IOS “Voice Memo” Audio Files Originality Authentication

Yongxue Wei*, Yi Ding, Siyuan Zhang, Guihao Pan, and Tianqi Jiang

Abstract—**In this paper, we modify the audio files recorded by IOS “Voice Memo” by cutting and erasing them in the software, and use various software and websites to authenticate the audio files originality by taking the method of graphical analysis and metadata feature analysis of the files. By comparison, we find out the differences of the original and modified audio files and the general method of judging the originality of audio files. The results of the experiments show that the audio files generated by the IOS system's own voice memo are in M4A format by default and that the audio files generated in the IOS system environment have the important feature of system version information encoding.**

Index Terms—**Audio files, audio modification, IOS voice memo, originality authentication**

I. Introduction

In 1982, China first stipulated audio-visual materials as an independent type of evidence in the Civil Procedure Law. Audiovisual material is an independent type of evidence stipulated in the Civil Procedure Law of China. It refers to evidence that proves the facts of a case through sound, language, image or behavior recorded in audio and video recordings. With the development of modern science and technology and the popularization of electronic products, people have more and more scientific and technological products available for recording and video recording. It is easier to collect evidence by recording and video recording. The parties often submit audio-visual materials to prove the facts of the case in civil litigation, and audio-visual materials have gradually become one of the most common evidences in civil litigation .Since the 1990s, the Civil Procedure Law, promulgated and implemented in 1991, continued to stipulate audio-visual materials as independent types of evidence, and made clear provisions on how the court should examine and adopt such evidence. Since then, the practice of evidence collection has been improved through the Rules of Evidence in Civil Procedure and the Civil Judicial Interpretation. At present, under the norms of the Criminal Procedure Law of the People's Republic of China, the Civil Procedure Law of the People's Republic of China and the Administrative Procedure Law of the People's Republic of China, the use of audio-visual materials in judicial courts has become increasingly prominent. Among them, voice evidence, as one of the main contents of audio-visual evidence, is an important form of evidence. With the arrival of the information and digital era, the development of communication technology and network technology, its application is becoming more and more extensive [1]. Despite the endless emergence of recording equipment, people generally prefer mobile phone recording. The mobile phone is very small, very convenient to carry, really let us do anytime and anywhere can be taken out as a recording equipment and work. Apple mobile phone, as a mature brand, is favored by consumers, and it has a deep consumer market .The audience is wide and the application is high， and the IOS system it uses is also a relatively perfect operating system developed for many years. So the research results of the characteristics of the systematic audio file generation method can be used to the maximum. IOS is a mobile operating system developed by Apple. Apple first announced this system at the Macworld Conference on January 9, 2007. It was originally designed for iPhone use, and later applied to iPod touch and iPad. IOS, like Apple's MacOS operating system, belongs to a Unix like commercial operating system. The original name of this system was iPhone OS. Because the iPad, iPhone, and iPod touch all use iPhone OS, the 2010 Apple Global Developer Conference announced that it was renamed iOS (iOS is a registered trademark of Cisco's network device operating system, and Apple has been authorized by Cisco). After years of development and constant research, the IOS system has made great progress in the use of users. Not only that, the inspection and research on the audio file generation mode of the IOS system also involves the research on the recording authenticity verification technology in general sense, and is also closely related to the research on the characteristics of the iPhone and its IOS operating system [2-4]. Voice memos on the iPhone or iPad can be cut, shared, saved, and modified by a file name. When it is necessary to submit the recording as evidence, it must ensure the originality and authenticity of the audio documents, and the relevant evidence should be provided to the court as far as possible to form the evidence and enhance the proof power of the recording evidence. Otherwise, the recording evidence as an isolated evidence is difficult to achieve the purpose of proof. In practical cases, we often face the cases in which questioned recordings are recorded by using the iPhone. The recordings on the iPhone are often presented in court served as supportive ev1dence. The expert witnesses have to judge on the authenticity and integrity of the submitted questioned recordings ,and have to answer whether or not the questioned recordings are primitive or tampered. Therefore, there search on the techniques of authentication of the iPhone 's recordings has important value. The research on authentication of the iPhone' s recordings involves the general techniques of audio authentication examination, at the same time, it is closely related to the characteristics of the iPhone and its operating system. Currently, the techniques for authentication examination of iPhone recordings are still unnoticed, which are the aim of this paper. This project analyses the originality and authenticity of voice memo files by finding a test procedure to verify that they are original recordings and have not been manipulated. When the recording needs to be submitted as evidence, the parties cannot clearly remember the specific source of the recording, making the proving power of the recording evidence unrecognized, and at the

Manuscript received November 11, 2022; revised December 10, 2022; accepted December 28, 2022.

