

104 – Fake Voice

Team Information

Team Name : ForensicGPT

Team Member : Eungchang Lee, Donghyun HA, Hyunwoo Shin, Jinhyeok Son

Email Address : forensicgpt@googlegroups.com

Instructions

Description Here is an audio file that supposedly added a fake voice intentionally. As a digital forensic investigator, solve the following questions.

Target	Hash (MD5)
original.wav	e3b2f872e7a93f1e39d831b07b88f3ae
modified.wav	5f90258daf2fa7dbf758854f3e6f5bd4

Questions

- 1) At what time does the fake voice play? (10 points)
- 2) Provide evidence to support the answer. (90 points)

Teams must:

- Develop and document the step-by-step approach used to solve this problem to allow another examiner to replicate team actions and results.
- Specify all tools used in deriving the conclusion(s).

Tools used:

Name:	HashTab	Publisher:	Implbits Software
Version:	6.0.0		
URL:	https://implbits.com		

Name:	Audacity	Publisher:	Audacityteam
Version:	3.2.1		
URL:	https://www.audacityteam.org/download/		

Step-by-step methodology:



[그림 1] original.wav 파일 해시 값 확인



[그림 2] modified.wav 파일 해시 값 확인

분석에 앞서, 두 파일에 대한 해시 값을 산출하여 MD5 해시 값이 일치함을 확인하였습니다.

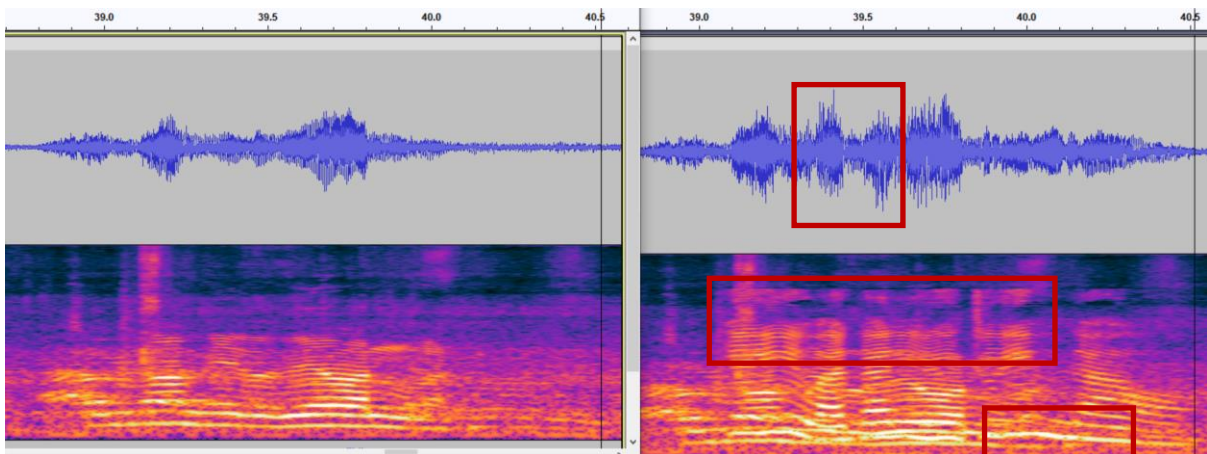
1) At what time does the fake voice play? (10 points)

두 wav 파일의 재생 길이는 54초입니다. Original.wav 파일을 재생하여 음성을 듣고, modified.wav 파일을 재생하여 음성을 들어보았을 때, fake voice가 **39초(약 39.1 초)부터 재생되어 40초(약 40.5초)까지** 들리는 것을 확인하였습니다.

해당 fake voice에서는 "Hey, What are you doing now?" 라는 문장을 식별할 수 있었습니다.

