

유저 이동경로 유사도 분석을 통한 매칭 서비스

고병후, 조진성(지도교수)
경희대학교 컴퓨터공학과
bhwkd@naver.com, chojs@khu.ac.kr

Matching Service based on Trajectory Similarity Analysis

Byeonghu Go, Prof. Jinsung Cho(Advisor)
Department of Computer Science and Engineering, KyungHee University

요 약

현재 스마트폰은 대중들의 일상 대부분을 함께하는 도구가 되었다. 이에 따라 mobile device들은 사용자의 일상 생활을 가리키는 지표가 될 수 있는 데이터들을 다수 저장한다. 본 연구에서는 모바일 디바이스를 이용하여 이러한 데이터를 수집하고 정제하고 가공하여 유의미한 information을 만들어 사용자에게 서비스의 형태로 제공되는 일련의 과정을 소개한다. 이를 바탕으로 스마트폰을 활용해 gps, wifi연결시간, 선형가속벡터값의 data를 수집하여 일상생활속 이동패턴을 파악하고 이를 통해 유사한 패턴을 보이는 user를 상호 매칭하는 서비스를 기획하고 설계한다.

1. 서 론

현대사회에서 대부분의 사람들에게 스마트폰은 필수적인 도구가 되었다. 대중들은 mobile device와 일상에서 끊임없이 상호작용한다. 이 결과 스마트 폰은 user의 일상 생활과 관련된 많은 data를 생산하고 저장하게 되었다. 해당 data들을 수집하고 가공하여 사용자의 일상의 패턴을 파악하는 유용한 information을 추출할 수 있다면 이를 여러 서비스의 형태로 확장함으로써 사용자 편의를 제공할 수 있다.

본 논문에서는 특히 LBS(Location Based Service)에 초점을 맞추어 연구를 진행한다. 대부분의 사람들은 출퇴근, 등하교 등 일상 속에서 다소 반복적인 이동패턴을 보인다. 이런 이동패턴을 나타낼 수 있는 data(gps, wifi연결시간, 선형가속벡터)를 mobile device를 이용해 수집하고 가공한다. 이를 통해 중앙서버에서 사용자 이동패턴을 분석하여 유사한 이동패턴을 보이는 user를 상호 매칭하는 service를 설계한다. 이를 통해 data를 수집하여 정제하고 가공하여 유의미한 정보를 만들어 사용자에게 서비스의 형태로 제공되는 일련의 과정을 소개하고 제시한다.

본 논문은 2절에서 관련된 기술과 선행된 연구를 조사한다. 이를 바탕으로 3절에서 서비스의 구체적인 설계와 방법을 제시하고 4절에서는 결론 및 향후 연구 방향에 대해 기술한다.

2. 관련 연구

2.1 Android& Android Studio

Android란 주로 Mobile Device를 대상으로 하는 linux kernel기반 운영체제이다. application 실행시 VM(virtual machine)인 Android RunTime을 통해 별도의 process에서 수행하도록 한다는 특징이있다. 이러한 Android는 현재 스마트폰 시장에서 높은 점유율을 보이고 있으며 application개발을 비교적 유연하게 제공해 주기에 해당 프로젝트의 플랫폼을 Android로 선정하였다.[1]

Android Studio란 Android application개발을 위한 공식IDE(Integrated Development Environment)이다. 여러 SDK(Software Development Kit)와 API를 지원하여 개발이 용이하기에 본 프로젝트에서는 android application 개발을 위해 이를 사용한다.[2]

2.2 GPS

GPS(Global Positioning System)는 위성을 통해 위치(Positioning) 및 항법(Navigation), 시각(Timing) 정보(PNT)를 제공해주는 위성 항법시스템이다.

GPS는 4개 이상의 인공위성으로부터 수신한 Radio 신호를 이용하여 정지 및 운동하고 있는 물체의 위치를 결정한다.

기본적으로 약 10m의 오차범위를 가지며 실내에서는 그 범위가 커진다는 한계점이 있다.[3] 본 연구에서는 사용자의 위치 좌표를 측정하기 위해 gps를 이용한다.

2.3 선형가속 센서

현대의 일반적인 Mobile device는 여러가지 동작센서를 지원한다. 동작센서란 기기의 동작을 모니터링 할 수 있는

센서이다. 센서의 종류는 센서값 도출의 기반에 따라 분류한다. hardware 기반 센서와 이를 바탕으로 가공한 데이터인 software기반의 센서가 있다. 본 연구에서는 흔들림을 감지하기 위해 중력 값을 제거하고 순수하게 기기에 가해진 가속도만 측정하는 software기반의 선형가속 센서를 이용한다.