The authors are with Department of Forensic Science and Technology, Jiangsu Police Institute, Nanjing, 210031 China.

*Correspondence: 1942287645@qq.com

DOI: 10.7763/IJET.2023.V15.1214

same time, their mobile phone system and application have been upgraded and iterated many times, further complicating the detection. Today ZHOU Juan, YI Shuang and others have conducted a comparative study of the metadata of audio files, the changes of information in the head and tail of the files during various operations, to generate the current technical means of authenticating audio recordings are the existing technical methods for authenticating audio recordings including recording metadata inspection, speech semantic analysis, audio signal analysis and electronic data inspection techniques. Electronic data, file MD5 value, etc. will be changed, but can not be used to determine the authenticity of the file, there is controversy, and need more theory to support the results of the discriminatory, most of today's research based on the differences in the file metadata and the differences in the electronic data at the end to, UUID number written and different to start research, but due to different equipment models, system versions, etc., for the data itself also has a certain .The results of today's research have been broadly understood in terms of which factors have an impact on the data and which factors do not have an impact, simplifying the identification process for the future, i.e. the metadata and electronic data of the audio files can be changed in a way that gives a reasonable basis for judging the way the audio files are generated and their flow process, according to which the results can provide more convenient support for future authenticity checks.

UUID (universal unique identifier) is used to identify data records in distributed systems. UUID is short for the universal Unique Identification Code (Universally Unique Identifier), a standard for software construction, and a part of the Open Software Foundation organization in the field of distributed computing environment. The purpose is to make all the elements in a distributed system have unique identification information, without having to specify the identification information through the central control end. Everyone can then create a UUID that does not conflict with others. The UUID defined in RFC 4122 is designed as a 128 bits binary number, expressed in hexadecimal, divided into 5 segments. Segments are connected with “-” and recorded as “*** - X * * - Y * * - *****************”. Where X is the version field, and UUID has five versions at present, so X values are 1, 2, 3, 4, and 5, respectively corresponding to the corresponding version. Y is a variable field, currently defined as 10 * *, so the values are 8, 9, a, b UUID-1 and UUID-2 are generated by random number, time and physical address of the host; UUID-3 and UUID-5 are generated by the hash value of the namespace; UUID-4 is generated from random numbers. The generating elements of UUID-1 and UUID-2 contain the physical address of the host, which is easy to expose customer information UUID-3 and UUID-5 generate the same UUID on the same user host, so they are not applicable to the database cloud platform. At the same time, the UUID method is based on random numbers, and the UUID generated is an unstructured and irregular symbol. In the database cloud platform, a large number of tasks need to be indexed. For example, the task ID generated by the UUID algorithm is an unordered number, resulting in inefficient indexing and task management [5].

MD5 is Message-Digest Algorithm 5 (Information-Summary algorithm 5), which is used to ensure complete and consistent consistency of information transmission. It is one of the widely used computer miscellaneous algorithms (and translated summary algorithm, hash algorithm), and the mainstream programming language has been generally implemented in MD5. It can generate a 128 bit (16 byte) hash value to ensure complete and consistent information transmission. The role of MD5 is to allow large-volume information to be “compressed” into a secret format (transforming a byte string of any length into a long 16 decimal digital string). The MD5 algorithm was designed by Ronald Livist and published in 1992 to replace the MD4 algorithm. The program of this algorithm is standardized in RFC 1321. The MD5 algorithm converts a string of any length (i.e. plaintext) into a 128 bits binary number (i.e. ciphertext). This conversion process is one-way and irreversible, which means that MD5 algorithm can only convert plaintext to ciphertext, and ciphertext cannot be decrypted to plaintext. At the same time, the algorithm is also unique, that is, the ciphertext encrypted by different plaintext is always different, and the ciphertext encrypted by the same plaintext is always unchanged [6].

