

모바일 소셜 네트워크 서비스를 위한 사용자 컨텍스트의 확률추론

(Probabilistic Inference of User's Context for Mobile Social Network Services)

민준기* 장수형**
(Jun-Ki Min) (Su-Hyung Jang)

조성배***
(Sung-Bae Cho)

요약 소셜 네트워크는 관심사가 유사한 사람을 찾거나 정보를 교환하는 주요한 수단으로 최근 각광받고 있다. 특히 페이스북(Facebook)이나 싸이월드 등과 같이 인터넷 상에서 제공되던 소셜 네트워크 서비스는 이제 휴대폰 영역으로 확장되고 있다. 휴대폰은 개인이 항상 소유하고 다니기 때문에 개인정보의 취득이 쉬우며, 센서 기술의 발달로 다양한 종류의 데이터를 함께 수집할 수 있다. 따라서 이를 모바일 소셜 네트워크 서비스에 활용하면 기존의 인터넷기반 소셜 네트워크 서비스와의 차별화를 이룰 수 있다. 이를 위해 본 논문에서는 확률추론 모델인 베이저안 네트워크를 이용하여 사용자와 주변 사람들간의 친밀도나 관계 등의 모바일 소셜 컨텍스트를 추출하는 방법을 제안한다. 이는 모바일 소셜 네트워크 서비스에 중요한 정보로 활용 가능하며, 본 논문에서는 실제 수집한 데이터를 기반으

로 그 가능성을 검증하였다.

키워드 : 모바일 소셜 네트워크, 베이저안 네트워크, 사용자 컨텍스트

Abstract In recent years, social networks come into the spotlight as the important ways to find people or share information. Especially, existing social network services based on the Internet, such as Facebook.com and Cyworld.com, are now extended into the mobile environment. A mobile phone easily collects the personal information since it is carried by the user at all times, and various types of data can be gathered together with the advance of sensor technologies. These features differentiate the mobile social network services from the previous Internet-based services. In this paper, we estimate the user's mobile social contexts like closeness and relationship between the user and surrounding people using Bayesian networks. The mobile social contexts can be employed as important information for providing mobile social network services, and experimental results on real world data have verified their possibilities.

Key words : Mobile Social Network, Bayesian Networks, User's Context

1. 서론

소셜 네트워크는 관심사, 취미, 친분 등의 공통점을 가지고 연결되어있는 사람들의 그룹을 의미한다. 소셜 네트워크 서비스는 인터넷의 발달과 함께 발달하였으며, 대표적인 예로 해외의 마이스페이스(myspace.com), 페이스북(facebook.com), 그리고 국내의 싸이월드(cyworld.com) 등이 있다.

최근 기술발달로 인한 고성능 휴대폰들이 출시됨에 따라 휴대폰의 주된 기능이 기존의 통화에서 사진, 동영상 등의 다양한 미디어에 대한 기능이 주목 받으며, 이를 이용한 서비스에 대한 관심이 높아지고 있다. 또한 휴대폰은 전화통화나 문자메시지를 주고받는 기존의 통신기능 이외에 인터넷을 통한 채팅, 이메일, 블로그 등의 기능들을 제공하게 되어 소셜 네트워크의 중심적인 매체가 되고 있다[1].

그러나 현재 휴대폰을 이용한 소셜 네트워크 서비스는 대부분이 기존 인터넷에서 제공되어 온 시스템의 확장 개념으로서 한정된 서비스를 제공하고 있다. 휴대폰은 언제 어디서나 사용자가 지니고 다니기 때문에 개인화된 정보의 활용이 가능하며, 특정 장소 및 시간대에 누군가와 통화한 기록은 상대방과 사용자와의 관계를 추론하는데 중요한 정보가 되어 모바일 소셜 네트워크 서비스에서의 활용이 가능하다[2].

