalenia, czy czas ten ma wpływ na sygnatury czasowe plików, wykonano dodatkowe ciągłe nagrania testowe, podczas których od chwili zakończenia nagrywania do chwili potwierdzenia i zapisu odczekano, aż zegar smartfonu wskaże następną minutę. Ustalono, że czas modyfikacji zarejestrowanych w ten sposób plików odpowiadał czasowi ich zakończenia w przypadku aplikacji z pozycji 4, 6 i 7, tj. *Voice Recorder* (recorder&smart apps), *Voice Recorder* (Appliqato Software) i *Voice Recorder* (thinksimple app). Natomiast dla aplikacji *Samsung Voice Recorder* w obu testowanych wersjach (pozycje 1 i 10) czas modyfikacji pliku odpowiadał dopiero chwili potwierdzenia nazwy i zapisu, co użytkownik może wykonać po różnym okresie.

Przeanalizowano również wpływ edycji nagrań w aplikacjach mających taką opcję na wpływ sygnatury czasu modyfikacji. Ustalono w ten sposób, że wszystkie pliki po wykonanej edycji opisanej w tabeli 1 odznaczały się zmienionym (późniejszym) czasem modyfikacji, przy czym dla aplikacji z pozycji 4, 8 i 9, tj. *Voice Recorder* (recorder&smart apps), *EZ Voice Recorder* (Top 1) oraz *Tape-a-Talk Voice Recorder* nowa sygnatura modyfikacji odpowiadała chwili zapisu nagrania po zmianach. Co bardzo ciekawe, sygnatura czasu modyfikacji dla plików poddanych edycji w aplikacji *Samsung Voice Recorder* w obu wersjach zmieniała się już w chwili wykonania modyfikacji nagrania, jeszcze przez potwierdzeniem zmian, a fakt zapisu pliku nie wpływał na sygnaturę modyfikacji.

Jak można zauważyć, wykorzystanie trzech metod podanych w tabeli 6 umożliwia odczyt sygnatur czasu modyfikacji w zaokrągleniu do pełnej minuty. Należy również dodać, że wymienione w tabeli nazwy plików ze wskazaniem czasu rozpoczęcia nagrywania oraz sygnatury ich modyfikacji zależą od ustawienia zegara czasu wbudowanego w smartfon. Zazwyczaj menu systemu Android umożliwia zmianę ustawień tego zegara i wówczas wskazane wyżej informacje o czasie mogą różnić się od czasu rzeczywistego. W konfiguracji domyślnej smartfony mają jednak włączoną opcję automatycznej aktualizacji czasu dostarczanego od operatora sieci, co jest istotne dla badań autentyczności, gdyż wtedy czas ustawiony w urządzeniu odpowiada czasowi rzeczywistemu. Jak już wspomniano, zwykle nazwy dowodowych plików dźwiękowych z pamięci smartfonów odpowiadają nazwom domyślnym, jednak menu wszystkich analizowanych dotąd aplikacji, zarówno w urządzeniach dowodowych, jak i w ramach niniejszej pracy, umożliwiało zmianę nazwy plików z nagraniami. Opisane metody odczytu informacji o czasie rejestracji nagrania i jego ewentualnej edycji na podstawie nazw plików i sygnatur czasowych utworzenia lub modyfikacji są pomocne w procesie analizy autentyczności nagrań, jednakże

powinny być zawsze weryfikowane innymi sposobami, takimi jak metoda wykorzystująca sygnał przydźwięku sieciowego lub analiza metadanych pliku.

Analiza metadanych

Wizualizacja oraz analiza struktury plików i informacji zapisanych w metadanych jest aktualnie jedną z podstawowych metod badania autentyczności nagrań cyfrowych (Ho, Li, 2015; Kajstura i in., 2017; Korycki, 2016; SWGDE, 2018). Metadane, czyli dodatkowe informacje opisujące dane z zakodowanym dźwiękiem, mogą być bardzo istotne i zawierać parametry zapisu, sygnatury czasowe, dane odnośnie do urządzenia i systemu operacyjnego oraz wiele innych w zależności od formatu i kontenera multimedialnego (Gloe i in., 2014; Koenig, Lacey, 2012; Korycki, 2016; Michałek, 2018). W przypadku rzeczywistych spraw wykonanie nagrań testowych badanym urządzeniem umożliwia analizę porównawczą struktury ich plików ze strukturą pliku dowodowego. Przeprowadza się ją z wykorzystaniem oprogramowań umożliwiających wizualizację plików w postaci kodu szesnastkowego i ASCII.