II. Processing of audio Files

With the development of computer technology, especially the realization of large capacity storage devices and large capacity memories on PC, the digital processing of audio media becomes possible. The core of digital processing is the sampling of audio information. By processing the collected samples, various effects can be achieved, which is the basic meaning of audio media digital processing. From the perspective of audio file processing, we can generally analyze and process from the following perspectives: 1. Sampling 2. Sampling frequency 3. Bit rate 4. Number of channels. The recording function in the voice memo built into iOS 14 can pause and resume the recording before saving the recording. That is to say, since users can operate the recording during the recording process, it is difficult to determine whether the recording has been operated using the audio signal feature analysis method mentioned above. In addition, even though the time of manipulation can be determined through the analysis of metadata, including timestamps and file structures, the location of manipulation in the recording cannot be revealed using traditional metadata based inspection methods. In particular, using the characteristics of audio delay time can help detect whether the beginning of audio recording has been tampered with. However, it is impossible to detect the manipulated position with the existing methods for any changes other than modifying the beginning of the audio. That is to say, the traditional methods mentioned above can only determine whether the questioned audio record has been tampered, but can not determine the position of any tampering in the audio record.

Audio identification refers to professional judgment on the originality, continuity and integrity of questionable recordings by means of auditory test, acoustic spectrum analysis, metadata analysis and digital data analysis. The original test of recording is mainly to judge whether the recording is original through metadata analysis and digital forensics technology. In this experiment, we used a simple and universal audio processing method, namely, cut and splice, to verify the originality. Common audio processing software on the market include audio processing masters, Audacity, etc. Our research focuses on the voice memo of the IOS system. The authenticity of the recording means that the file can be copied, moved and the filename changed, but the content of the file cannot be changed and the MD5 code of the file must be the same as the original recording. So far, the files in IOS Voice Memo only support the cut function, and there are only a few ways to modify audio files. In our experiments, we have recorded an audio file on Voice Memo, cut and modified it so that its meaning is completely changed, but it is difficult to distinguish the authenticity of the audio file by the human ear.

III. Methods for Checking the Originality of Audio Files

When recording an audio, users can name the audio recording, but in the application, the memos are named by using their creation time with the suffix of ".m4a", that is, the file name indicates the creation time of the audio recording.

The meta-data information of the audio files contain the time related information, such as recording time, duration, encoding time, tagged time, etc. The recording time is the starting time of audio recordings displayed in the time zone set in the 'iPhones' iOS. The encoding and tagged times are the last save time of audio recordings, which are displayed in the UTC format. The recording time is included in the audio meta-data in the iOS 8, but not in the iOS 10 or later iOS versions.

The website www.metadata2go.com is a free online EXIF viewer which allows you to view the metadata information of the three audio files and compare them on the differences.

Adobe Auditions is a professional audio editing and mixing environment developed by Adobe, providing advanced audio mixing, editing, control, and effects processing features for recording and mixing audio video, podcasts, and sound effects design. In this article, we use the Adobe audition cc version 2020 to check the audio properties by importing three audio file clips and comparing the XMP information in the metadata field.

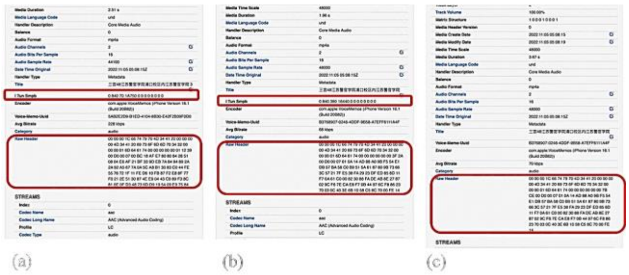


Fig. 1. Comparison of the metadata of the modified and original audio files. (a) Modified file with content "You are wrong"; (b) Modified file with content "You are right"; (c) Original file with content "Are you right, or are you wrong?"

IV. Experiments and Analysis

We first use the website www.metadata2go.com to view metadata of the audio files, which shown as Fig. 1. As can be seen from the figure, the data describing the essential properties of the audio file such as I Tun Smpb and Raw Header before and after editing are completely different because the material is two different pieces of audio, even if the original file is not shown after detection I Tun Smpb is not shown, other items such as encoder and other criteria that can show the characteristics of the audio file itself are the same.

Then we detect the relevant version information of apple iTunes with WinHex, shown as Fig. 2. In the default state, the audio files are named after the recording location and saved with the default .m4a suffix, but the file name and format remain the same after editing. All the files were opened in Winhex and the search for "apple" in both the ANSI and ASCII encoding fields revealed the version information for apple iTunes.