본 논문에서는 사용자의 위치정보와 통화로그, 문자로

* 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 2009-0083833)

* 이 논문은 2009 한국컴퓨터종합학술대회에서 '모바일 소셜 네트워크 서비스를 위한 사용자 컨텍스트의 확률추론'의 제목으로 발표된 논문을 확장한 것임

† 학생회원 : 연세대학교 컴퓨터과학과
loomlike@sclab.yonsei.ac.kr

** 학생회원 : 연세대학교 인지과학과
neogates@sclab.yonsei.ac.kr

*** 종신회원 : 연세대학교 컴퓨터과학과 교수
sbcho@cs.yonsei.ac.kr

논문접수 : 2009년 8월 13일

심사완료 : 2009년 11월 18일

Copyright©2010 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제16권 제3호(2010.3)

그 등의 모바일 데이터로부터 확률추론 모델인 베이저안 네트워크(Bayesian Network)를 이용하여 모바일 소셜 컨텍스트를 추론하고, 이를 이용하여 소셜 네트워크를 모델링하는 방법을 제안한다.

2. 관련연구

센서와 무선통신 기술의 발달로 우리 주변의 많은 정보들을 정밀하게 수집하는 기술이 발전하고 있다. 이러한 기술들은 모바일기기에 적용되고 있는데, CMU에서 개발한 Sensay[3], 핀란드 VTT연구소의 CyPhone[4] 등이 센싱 기술을 휴대폰에 적용한 대표적 사례이다. 이와 같은 모바일 컨텍스트 기반 상황인지 서비스는 사용자에게 편리성을 제공해준다. Cao 등은 모바일 여행 가이드 서비스를 제안하였는데, 이는 커뮤니티 컨텍스트인 사용자의 여행 히스토리, 프로필, 그리고 모바일 기기를 통한 위치정보 등을 활용하여 규칙기반으로 서비스를 제공하였다[5]. Lee 등은 모바일 기기를 이용하여 사용자의 감정을 소셜 네트워크 상의 지인들에게 전송해주는 서비스를 제안하였다. 이들은 심장박동, 피부 온도 등을 측정할 수 있는 모바일 센서를 사용하여 사용자의 신체정보를 직접 수집하고 이로부터 규칙 기반으로 사용자의 감정을 추론하였다[6].

모바일 환경에서의 사용자 컨텍스트는 불확실성을 내포하기 때문에 단순한 규칙기반 모델로는 정확한 컨텍스트 정보의 추출이 어렵다[7]. 따라서 본 논문에서는 사용자의 모바일 컨텍스트를 확률기반 모델인 베이저안 네트워크를 활용하여 추론하며, 특히 소셜 네트워크 서비스의 중요한 정보로 활용 가능한 사용자의 사회적 상태(소셜 컨텍스트)를 추론한다.

3. 모바일 소셜 네트워킹

사용자의 소셜 컨텍스트는 소셜 네트워크상의 사람들을 그룹화 하는 기준이 되며, 이를 바탕으로 그룹추천 및 그룹정보공유 등의 소셜 네트워크 서비스를 제공할 수 있다. 그림 1은 제안하는 모바일 소셜 네트워크 생성 방법의 전체 흐름을 보여준다. 본 논문에서는 삼성

SPH-M4650 스마트폰과 Bluetooth로 연동되는 외장형 GPS기기를 이용하여 모바일 로그를 수집하였다.

3.1 로그 수집 및 전처리

휴대폰은 음성통화, 문자메세지, 데이터 전송 등과 같은 다양한 종류의 통신을 제공해준다. 또한 Bluetooth로 상대방의 모바일기기 ID를 인식하여 상대방과 만났는지 여부도 알 수 있다. GPS 정보로부터 사용자가 방문한 장소에 대한 정보를 얻을 수 있으며, 그 밖의 다양한 사용자 상황정보(컨텍스트)도 취득 가능하다. 이와 같은 정보와 함께 사용자가 어떤 사람과 연락하거나 만났는지 여부를 파악하면 사용자와 상대방간의 관계를 추출할 수 있다. 표 1은 본 논문에서 수집한 모바일 로그 정보를 나타낸다.

