

문제정의서(연구계획서)

과제명	무선 AP 펌웨어수정을 통한 데이터 수집 모듈 개발
-----	------------------------------

조	졸업할 수 있을까 조
지도교수	원유재 교수님 (서명)
조원	201502038 김정우 201502006 강영균 201502032 김승기

1. 연구의 필요성

현재 단말기의 위치 추적을 하기 위해서는 cell방식의 위치추위 기술을 사용하고 있다. 이동사의 전파기지국을 기반으로 위치값을 측정하는 방식이다. 철저히 이동사 의존적이며, 200m~1,500m라는 큰 오차범위를 갖고 있다. 그럼에도 널리 사용되고 있는데는 실내에서도 측위가 가능하고, 단말기에 별도의 하드웨어적 소프트웨어적 추가없이 이용이 가능하기 때문이다. 또한 GPS를 통한 광역추적방식이 존재한다. Gps는 Accuracy가 일반적으로 뛰어나지만, GPS 위성 신호가 도달하지 못하는 건물 안과 같은 Indoor에서는 위치 계산이 불가능하다. 또한 단말에 GPS Chip이 내장되어야 하므로 단말기 에 대한 Impact가 크고, GPS 위성 Searching에 많은 소간이 소요된다. 특히, Cold start일 경우에는 이론적으로 12.5 분이 소요된다는 단점이 존재한다.

그러나 이와 구분되는 단말기 식별방식이 있다. 바로 무선 액세스 포인트 이며, 현대사회에는 인터넷 사용자가 무선 액세스를 사용하는 경향이 증가하고 있다. 무선 네트워크의 경우, IP 주소는 종종 사용자에게 동적으로 할당되기 때문에 조사자가 고유하고 일관된 태그로 완전히 신뢰할 수 없는 경우가 있다. 따라서, 개방형 무선 AP의 편재성은 사람들이 익명으로 행동 할 수 있는 새로운 길을 만들었다. 특히 많은 사업체와 거주지에 개방형 액세스 포인트가 있으며 이를 통해서 여러 무선 단말기들이 스캐닝(식별) 된다. 그러나 유선 네트워크에 비해 무선 네트워크는 주위 환경에 민감하고, AP에 진입하는 디바이스를 식별하기 위해서는 정밀한 값을 계산하기 위한 고가의 무선 신호 수집 장비가 필요하다. 또한 무선 통신 디바이스의 기초 식별자인 SSID와 MAC address는 외부에 쉽게 노출될 수 있다.

이를 위해 기계학습기반의 디바이스 고유의 지문을 feature로 하여 인공지능 기반의 오차에 강한 AP 및 단말 식별 모델을 구축하고, 식별된 단말기가 파악된 AP의 신호 세기와 측위 시간을 바탕으로 하여 기존 위치 측위 기술과 다른 사용자 단말기의 위치추적 기술을 연구자료가 있다. 서버에 기록된 단말기들의 정보를 머신러닝 기술을 통해 확인하고, 수집하여 해당 위치, 시간을 설정하여 단말기들의 위치와 특정 단말기를 지정하여 단말기의 이동 경로를 추적할 수 있도록 오픈소스 기반으로 무선AP에서 서버로 데이터를 전송하는 데이터 수집 모듈을 만들고자 한다.

