

# 디지털 오디오 파일의 편집 여부 분석 절차에 대한 연구

박남인\*, 심규선\*, 전옥엽\*\*  
국립과학수사연구원 디지털분석과 (공업연구사)\* (공업연구관)\*\*

## A Study on Authentication Analysis Procedure of Digital Audio Files

Nam In Park\*, Kyu-Sun Shim\* and Oc-Yeub Jeon\*\*  
Digital Technology and Biometry Division, National Forensic Service (NFS)  
(Researcher)\* (Senior Researcher)\*\*

### 요 약

최근 스마트폰을 비롯한 스마트 기기의 대중화 등의 영향으로 사진이나 동영상 촬영 오디오 녹음이 빈번하고 쉽게 이루어짐에 따라, 이러한 형태의 디지털 파일이 증거물로써 활용되는 사례도 증가하고 있다. 이러한 디지털 파일이 디지털 증거물로써 법적 효력을 갖추기 위해서, 디지털 증거물에 대한 원본 확인이 필수적이다. 특히, 디지털 오디오 파일의 특성상 원본파일로부터 대량 복제 및 디지털 편집 프로그램 등을 통해 정교한 위변조가 가능하며, 실제로 디지털 오디오 파일을 조작하는 사례가 증가하고 있다. 본 논문에서는 다양한 실제 사례들을 분석한 후, 디지털 오디오 파일에 대한 오디오 신호 및 파일 포맷 분석을 통한 편집 여부 분석 절차에 대해 제안하였다. 실제 사례들을 분석한 결과, 편집이 이루어진 파일에 대해서 디지털 오디오 파일의 조작 과정에서 발생하는 조작 흔적을 통해 원본이 아님을 제한적으로 증명할 수 있음을 확인하였고, 일부 특정 조건에서는 제안한 편집 여부 분석 절차로도 위변조 여부가 검출되지 않을 수 있다는 것이 확인하였다.

주제어 : 디지털 녹음, 오디오 포렌식, 디지털 인증, 오디오 편집 여부

### ABSTRACT

Recently, due to the popularization of smart devices such as smartphones and the like, photography and video recording audio recordings are frequently and efficiently performed. Thus the use of digital files of those types is increasing. For these digital files to be legally valid as digital evidence, the proof of origin for digital evidence was essential. Mainly, due to the characteristics of the digital audio file, it is possible to make a sophisticated forgery and falsification through the mass copying and digital editing program from the original record. In this way, it has been increasing more and more cases of manipulating digital audio files. In this paper, we propose the procedure of analyzing the file format of digital audio files and the audio signal after the analysis of various actual cases. As a result, it has confirmed that the digital audio files can be not original through traces generated during the manipulation of the digital audio file in case of the manipulated data. Moreover, it has confirmed that the proposed authentication analysis procedure may not clarify whether or not the digital audio file has been a forgery in some specific condition.

**Key Words** : Digital Recording, Audio Forensic, Digital Authentication, Audio Forgery Detection

## 1. 서 론

스마트폰, 태블릿PC, 및 스마트워치 등과 같은 휴대용 멀티미디어 장비의 대중화 경향은 시간에 따라 완만한 증가세를 보이는 것이 아닌 폭발적인 증가의 특징을 보이고 있으며, 이와 더불어 디지털 데이터 또한 급증하고 있는 추세이다 [1]. 또한, 범죄와 관련하여 사건에 대한 증거 확보를 위해 주변에서 쉽게 구할 수 있는 휴대용 멀티미디어 장비로 촬영하거나 녹음하여 수사기관에 제출하여 디지털 증거물로 활용하는 경우가 증가하고 있다. 이러한 디지털 증거물이 법정에서 증거물로써 효력을 가지기 위해서는 디지털 데이터에 대한 무결성 및 원본성이 보장되어야 한다. 최근 들어 언론에 보도되고 있는

※ 이 논문은 행정안전부 주관 국립과학수사연구원 중장기과학수사감정기법연구개발(R&D)사업의 지원을 받아 수행한 연구임 (NFS2019D TB01).

• Received 07 October 2019, Revised 19 October 2019, Accepted 17 December 2019

• 제1저자(First Author) : Nam In Park (Email : namin.park@gmail.com)

• 교신저자(Corresponding Author) : Oc-Yeub Jeon (Email : yeubjeon@korea.kr)

다수의 사건에서도 디지털 증거물에 대한 무결성이 화두가 되면서, 디지털 증거물의 위변조 분석 방법에 대해 관심이 높아지고 있다. 디지털 포렌식 기술이 발달하면서 디지털 증거물 확보의 중요성이 증가하였고, 자신에게 불리한 증거를 삽입, 삭제 및 복사 등의 훼손하는 안티 포렌식(anti-forensics)의 기법의 확산으로 인해 특정 영상 혹은 음성을 제어하는 방식과 같은 위변조가 가능하게 되었다 [2]. 이와 같이 디지털 파일에 대한 위변조가 가능하고 최근 딥러닝 기술을 활용한 가짜 뉴스와 같은 실제 위변조 사례가 발생하면서, 디지털 파일의 위변조 검출에 대한 연구가 많이 진행되고 있다 [3].

수사기관이 디지털 오디오 파일에 대해 증거물로 활용하기 위해, 법과학 감정기관 등을 통하여 해당 파일의 위변조 여부에 대한 확인을 요청하게 되면, 법과학 감정기관 등에서는 의뢰사항에 대한 분석 결과를 회신하게 되며, 이후 경우에 따라 재판과정에서 분석 방법 및 결과에 대해 법정 증언을 하게 된다. 최종적으로 디지털 오디오 파일의 위변조 여부에 대한 결과를 종합하여 증거 채택에 대한 결정은 법원에서 이루어진다 [4]. 여기서, 법에 명시된 디지털 오디오 파일의 원본성에 대한 정의가 무엇인지 알아볼 필요가 있으며, 이에 대한 디지털 오디오 파일에 대한 편집 여부에 대한 절차를 수립하는 것이 중요하다. 디지털 오디오 파일의 원본성의 정의에 대해 판례를 통해 확인해보면, 다음과 같이 기술되어 있다.

“피해자의 진술에 의하여 녹음테이프에 녹음된 피고인의 진술내용이 피고인이 진술한 대로 녹음된 것이 증명되고 나아가 그 진술이 특히 신빙할 수 있는 상태하에서 행하여진 것이 인정되어야 할 것이고, 녹음테이프는 그 성질상 작성자나 진술자의 서명 혹은 날인이 없을 뿐만 아니라, 녹음자의 의도나 특정한 기술에 의하여 그 내용이 편집, 조작될 위험성이 있음을 고려하여, 그 대화내용을 녹음한 원본이거나 혹은 원본으로부터 복사한 사본일 경우에는 복사과정에서 편집되는 등의 인위적 개작 없이 원본의 내용 그대로 복사된 사본임이 입증되어야만 하고, 그러한 입증이 없는 경우에는 쉽게 그 증거능력을 인정할 수 없다.” [5]

\*출처: [대법원 2005. 12. 23. 선고, 2005도2945, 판결] 국가법령정보센터 판례

대법원 2005도2945[5]에서 확인되는 바와 같이 디지털 오디오 파일의 원본이라 함은 “녹음한 원본 혹은 원본으로부터 복사한 사본일 경우에 원본의 내용 그대로 복사된 사본이어야 한다.

[대법원, 2008. 3. 13., 2007도10804][6]에서는 보이스펜으로 일컬어지는 녹음기와 저장매체가 일체형인 기록장치와 해당 장치로 최초 녹음된 정보 복사한 카세트 테이프, 카세트 테이프에 녹음된 내용을 음질개선했어 저장한 재복제본의 기록물이 저장된 CD가 다루어졌으며, 최초의 기록장치로부터 복제된 사본, 재복제본의 생성과정이 증명되어 증거능력을 인정받은 사례가 확인되고 있다. 또한, [서울고등법원, 2014. 8. 11., 2014노762][7]에서는 디지털 오디오 파일에 대하여, 사본에 대한 해시값의 연속성에 대하여 확인할 수 없다고 하더라도, “녹음파일의 수집, 보관·복제 등의 절차에 관여한 사람의 증언이나 진술 또는 녹음파일에 대한 검증 및 감정 결과 등 제반 사정을 종합하여 그 원본성이나 원본과의 동일성 유무를 판단할 수도 있다. 그러므로 반드시 원본이 보존되어 있어야 한다거나, 디지털 포렌식 절차를 준수하였어야 한다거나, 봉인조치에 의해 보관의 연속성이 증명되어야 한다고 볼 것은 아니다”라고 판시하고 있다. 디지털 오디오 파일이 위변조 되지 않았음에 대한 전제는 디지털 파일로서 해시값의 변화가 없는 것만을 전제[5]하는 것뿐만이 아니라, 감정물의 최종상태에 이르는 과정이 증명될 수 있는 점 또한 중요하게 다루어지고 있다.