표 1 소셜 컨텍스트 추론을 위해 수집한 모바일 로그

종류	정보
주소록	전화번호, 주소, 직장 등
통신	통화, 문자메세지 등
장치	장소(GPS), 시간, 날짜, 가속도 등

수집한 데이터는 컨텍스트 추론 베이저안 네트워크의 증거값으로 이용된다. 이를 위해 GPS를 통해 수집한 경도와 위도 정보를 의미적 위치 레이블로 변환하였다. 본 논문에서는 위치 레이블을 {집, 직장, 그 외 장소} 3 가지 경우로 간단히 정의하였으며, 집과 직장의 좌표는 사용자가 입력하였다.

3.2 소셜 컨텍스트 추론

베이저안 추론모델은 불확실한 정보를 처리하는데 유용한 모델이며, 전문가의 지식을 네트워크 구조나 확률 테이블 값 모델링에 적용할 수 있다는 장점이 있다[8]. 따라서 본 논문에서는 베이저안 네트워크를 이용하여 소셜 컨텍스트를 추론하였다. $G(B_S, \theta)$ 를 베이저안 네트워크 모델이라 하였을 때, B_S 와 θ 는 각각 네트워크의 구조와 파라미터를 나타낸다. θ 는 조건부확률테이블 B_ϕ 와 조건부확률분포 B_p 로 구성된다. 조건부 확률 $P(\theta)$ 에 대해, 추론과정은 다음의 식 (1)과 같다.

$$P(Z_T, Y_T, \theta) = P(Y_T | Z_T, B_\phi) P(Z_T | B_p) \quad (1)$$

식 (1)에서 $Z_T = \{z_1, z_2, \dots, z_T\}$ 는 T 개의 상태를 갖는 변수를 나타내고, Y_T 는 해당하는 관측 값을 나타낸다. 본 논문에서는 표 2와 같이 친밀도, 관련행동, 관계의 3 가지 소셜 컨텍스트를 정의하고, 이를 인식하는 베이저안 네트워크를 도메인 지식을 기반으로 설계하였다.

베이저안 네트워크 설계의 구체적 기준은 다음과 같다. 친밀도: 상대방과 사용자가 친한지 여부를 나타내는 값으로, 자주 연락하거나 오랜 시간 동안 접촉을 하면 친한 관계일 확률이 높은 것으로 간주하였다. 또한 일반

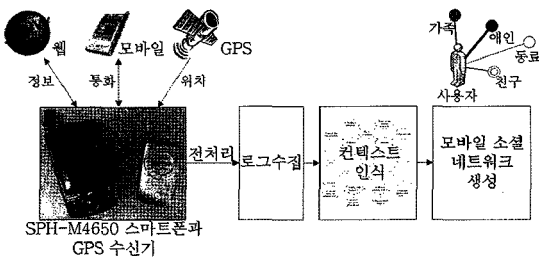


그림 1 모바일 소셜 네트워크 추론방법

표 2 소셜 컨텍스트

컨텍스트 종류	BN노드 상태 종류
친밀도	가까움, 멀
관련행동	이동, 일, 공부, 휴식, 놀기, 먹기, 기타
관계	가족, 애인, 친구, 동료, 지인, 기타

적으로 친한 사람들은 주로 주말이나 저녁과 같이 사적인 시간에, 혹은 휴식을 취하거나 여가를 즐길 때 만나거나 연락한다는 가정을 하였다.

관련행동: 상대방과 연락 중의 사용자 행동을 나타내는 것으로, 어떤 행동 중에 상대방과 연락하는지에 따라서 그 사람과 사용자간의 관계나 친밀도를 파악할 수 있다. 예를 들어 업무 중에 만나거나 연락하는 사람은 동료일 가능성이 크다. 사용자의 행동은 현재 시간, 요일, 장소로부터 추론한다.

