디지털 포렌식을 위한 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템 설계

장민희[1], 권기협[2], 이재민[3], 김동성*
금오공과대학교{IT융복합공학과[13*], ICT융합특성화연구센터[2]}
{unimini27[1], navkwon[2], ljmpaul[3], dskim*}@kumoh.ac.kr

Design of Mobile Device Detection System based on Edge Computing For Digital Forensic

Min-Hui Jang[1], Ki-Hyup Kwon[2], Jae-Min Lee[3], Dong-Seong Kim*
{Dept. of IT Convergence Eng.[13*], ICT-CRC[2]}Kumoh National Institude of Technology

요 약

본 논문은 군사시설, 정보기관 및 정부 기관이나 기업의 기술 연구개발 부서 등과 같이 높은 보안을 요구하는 기관에서 모바일 디바이스의 보안검증에 활용 및 판별의 신뢰성을 향상시키기 위한 시스템을 제안한다. 모바일 디바이스의 접근성으로 인한 기업 정보 유출 문제를 방지하고자 높은 보안을 요구하는 기관에서는 출입구에서 보안 스티커 부착 및 보안 애플리케이션의 설치를 권장한다. 본 논문에서는 앞서 말한 과정 대신 DCNN(Deep Convolution Neural Network) 및 SVM(Support Vector Machine)을 적용한 디지털 포렌식을 위한 엣지 컴퓨팅 기반 모바일 디바이스 검출을 위한 시스템을 제안한다. 모바일 디바이스 사용 제한이 필요한 기관의 출입구에 배치되어 동영상을 촬영하고 실시간으로 얼굴을 인식하여 출입 제한 및 녹음되는 전체 사운드에서 모바일 디바이스의 노이즈 데이터를 추출하여 등록된 기기 외의 출입을 제한한다.

Ⅰ. 서론

다수의 기업이 업무에 직원들의 모바일 디바이스 사용을 허용하면서 기업 정보의 유출 위험 부담은 증가하고 있다. 분실과 도난의 위험이 있고 고기능의 대용량 스토리지를 탑재하고 있는 스마트폰의 특징이 그 이유이다. 따라서 일부 기업에서는 보안 정책상 회사 인근에 위치하면 스마트폰의 기능을 제한하는 MDM(Mobile Device Management) 애플리케이션의 설치를 권장한다. 그러나 이와 같은 애플리케이션을 사용하지 않는 외부인 또는 식별 및 허가되지 않은 모바일 디바이스를 모두 통제하거나 검출하기에는 역부족이다. 따라서 기존 연구에서는 동영상의 음성 데이터의 노이즈를 추출하여 모바일 디바이스의 노이즈를 분석하여 판별하는 기법을 제안한다[1, 2]. 이러한 기법은 사용자가 녹화된 동영상에서 직접 사운드 데이터를 추출해야 하므로 실시간 학습 및 판별이 불가능하다. 또한 이러한 과정은 단순 모바일 디바이스의 노이즈 추출 및 판별과정을 위한 기법으로 데이터 관리 및 분배에 대한 고려사항이 부족하다.

본 논문에서는 높은 보안을 요구하는 기관에서 개인 모바일 디바이스의 보안 위협으로부터 보호하기 위하여 엣지에 판단 및 검출 시스템을 배치함으로써 실시간적이고 분산 작업이 가능한 모바일 노이즈 데이터 수집, 학습, 판별, 대응 방안을 제안한다. 해당 시스템은 출입구에 설치되어 DCNN(Deep Convolution Neural Network)을 통한 FR(Face Recognition)과 SVM(Support Vector Machine)을 통한 모바일 디바이스 판별과정을 통하여 기존에 등록되지 않은 기기의 출입을 제한한다[3, 4]. 이러한 엣지 기반 시스템은 데이터를 로컬에서 분산 컴퓨팅 즉, 수집 및 분석하여 처리 및 응답속도를 최소화하고 실행 부하와 지연을 감소시킨다. 또한 전체 데이터가 아닌 중요 데이터만 취급하여 클라우드에 저장함으로서 클라우드 자원 할당 문제를 해결한다. 한편 모바일 기기로 인한 보안 유출 발생 시 데이터베이스에 저장된 검출 데이터는 모바일 디바이스에서 발생하는 순수 노이즈로서, 무결성을 확보할 수 있으므로 디지털 포렌식 과정의 증거 자료 및 역학조사에 활용할 수 있을 것이다.

Ⅱ. 기존의 문제점 분석

기존 연구에서는 녹화된 동영상의 음성 데이터에서 모바일 디바이스의 노이즈 특징 데이터만을 추출하여 다층 신경망에서 학습 및 판별하였다[1, 2]. 또한 사용자의 사용습관 및 사용 환경에 따라 모바일 디바이스의 컨디션이 달라짐을 착안하여 이를 고려한 디바이스 판별에 주요한 목적을 두었다. 그러나 기존 연구는 모바일 디바이스 판별 과정에 필요한 학습 데이터를 연구자가 직접 추출하여야 한다. 즉, 촬영된 동영상을 이용하여 연구자가 각 모바일 디바이스 별 노이즈 데이터를 추출해 학습시켜야 한다. 따라서 촬영된 동영상을 이용하고 학습을 하기 위해서는 시간적 제약이 존재하기 때문에 실시간 모바일 디바이스 노이즈 데이터 처리와 판별이 불가능하다. 또한 이러한 과정은 단순 모바일 디바이스 노이즈 추출 및 판별과정만을 수행하고 있으며, 데이터 관리 및 분배가 불가능하다.