디지털 오디오 파일의 무결성을 확인하기 위한 방법으로 디지털 데이터에 저작권 정보와 같은 비밀 정보를 삽입하는 디지털 워터마킹 기술들이 연구되어 있으나, 이러한 워터마킹 기술은 음악 저작권 관련 사업, 인터넷 이미지 판매 사업, 문서 보안 관련 사업 등에 한정하여 적용되고 있다 [8]. 최근 들어 머신러닝 기술의 발달로 디지털 데이터의 위변조 구간을 자동적으로 검출하는 연구가 활발히 이루어지고 있다 [9]. 특히 양일호 등은[4] GMM(Gaussian Mixture Model) 기반의 확률 모델과 음성 프레임 간의 거리를 통해 음성이 외부에서 삽입된 것인지 여부를 자동으로 검출하는 연구를 진행하였다. 또, 박세진 등은[2] 음성 데이터의 저대역 내에 존재하는 ENF(Electric Network Frequency)를 검출하여 해당 신호의 불연속성을 검출하는 방법을 제안하였다. 그러나, 확률 모델 기반 및 ENF 기반 프로세스는 디지털 오디오 파일의 편집여부 검출을 위해 실제 범죄수사에 적용하기 어렵다. 왜냐하면, 확률 모델 기반의 방법은 음성의 삽입으로 인한 배경잡음의 특성만을 고려하기 때문에 음성 구간이 삭제된 경우에 대해서는 적용하기 어렵고, ENF 기반의 방법은 최근 스마트폰에서의 통화 녹음에서는 전처리를 통해 저대역 신호를 제거하기 때문에, 실제 녹음된 파일에서 ENF 신호를 검출한다는 것이 거의 불가능하다. 따라서, 디지털 오디오 파일의 편집 여부에 대해 실제 범죄수사에 적용 가능한 구체적인 프로세스가 요구되며, 특히, 대조파일이 제시되었을 때, 디지털 녹음파일과의 비교 분석을 통해, 녹음 가능한지 여부에 대해 판단할 수 있는 절차가 필요하다.

본 논문에서는 실제 범죄 수사에 적용을 위한 디지털 오디오 파일의 편집 여부 검출 방법에 대해 오디오 파일 포맷과 오디오 신호 관점에서 기술한다. 실제 다양한 디지털 증거물에서의 편집 사례를 기반으로 오디오 편집 여부 기법에 대해 설명하고 이를 기반으로 편집 여부 분석 절차에 대해 제안한다.

논문은 다음과 같은 형식으로 구성되어 있다. 2장에서는 디지털 데이터 인증 시스템 및 파일 시스템 분석 방법과 그 한계에 대해서 소개하고, 3장과 4장에서 실제 사례 중심의 디지털 오디오 편집 여부 분석 기법과 제안한 편집여부 분석 절차에 대해 기술하면서 5장에서 결론을 맺는다.

## II. 현재 서비스 중인 디지털 증거물 관리 및 분석 방법

상당수의 디지털 파일이 증거물로 활용되는 판례에서 보는 바와 같이, 범죄 수사 과정에 증거물 채증 과정과 그 과정을 증명하는 것에 상당한 자원이 동원되어 수사력과 행정력의 낭비가 발생하는 점이 있어, 이러한 점을 보완할 수 있는 방법으로 디지털 증거물 인증에 대한 요구가 늘어나고 있다. 디지털 증거물 인증시스템은 이러한 수사력과 행정력의 불필요한 낭비를 줄일 수가 있어, 범죄수사와 법과학 분석, 행정의 과정이 보다 효율적으로 이루어 질 수 있다. 실제 범죄 수사에 적용되는 디지털 데이터 인증 시스템 및 파일 시스템 분석 방법을 확인해 보았다.

## 2.1. 디지털 증거물 인증 시스템 [10]

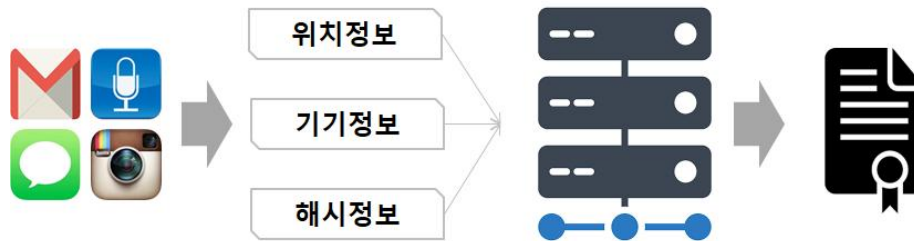


그림 1. 디지털 증거물 인증 시스템 개념

Fig. 1. The concept of digital evidence authentication system.

그림1은 국립과학수사연구원에서 현재 서비스하고 있는 디지털 증거물 인증 시스템의 개념도이다. 수사기관에서 범죄 현장에서 스마트폰으로 디지털 증거물 인증 전용 앱을 구동하여 사진, 동영상, 오디오 등의 디지털 증거를 수집한 후, 수집된 디지털 증거물에 대한 해시값 및 각종 메타정보(위치 정보, 기기 정보 및 시간 정보 등)를 인증 서버로 전송한다. 그러면, 인증 서버는 인증데이터를 저장하게 되고, 인증 서버에 대한 조작을 방지하기 위해 매일 하루 동안 전송받은 인증 데이터에 대한 2차 전자지문을 생성하여 오프라인으로 관보에 게재하고 있는 방식으로 이루어지고 있다.

검찰 혹은 법원 재판 중에 디지털 증거물에 대한 인증서를 요청할 경우, 인증 서버에서 해당 디지털증거물에 대한 원본 인증서를 발급하여 무결성을 검증할 수 있게 한다.

## 2.2. 파일시스템 기반 분석 방법

디지털 증거물을 분석하는데 EnCase 혹은 FTK와 같은 포렌식 소프트웨어를 주로 사용한다. 먼저 증거물의 손상을 막기 위해 분석 대상에 대해 쓰기방지 장치, 복제 장비 등의 포렌식 하드웨어를 통해 포렌식 소프트웨어를 이용하여 이미징(Imaging)데이터를 획득한다. 디지털 파일에 대해 분석할 경우, 해당 디지털 파일이 생성되는 위치 경로를 확인한다. 또한, 파일 시스템에 의해 기록된 디지털 파일의 파일 생성 시간 및 최종 수정 시간, 파일 재생 시간 등을 확인한다. 일반적으로 스마트폰에서 음성/오디오에 대해 녹음을 수행한다면, 파일 시스템에서 해당 파일의 파일 생성 시간 및 최종 수정 시간은 동일한 것으로 기록된다. 그런데, 스마트폰에서 제공하는 오디오 파일 편집 기능의 사용으로 인해 파일 생성 시간과 최종 수정시간의 시간 차이가 발생할 수 있고, 외부에서 편집한 파일이 원본의 위치 경로와 다른 경로에 저장될 수 있다. 이렇게 파일 생성 시간과 최종 수정 시간의 시간차이가 발생하거나, 위치 경로가 다르다면, 해당 파일의 위변조 가능성을 의심할 수 있다.

## 2.3. 디지털 증거물 인증 시스템 및 파일 시스템 기반 분석 방법의 한계

국립과학수사연구원의 디지털 증거물 인증 시스템은 현재 수사관 등의 한정된 권한을 가진 사용자만 사용할 수 있기 때문에, 범죄가 이뤄질 당시의 피의자 혹은 피해자에 의해 생성된 디지털 증거물과 관련하여 일반적인 피의자 또는 피해자가 직접적으로 해당 시스템을 사용하여 데이터에 대한 인증을 받는 것은 어렵다. 또한, 디지털 파일의 특성상, 파일 시스템 기반의 분석 방법도 정교한 위변조가 가능하며 특히, 파일 생성 시간 및 최종 수정 시간과 같은 시간 정보에 대해서는 시스템 시간에 의존적이기 때문에 파일 생성 시간이 다르다는 이유만으로 위변조 가능성을 의심하는 것은 무리가 있을 수 있다. 따라서, 디지털 오디오 파일은 파일 포맷 구조 및 오디오 신호 정보 등을 추가적으로 분석해야 하는 점이 있으며, 이러한 특성은, 기존의 판례에서도 확인된다 [6][7].

## Ⅲ. 디지털 녹음 파일에 대한 편집 여부 분석법: 분석 사례를 중심으로

디지털 오디오 신호의 압축 방법은 크게 음성 코덱과 오디오 코덱으로 구분된다. 음성 코덱과 오디오 코덱을 분리하는 요소 중의 하나는 대역폭이다. 만약 말소리와 같이 대부분의 음역대가 4kHz 혹은 8kHz 이하로 분포하면 GSM, AMR 등

과 같은 음성 코덱으로 압축을 하며, 20kHz 이하 구간에 정보가 존재하면 MP3, AAC 등과 같은 오디오 코덱으로 압축한다. 디지털 녹음기 입장에서 보면 오디오 신호를 압축을 하면 정해진 규격과 방법으로 오디오 신호를 프레임이라는 작은 단위로 만들어 인코딩하게 되고, 사용자 입장에서는 인코딩된 파일을 디코딩하여 스피커를 통해 듣게 되는 것이다. 본 논문에서 기술한 편집 여부에 대한 분석은 인코딩 직후부터 디코딩 직전까지 어떠한 인위적인 조작이 있었는지 여부에 대해 분석하는 작업을 의미한다.