관계: 휴대폰 상의 주소록으로부터 일반적인 관계를 알 수 있다. 예를 들어 성이 같거나 전화번호의 뒷자리가 같으면 가족일 확률이 높고, 직장주소나 email등을 통해 동료인지 여부를 판단할 수 있다. 그 밖에도 주로 어떤 행동을 함께 하는지, 어느 정도로 가까운 사이인지 등의 정보에 따라서 좀더 정확한 관계를 파악할 수 있다.

수집한 데이터 로그를 컨텍스트 추론 베이지안 네트워크의 증거값으로 사용하기 위해서는 노드별 상태에 따라 연속된 관측 값을 이산화해야 한다. 예를 들어 그림 2의 “최근연락” 노드는 {3일 이내, 3~6일 사이, 7일 이상} 등과 같이 이산화 될 수 있다. 이 경우 사용자가 하루만에 연락하면 노드의 확률테이블은 증거값을 설정함에 따라 {1, 0, 0}이 되며, 2일만에 연락할 경우에도 마찬가지로 {1, 0, 0}이 된다. 일반적인 방법은 이와 같이 정보 손실을 발생시키며, 이를 해결하기 위해 본 논문에서는 관측된 통계적 수치 혹은 미리 정의한 퍼지 함수를 이용하여 연속된 실수 값으로 파라미터를 변경한다. 앞의 “최근연락” 노드를 예로 들면, 우선 노드가 {최근, 오래전}의 두 상태를 갖도록 설계한 후, “최근”과 “오래전” 값을 퍼지 함수로 모델링하여 최근에 연락한 경우에는 노드의 확률값을 {0.8, 0.2}, 근래에 연락하였을 경우에는 {0.6, 0.4} 등과 같이 변경한다. 이 경우에는 해당 노드가 부모노드이기 때문에 변경한 확률테이블값이 추론 과정에서 자식노드로 전파된다. 하지만 그림 2의 “시간”(흰색 노드)과 같이 증거노드가 자식일 경우에는 증거값을 설정해야 부모노드로 확률이 전이된다. 이를 해결하기 위해 {Yes, No}의 이진 상태를 갖는 가상노드를 추가하는 방법을 적용하였다[9]. 표 3은 “시간”노드의 증거값을 실수값으로 설정하기 위한 가상노드(그림 2에서의 열은 회색 노드)의 확률값 설정 예를

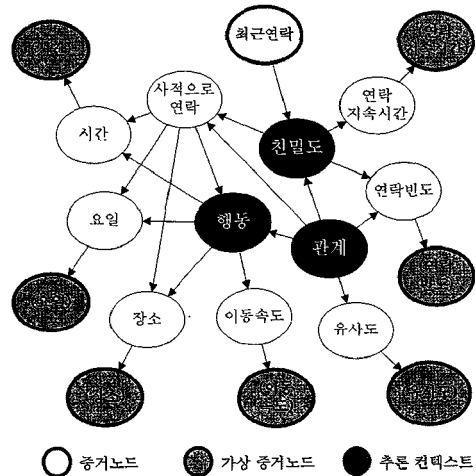


그림 2 소셜 컨텍스트 추론을 위한 베이지안 네트워크 구조

표 3 가상노드 확률 테이블 설정 값(사용자가 상대방과 10번 중 두번은 저녁에, 한번은 밤에 연락할 경우의 시간 가상노드 예)

시간 가상노드 상태	부모노드(시간 노드)의 상태			
	아침	점심	저녁	밤
Yes	0.05	0.05	0.8	0.1
No	0.95	0.95	0.2	0.9

보여준다. 관측된 통계값 혹은 퍼지값을 표 3과 같이 조건부확률로 설정한 후에는 Yes상태를 증거값으로 설정하여 부모노드로 확률이 전이되게 한다. 이때, 일반적인 확률테이블의 경우 노드 상태가 확률분포를 이루어야 하기 때문에 열의 합이 1이 되지만, 가상노드의 경우 Yes상태에 해당하는 값들이 확률 분포를 이뤄야 하기 때문에 Yes상태 행의 합도 1이 되도록 설정한다.