해당 프로젝트에서는 선형가속센서 관련 Android studio API인 TYPE_LINEAR_ACCELERATION를 이용하여 기기에 가해진 힘(흔들림)을 측정한다.[4]

3. 시나리오 제안 및 구현

3.1 서비스 시나리오

Mobile Device로 data를 수집하여 이를 바탕으로 이동경로 패턴을 도출한다. 해당 이동경로 패턴을 이용해 타인과 유사도를 검사하여 가장 높은 유사도를 보이는 user를 매칭하는 서비스를 기획하고 구현한다.

이때 수집되는 data는 GPS좌표(위도, 경도), 선형가속벡터 값(x,y,z값), wifi연결시간(연결시간, 해제시간) 이다. 해당 데이터를 Mobile device에서 background program을 통해 지속적으로 수집하여 server로 전송한다. server에서는 수집된 raw data를 정제, 가공하여 이동패턴이라는 유의미한 information을 생성한다. 이동패턴에는 이동 궤적, 정지장소, 이동수단 등이 포함된다. 해당 이동패턴을 다른 user와 상호비교하여 유사도를 도출하고 해당 유사도가 가장 높은 유저를 추천하여 상호 매칭한다.

각 단말의 동작 시나리오는 아래와 같다

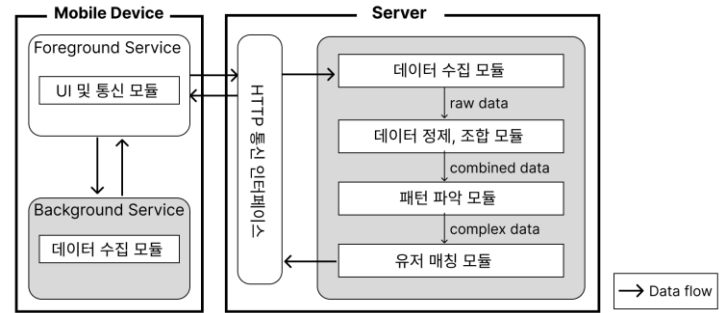
◦ Mobile Device

background에서 지속적으로 data를 수집한다. 해당 data는 server로 전송되고 local storage에서는 삭제된다.

foreground에서 로그인을 비롯한 UI가 출력된다. user의 도출된 이동경로 패턴 결과와 추천받은 유사한 이동경로의 user와 매칭여부를 결정한다.

◦ Server

Mobile Device로부터 전달받은 data를 정제, 가공한다. 또한 정제된 data를 바탕으로 이동경로 data를 도출하고 해당 data를 바탕으로 user간 유사도를 연산한다. 해당 유사도를 바탕으로 user를 상호 매칭하여 결과를 Mobile단으로 전송한다.



[그림1] Architecture overview

3.2 시스템 설계

시스템은 그림1과 같이 구성된다. 해당 절에서는 data flow를 중심으로 data가 처리되는 과정을 기술하며 시스템의 전체적인 구조와 흐름을 제시한다.

3.2.1 데이터 수집

user_id	time	latitude	longitude
test_user_01	2023-04-19 07:04:17	37.2459902	127.076736
test_user_01	2023-04-19 07:09:17	37.2482411	127.0781278
test_user_01	2023-04-19 07:14:17	37.2509739	127.0792972
test_user_01	2023-04-19 07:19:17	37.2538553	127.0753253

acceleration_x	acceleration_y	acceleration_z	onwifi
3.37854	0.00340271	-3.41475	0
8.37893	0.0044108	-1.41443	0
0.287939	1.47741	0.41453	0
0.107279	0.164454	-3.41453	0

[그림2] Data sample

Mobile Device의 background service에서 일정주기로 data를 수집한다. 기본적으로 정해진시간 t 마다 주기적으로 data를 수집하며 아래 기술한 특정 data가 수집되는 조건발생시에도 수집을 진행한다.

- ① GPS좌표(위도,경도) : 시간 t 경과마다 수집.
- ② 선형가속벡터(x,y,z) : 시간 t 경과마다 수집.
- ③ wifi접속 여부 : 시간 t 경과시마다 수집. 혹은 wifi 연결, 해제 이벤트 발생시에 수집.

하나의 data가 수집되는 조건이 발생하면 다른 data들도 동시에 수집되어 저장된다. 그림2의 test에서는 t 를 5분으로 설정하였다.

해당 data는 sever로 전달되어 그림 2와 같이 DataBase에 저장된다. 이 Raw Data는 data 정제 단계로 넘어간다.

3.2.1 데이터 정제

Mobile Device의 sensor로부터 수집된 raw data는 noise를 내포하기에 filtering 과정을 거쳐야 한다.

◦ GPS좌표(위도,경도)

① location filter : gps는 일정정도의 오류를 내포하며 실내에서는 오류의 정도가 커짐으로 이를 wifi 연결로 인해 확인되는 위치 정보 등으로 보정한다. sensor detect와 동시에 이루어져야 한다.