본 논문에서는 엣지에 판단 및 검출 시스템을 배치함으로써 실시간적이고 분산 작업이 가능한 모바일 노이즈 데이터 수집, 학습, 판별, 대응 방안을 제안한다. 이로써 모바일 디바이스 이용 제한이 필요한 기관의 출입구에서 등록되지 않은 디바이스를 감지하여 출입을 통제할 수 있다.

Ⅲ. 제안하는 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템

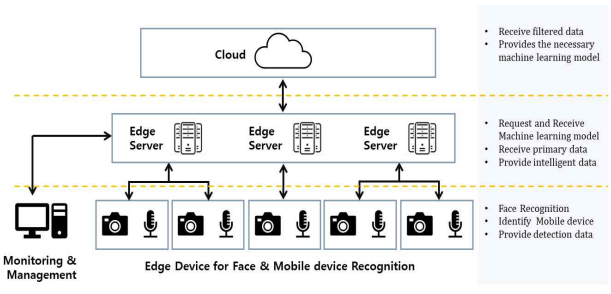

그림 1 제안하는 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템 아키텍처

본 논문에서 제안하는 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템은 모바일 디바이스 사용의 제한이 필요한 장소의 출입구에 배치한다. 배치된 시스템은 동영상을 이용한 얼굴 인식 및 동영상 전체 사운드의 노이즈를 추출하여 등록되지 않은 모바일 디바이스의 출입을 통제하는 것을 목표로 한다. 제안하는 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템 아키텍쳐는 그림 1과 같으며 클라우드, 엣지 서버 및 디바이스, 통합 모니터링 시스템으로 이루어진다.

클라우드는 본 논문이 제안하는 시스템을 사용하고자 하는 기관 관계자의 모든 얼굴 및 각 보유 모바일 디바이스 정보를 저장/관리하고, 건물별로 필요한 데이터를 엣지 서버에 송신한다. 또한 엣지에서 수집, 분석한 얼굴 및 특징 노이즈 데이터를 수신하고 저장한다. 얼굴 인식 및 모바일 디바이스 검출을 위한 엣지 디바이스는 촬영, 분석, 알림, 전송기능을 가지며 건물 혹은 층 별로 위치한 출입구에 설치된다. 먼저 클라우드에서 제공한 데이터를 이용하여 해당 건물 내 관계자를 리스트화 하고 학습한다. 이후 카메라를 이용하여 DCNN을 통한 얼굴 인식 그리고 SVM을 통한 모바일 디바이스 검출 과정을 거친다. DCNN을 통한 얼굴 인식은 사전에 각 관계자 별 얼굴 이미지 데이터를 이용하여 얼굴의 특징을 추출하고 학습한다. 이후 시스템이 설치된 곳을 지나칠 때, 카메라로 촬영한 이미지에서 특징을 추출하여 기존 학습된 데이터와 비교하여 관계자 임을 확인하고, 보유 모바일 디바이스 리스트를 확인한다. SVM을 통한 모바일 디바이스 검출 과정은 4종류의 커널함수(Linear, RBF, Sigmoid, Polynomial)을 적용한다. 디바이스의 판별이 완료되면 해당 관계자가 등록한 모바일 디바이스 모델인지 확인한다. 만약 기존 데이터에 없는 얼굴 이미지 데이터 및 모바일 디바이스 노이즈 데이터가 검출되면 알림을 발생시키고 출입을 제한하고 클라우드와 통합 모니터링 시스템에 해당 데이터를 전송한다. 관계자 및 모바일 디바이스 검출 통합 모니터링 시스템은 각 건물에 위치한 엣지 디바이스의 네트워크 연결 관리, 각 출입구의 모니터링 기능을 제공한다. 해당 시스템은 LoRa 등의 네트워크 프로토콜을 통해 각 엣지 디바이스에서 데이터를 수신한다.