이러한 디지털 오디오 파일의 편집 여부 분석을 위해 필수적인 요소는 대조 녹음기 제공 여부 혹은 동일한 조건으로 녹음된 대조 파일이 있는지 여부이다. 예를 들어, 동일한 제조사의 스마트폰이라고 할지라도 펌웨어 버전에 따라 파일 구조 및 유효 주파수 대역이 다를 수 있기 때문에 대조자료에 의한 비교 분석이 필요하다. 편집 여부 분석에 필요한 대조 자료가 제공되지 않는다면, 정확한 분석 결과를 도출할 수 없다.

본 절에서는 녹음장치에서 녹음된 파일이 나타내는 파일 형식으로 갖추어진 외형적 특징과 파일 내에 기록된 주파수 정보에서 확인할 수 있는 특징을 나누어 분석한 사례로부터 편집여부 분석에 필요한 기준을 확인해 보았다.

### 3.1. 녹음장치에 따른 오디오 파일 포맷 구조 특징 확인

[11]에서는 동영상의 편집여부를 분석하는데 멀티미디어 컨테이너 파일 포맷 내부의 구조적인 특성으로 분석하는 방법에 대해 소개하였다. 멀티미디어 컨테이너 파일 포맷 구조 분석 방법을 오디오 파일에 적용하여 녹음 장치에 따른 오디오 파일 포맷에 대한 구조를 분석한다.

#### 3.1.1. 스마트폰으로 녹음한 파일에 대한 오디오 파일 포맷 특징 분석 사례

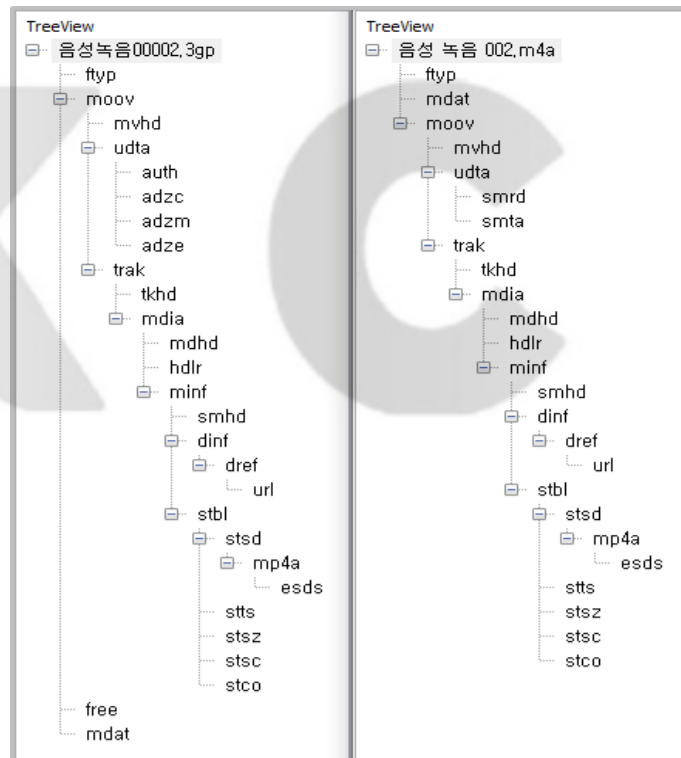


그림 2. (좌)LG 스마트폰과 (우)삼성 스마트폰으로 녹음된 파일에 대한 파일 포맷 구조 (안드로이드 4.2)  
Fig. 2. The file format structure (android version 4.2) of digital audio file recorded from (left) LG smart phone and (right) Samsung smart phone.

디지털 오디오 신호를 인코딩하는 방법에 따라 기록되는 부가정보들이 다양하며 특징이 존재한다 [12]. 스마트폰을 사용하게 되는 경우, 특히, 해당 스마트폰에 기본으로 탑재된 프로그램을 사용하게 될 경우, 대체로 MPEG-4 형식으로 오디오 파일이 생성된다. 물론, 기본으로 탑재된 프로그램이 아닌 애플 앱스토어나 구글 플레이 등에서 확인될 수 있는 프로그램들을 사용하게 될 경우는 다른 형식으로도 기록 가능하다 [13][14]. 우선 안드로이드를 운영체제로 사용하는 스마트폰(이하, 안드로이드 스마트폰)으로 녹음된 파일을 예로 들어보자.

그림2는 안드로이드 4.2 버전으로 구동된 안드로이드 스마트폰에 기본 탑재된 녹음 기능을 통해 녹음한 파일에 대한 파일 포맷 구조를 나타낸 것이다. 그림2에서 보는 바와 같이 오디오 파일 포맷이 각각 3gp 및 m4a인 것으로 확인되고, 각 파일 포맷 안에 기록된 'mdat', 'mvhd', 'udta', 'trak'와 같은 구성요소(box)의 배열 순서와 각 내부 구조 등이 서로 상이하다. 또한, 그림3에서 보는 바와 같이 LG전자 제조의 안드로이드 스마트폰으로 녹음 받은 파일은 'auth' 구성요소에 'LGE.'라는 제조사를 식별할 수 있는 특징 정보가 기록되어 있고, 삼성전자 제조의 안드로이드 스마트폰으로 녹음 받은 파일

은 'smrd' 구성요소에 'TRUEBLUE'라는 특징 정보가 기록되어 있으며, 각 파일들에서는 ISO 기반의 표준 MPEG-4 파일 구조에서 정의된[15] 파일의 생성 시간 등의 정보가 확인된다. 그림4에서 보는 바와 같이 LG 스마트폰에 기록된 오디오 파일 내에 'mvhd'문자열에 파일 생성 시간과 수정 시간이 [2015-12-05, 08:46:31]으로 설정되어 있으며, 삼성 스마트폰에 기록된 오디오 파일에서는 [2014-08-03, 15:40:49]으로 설정되어 있다. 이 시간은 협정 세계시를 기준으로 표시된 것이며, 한국표준시로 변환하면 각각 [2015-12-05, 17:46:31]과 [2014-08-04, 00:40:49]이 된다.

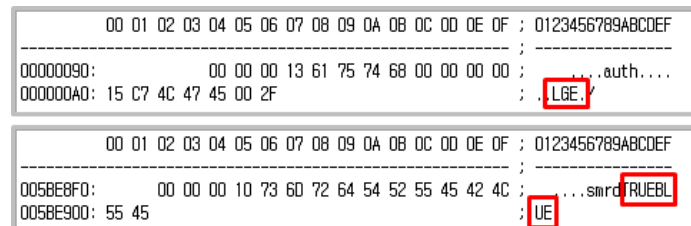


그림 3. (상)LG 스마트폰과 (하)삼성 스마트폰으로 녹음된 파일에 대한 특징(I)

Fig. 3. The characteristics (I) of digital audio file from (top) LG smart phone and (bottom) Samsung smart phone.

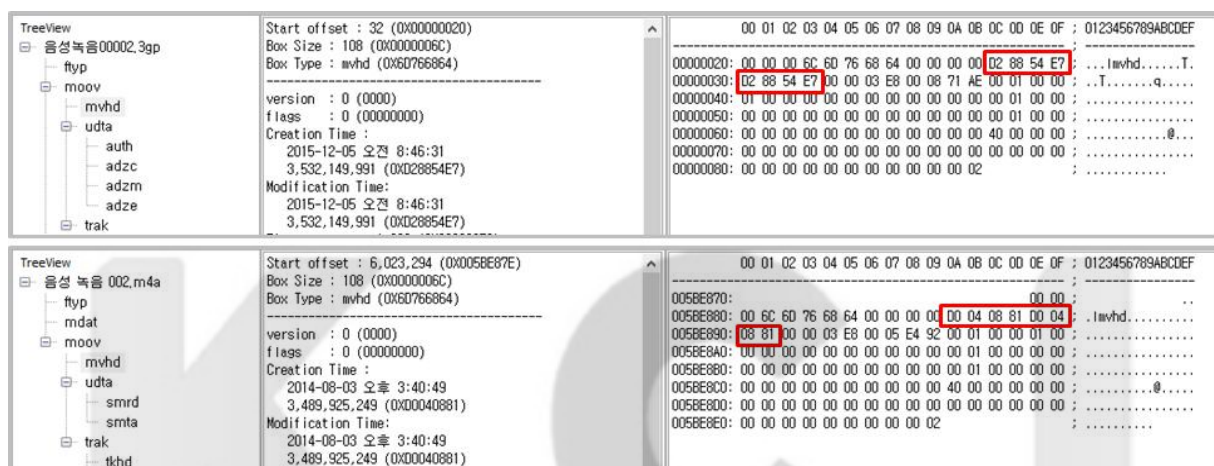


그림 4. (상)LG 스마트폰과 (하)삼성 스마트폰으로 녹음된 파일에 대한 특징(II)

Fig. 4. The characteristics (II) of digital audio file from (top) LG smart phone and (bottom) Samsung smart phone.