2) Provide evidence to support the answer. (90 points)

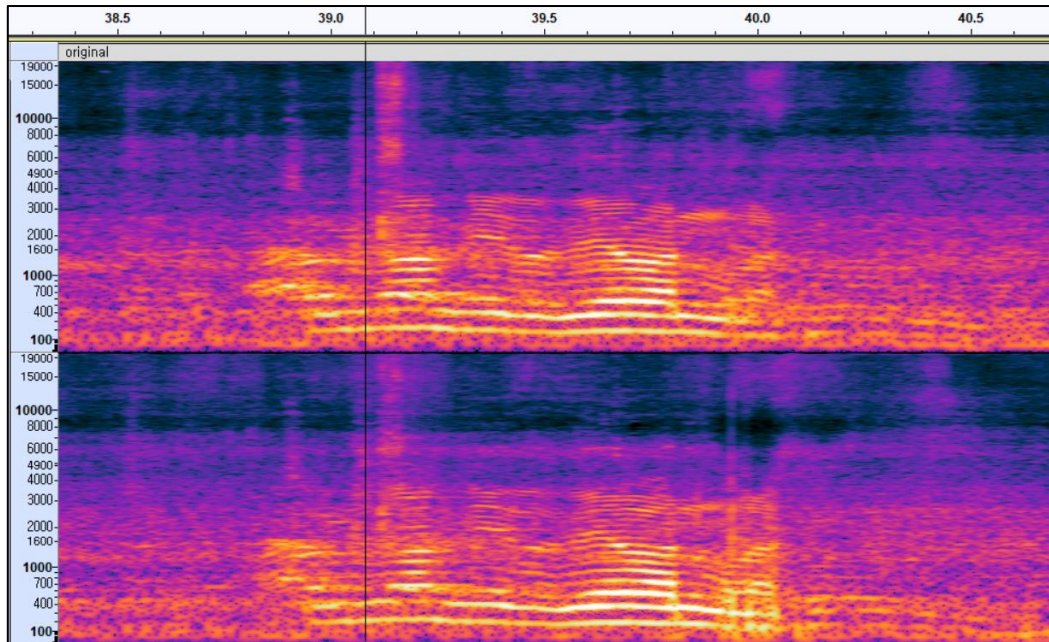
- 왼쪽 : Original.wav, 오른쪽 : Modified.wav



[그림 3] 개략적인 파형 및 스펙트로그램 분석

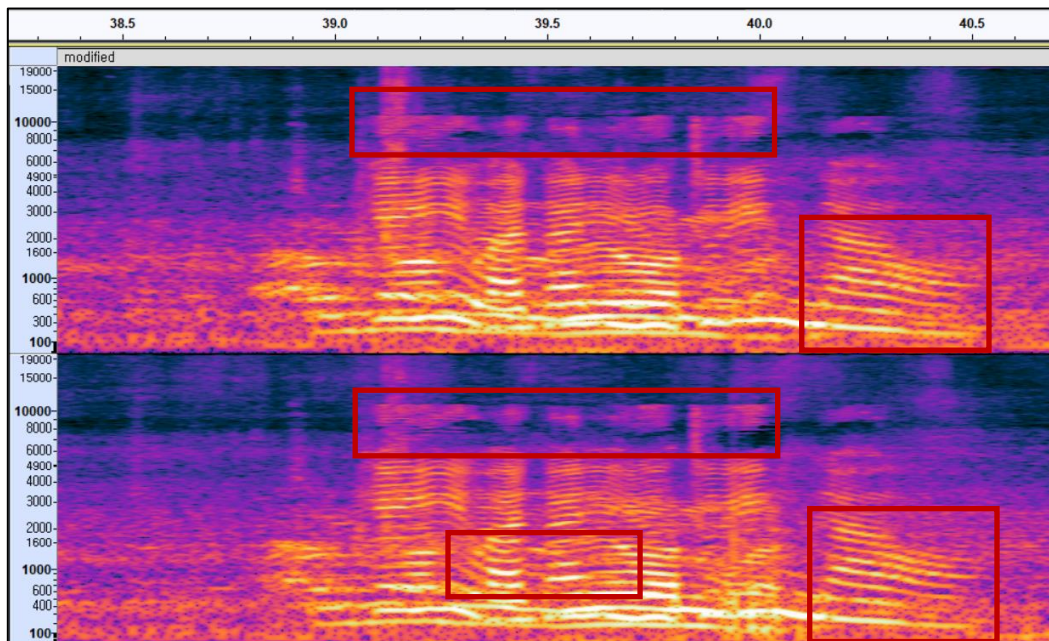
Audacity 도구를 통해 두 wav 파일을 비교 분석해보았습니다. 파형과 스펙트로그램에 대해 개략적으로 보았을 때 1에서 인식한 fake voice의 구간에 대해 차이가 나는 부분이 확실히 보였습니다.