3.1. 모바일 디바이스 별 특징 노이즈 추출

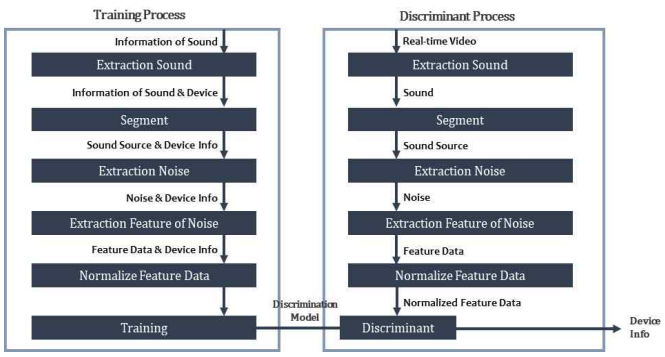

그림 2 모바일 디바이스 노이즈 추출 및 검출 과정

모바일 디바이스 노이즈 추출 및 검출 과정은 그림 2와 같다. Segment 단계에서는 동영상의 사운드 추출 및 사운드 세그먼트. 노이즈 추출을 위해 동영상의 전체 사운드를 추출하여 실시간으로 분석 시스템에 전송한다. 이후 영상에서 추출한 사운드를 세그먼트 한다. Extraction Noise 단계에서는, 위너 필터를 사용한 노이즈 추출. 위너 필터를 사용하여 촬영 영상의 전체 사운드에서 발생한 노이즈 중 모바일 디바이스에서 발생하는 노이즈를 추출한다. 이러한 데이터는 디바이스별 특징 노이즈를 분석하는 데 사용한다. Extraction Feature of Noise 단계에서는, 노이즈 특징 벡터 추출. 노이즈의 특징 벡터를 추출하기 위해 다양한 사운드 특징을 제공하는 MIRtoolbox를 사용한다. 위너 필터를 사용하여 추출한 노이즈를 MIRtoolbox의 1차 분류 특징인 Dynamic, Timbre, Rhythm, Tonality, Spectrum, Fluctuation의 노이즈 특징 데이터 SET을 추출한다. Normalize Feature data 단계에서는, MIRtoolbox를 통해 추출한 노이즈 특징 데이터들의 분포가 광범위하게 분산되어 있을 때 판별률이 낮아짐을 고려하여 수집한 특징 SET을 0부터 +1 범위에서 정규화하여 특징 벡터를 획득한다. 이후 디바이스를 판별한다. 디바이스의 판별 단계에서는 커널 Linear, RBF, Sigmoid, Polynomial을 활용한다. 데이터들의 분포 형태가 비선형적이라는 것을 고려해 4종류의 커널함수를 적용한 SVM을 사용하여 정규화된 특징 데이터들을 고차원 평면 공간에 사상하여 초평면함수로 데이트들을 분류하여 학습하고 판별한다.

Ⅵ. 결론

본 논문에서는 높은 보안을 요구하는 기관에서 개인 모바일 디바이스의 보안 위협으로부터 보호하기 위한 엣지 컴퓨팅 기반 모바일 디바이스 검출 시스템을 제안한다. 모바일 디바이스 검출은 클라우드에 등록한 관계자 얼굴 이미지 데이터와 보유 모바일 기기 정보를 활용한다. 이후 출입구에 설치된 엣지 디바이스는 DCNN을 이용한 얼굴 인식과 SVM을 이용한 모바일 디바이스 노이즈 추출을 통하여 관계자 별 등록되지 않은 모바일 디바이스의 출입을 제한한다. 이러한 엣지 시스템 구조는 데이터를 로컬에서 분산 컴퓨팅하여 처리 속도를 최소화하고 시행 부하와 지연을 감소할 수 있다. 또한 중요 데이터만 추출하여 클라우드에 저장함으로써 Data Storage 할당 문제를 해결한다. 그리고 모바일 기기로 인한 보안 유출 발생 시 클라우드에 저장된 검출 데이터를 디지털 포렌식 절차에서 활용할 수 있을 것이다.

향후 연구로는 디지털 포렌식 절차에서 무결성을 보장하기 위하여 수집한 데이터의 변조 및 손상이 일어나지 않음을 보증하는 방안을 고려할 것이다. 또한 시스템의 보안성을 향상 시키기 위하여 암호화 방안을 고려해볼 것이다.

ACKNOWLEDGMENT

본 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 대학중점연구소 지원사업으로 수행된 연구임(2018R1A6A1A03024003).

참 고 문 헌

[1] G.H. Kwon, C.B. Moon, J.M. LEE, D.S. Kim, “Identification Method of Military Mobile Device Using for C3I”, International Conference on Information and Communication Technology Convergence (ICTC), Oct. 2019

[2] C.B. Moon, H.S. Kim, M.H. Yi, H.A. Lee, B.M. Lim, “Audio recorder identification using noise property”, Information Journal Japan, Vol. 16, No. 11, pp. 8261-8269, Nov. 2013

[3] Meng Shen, Jie Zhang, Liehuang Zhu, Ke Xu, Xiangyun Tang, “Secure SVM Training Over Vertically-Partitioned Datasets Using Consortium Blockchain for Vehicular Social Networks”, IEEE Transactions on Vehicular Technology, Vol. 69, pp. 5773-5783, Dec. 2019

[4] An-Ping Song, Qian Hu, Xue-Hai Ding, Xin-Yi Di, Zi-Heng Song, “Similar Face Recognition Using the IE-CNN Model”, IEEE Access, Vol. 8, pp. 45244-45253, Mar. 2020