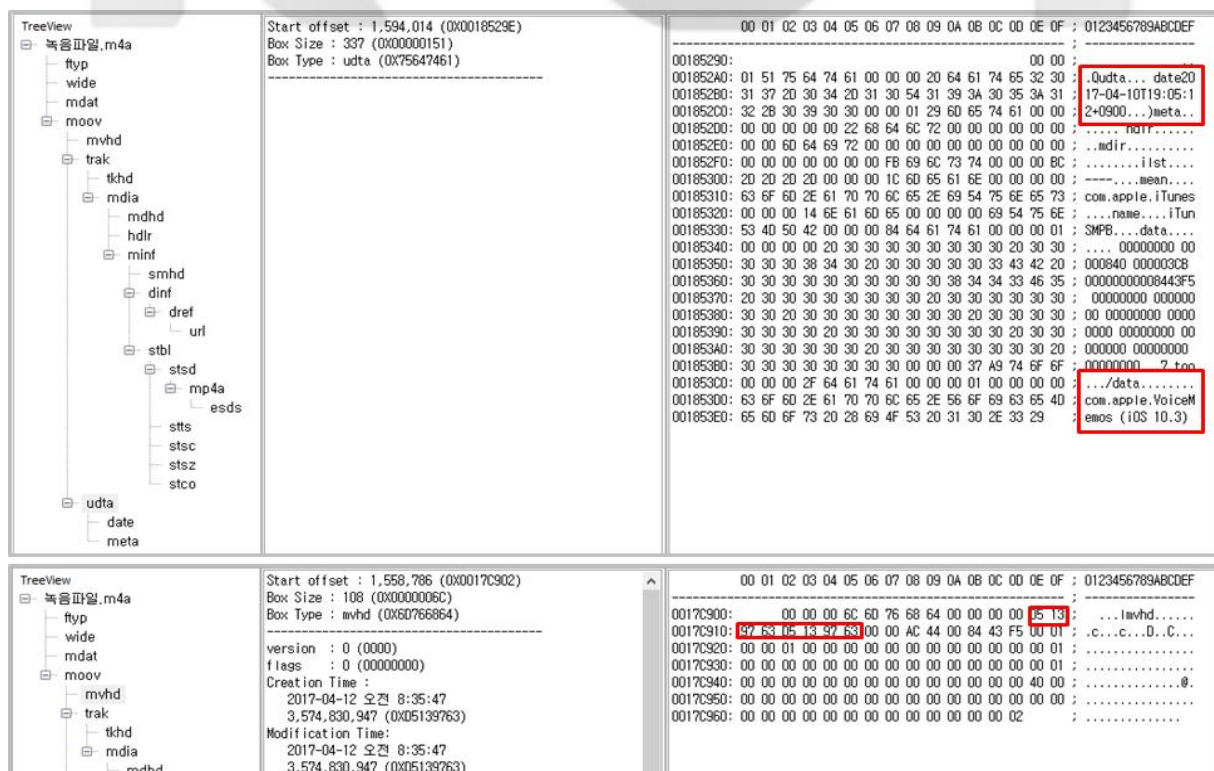


그림 5. 아이폰으로 녹음된 파일에 대한 파일 구조 및 특징.

Fig. 5. The file format structure and characteristics for audio file recorded by iphone.



아이폰의 경우, 그림5와 같이 파일 내에서 확인되는 MPEG-4를 구성하는 box의 배열 구조가 안드로이드 스마트폰에 의한 것과 다르다. 파일 포맷 구조 내에 'udta'의 'meta' 구성요소에 'com.apple.VoiceMemos (iOS 10.3)'와 같이 사용된 앱과 운영체제 정보가 기록되어 있고, 'udta'의 'date' 구성요소에 '2017-04-10T19:05:12+0900'이라는 정보는 아이폰의 음성메모 녹음 시작 시점을 의미한다. 아이폰에서 해당 음성 파일에 대한 파일 생성 시간 및 수정 시간은 [2017-04-12 08:35:47]로 기록되어 있고, 이는 협정 세계시에 해당하므로 한국표준시로 변환하면, [2017-04-12 19:35:47]이 된다.

### 3.1.2. 일반 녹음기로 녹음한 파일에 대한 오디오 파일 포맷 특징 분석 사례

스마트폰이 아닌 일반 녹음기기의 경우, MPEG-4 뿐만 아니라, 아닌 다양한 파일 포맷이 존재한다. SONY사의 녹음기 경우, 녹음된 오디오 파일은 mp3로 저장될 수 있는 경우가 있다. 일반적으로 mp3형식의 파일에는 ID3 태그, Xing 헤더 및 Lame 헤더 등의 규칙성 있는 형식의 메타정보로 가수, 앨범명, 가사, 장르 등과 같은 정보를 기록하게 될 수 있다 [14]. SONY사의 녹음기 중, ICD-UX543F의 경우, 메타 정보로 그림6과 같이 녹음 장비명(ICD-UX543F) 및 파일이름(150327\_001) 등이 ID3에 기록되는 것을 알 수 있다 [16]. 이때, 파일 이름은 녹음기에 설정된 시간으로, 해당 녹음기의 경우, 설정 시간에 따라 파일 이름이 변경될 수 있는 특징이 있다 [17].

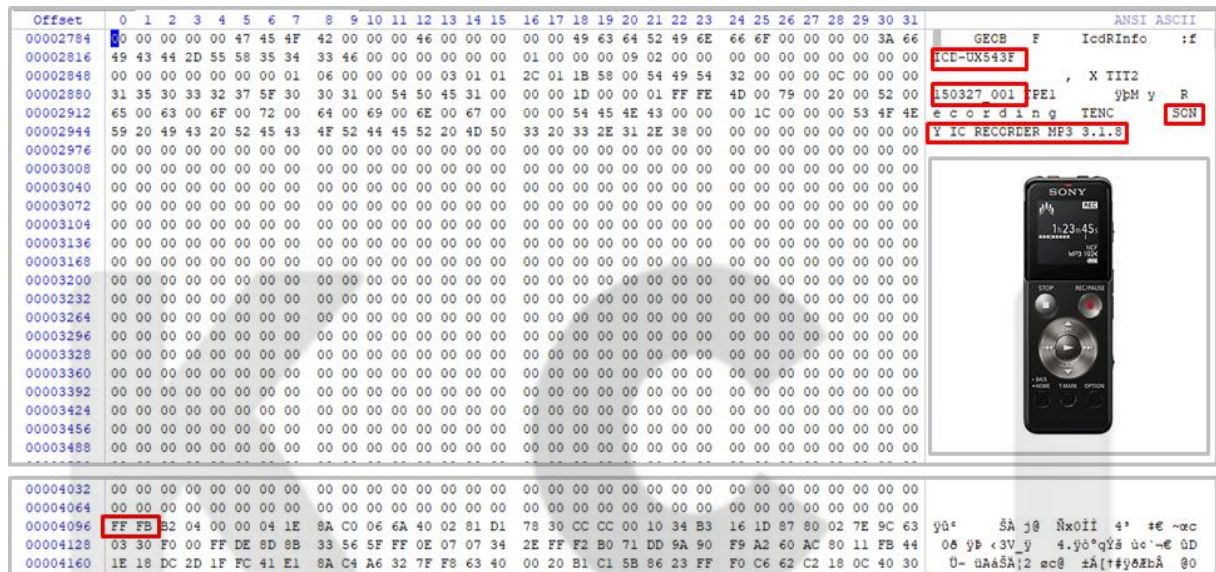


그림 6. 소니 녹음기로 녹음된 파일에 대한 파일 특징.

Fig. 6. The file format structure of audio file recorded from Sony Recorder.

그림7은 USB타입의 녹음기기이며, 이 녹음기기의 파일 구조는 그림8과 같다. 스마트폰의 파일 포맷은 m4a, 3gp, mp4 등의 형태로 구성되어 있는 반면, 그림7의 녹음기기로 녹음할 경우 WAV 형태로 저장된다. 일반적으로 WAV 포맷에는 파일 생성 시간 및 수정 시간 등의 메타 정보가 거의 기록되지 않는다. 하지만 상용 편집툴을 사용해서 WAV포맷으로 저장할 경우, 그림8과 그림9에서 보는 바와 같이 WAV 헤더 위치와 데이터 위치의 오프셋(offset) 정보가 달라질 수 있기 때문에 오프셋 정보는 WAV 포맷의 특징이 될 수 있다.