- Original.wav



[그림 4] Original.wav - 스펙트로그램 분석(mel)

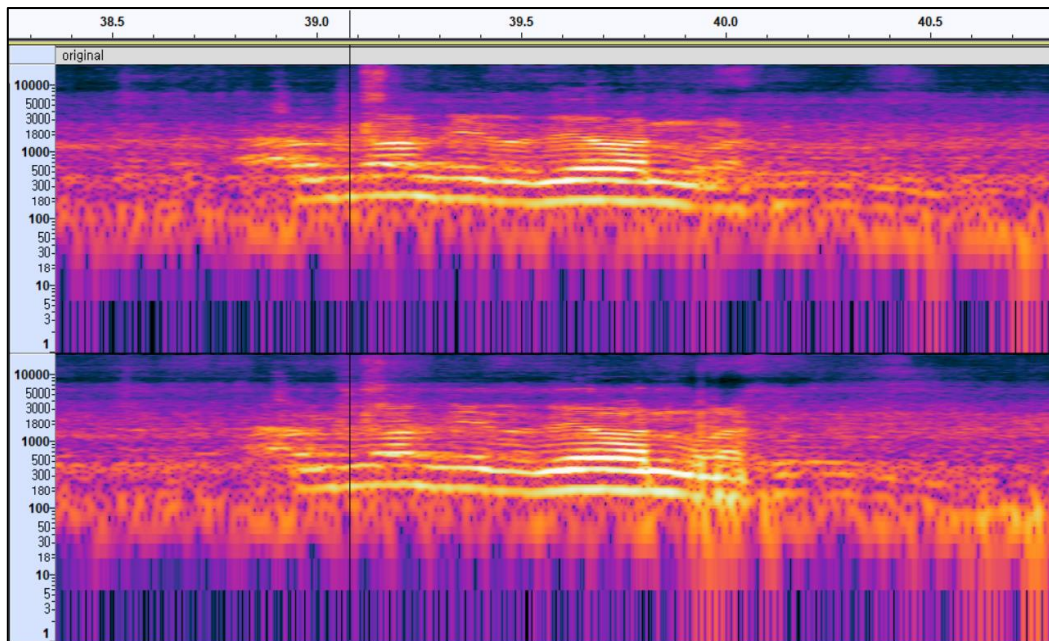
- Modified.wav



[그림 5] Modified.wav - 스펙트로그램 분석(mel)