② vector filter : 수집된 data를 시간을 기준으로 정렬하면 인접한 index사이에 gps좌표간 euclidean distance를 구할 수 있고 이를 해당 index사이의 시간 차로 나누어 속도를 구한다. 해당 속도가 지나치게 큰 값이라면 해당 data를 error로 간주하여 삭제한다.[5]

◦ 선형가속벡터(x,y,z)

① absolute filter : 가속도의 값이 지나치게 큰 값으로 측정된 data는 오류로 간주하여 인접한 시간의 vector의 산술평균으로 값을 바꾼다.

이렇게 정제된 data는 이동경로라는 정보로 도출되고자 다음 단계로 넘어간다.

3.2.1 데이터 조합

오류가 보정된 raw data를 바탕으로 이동경로라는 data를 도출한다. 이동경로 data는 일종의 graph로 구성된다. vertex는 상태(정지 혹은 이동)와 머문시간(도달시간, 출발시간)으로 구성된다. edge는 이동수단이 저장된다. 해당 data를 도출하는 방법은 아래와 같다.

① 시간으로 정렬했을 때 인접한 gps간의 (거리/시간)으로 도출한 속도가 일정 값 STOP_DECISION 이하인 연속한 data들을 합쳐서 하나의 vertex를 구성한다. 상태는 정지로 판단된다. 합쳐진 data들의 시간 값의 배열을 T 라고 했을 때 도달시간은 $\min(T)$ 로 출발시간은 $\max(T)$ 로 구성된다.[6] wifi연결이 True인 data도 정지로 판단한다.

② 인접한 data간 속도가 STOP_DECISION이상인 data는 이전의 data와 합쳐지지 않고 독립적으로 구성되어 상태는 이동중으로 저장된다.

③ 위의 ①, ②과정이 끝난 vertex간의 가속벡터(흔들림), 속도에 따라 decision value를 두어 도보인지 자동차인지 지하철인지 판단하여 edge에 이동수단을 저장한다.

위의 방식으로 도출된 이동경로 data는 전체 데이터가 연결되어 있다. 그러므로 유사한 위치 값을 갖는 정지상태 vertex를 기준으로 회귀주기가 반복됨을 판단한다면 해당 주기 별로 데이터를 자르거나 혹은 24시간을 기준으로 자를 수 있다.

이렇게 구성된 combined data는 패턴파악 모듈로 전달된다.[7]

3.2.1 패턴 파악

user마다 이동경로 data가 복수 개 저장 되어있다. 일반적으로 요일마다 이동패턴이 반복됨으로 요일 별로 이동경로 data를 분류한다. 가장 빠른 시간의 정지vertex를 출발지로 가장 느린시간의 정지vertex를 도착지로 판단한다. 요일별 이동경로 데이터들에 대해 정지부터 도착까지 vertex간의 좌표의 평균을 구해 요일마다 하나의 대표 이동경로 패턴을 도출한다.

3.2.1 의미 있는 정보 추출

user마다 갖는 대표이동경로(패턴)을 바탕으로 상호간에 유사도를 측정한다. 유사도 측정 방식은 rank를 계산하는 것이다. 출발, 도착지의 거리가 가까운지, 이동시간이 유사한지, 이동수단이 같은지에 대해 각각 점수를 배정해 해당 점수에 따라 가장 큰 유사도를 갖는 user를 찾는다.

4. 결론 및 향후 연구

본 논문에서는 Mobile device를 통해 user의 data를 수집하고 이를 이용하여 이동경로라는 information을 도출해 해당 이동경로가 유사한 user를 매칭해 주는 서비스를 제안하였다. 해당 기능은 확장하여 차량 공유 서비스, 택시비 절감 등 여러 유용한 응용이 가능하다. 향후에는 해당 기능을 바탕으로 보다 생활에 직접적이고 유용한 서비스로의 구현을 수행할 것이다.

참고 문헌

- [1] <https://developer.android.com/guide/platform>
- [2] <https://developer.android.com/studio/intro>
- [3] 한진희, “GPS 수신기 설계 기술 및 동향”, 2007
- [4] <https://developer.android.com/guide/topics/sensors/>
- [5] 양승국, 백혜정, 김제민, 전명중, 박영택 “개인 경로 학습을 위한 GPS 좌표 기반 이동 궤적 추출”, 2011
- [6] 김선용, 박범준, 정재진 “모바일장치의 GPS를 이용한 사용자 이동경로 분석 및 이동경로 추천 시스템”, 2011.2
- [7] 양혜진 “모바일의 위치데이터를 활용한 사용자 조사 방법”, 2015.5