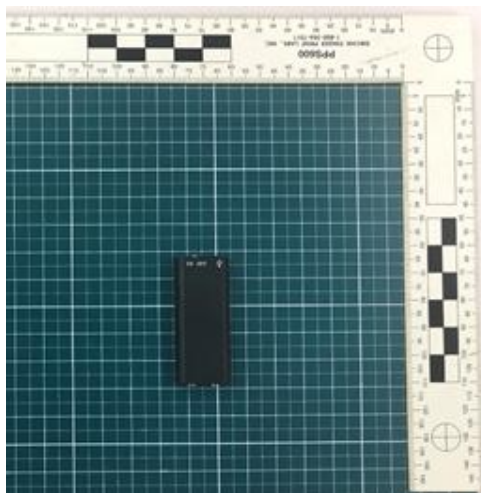


그림 7. USB 타입의 녹음 장치.

Fig. 7. The recording device of USB type.

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	ANSI ASCII				
00000000	52	49	46	46	F8	01	00	00	57	41	56	45	66	6D	74	20	E4	01	00	00	11	00	02	00	C0	5D	00	00	C0	5D	00	00	RIFF	WAVEfmt	ä	ä	ä
00000032	10	08	04	00	02	00	F9	07	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000064	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000116	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000480	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000512	5F	FE	00	00	FA	FD	00	00	FF	78	13	33	FF	78	13	33	B2	F9	AA	18	B2	F9	AA	18	11	23	70	03	11	23	70	13	00	00	00	00	

그림 8. USB 타입의 녹음 장치로 녹음한 파일의 파일 구조.

Fig. 8. The file format structure of audio file by recording device of USB type.

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	ANSI ASCII
00000000	52	49	46	46	34	70	06	00	57	41	56	45	66	6D	74	20	E4	01	00	00	11	00	01	00	80	BB	00	00	E9	5D	00	00	RIFF4p WAVEfmt ë
00000032	10	08	04	00	02	00	F9	0F	66	61	63	74	04	00	00	00	5E	DA	0C	00	54	61	74	61	00	70	06	00	01	00	00	00	00
00000064	19	09	00	10	09	00	91	91	09	01	91	00	90	01	00	19	00	91	00	90	91	01	19	00	91	91	00	19	00	00	00	00	00

그림 9. 상용 편집툴을 사용해서 WAV 포맷 형태로 저장한 파일의 파일 구조.

Fig. 9. The file format structure of audio file saved in WAV format using commercial editing tools.

이처럼, 오디오 파일 포맷 구조 분석 방법은 동일한 조건의 대조 녹음기 혹은 대조 파일을 통해 파일 포맷을 비교 분석함으로써, 제시된 디지털 증거물이 생성될 수 있는 가능성이 있는지 여부에 대해 분석하는 기법이다.

### 3.1.3. 스마트폰 기본 앱을 사용해서 편집한 파일의 파일 포맷 구조 변화 특징

만약 스마트폰에서 편집 기능을 사용해서 편집할 때와 로컬 PC로 파일을 복사하여 편집툴을 사용하여 편집할 때, 파일 구조의 변화가 생기는지 알아보자.

삼성스마트폰(안드로이드버전 9)을 통해 녹음한 파일을 스마트폰에 탑재된 기본 앱에서 제공하는 편집툴로 편집했을 경우, 그림10과 같이 편집툴로 편집하기 전과 비교해 보면 파일 포맷 구조가 변화된 것을 확인할 수 있다. 특히, 안드로이드 버전 정보가 기록된 'meta' 구성요소가 편집툴을 사용한 후에 삭제되었으며, 'tkhd' 구성요소에서 보는 바와 같이 편집툴을 사용하게 되면 편집 후 저장하는 시점의 파일시스템 시간으로 수정 시간이 변경되었다.

TreeView	Start offset : 491,242 (0X00077EEA)	Box Size : 92 (0X0000005C)	Box Type : tkhd (0X74686864)
ftyp	version : 0 (0000)	flags : 7 (0X00000007)	
mdat	Creation Time : 2019-04-23 오전 6:43:59		
moov	3,638,846,639 (0XD8E464AF)		
mvhd	Modification Time : 2019-04-23 오전 6:43:59		
udta	3,638,846,639 (0XD8E464AF)		
vrtd	Track ID : 1 (0X00000001)		
metd	Reserved : 0 (00000000)		
ampl	Duration : 28,491 (0X00006F4B)		
book	Reserved[0] : 0 (0000)		
SDLN	Reserved[1] : 0 (0000)		
smrd	Video Layer : 0 (000000)		
smta	Alternate : 0 (000000)		
meta	Audio volume : 256 (0X0100)		
trak	Reserved : 0 (000000)		
tkhd	Matrix[0] : 65,536 (0X00010000)		
mdia	Matrix[1] : 0 (0000000000)		
mdhd	Matrix[2] : 0 (0000000000)		
hdlr	Matrix[3] : 0 (0000000000)		
minf	Matrix[4] : 65,536 (0X00010000)		
smhd	Matrix[5] : 0 (0000000000)		
dinf	Matrix[6] : 0 (0000000000)		
dref	Matrix[7] : 0 (0000000000)		
url	Matrix[8] : 1,073,741,824 (0X40000000)		
stbl	Video Width(0): 0 (0000000000>>16)		
stsd	Video Height(0): 0 (0000000000>>16)		
mp4a			
esds			
stts			
stsz			
stsc			

그림 10. 삼성 스마트폰으로 녹음된 파일에 대해 편집 (좌)전/(우)후 파일 구조 및 특징 변화.

Fig. 10. The difference in file format and characteristics of the audio file (left)before and (right)after manipulation by Samsung smart phone.

다음은 아이폰(iOS 10.3.3)을 통해 녹음된 파일을 아이폰에 음성 메모 앱에서 제공하는 편집툴로 편집했을 경우, 편집툴로 편집하기 전과 비교해 보면 파일 포맷 구조에서의 변화는 관찰되지 않았으며, 녹음 파일을 편집 후 다시 저장하는 시점의 아이폰 파일시스템 시간으로 파일 생성 시간과 수정 시간이 모두 변경되었다.

### 3.1.4. 로컬 PC에서 편집한 파일의 파일 포맷 구조 변화 특징 분석 사례



로컬PC로 복사된 녹음된 파일을 편집한 후에 Freeware Advanced Audio Coder (FAAC)통해 m4a 파일 포맷으로 저장할 경우, 그림2 및 그림5에서와 같이 안드로이드 계열 스마트폰 및 아이폰의 기본 앱으로 녹음한 파일 비교해 보면 그림11에서의 파일 포맷 구조가 상이한 것이 확인되며, 'mvhd' 구성요소의 파일 생성 시간 및 수정 시간은 FAAC를 통해 인코딩 되는 시점의 로컬PC의 파일 시스템 시간으로 설정된다. 또한, 'udta'의 'date' 구성요소에 'FAAC 1.28'과 같이 FAAC의 버전 정보가 기록된다 [18].

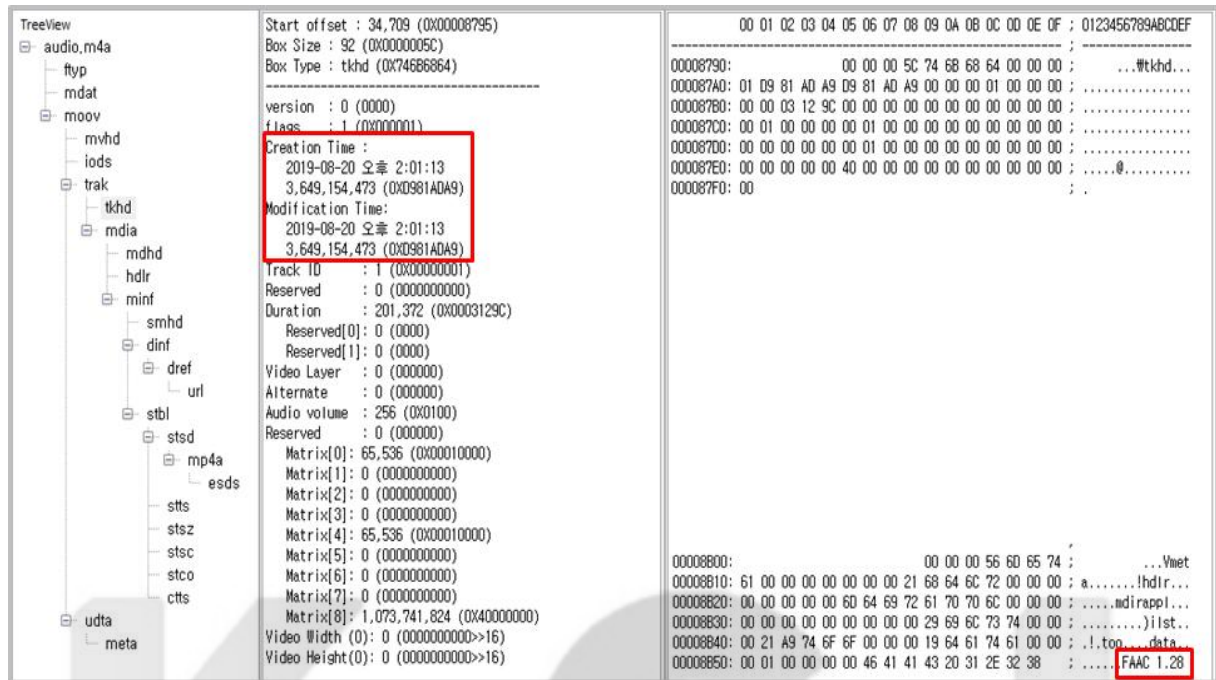


그림 11. FAAC로 인코딩 후 파일 구조 및 특징.