2. 연구의 목표 및 내용

본 연구의 궁극적인 목표는 과제의 목표 사용자 위치 추적 기술 연구 등에 필요한 무선신호를 수집하는 것이다. 무선 AP의 펌웨어를 수정해서 원하는 무선신호 패킷을 수집하고, 서버로 전송한다. 수집된 무선신호 패킷을 분석해서 단말기를 인식하고 해당 단말기와 무선 AP와 통신 되는 신호 세기를 통해서 단말기 사용자의 이동 경로를 추적한다. 이를 위한 무선 AP신호 수집 소프트웨어와 무선 스캐닝 패킷을 수집하고 분석하고 시각화 해줄, 웹 어플리케이션과 통신할 서버를 구축한다. 무선 신호를 수집하는데 있어서 무선AP와 단말기가 통신하는 능동, 수동 스캐닝방식에 대한 이해가 우선적이다. 비콘프레임, 프로브패킷에 대한 분석을 하고, 이에 대한 정보를 기반으로 정확하고 빠르게 무선신호를 수집, 구축된 서버로 전송하는 것이 목표이다. 웹 어플리케이션의 높은 성능을 위해 실시간 데이터 처리와 정확성을 가장 중요한 성능 지표로 하여 소프트웨어를 구현한다.

목표에 대한 구체적인 방안은 다음과 같이 나타낼 수 있다. 우선 프로그래밍이 가능한 무선공유기를 통해서 단말기가 스캐닝하는 방식에 대한 이해가 필요하다. 스캐닝되는 전체 네트워크 패킷 프레임이 확인이 되면, 전체 각 패킷 항목에 대한 파악을 해야하고 이는 비콘프레임,프로브리스폰,프로브리퀘스트 이렇게 3개 종류의 패킷이다. 각 패킷에 대해서 구조,길이,비트에 대한 조사를 하고 각 패킷의 항목들의 기능을 분석해서 활용될 수 있는 방안을 마련한다. 이를 바탕으로 수집된 데이터를 전처리하여 서버로 전송한다.

프로젝트의 실패기준은 다음과 같다. 단말기 인식 실패 및 낮은 정확도, 무선AP 정보 기록 및 서버 전송 실패, 전송된 단말기 이동경로 추적 실패 이렇게 3가지 사항에 부합하면, 프로젝트 실패로 간주한다.

해당 프로젝트의 활용방안의 예제이다. 각종 전염병(COVID-19)에 대한 감염자 이동경로 추적, 위조 단말기 기반 네트워크공격자 이동경로 추적, 특정 단말기에 대한 이동경로 추적, 무선AP 정보를 통한 사용자 밀집지역 파악 이렇게 4가지를 뽑을 수 있다.

개발된 모듈은 독립성을 가지고 무선AP에 연동되어 다양한 어플리케이션 개발에 이용될 수 있다. 단말기에 대한 위치 추적을 가능할 뿐 아니라 다양한 패킷에 대한 수집을 진행할 수 있다.

무선 통신 디바이스 식별 시스템을 구축하기 위해서는 디바이스 정보를 식별에 필요한 고유지문을 추출해야한다. 구축된 시스템은 Wi-Fi 통신의 핵심 구성 요소인 AP 와 무선 단말의 존재를 인식하고 디바이스 고유 지문을 통해 무선 통신 디바이스를 구별 한다. AP를 식별하기 위한 지문은 SSID, MAC address, Clock skew, Channel, RSS, Duration 이며 기계학습 분류 알고리즘을 이용하여 학습하고 분류한다. 무선 단말을 식별하기 위한 지문은 MAC address, Sequence Number 의 증가율, 패킷의 길이가 있다. 해당 정보는 수집되는 비콘프레임 혹은 프로브패킷에서 가져오며 이에 대한 이해를 해야한다. 수집된 정보는 리눅스 기반의 실시간 서버로 전송된다. 전송된 정보를 관리하기 위한 시스템을 구축해야 하며, 클라우드의 형태로 서버를 이용가능하게 하는 것이 목표이다. 해당 서버를 기반으로 하여 이를 가지고 웹 어플리케이션 구축을 한다.

지도 교수님 : 원유재
연구진 :
김정우 201502038
강영균 201502006
김승기 201502032

[illegible]

- 참고문헌(Reference)

[1] 서정훈, “기계학습을 이용한 무선 통신 디바이스 식별 시스템”(2020)

[2] Wireless Device Identification Based on RF Oscillator Imperfections Adam C. Polak and Dennis L. Goeckel, *Fellow, IEEE*.