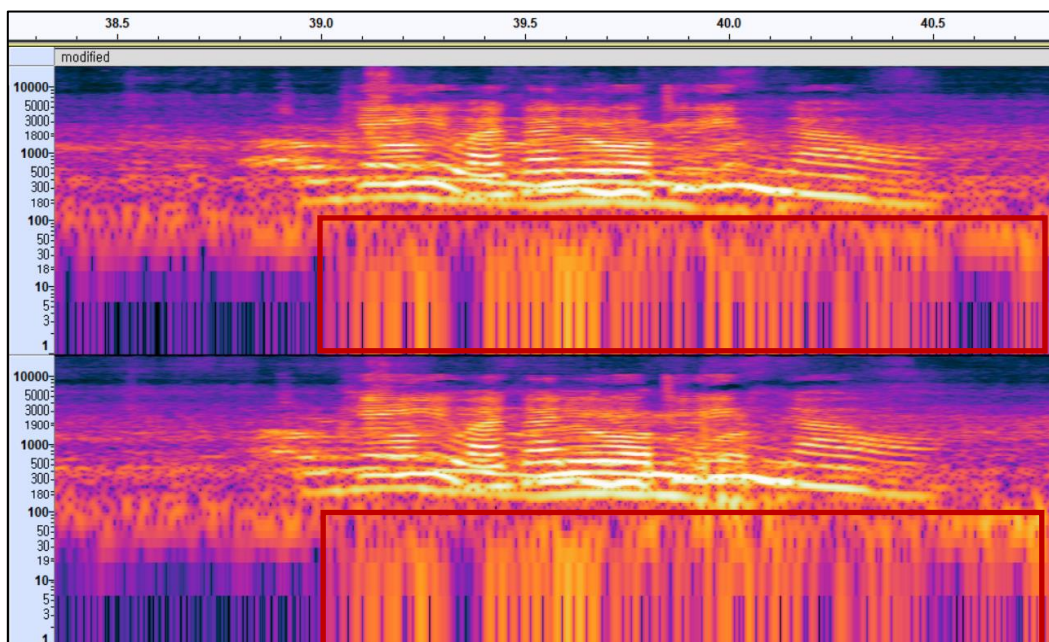
스펙트로그램을 스테레오에서 mel 방식으로 살펴보았을 때, Modified.wav 파일에서 약 39초에서 약 40초 경에 주파수 10000Hz 부분에서 원본 파일과의 진폭 차이를 보였으며, 약 39.2초부터 39.6초까지도 약 600Hz 부터 2000Hz에 해당하는 구간에서 원본 파일에서 보다 더 선명한 진폭을 보였습니다. 또한, 40.1초부터 40.5초 구간까지 수평선의 진폭이 생성된 것을 확인하였습니다.

- Original.wav



[그림 6] Original.wav – 스펙트로그램 분석(Linear)