Na podstawie analizy wszystkich nagrań testowych ustalono, że relatywnie najwięcej metadanych znajduje się w plikach zapisanych w kontenerze MPEG-4. Jak ustalono, jest on też najczęściej wykorzystywany przez analizowane aplikacje, co wskazano w tabeli 5. W strukturze tego rodzaju plików testowych ujawniono także ślady jednoznacznie wskazujące na edycję nagrania dźwiękowego za pomocą zainstalowanych aplikacji.

Wewnętrzna struktura testowych plików dźwiękowych w kontenerze MPEG-4 opiera się na wytycznych wskazanych w zestawie norm ISO/IEC 14496 (ISO/IEC, 2003; ISO/IEC, 2010; ISO/IEC, 2015). W jej obrębie wyróżnić można tzw. boksy, czyli wyodrębnione i autonomiczne części pliku zawierające pola z metadanymi (informacjami) lub zakodowanymi danymi multimedialnymi, które tworzą uporządkowaną i hierarchiczną strukturę pliku. Dane w boksach zapisywane są w konwencji *big-endian*, czyli najbardziej znaczący bajt występuje jako pierwszy.

Analiza struktury wszystkich testowych plików dźwiękowych zapisanych w kontenerze MPEG-4 wykazała, że każdy z nich zawiera 3 główne boksy: *File Type* (identyfikator w strukturze: *ftyp*) z informacjami o kompatybilnych specyfikacjach do odtworzenia pliku, *Media Data* (*mdat*) z zakodowanymi danymi multimedialnymi oraz *Movie Box* (*moov*) z metadanymi. Analiza boksu *File Type* dla plików testowych wskazuje, że aplikacja *Samsung Voice Recorder* w obu wersjach korzysta ze specyfikacji podstawowej 3GPP w wersji 4, natomiast pozostałe aplikacje zapisujące pliki w MPEG-4 wykorzystują specyfikację MP4 w wersji 2. W żadnym pliku testowym w obrębie boksu *Media Data*, oprócz jego nagłówka, nie ujawniono dodatkowych metadanych. W kontenerze MPEG-4 najistotniejszym obiektem do badań jest boks o nazwie *Movie Box*, w którym zawarte są metadane dla zakodowanych mediów w obrębie licznych boksów wewnętrznych, z których istotne dla niniejszej pracy zostaną omówione poniżej. Według zaleceń norm ISO powinien on być zlokalizowany na końcu całego pliku (ISO/IEC, 2015).

W przypadku wszystkich zainstalowanych aplikacji pliki testowe w formacie M4A, MP3 i AAC w obiekcie *Movie Box* zawierają 28 boksów wewnętrznych. W obu wersjach aplikacja *Samsung Voice Recorder* w plikach M4A zamieszcza dodatkowe 4 boksy informujące o: zakładkach umieszczonych przez użytkownika, ich liczbie, położeniu i ewentualnym opisie. Aplikacja ta dodaje jeszcze istotny boks oznaczony identyfikatorem *vrdt* z ciągiem znaków ASCII *com.sec.android.app.voicenote.common.util.VoiceRecorder* informującym, że plik zarejestrowano aplikacją *Voice Recorder* marki Samsung, przeznaczoną dla systemu operacyjnego Android. Pliki testowe zapisane w formacie 3GP zawierają w obiekcie *Movie Box* 29 wewnętrznych boksów. Wszystkie analizowane pliki testowe w kontenerze MPEG-4 zawierają boks *User Data* (*udta*), w obrębie którego znajdują się boksy *SDLN*, *smrd* i *smta* wskazujące na to, że plik został zapisany z wykorzystaniem urządzenia Samsung. Istotnym obiektem we wszystkich analizowanych plikach MPEG-4 jest również boks *meta*, w którym wpisy *com.android.version* z wartościami *8.0.0* albo *9* wskazują na system operacyjny Android i jego wersję. Przykłady takich metadanych dla nagrań wykonanych aplikacjami *Samsung Voice Recorder* (pozycje 1 i 10 w tabeli 1) zainstalowanych w obu wymienionych systemach przedstawiono na ryc. 1.