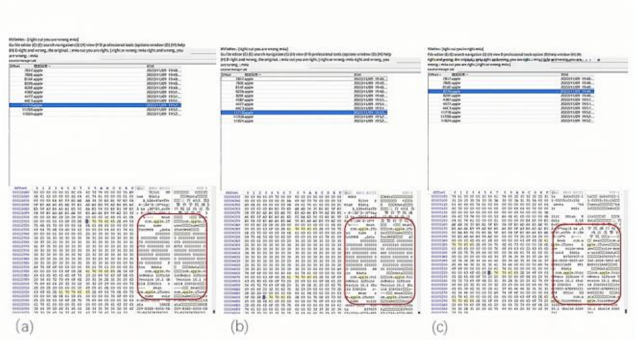


Fig. 2 Comparison of data after Winhex manipulation of the modified and original audio files. (a) Modified file with content "You are wrong"; (b) Modified file with content "You are right"; (c) Original file with content "Are you right, or are you wrong?"

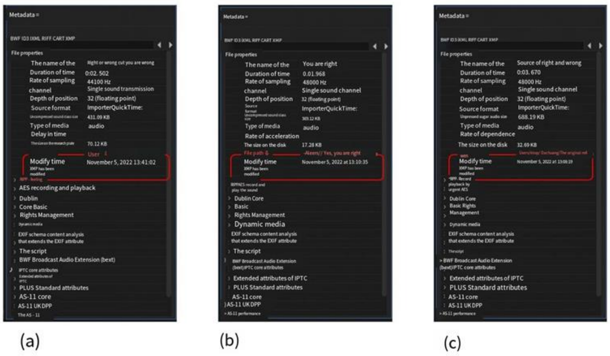
iTunes backup in can be obtained through a search warrant, so that the backup and lockdown files can be obtained. The IOS backup file forensics mainly involves the analysis of offline backups generated by iPhones, iPads, iPod touch or Apple Watch. The data of Apple Watch will be included in the synchronized iPhone backups.

When the physical extraction and file system extraction of IOS devices are not feasible, and the logical extraction cannot meet the forensic requirements, iTunes backup extraction can also play a role. In this case, the investigation and forensics personnel only need to make a backup of the equipment and use the forensics software to analyze it. Therefore, investigators and forensics personnel must fully understand the backup process and the tools involved, so as to ensure that they can make forensics backups and prevent other data from polluting the data in iTunes.

iTunes, an app launched by Apple Inc. to connect IOS devices and computer hosts, is a free digital media playback app for Mac and PC that can manage and play digital music and video. Its "backup" function can capture most of the data from IOS devices, such as call records, online records, instant messaging, mobile wallet, payment application information, etc., so forensics people often use it to extract data from IOS devices. I Tunes provides an encrypted backup option, but by default, it creates an unencrypted backup when you synchronize your I Phone. After decrypting the encrypted backup, you can have access to the additional data stored on the IOS device [7].

Winhex is a hexadecimal editing software running under Windows. It has sound file management and partition management functions. It can automatically analyze the file cluster chain and partition chain, backup the hard disk in different degrees and ways, and even clone the entire hard disk; The binary content of any file type (displayed in hexadecimal system) can be edited. The disk editor of this software can edit any sector of logical disk or physical disk. It is an excellent software for manual data recovery [8]. Winhex integrates powerful tools, including disk editor, Hex converter and RAM editing tools, and can conveniently call common tools of the system, such as calculator, notepad, browser, etc. In the unregistered version, you can edit, but you cannot save files larger than 512K, and you can only browse, but not modify, the RAM area. Winhex is simple to use and powerful. It can facilitate your program debugging, text editing, scientific computing and system management. I believe you will like it. If you want to delete Winhex software, simply delete the entire directory. As a hexadecimal file editing and disk editing software. WinHex is famous for its small file size, fast speed and powerful functions. Even ZDNetSoftwareLibrary gave him the highest rating of 5 stars. It is capable of editing and modifying Hex and ASCII codes, searching and replacing multiple files, general and logical operations, automatic search and editing of disk regions (supporting FAT16, FAT32 and NTFS), file comparison and analysis, and editing of data in memory.