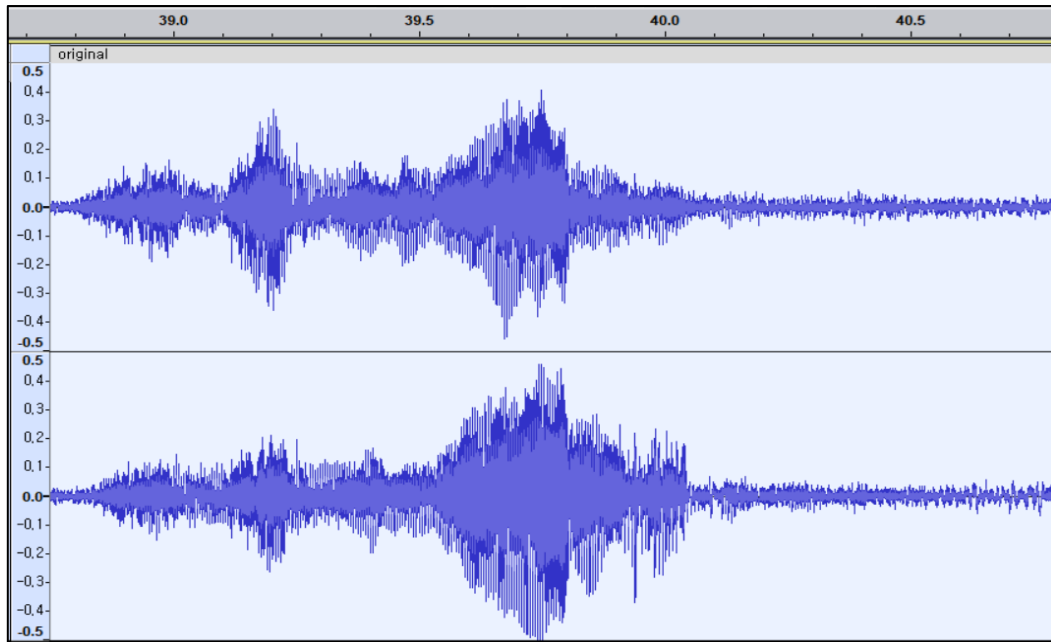
- Modified.wav



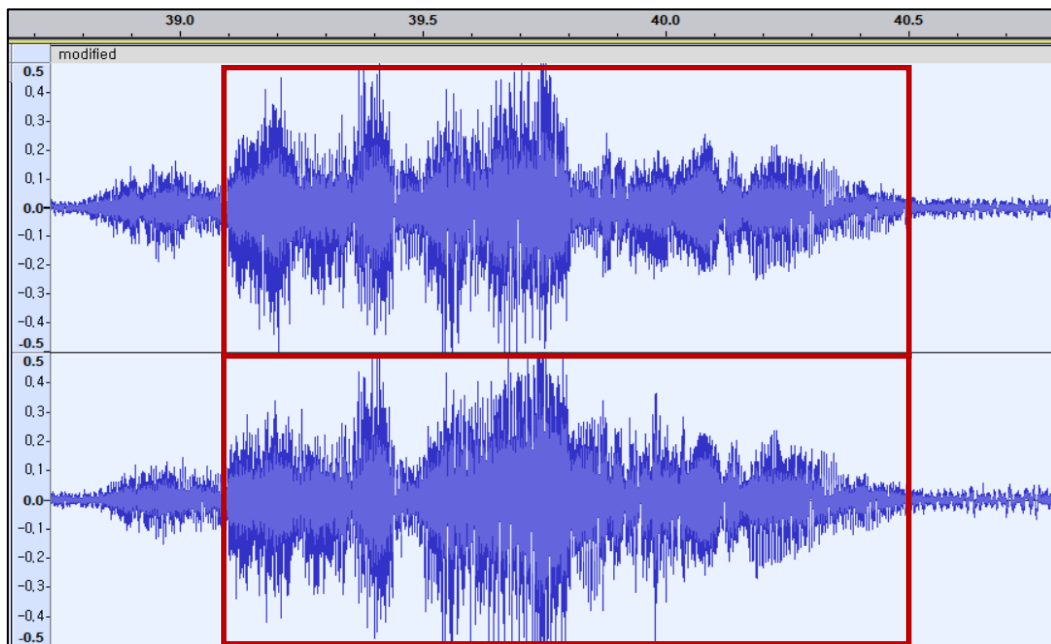
[그림 7] Modified.wav – 스펙트로그램 분석(Linear)

또한, 앞선 mel을 통한 분석과는 다르게 Linear형식으로 스펙트로그램을 살펴보았을 때, 약 39초부터 약 40.7초까지의 시간대와 약 1Hz ~ 약 100Hz에 해당하는 주파수 구간에서 fake voice 중첩에 따른 확연한 진폭 농도 차이를 Original.wav와 비교해서 확인할 수 있었습니다.

- Original.wav



[그림 8] Original.wav – 파형 분석(배율 x2)



[그림 9] Modified.wav – 파형 분석(배율 x2)

또한, 파형 분석도 수행해보았을 때, 변조된 파일에서는 원본파일과 달리 의심되는 약 39.1초 구간부터 40.5초 구간까지 fake voice 중첩으로 인해 진폭의 크기가 더욱 커진 것을 가시성있게 확인할 수 있었습니다. 따라서, 파형과 스펙트로그램 분석을 통해 **modified.wav** 파일에서 약 39초(39.1초)부터 약 40초(40.5초)까지 fake voice가 의도적으로 추가되었음을 판단하였습니다.