W plikach MPEG-4 odnotowano również boks *stsd* zawierający dane o kodeku i parametrach zapisu dźwięku: dla formatów M4A, MP3 i AAC są one w wewnętrznym boksie *esds*, natomiast dla formatu 3GP w boksie *sawb* albo *samr* (co odpowiada AMR-WB dla aplikacji z pozycji 2 albo AMR-NB dla aplikacji z pozycji 3) oraz w boksie *damr*. Informacje zdekodowane z obiektu *stsd* pozwoliły na weryfikację kodeków i parametrów nagrania odczytanych za pomocą menu aplikacji i narzędzi wymienionych w rozdziale 3.2 niniejszej pracy.

W obiekcie *Movie Box* w boksach wewnętrznych *mvhd*, *trak\tkhd* oraz *mdia\mdhd* metadane zawierają łącznie 6 wpisów pozwalających zdekodować sygnatury czasowe utworzenia i modyfikacji prezentacji, ścieżki i mediów w zaokrągleniu do 1 sekundy dla nagrania zapisanego w kontenerze MPEG-4. Ustalono, że dla każdego z nagrań niepoddanych edycji zdekodowane w ten sposób sygnatury utworzenia są takie same jak sygnatury modyfikacji i związane są one z zapisem nagrania do pliku. Odpowiadają one również wyznaczonym uprzednio systemowym sygnaturom czasu modyfikacji

przy uwzględnieniu różnicy czasu UTC stosowanego w metadanych z czasem systemowym i możliwym zaokrągleniu do 1 minuty. Istotne dla badań autentyczności jest to, że sygnatury czasowe w metadanych w pliku MPEG-4 nie zmieniają się po wykonaniu kopii pliku na inny nośnik. Zależą one jednak w chwili zapisu nagrania od ustawienia zegara smartfonu. Metadane w boksach *mvhd*, *trak* i *mdia* zawierają również czas trwania dla nagrania dźwiękowego w zaokrągleniu do 1 ms. Dodatkowo wszystkie zarejestrowane pliki testowe w kontenerze MPEG-4 w boksie *hdlr* posiadają wpisy wskazujące na zawartość jedynie ścieżki dźwiękowej.

Testowane aplikacje, które zapisują swoje nagrania w formacie WAV, stosują kodek PCM bez kompresji sygnału. Nagrania tego formatu umieszczane są w kontenerze RIFF i zawierają bloki danych nazwane *chunks* zapisywane w konwencji *little-endian*. Analiza struktury testowych plików RIFF wykazała, że wszystkie one zawierają podstawowe metadane dla tego rodzaju plików w postaci kanonicznej: identyfikator *RIFF* informujący o typie kontenera, ciąg znaków *WAV* oraz chunki *fmt* z parametrami zapisu i *data* z zakodowanym nagraniem. Wśród aplikacji zapisujących nagrania w tym kontenerze wyróżnia się *Smart Recorder* (Smart Mob), który zamieszcza dodatkowy chunk *list* z wewnętrznymi obiektami *INFO* i *INAM*, które zawierają ciągi znaków z nazwą pliku. Co istotne, zapisana tym sposobem nazwa pliku w metadanych nie zmienia się po ewentualnej zmianie nazwy w systemie plikowym.

W strukturach plików w formacie MP3 (aplikacje z pozycji: 4, 7 i 8 w tabeli 1) zapisane są informacje o zastosowanej wersji kodeka LAME. W strukturze tego rodzaju plików wyłącznie aplikacja *Voice Recorder* (thinksimple app) umieszcza metadane z identyfikatorem *TAG*, które zawierają ciąg 4 znaków z rokiem rejestracji nagrania.

Jak ustalono, wśród zainstalowanych aplikacji jedynie *Voice Recorder* (Appliqato Software) zapisuje nagrania w formatach AMR i AAC w sposób „klasyczny". Analiza struktur plików testowych w tych formatach nie wykazała obecności żadnych indywidualnych metadanych, a jedynie standardowe i podstawowe informacje o parametrach zapisu niezbędne do odtworzenia nagrania.