Fig. 11. The file format structure and characteristics after encoded by FAAC.

### 3.2. 녹음 장치 및 방법에 따른 오디오 신호 특징

#### 3.2.1. 오디오 파일의 유효주파수 대역 분석 사례

오디오 파일 포맷 구조 분석이 오디오 기록 방법에 대한 것이라면, 오디오 신호 분석은 기록된 신호를 디코딩하여 주파수 분포 등을 해석하는 것이다. 주파수 분석이 가능한 상용 편집툴인 Adobe사의 "Audition"을 사용해서 안드로이드 계열 스마트폰 및 아이폰에서 녹음된 오디오 파일을 확인하면 그림12와 같다. 각각의 녹음파일들은 샘플링레이트가 44.1kHz, 채널 수도 모노로 동일하다. 그러나, 그림에서 흰색선으로 표기된 바와 같이 실제 오디오 신호가 기록되는 유효대역폭은 각각 다른 것으로 확인된다. 또한, 동일한 기종의 스마트폰이라고 할지라도, 운영체제 및 펌웨어 버전에 따라 녹음 방식의 차이가 있을 수 있으며, 만약 유효대역폭이 해당 녹음기기가 가져야할 대역폭과 다르다면 편집을 의심할 수 있다.

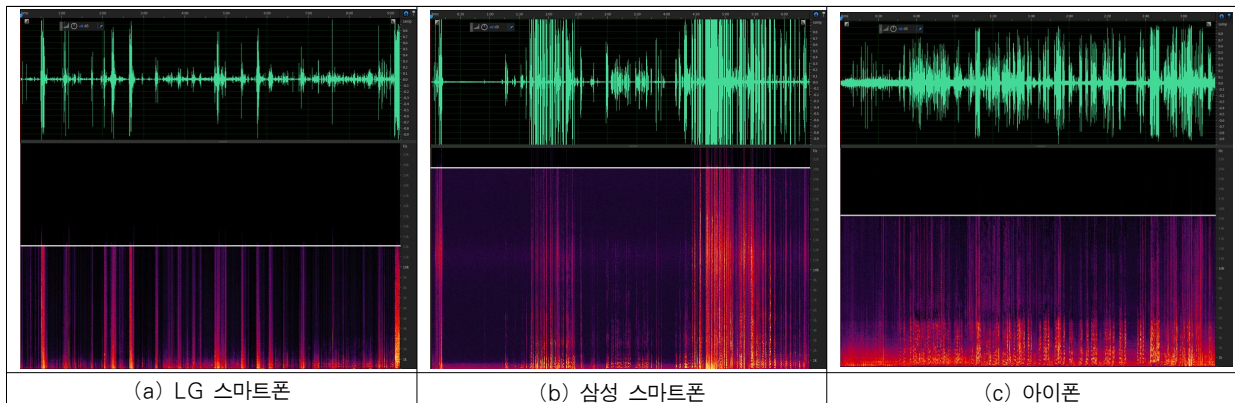


그림 12. 오디오 파형 및 스펙트로그램

Fig. 12. The audio waveform and spectrogram.

#### 3.2.2. 녹음기의 고유 특성 분석 사례



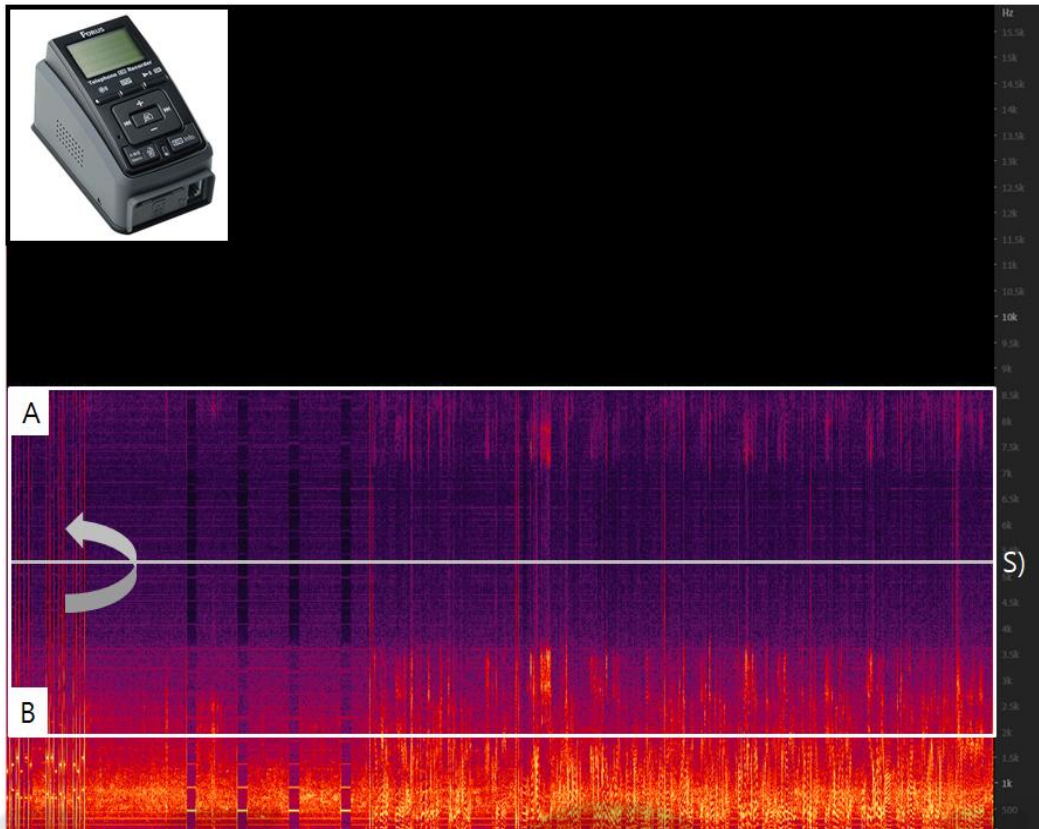


그림 13. 특정 녹음기로 녹음된 파일에 대한 오디오 파형 및 스펙트로그램

Fig. 13. The audio waveform and spectrogram of audio file recorded by particular device.

오디오 신호의 주파수 정보에서 그림13과 같이 녹음기기를 특정할 수 있는 중요한 요소를 검출할 수도 있다. 그림에서 보는 바와 같이 'FSC-1000'이라는 통화녹음 장치로 유선전화기에 연결하여 통화내용을 녹음하는 장치이다. 'FSC-1000'을 사용하여 녹음할 경우, 그림13의 S) 주파수 대역을 기준으로 그림15의 A)영역과 B)영역이 서로 미러링되어 있는 것을 확인할 수 있다. 이는 오디오신호의 대역폭 확장 방법이라는 기법이 녹음기에 적용되어 음질을 개선하기 위한 방법 중 하나다 [19].

### 3.2.3. 주변잡음을 활용한 불연속 구간 검출 사례

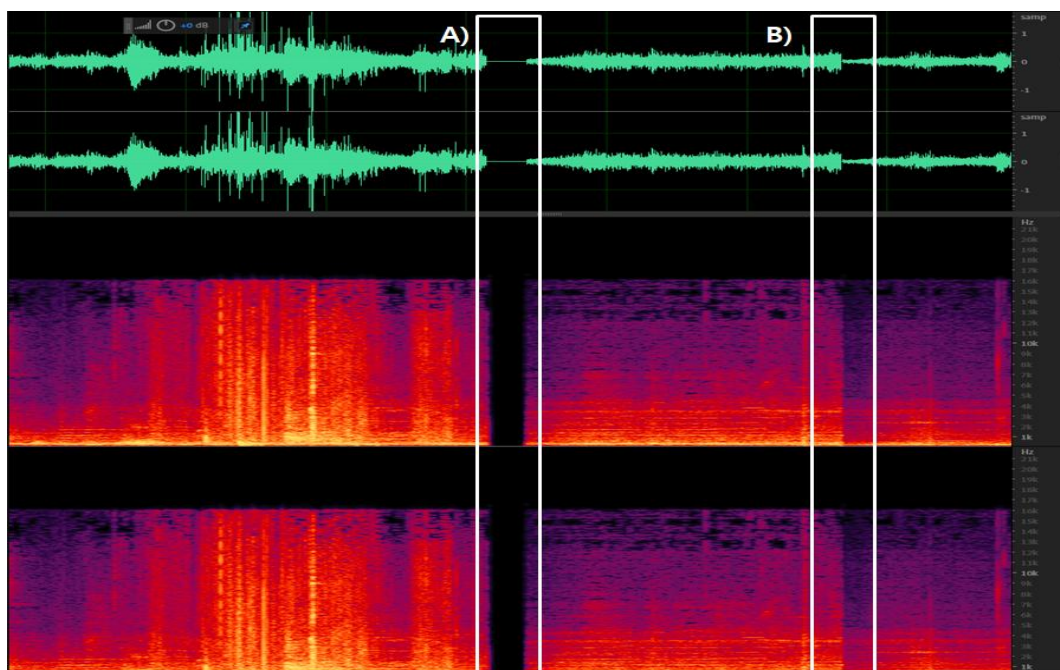


그림 14. 불연속구간이 존재하는 오디오 파형 및 스펙트로그램

Fig. 14. The audio waveform and spectrogram including the discontinuity period.

