

모바일의 위치데이터를 활용한 사용자 조사 방법

User research method by utilizing the location data of the mobile

양혜진

국민대학교 테크노디자인전문대학원

Yang hye-jin

Graduate School of Techno Design, Kookmin University

요약

개인의 모바일에서 자동으로 축적되는 데이터의 종류와 유형이 점점 더 다양화되고 있다. 본 논문에서는 UX를 위한 사용자 조사를 진행할 때, 사용자의 모바일에 자동으로 기록되는 '위치데이터'를 활용하는 방법론을 제시한다. 사용자 조사의 양과 질적인 효과를 증대시키기 위하여 조사자가 수행해야하는 과정에 대해 살펴보았으며, 주로 사용자 조사의 수행 주체가 되는 콘텐츠스(UX) 디자이너를 위한 접근으로서 간편하고 쉽게 실행할 수 있는 방안에 중점을 두었다. 이는 정성적 조사가 지니는 한계를 일정부분 보완해줄 수 있었으며, 기존의 방법과 병행했을때 그 효과가 두드러지는 것으로 나타났다.

I. 서론

사용자 조사방법은 크게 정성 조사와 정량 조사로 나눌 수 있는데, 정성 조사는 주로 소수의 표본을 대상으로 심층적인 내용을 이끌어 내는 방법으로 인간의 행동과 사건의 의미에 대해 해석을 통해 이해하고 접근하기에 유용하다. 하지만 사용자가 직접 의사표현을 하는 경우에 사용자의 주관적인 느낌과 의견을 토대로 작성되고, 조사자 또는 조사대상자에 의해 변수(noise) 발생 여지가 많기 때문에 데이터의 왜곡가능성이 높아진다. 반대로 정량 조사는 적은 시간동안 다수의 표본을 조사할 수 있지만 조사대상자가 응답하는 과정에서 발생하는 변수를 잡아내거나 거짓이나 성의 없는 응답을 판별하기 어려우며, 문항 자체의 객관성이 결여된 경우 조사결과에 절대적인 영향을 미치게 된다. 이처럼 두 가지 방법은 각각 장단점을 지니고 있으며, 사용자에게 대부분을 의존하는 방식으로 수행된다.

본 연구에서는 기존의 사용자 조사방법이 지니는 한계점을 보완하기 위해 자동으로 기록되는 데이터를 수집하여, 조사의 객관성을 높이고 풍부한 정보를 확보할 수 있도록 스마트폰의 로그데이터를 활용하는 방법을 강구하였다. [1]

II. 스마트폰 내(內)의 데이터

1. 스마트폰의 데이터 수집 방법

개인의 스마트폰 위치데이터와 같이 디바이스에서 자체적으로 발생하는 내부 데이터(internal data)를 추출하는 방법이 시스템에 따라 달라지므로 본 연구에서는 시

장점유율이 가장 높으며, 데이터가 오픈되어 있는 안드로이드 OS를 대상으로 진행하였다. 안드로이드 OS에서 데이터 수집 방법은 4가지가 있다.[2]

표 1. 안드로이드 OS의 모바일 데이터 수집 방법 (안영권, 2011)

방법	설명
SD카드 분석	SD카드에 저장되는 데이터 분석
논리적 분석	안드로이드 애플리케이션을 이용한 분석
물리적 분석	리눅스의 dd 명령어를 이용하여 NAND 메모리의 덤프 이미지 분석
Chip-off	플래쉬 메모리 분리 후, 하드웨어적으로 분석

2. 스마트폰 내의 데이터 유형

스마트폰에는 두 가지 유형의 데이터가 생성된다. 첫째는 사용자가 직접 생성하고 기록하는 데이터로 캘린더의 일정, 다이어리, 가계부 등이 이에 해당된다. 둘째는 스마트폰 이용 과정에서 자동으로 생성되는 데이터로 이는 다시 휘발성과 비휘발성으로 분류할 수 있다.

표 2. 자동 생성형 데이터의 분류

분류	설명
휘발성	자동으로 생성되지만 기록되지 않고 사라지는 데이터 (ex. 접근flow, 터치 정보 등)
비휘발성	정량적인 기록 형태로 누적되어 수집가능한 데이터 (ex. 위치정보, 사용량, 조작시간, 체류 시간 등)

휘발성 데이터는 주로 태스크가 수행 과정에서 발생하는 flow 또는 화면 밖에서 수행되는 인터랙션에 관한 정보들이 많다. 반면 비휘발성 데이터는 과정보다는 결과로 인해 발생하는 정보가 주를 이룬다. 본 연구에서 다루고자 하는 위치데이터는 비휘발성 데이터에 속하는데, 이는 언뜻 보기에는 '과정' 같아 보이지만 정해진 시간에 위치하고 있는 곳의 정보 '결과'를 기록하기 때문이다.

