

미래인터넷포럼(FIF) - 서비스WG 회의

시간: 7월 19일(목) 오후 3시

장소: 서울역 KTX별실 (4F)

1. 참여자

- A. 이동만(KAIST), 이영희(KAIST), 박주용(경희대), 유재필(인터넷진흥원), 박종민(에이콘테크), 최덕재(전남대), 이민석(KAIST), 이대규(KAIST), 신병헌(KAIST)

2. 세미나 발표: Human Motion Modeling and Future Data Studies

- A. 발표자: 박주용 교수님 (경희대 물리학과, KAIST 문화기술대학원)

- B. 내용

- 아이폰: 처음으로 위치기반 서비스를 가능하게 한 획기적인 제품으로 평가된다. (예) 렌터카가 도시 이곳 저곳에 퍼져서 주차돼 있고, 사용자가 아이폰 위치기반 서비스를 통해 차 근처에 가면 스마트키를 아이폰으로 받아서 그 차를 탈 수 있도록 하는 아이폰 서비스 출현.

- 모바일 환경에 대한 관점의 변화: Mobile is the Future → Mobile is the Present. 모바일 환경에서의 부산물은 대용량의 데이터이고, Big Data의 일종으로 볼 수 있다. 이러한 모바일 데이터를 이용하면 지금껏 생각해보지 못한 스케일로 사람의 행동이 측정 가능하다. 그리고 사용자의 행동에 대한 더 나은 이해를 기반으로 사람이 필요로 하는 것을 알 수 있다.

- 사람과 기술의 관계: 사람 ↔ 기술 (사람의 본능을 기술이 어떻게 충족할지 연구/개발하면 사람 또한 그 기술에 대해 반응하며, 이는 또다시 새로운 기술 개발에 대한 동기가 됨)

- Data and Analysis

- 기존 통신사의 모바일 데이터: 어느 시간에, 어느 타워에서 (어느 위치) 통화를 했는지 등

- 데이터가 있다고만 해서 좋은 것은 아니다. 깨끗하게 잘 정리된 데이터가 가치가 있지만 사람에 대한 빅데이터는 messy (비정형); 또한 데이터는 현상만을 나타낼 뿐인데 그런 현상이 나타나게 된 원인과 방법은 다양하다. 따라서 데이터 분석(analysis)과 모델링(modeling)이 필요하다.

- 모델링: 분석을 통해서 이 데이터는 어떤 것이 핵심적이고 눈에 띄는 성질임을 나타내고 이용

하는 것.

- (예) 한 사람이 수행한 통화들 사이의 물리적 거리: 출장을 갔을 때와 일상생활에서 그 거리는 다르다. (예2) 동물이 먹이 찾기 위한 움직임.

- Markov Chain으로 표현 가능: 사람의 움직임이 점점 diverge하는 것 같지만 어느 시점에서 더 이상 범위가 넓어지지 않고 자기만의 typical 영역으로 나타난다. 빅데이터 분석에 있어서 Markov Chain이 설득력이 있음. 하지만 random walk 모델을 쓰면 무한대로 diverge하기 때문에 적용되지 않음.

- 예: 집에서 일주일에 5일 일터로 감. 가끔 출장. 일터에서 일 마치고 회의, 운동, 마트 중에 하나로 이동. 빅데이터 분석을 통해 집→일터→마트→집의 움직임이 높은 확률로 나타나는 현상이고, 이것이 dominant.

- 휴대폰 모바일 데이터 분석 결과

- 사람의 움직임은 질서가 없는 게 아니라 주기를 가지고 반복적으로 움직이는 패턴을 보인다. 그 패턴을 빨리 알게 되면 그 사람들의 움직임에 도움을 주기 위해서 분석을 하거나 서비스 제공에 유용할 것이다. (예: demographic information으로 활용)

- Leading eigenmodes of the matrix 이용해서 dominant한 것을 찾아내는 식으로 fast detection of pattern 찾아낼 수 있다.

- 일반적인 사람들은 2~3곳의 장소들을 많이 오가는 하는 패턴을 보인다.

- How to build upon these?

- strong correlation between locations (for many reasons): → highly correlated locations 사이에 infrastructure 구축/보완을 통해 사람들에게 더 나은 서비스 지원

- social distance vs. physical distance

세계적 관심: BarabasiLab (Northeastern Univ. & MIT Civil Eng.), Oxford & London Underground, DTU

- 질의응답

*[질문] (이영희) 실제 데이터로부터 modeling이 되는가? [질문] (이동만) General modeling은 아직 힘들 것 같고, type별로 특성을 나눠서 modeling은 가능하지 않을까?

- 도시 안에서 체육관을 갔다가 음식점을 갔는지까지는 모바일폰 데이터 가지고는 부족하며, foursquare 서비스를 이용해서 실험을 수행하면 좋은 정보가 될 것 같다.

*[질문] (이동만) 현실세계와 가상세계 사이의 연관(correlation)이 어떻게 나올까?

- 하나의 예로써 Facebook 친구들 사이에는 현실의 물리적 거리와 사회적 거리의 연관이 깊다.

*[질문] (최덕재) 데이터는 휴대폰 기지국 정보인지?

- 맞습니다.

[질문] (이동만) 데이터 분석을 위해 통신사로부터 제공받을 때 privacy가 문제가 있고, guideline이 없어서 함부로 활용하고자 하는 생각을 갖기 어렵다.

- 요즘 분위기가 바뀌어 가고 있어서 가능해질 것이다. (온라인 게임 회사로부터 캐릭터 움직임 데이터 얻는 것)

[의견] (박정민) 위치정보보호법 등 민감한 이슈가 있어서 머뭇거리게 되는 경우가 있다. 약관에 넣어도 사람이 못보고 동의했을 수 있다고 해서 이슈가 될 수 있다. → 실험 목적을 명시해서 데이터를 수집할 필요가 있음.

3. 다음 회의: 9 월, HTML5 와 클라우드를 주제로 세미나 발표 예정

4. FIF Day 행사 발표내용 논의

A. GFI Week 2012 에서 첫째날 (9/10)에 FIF Day 있음. 각 WG 별로 발표를 하기로 했다. 어떤 주제로 발표하는 것이 좋을지? (WG 별로 40 분 배정. 30 분 발표하면 될 듯)

B. [의견] (이동만) 활동내역 + 기술적 내용도 발표하는 것이 좋다. 요즘 트렌드 등을 소개하는 것으로 시작하면 좋겠다.

C. 주제: Trends of Linked Open Data