오디오 신호에 대해 주파수 분석과 동시에 오디오 파일 전체를 청취함으로써, 오디오 신호의 불연속성을 확인할 수 있는 신호가 있는지 확인해야 한다. 그림14의 A)에서 보는 바와 같이 오디오 파일의 중간 부분에 '0'의 데이터를 갖는 부분과 그림14의 B)에서와 같이 주변 잡음의 주파수 분포가 급격하게 변하는 지점을 불연속 구간으로 볼 수 있다. 이와 같이 만약 음악이 존재하는 카페 안에서 오디오 파일이 녹음되었다면, 해당 파일에 음악의 불연속 구간이 있는지 여부를 확인함으로써, 편집 여부에 대해 한정할 수 있다.

### 3.3. 오디오 파일 포맷 구조 분석 및 오디오 신호 분석의 한계

위에서 언급한 바와 같이 오디오 파일 포맷 분석을 통해 스마트폰으로 녹음된 파일의 경우에는 오디오 파일을 녹음한 시점 및 녹음파일을 저장한 시점에 대해 제한적으로 확인할 수 있다. LG, 삼성, 애플에서 제조한 스마트폰에 대해서는 제조사의 특징 정보가 녹음파일에 기록되어 있기 때문에 분석이 가능하다. 그러나 오디오 파일 구조 분석은 대조파일이 제시되어야 하며, 분석 대상의 오디오 파일과 대조 파일 간의 파일 구조가 동일하다고 해서 반드시 편집이 되지 않았다고 볼 수는 없다. 왜냐하면, 분석 대상의 오디오 파일에 대한 인코더가 있다면, 어떠한 오디오 파일도 정교하게 조작한 후에 파일 포맷 구조로 동일하게 만들어 낼 수 있기 때문이다. 그러므로 비압축 방식의 녹음기를 사용할 경우, 파일 포맷 구조 분석 및 오디오 신호 분석을 통해 편집 여부를 검출할 수 없다. 또한, 연속성을 확인할 수 있는 주변 잡음없이 대화자의 음성만 녹음되었을 경우, 문장 혹은 어절 단위의 구간을 삭제하게 되면, 오디오 신호 분석 방법으로는 불연속 구간의 검출이 불가능하다. 디지털 파일의 무결성 검증을 위한 편집 여부에 대한 과학적 이론을 바탕으로 확실성을 제시하는 것은 매우 어려운 일이다. 표1은 3장에서 언급한 편집 여부 분석을 위해 검사해야 될 녹음 장비 및 디지털 오디오 파일에 대한 특징 요소들을 나타낸 것이다. 표1의 측정 요소는 편집 여부 분석을 위한 모든 검사 요소를 나타낸 것은 아니지만, 오디오 파일에 대한 무결성 혹은 녹음 과정에 대한 과학적인 결론을 도출하는데 유용하다 [20]. 표1에 나열한 검사 방법 및 측정 요소들은 추후 수정, 확장, 및 대체될 수 있다.

표 1. 디지털 오디오 편집여부에 대한 프레임워크.  
Table 1. The framework of digital audio authentication

검사 방법	측정 요소
녹음 장비 사용 여부	녹음기에서 제공하는 기능
디지털 오디오 파일 포맷 구조 분석	파일 압축 방식
	파일 컨테이너 구성 및 순서
	메타 데이터 내용
디지털 오디오 신호 분석	유효 주파수 대역
	주파수 분포
	배경 잡음 및 주기성 신호

## IV. 디지털 오디오 파일의 편집 여부 분석 절차

3장에서 설명한 바와 같이 오디오 파일 포맷 구조 분석 및 오디오 신호 분석을 통해 편집여부에 대해 검출할 수 있는 절차를 그림15와 같이 도식화한다. 먼저 편집 여부가 의심스러운 디지털 오디오 파일이 주어지면, 오디오 신호와 파일 포맷 구조를 분석한다. 이 때, 녹음기기를 통해 대조파일을 획득할 수 있는지에 대한 확인 절차가 먼저 선행되어야 한다. 만약 녹음기기 혹은 대조파일이 제시되지 않으면, 파일 포맷 구조 분석은 수행할 수 없기 때문에, 주기적인 신호(의미있는 신호) 및 음악 등과 같은 주변 잡음(주변 신호)에서 불연속 구간이 확인되는지 여부에 따라 편집 여부에 대한 가능성을 판단하게 된다. 왜냐하면, 디지털 오디오 파일을 비교하기 위한 녹음기는 제시되지 않았지만, 만약 최초 녹음 과정에서 사용했던 녹음기가 녹음 중 멈춤이 가능한 경우라면, 의미있는 신호 및 주변 잡음 등에서 불연속 구간이 확인되었다고 해서 편집으로 볼 수 없기 때문이다. 본 논문에서 정의하는 디지털 오디오 파일의 원본에 대한 정의는 녹음 과정을 최초로 저장한 파일을 의미한다. 따라서 녹음 도중에 멈춘 후, 다시 녹음을 재개할 경우, 불연속 구간이 발생할 수 있기 때문에 선불리 조작된 것으로 단정할 수 없다.

만약, 대조파일을 획득할 수 있는 녹음 기기가 제시되었다면, 제시된 디지털 증거물과 동일한 방법으로 대조 파일을 만든 다음, 오디오 파일의 압축 방식을 확인한다. 오디오 신호가 PCM(Pulse Code Modulation)과 같은 비압축 방식으로 저장되어 있다면, 다른 메타 정보를 기록되지 않기 때문에 파일 포맷 구조 검사를 수행하지 않고, 오디오 신호만 분석한다. 오디오 신호가 압축되어 인코딩되어 있는 것이라면, 파일 포맷 구조 비교, 녹음기의 하드웨어적인 특성 등을 고려해서 조작 가능성을 판단할 수 있다. 그림15에서와 같이 제시된 디지털 증거물과 대조 파일 간의 비교 분석하여 특이 사항이 발견되면, 조작으로 판단할 수 있지만, 제안한 편집여부 분석 절차에 따라 수행하였을 때, 디지털 파일의 특성상 정교한 위변조가 이론상 가능하기 때문에 특이사항이 발견되지 않았다고 해서 원본이라고

보장할 수 없다.