Finally we viewed the file information with Adobe Audition, shown as Fig. 3. All files are opened in Audition with the metadata information shown in the image. Within the XMP options there are two lines of information at the bottom, one for the modification information and the other for the XMP modification information. The modification time for the original file is when the file was created, and the modification time for the modified file is when the editing operation was completed [9–10]. Adobe applications use the Extensible Metadata Platform (XMP) to store metadata. XMP is built on XML, which facilitates the exchange of metadata across various applications and publishing workflows. Most other formats of metadata (for example, Exif, GPS, and TIFF) are automatically transferred to XMP, making it easier for you to view and manage them.



(a)

Metadata ≡

BWF ID3 iXML RIFF CART XMP

File properties

The name of the	Right or wrong cut you are wrong
Duration of time	0:02. 502
Rate of sampling	44100 Hz
channel	Single sound transmission
Depth of position	32 (floating point)
Source format	ImporterQuickTime:
Uncompressed sound class size	431.09 KB
Type of media	audio
Delay in time	
[illegible]	70.12 KB
	User
Modify time	November 5, 2022 13:41:02
XMP has been modified	

AES recording and playback
Dublin
Core Basic
Rights Management
Dynamic media
EXIF schema content analysis that extends the EXIF attribute
The script
BWF Broadcast Audio Extension (bext)
IPTC core attributes
Extended attributes of IPTC
PLUS Standard attributes
AS-11 core
AS-11 UK DPP
The AS - 11

(b)

Metadata ≡

BWF ID3 iXML RIFF CART XMP

File properties

The name of the	You are right
Duration of time	0.01.968
Rate of sampling	48000 Hz
channel	Single sound channel
Depth of position	32 (floating point)
Source format	ImporterQuickTime:
Uncompressed sound class size	303.12 KB
Type of media	audio
Rate of acceleration	
The size on the disk	17.28 KB
File path	-Aleers// Yes, you are right
Modify time	November 5, 2022 at 13:10:35
XMP has been modified	

BPPAES record and play the sound
Dublin Core
Basic
Rights Management
Dynamic media
EXIF schema content analysis that extends the EXIF attribute
The script
BWF Broadcast Audio Extension (bext)IPTC core attributes
Extended attributes of IPTC
PLUS Standard attributes
AS-11 core
AS-11 UK DPP
AS-11 performance

(c)

Metadata ≡

BWF ID3 iXML RIFF CART XMP

File properties

The name of the	Source of right and wrong
Duration of time	0:03. 670
Rate of sampling	48000 Hz
channel	Single sound channel
Depth of position	32 (floating point)
Source format	ImporterQuickTime:
Unpressed sugar audio size	688.19 KB
Type of media	audio
Rate of dependence	
The size on the disk	32.69 KB
	Users/bopp/Dachuang/The original ref
Modify time	November 5, 2022 at 13:08:19
XMP has been modified	

XMP. Record playback by urgent AES
Dublin Core
Basic Rights Management
Dynamic media
EXIF schema content analysis that extends the EXIF attribute
The script
BWF Broadcast Audio Extension (bext)IPTC core attributes
Extended attributes of IPTC
PLUS Standard attributes
AS-11 core
AS-11 UK DPP
AS-11 performance

Fig. 3. XMP information in Adobe Audition of the modified and original audio files. (a) Modified file with content "You are wrong"; (b) Modified file with content "You are right"; (c) Original file with content "Are you right, or are you wrong?"

In most cases, XMP metadata is stored directly in the source file. However, if a file format does not support XMP, metadata is stored in a separate sidecar file. Project resources without corresponding files do not support XMP.

V. Conclusion

Through the above experimental operations and data collection, sorting and analysis, we can get the following experimental results: the audio file format generated by the voice memo attached to the IOS system is M4A by default. If it is edited and modified accordingly, the metadata information edited will be completely different, and the default name is the location of the audio recording. In addition, the audio file generated by "Voice Memo" shows the modification time of XMP information in the audition. When it is edited in IOS, and then opened in audition, we can find a fixed fact that the modification time of XMP information is displayed as the time when the audio file was first edited and processed. In addition to the above findings, we also found that another special feature of the IOS audio file generation method is that although the original audio file can only be detected through the system's own voice memo recording, it can be detected as long as the IOS system information is modified. Therefore, the audio file generated in the IOS environment has an important feature of the version information of the encoding system. Using this feature, we can accurately and quickly judge whether the audio file in the IOS voice memo has been artificially modified, which will provide a strong entry point for us to verify the originality of the audio file. With the development of science and technology, there will be more scientific means to make more covert modifications to audio-visual documents. In order to maintain the legitimacy and accuracy of evidence, we should make more effective research on the original verification of audio documents.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Yongxue Wei conducted the research and analyzed the data; Yongxue Wei, Tianqi Jiang, Siyuan Zhang, Guihao Pan and Yi Ding wrote the paper; all authors had approved the final version.