Wpływ edycji nagrań na metadane

W ramach prac przeanalizowano również struktury plików z nagraniami po edycji z uwzględnieniem wszystkich możliwych opcji wskazanych w tabeli 1. Spośród aplikacji mających taką możliwość trzy z nich pozwalają na edycję plików M4A (pozycje: 1, 8 i 10 w tabeli 1) i po dwie aplikacje na edycję plików MP3 (pozycje 4 i 8) oraz plików WAV (pozycje 4 i 9).

Analiza porównawcza plików z nagraniami źródłowymi względem plików z nagraniami po edycji dla formatów WAV i MP3, które poddano badaniom w ramach niniejszej pracy, nie wykazała zmian w strukturze tego rodzaju plików oraz zawartości ich metadanych.

Istotne zmiany w plikach po edycji odnotowano natomiast dla nagrań w formacie M4A umieszczonych w kontenerze MPEG-4, co podsumowano w tabeli 7.

Jak można zauważyć, edycja nagrania aplikacjami *Samsung Voice Recorder* w obu wersjach oraz *EZ Voice Recorder* (Top 1) skutkuje reorganizacją struktury plików M4A: zawartość metadanych jest modyfikowana, usuwane są fragmenty lub wręcz całe boksy, a ich położenie albo rozmiar zostają zmienione. Opisane w tabeli 7 modyfikacje metadanych odnotowano zarówno w plikach zapisanych po edycji jako nowe, jak również w plikach po edycji nagrania zapisanych jako źródłowe.

W boksach wewnętrznych *mvhd*, *tkhd* i *mdhd* znajdują się metadane pozwalające zdekodować czas utworzenia i modyfikacji prezentacji, ścieżki oraz mediów zapisanych w ścieżce. Pola z sygnaturami czasowymi zawierają wartości oznaczające liczbę sekund od północy 1 stycznia 1904 roku w formacie czasu UTC do momentu zapisu albo modyfikacji nagrania. Wszystkie pliki testowe utrwalone w ramach niniejszej pracy zawierają ścieżki z zapisem dźwięku.

Przeanalizowano wpływ edycji nagrania na metadane w pliku M4A pozwalające zdekodować sygnatury czasowe. Z wykorzystaniem edytorów szesnastkowych odczytano zawartość pól oznaczonych jako *creation_time* i *modification_time* w boksach *mvhd*, *tkhd* i *mdhd* (ISO/IEC, 2015). Odczytano w ten sposób 6 wartości, które zdekodowano do czasu UTC, a następnie do czasu środkowoeuropejskiego.

Aplikacja *EZ Voice Recorder* (Top 1) po edycji nagrania umożliwia jego zapis wyłącznie do nowego pliku (tabela 2). Dla tej aplikacji analiza wymienionych 6 pól z boksów *mvhd*, *tkhd* i *mdhd* wykazała, że wszystkie zdekodowane sygnatury czasowe są takie same i odpowiadają one chwili zapisania nagrania po edycji do nowego pliku.

Z kolei dla aplikacji *Samsung Voice Recorder* w obu wersjach po edycji nagrania istnieje możliwość zapisu zmian w pliku źródłowym albo nowym, co ma wpływ na sygnatury czasowe w metadanych. Ustalono, że po wykonanej edycji nagrania i zapisaniu go:

– do nowego pliku: sygnatury czasu utworzenia i modyfikacji prezentacji oraz czasu modyfikacji ścieżki i mediów odpowiadają chwili wykonania edycji nagrania (przykładowo: usunięcia fragmentu), z kolei sygnatury utworzenia ścieżki i mediów odpowiadają czasowi zapisania nagrania źródłowego; ustalono też, że dalsze modyfikacje pliku powodują zmianę sygnatury utworzenia i modyfikacji prezentacji oraz

modyfikacji dla ścieżki i mediów, natomiast sygnatury utworzenia ścieżki i mediów pozostają bez zmian,
– w pliku źródłowym: sygnatury utworzenia i modyfikacji prezentacji oraz modyfikacji ścieżki i mediów odpowiadają chwili edycji nagrania, natomiast sygnatury utworzenia dla ścieżki i mediów związane są z zapisem nagrania do pliku przed jego edycją.