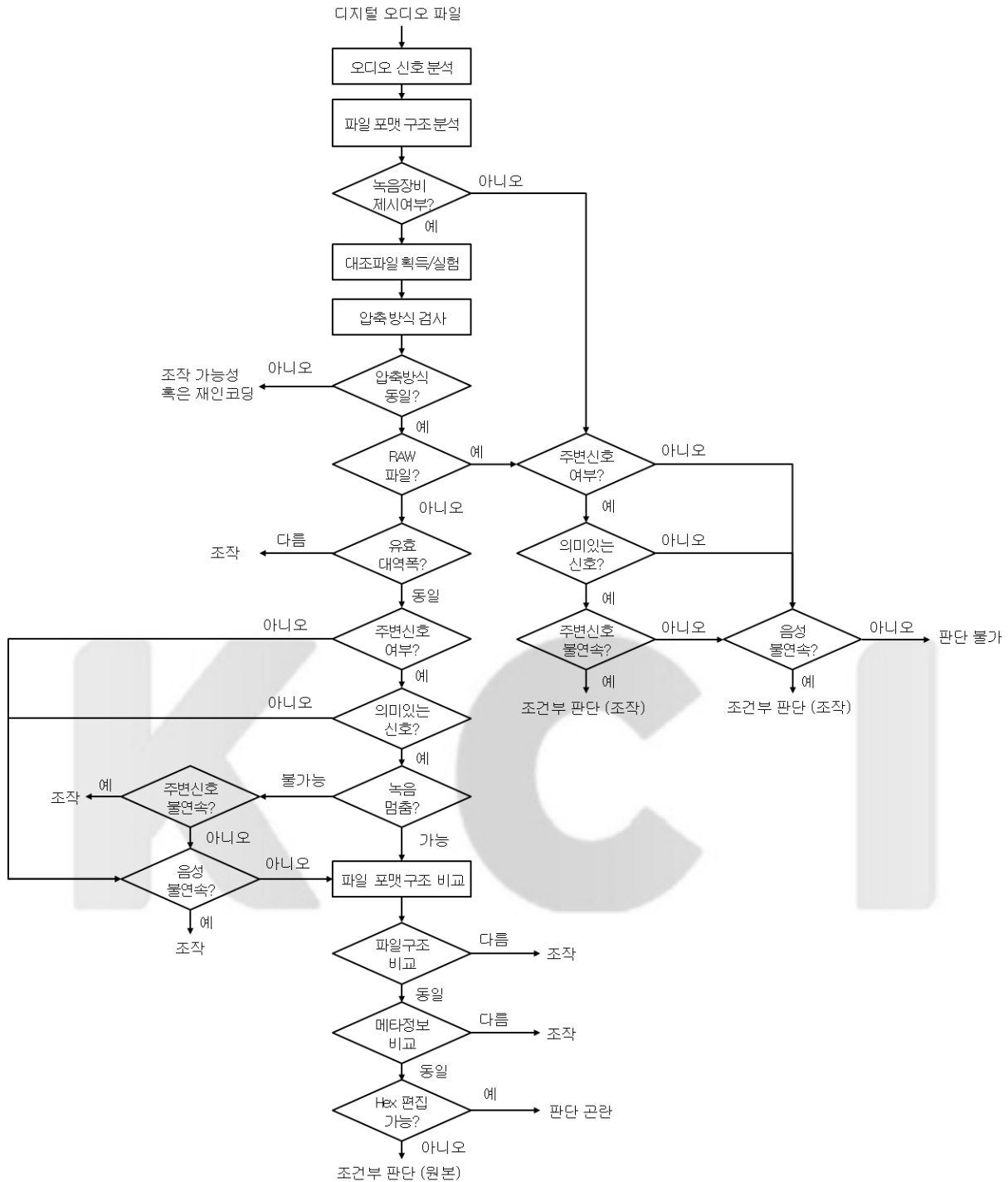


그림 15. 제안한 편집 여부 분석 절차

Fig. 15. The proposed authentication analysis procedure.

## V. 고찰

본 논문에서는 다양한 실제 사례들을 분석한 후, 디지털 오디오 파일에 대한 오디오 신호 및 파일 포맷 분석을 통한 편집 여부 분석 절차에 대해 제안하였다. 제안된 편집 여부 분석 절차에 따라 편집이 이루어진 파일에 대해서 디지털 오디오 파일의 조작 과정에서 발생하는 조작 흔적을 통해 원본이 아님을 제한적으로 판단할 수 있다. 그러나, 일부 특정 조건에서는 제안한 편집 여부 분석 절차로도 위변조 여부가 검출되지 않을 수 있으므로, 이를 해결하기 위한 정교한 분석 절차에 대한 연구가 필요하다.



## 참 고 문 헌 (References)

- [1] 전용수, 안상진. (2018). *4차 산업혁명 기술경쟁력 분석 및 시사점: 사물인터넷 중심으로*. 한국과학기술기획평가원.
- [2] 박세진, 윤지영, “이차 보간에 따른 ENF 기반의 위변조 디지털 파일 탐지 기법,” *정보과학회지*, 제45권, 제3호, 2018년, 3월.
- [3] David G. and Edward J. D., “Deepface Video Detection Using Recurrent Neural Networks,” in *proceedings of IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS)*, Nov. 2018.
- [4] 양일호, 김경화, 김명재, 백록선, 허희수, 유하진, “법음성학에서의 오디오 신호의 위변조 구간 자동 검출 방법 연구,” *말소리와 음성과학*, 제6권, 제2호, 2014년, 6월.
- [5] “공갈미수 [대법원 2005. 12. 23., 선고, 2005도2945, 판결].” (2019년 7월 20일). 국가법령정보센터 판례. <http://www.law.go.kr/precInfoP.do?precSeq=84232>.
- [6] “특정범죄가중처벌등에관한법률위반(뇌물) [대법원, 2008. 3. 13., 2007도10804, 판결].” (2019년 8월 24일). 국가법령정보센터 판례 <http://www.law.go.kr/precStmdInfoP.do?precSeq=85593>.
- [7] “내란음모·국가보안법위반(찬양·고무등)·내란선동 [서울고등법원, 2014. 8. 11., 2014노762 판결].” (2019년 8월 24일). 국가법령정보센터 판례. <http://www.law.go.kr/LSW/precInfoP.do?jsessionid=t3kT8WD-OPdIYIby58F8wfeO.LSW2?evtNo=2014%EB%85%B8762>.
- [8] 오병택, 문병주, 이동일, “디지털 워터마킹 기술동향 및 전망,” *전자통신동향분석*, 제17권, 제6호, 2002년 12월.
- [9] 이지우, 이정환, 심규선, 변준석, 나기현, 이중 “사진 위변조 검출 방법의 한계성 연구,” *디지털포렌식연구*, 제12권, 제1호, 2018년, 6월.
- [10] “디지털 증거물 위변조 원천 봉쇄한다 국과수, 수사기관 대상 디지털 증거물 검증 서비스 시행.” (2019년 8월 1일). 행정안전부 보도자료. [https://www.mois.go.kr/frt/bbs/type010/commonSelectBoardArticle.do?bbsId=BBSMSTR\\_0000000000008&nttId=47604](https://www.mois.go.kr/frt/bbs/type010/commonSelectBoardArticle.do?bbsId=BBSMSTR_0000000000008&nttId=47604).
- [11] Thomas G., Andre F., and Matthias K., “Forensic analysis of video file formats,” *Digital Investigation*, vol. 11, 2014.
- [12] 박기현. (2007). *이론과 실무의 조화 코덱의 세계로의 초대*. 서울:홍릉과학출판사.
- [13] “Apple Inc.” (2019년 8월 22일). *App Store - Apple (KR)*. <https://www.apple.com/kr/ios/app-store/>.
- [14] “Google Inc.” (2019년 8월 22일). *Google Play*. <https://play.google.com/store/>.
- [15] ISO/IEC 14496-12. (2019년 8월 24일). *Information technology -- Coding of audio-visual objects - - Part 12: ISO base media file format* <https://l.web.umkc.edu/lizhu/teaching/2016sp.video-communication/ref/mp4.pdf>.
- [16] “MP3” (2019년 8월 22일). *MP3 - Hydrogenaudio Knowledgebase*. <https://wiki.hydrogenaud.io/index.php?title=MP3>.
- [17] “SONY” (2019년 8월 24일). *IC Recorder ICD-UX543/UX543F/UX544F 사용설명서*. <https://helpguide.sony.net/icd/u54/v1/ko/index.html>.
- [18] “Freeware Advanced Audio Coder.” (2019년 8월 21일). *Sourceforge*. <https://sourceforge.net/projects/faac/>.
- [19] Jax P. and Vary P., “Artificial bandwidth extension of speech signals using MMSE estimation based on a hidden Markov model,” in *proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing* May 2003.
- [20] Catalin G., Daniel R. and Jeff M. S., “Analytical framework for digital audio authentication,” in *proceedings of Audio Engineering Society Conference: 46th International Conference: Audio Forensics* Jun. 2012.

## 저 자 소 개



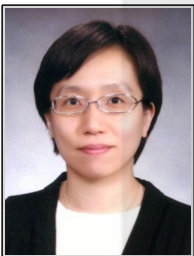
**박 남 인 (Nam In Park)**

2007년 2월 : 광운대학교 전자통신공학과 공학사  
2009년 2월 : 광주과학기술원(GIST) 정보통신공학과 공학석사  
2013년 8월 : 광주과학기술원(GIST) 정보통신공학과 공학박사  
2013년 8월~2014년 3월 : 광주과학기술원(GIST) 박사후연구원  
2014년 4월~현재 : 국립과학수사연구원 디지털분석과  
관심분야 : 화자인식, 음성합성, 음성/오디오 신호처리, 오디오포렌식, 머신러닝 및 딥러닝 등



**심 규 선 (Kyu-sun Shim)**

2009년 2월 : 고려대학교 정보통신대학 컴퓨터학과 공학사  
2011년 2월 : 고려대학교 정보통신대학 컴퓨터학과 공학석사  
2017년 2월 ~ : 고려대학교 정보통신대학 컴퓨터학과 박사과정  
2011년 11월~현재 : 국립과학수사연구원 디지털분석과  
관심분야 : 디지털 포렌식, 모바일 포렌식, 이미지 포렌식 등



**전 옥 엽 (Oc-Yeub Jeon)**

1998년 2월 : 부산대학교 물리학과 이학사  
2000년 2월 : 부산대학교 물리학과 이학석사  
2006년 8월 : 부산대학교 물리학과 이학박사  
2006년 6월~현재 : 국립과학수사연구원 디지털분석과  
관심분야 : 오디오 포렌식, 오디오 신호처리, 머신러닝 및 딥러닝 등

K C I