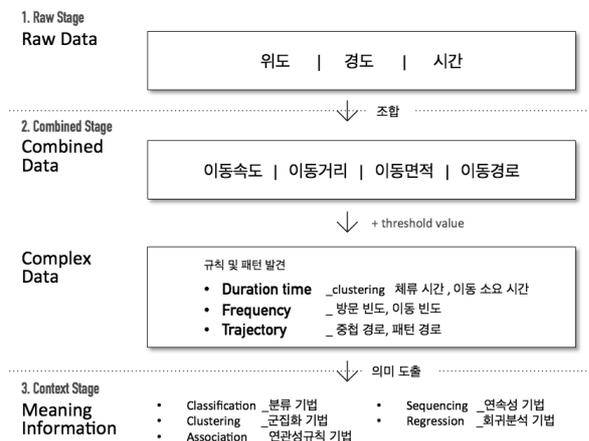
Ⅲ. 위치데이터

1. 위치 데이터 수집 방법

누구나 간단한 설정을 거쳐 접근할 수 있도록 구글의 위치기록 서비스를 활용하였다. 먼저 안드로이드에서 위치관련 설정을 활성화하고 조사기간 동안 유지한다. 기록된 위치정보는 'maps.google.com/locationhistory'에서 확인할 수 있으며, 'kml로 내보내기'를 하여 추출 할 수 있다. kml파일에는 '날짜', '시간', '위도', '경도', '고도' 정보를 포함한다. [3]

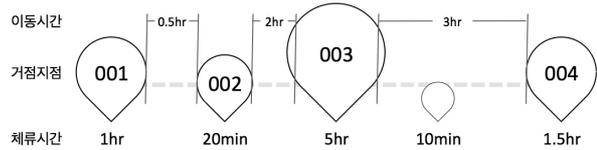
2. 위치 데이터 분석 프레임

위치데이터를 분석하는 과정을 3단계로 구조화하여 프레임 구성하였다. 먼저 Raw stage에서는 가공되지 않은 원시 데이터를 수집한다. 그리고 Combined Stage에서는 수집된 원시 데이터를 조합하여 구체적인 데이터를 얻을 수 있도록 1차 가공을 하고, 임계값을 입력하여 규칙 및 패턴을 발견할 수 있는 새로운 정보를 추출한다. 이 정보는 무의미한 데이터를 정보화하는 역할을 하며, 조사 목적에 따라 임계값의 기준과 데이터의 양이 달라질 수 있다. 마지막 Context Stage에서 해석을 통해 실질적으로 활용할 수 있는 정보를 뽑아낸다. 각 단계를 실체화 시키면 그림1 과 같다.



▶▶ 그림 1. 데이터 분석 프레임 구조

체류시간과 이동시간을 구분하기 위해서 군집화(Clustering)를 하여 거점이 되는 지점 추출하게 되는데 데이터 분석 과정에서 임계값을 적용 한 후 무의미한 데이터를 제거한다. 이렇게 추출된 체류시간을 기점으로 체류 종료시간과 다음 장소의 체류 시작시간을 계산하여 이동시간을 유추한다.



▶▶ 그림 2. 위치정보 코드화

이때 위치 정보를 코드화하여 패턴을 발견하기 쉽도록 할수 있는데, 패턴유형으로는 장소와 장소사이에 연관된 규칙 발견하는 연관성(association), 장소와 장소간의 연속되는 관계 발견하는 연속성(sequencing)등이 있다.

IV. 결론

본 연구의 실효성을 알아보기 위해 대학생 20명을 두 그룹으로 나누어 A그룹에게는 기존의 사용자 조사방법인 다이어리를 사용하고, B그룹에게는 다이어리와 위치데이터를 병행하여 2주간에 걸친 조사를 진행하였다. 그 결과 A그룹에서는 조사대상자가 기입한 내용에만 의존해야 하며, 오차범위가 발생해도 확인할 수 있는 방법이기 때문에 버려지는 정보가 많았으며 조사대상자가 제공한 내용의 65%만을 취할 수 있었다. 반면 B그룹에서는 조사대상자가 불확실하게 제공한 정보를 위치데이터를 통해 2차 점검을 수행할 수 있었으며, 조사대상자가 제공한 내용의 85%의 내용이 유의미하게 활용되었다. 또한 A 그룹에서는 알수 없었던 패턴과 정보를 찾는데 유용했으며, 실질적인 정보 획득에 영향을 주었다.

■ 참고 문헌 ■

[1] O'Reilly Media, 빅 데이터 어떻게 활용할 것인가, 제이펍, 2013
 [2] 안영건, "안드로이드 스마트폰의 데이터 수집 방법". 한국컴퓨터정보학회 동계학술대회 논문집 제19권 제1호, pp.97-99, 2011
 [3] <https://maps.google.com/locationhistory>