Na podstawie przedstawionych wyżej wyników można stwierdzić, że wykonanie edycji nagrania z wykorzystaniem zainstalowanych aplikacji powoduje zmiany w strukturze plików M4A oraz modyfikację istotnych metadanych. Zmiany te jednak pozwalają przyczynić się do oceny autentyczności takiego nagrania.

Przeprowadzone badania aplikacji i nagrań testowych wraz z otrzymanymi wynikami stanowią źródło informacji możliwych do wykorzystania w analizie autentyczności nagrań dowodowych. Zastosowanie dedykowanych programów i narzędzi systemowych oraz zaprojektowanego programu w środowisku MATLAB pozwoliło na zbadanie parametrów plików z nagraniami, ich sygnatur czasowych i metadanych. Zarejestrowane pliki dźwiękowe poddano globalnej analizie porównawczej i na tej podstawie ustalono, że każda z zainstalowanych aplikacji w obrębie swoich plików zamieszcza co najmniej jedną cechę dystynktywną. W tabeli 8 przedstawiono właściwości plików testowych ze wskazaniem tych cech, które dla danej aplikacji są wyróżniające względem pozostałych.

W każdym przypadku podstawowym materiałem do badań fonoskopijnych jest nagranie zapisane w pliku, a ujawnione cechy dystynktywne i ich porównanie z nagraniami testowymi mogą być pomocne w identyfikacji aplikacji, oprogramowania lub urządzenia zastosowanego do jego rejestracji.

5. Wnioski

Podsumowując wykonane badania, należy stwierdzić, że umożliwiły one poszerzenie wiedzy z zakresu najczęściej wykorzystywanych aplikacji przeznaczonych do rejestracji nagrań dźwiękowych instalowanych w systemie Android. Analiza wyselekcjonowanych aplikacji pozwoliła poznać ich możliwości oraz wykonać nimi nagrania testowe. Wyniki badań wykazały, że aplikacje te, choć bezpłatne, pozwalają na rejestrację dźwięku w relatywnie szerokim spektrum możliwości: w 6 różnych formatach zapisu, z predefiniowanymi albo dostosowywanymi parametrami, w wysokiej jakości plikach nieskompresowanych albo z kompresją dźwięku oraz z możliwością zamieszczania dodatkowych informacji w ich nazwach. Połowa analizowanych aplikacji umożliwia edycję polegającą na usunięciu fragmentu nagrania, jego skopiowaniu i wklejeniu w inne miejsce lub nadpisaniu nowym zapisem. Z wykorzystaniem narzędzi systemowych i programów dedykowanych oraz algorytmu opracowanego w środowisku MATLAB przeanalizowano właściwości zgromadzonych nagrań i struktury plików. Na tej podstawie stwierdzono, że możliwe było ustalenie parametrów zapisu nagrań testowych, ich sygnatur czasowych i zawartości metadanych. Wyniki przeprowadzonych badań jednoznacznie wskazują na możliwość ich wykorzystania w analizie autentyczności dowodowych nagrań dźwiękowych. Po pierwsze wykonane prace pozwoliły na zgromadzenie i zbadanie nagrań dźwiękowych utrwalonych różnymi aplikacjami mobilnymi, co umożliwi wykonanie analiz porównawczych ich właściwości z nagraniami dowodowymi przekazanymi w przyszłości do badań. Pozwoli to również na uzupełnienie ciągle rozwijanej przez autora bazy danych nagrań dźwiękowych, która stanowi dodatkowe narzędzie wspomagające badania autentyczności. Po drugie analiza metadanych plików testowych zapisanych w popularnym kontenerze MPEG-4 pozwala uzyskać wiele istotnych informacji o parametrach zapisu, sygnaturach czasowych, wykorzystanym urządzeniu czy aplikacji nagrywającej. Umożliwia również wskazanie różnic pomiędzy nagraniami źródłowymi a poddanymi edycji. Dodatkowo każda z zainstalowanych aplikacji odznacza się co najmniej jedną cechą odróżniającą jej pliki od pozostałych, co może być pomocne w identyfikacji konkretnej aplikacji wykorzystanej do zapisu badanego pliku dźwiękowego. Otrzymane wyniki mają istotne znaczenie, gdyż kwestie dotyczące autentyczności nagrań są często zawarte w postanowieniach organów zlecających opinie fonoskopijne.