Funding

This work paper was supported by Practical Innovation and Entrepreneurship Training Program for Students of Jiangsu Police Institute (Grant No.XJ202210329001).

Acknowledgment

All the authors thank the guidance and helpful discussions from Dr. Dong Ye and the support from Jiangsu Police Institute.

References

[1] H. Hollien, *The Acoustics of Crime: The New Science of Forensic Phonetics*, Springer Science & Business Media, 2013.

[2] J. Zhou and S. Yi, "Exploration into essentials about checking procreation ways of audio files saved with 'Voice Memo' APP under IOS," *Forensic Science and Technology*, vol. 46, pp. 642-646, Dec. 2021.

[3] J. Zhou, J. Wang, S. Yi, and H. Yuan, "Source-tracing into iPhone-stored audio files based on WeChat logs," *Forensic Science and Technology*, vol. 47, pp. 211-215, Apr. 2022.

[4] J. Zeng, J. Xi, W. Sun, and X. Qiu, “Authenticity authentication technology of Apple mobile phone recording,” *China Judicial Expertise*, vol. 5, pp. 93-97, Sep. 2020.

[5] W. Liu, “A new algorithm for UUID method on database cloud platform,” *Journal of Minnan Normal University (Natural Science)*, vol.35, pp. 44-47, June 2022.

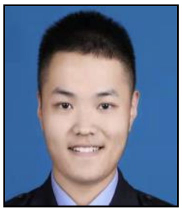
[6] H. Yang, J. Song, and C. Wang, “A brief introduction to MD5 encryption algorithm in network security,” *Network Security Technology & Application*, vol. 9, pp. 40, Sep. 2018.

[7] J. Liu, P. Qiu, and S. Su, “A comparative study on evaluation of three forensics methods of Apple Mobile IOS system,” *Journal of People's Public Security University of China (Science and Technology)*, vol. 4, pp. 62-67, Dec. 2020.

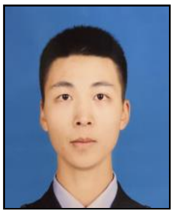
[8] X. Wu, “Winhex-based data recovery,” *Computer Knowledge and Technology*, vol. 16, pp. 42-43, Oct. 2020.

[9] N. I. Park, J. W. Lee, K. -S. Shim, J. S. Byun, and O. -Y. Jeon, “A method of forensic authentication of audio recordings generated using the Voice Memos application in the iPhone,” *Forensic Sciences International*, vol. 320, pp. 110702, Mar. 2021.

[10] N. I. Park, K.-S. Shim, J. W. Lee, J. -H. Kim, S. H. Lim, J. S. Byun, Y. J. Kim, and O.-Y. Jeon, “Advanced forensic procedure for the authentication of audio recordings generated by Voice Memos application of iOS14,” *Journal of Forensic Sciences*, vol. 67, pp. 1534-1549, July 2022.



Yongxue Wei was born in Nanjing, Jiangsu Province, China in 2003. He is currently studying at Jiangsu Police Institute.



Yi Ding was born in Suzhou, Jiangsu province in 2003. He is currently studying in the 2021 public security audio-visual technology major of Jiangsu Police Institute.

Mr. Ding has won a number of inside and outside the school awards, such as outstanding Student Advanced Individual, Advanced Individual for Quality Development, and Advanced Individual for learning.

Siyuan Zhang was born in Nantong, Jiangsu Province in 2003. He is currently studying at Jiangsu Police Institute. majoring in public security audiovisual technology in 2021. He is active, diligent in his studies, and friendly, and was named an active member of the party in the first semester of his sophomore year.

Mr. Zhang has won many on-campus and off-campus awards, such as the Outstanding Student Scholarship, Advanced Individual in Quality Development, and Advanced Individual in Policing.

Guihao Pan was born in Yangzhou, Jiangsu Province, China in 2002. He is currently studying at Jiangsu Police Institute.