지금까지 설명한 모바일 소셜 컨텍스트는 사용자와 상대방과의 관계를 나타내는 값이기 때문에 사용자가 해당 상대와 접촉할 때마다 추론 값이 갱신되어야 한다.

4. 실험 및 결과

제안하는 방법을 실제계에 적용하기 위하여 앞의 그림 1에서 소개한 모바일 기기(삼성 SPH-M4650와 Bluetooth GPS)를 이용하여 대학생 10명에 대해 각각 1달 분량의 모바일 사용로그를 수집하였다. 이 때 사용자 1~9는 통화로그와 위치정보를 함께 수집하였으며, 사용자 10은 통화로그만 수집하였다.

그림 3은 실제 수집한 데이터의 통화 수와 통화 상대 관계분포를 보여준다. 그림 4는 사용자 1과 주변 2명에 대한 소셜 컨텍스트 추론결과를 보여준다. 그림 4의 위에 사람은 사용자 1과 친밀하며, 사적이거나 공격적인 시

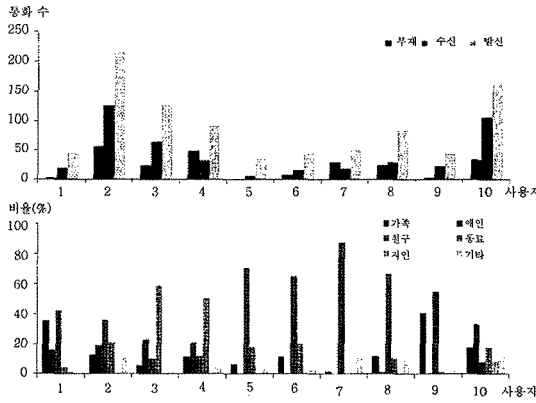


그림 3 수집한 모바일 데이터 분포

간에 상관없이 주로 식사와 같은 행동 중에 연락하는 관계인 것을 확인할 수 있으며, 유사도(주소록 상의 정보로부터 획득)가 높아 사용자 1과는 가족 같은 관계인 것으로 추론되었다. 그림 4의 아래 사람의 경우는 사적으로 연락을 하며 공부를 하거나 놀 때 연락을 했기 때문에 친구관계로 추론되었다. 하지만 학교 외의 장소에서의 연락빈도가 높고 친밀도가 낮기 때문에 안면만 있는 사람(acquaintance)이나 모르는 사람(etc)의 확률도 높은 것을 확인할 수 있다.

실제 사용자의 관계와 추론한 결과의 일치여부를 확인해보기 위해 사용자가 수동으로 태깅한 관계정보를 이용하여 정확도를 측정해보았다. 10명의 피험자들이 모두 학생이었기 때문에 동료(coworker)와 친구(friend)의 구분이 실제적으로는 모호하여 이를 동일한 관계로 보았을 때 사용자 1~9의 관계추론 정확도는 약 60% 정도였으며, 위치 컨텍스트를 사용하지 않은 사용자 10에 대해서는 45.5%의 정확도를 보였다. 현재 사용 가능한 데이터는 통화기록(통화시간, 길이, 횟수 등)과 위치정보 두 가지로, 정보의 종류가 적었기 때문에 정보가 누락된 경우의 정확도 손실이 컸다. 따라서 Bluetooth를 이용한 만남여부 정보나 일정, 연락처의 주소 정보 등 사용자 개인정보(PIMS)를 활용하여 추론모델의 안정성을 높일 필요가 있다.

관계의 추론결과가 틀린 경우의 대부분은 로그수집기간이 짧아서 기간 중에 연락하지 않은 친구나 동료를 기타(etc)로 분류한 경우였으며, 가족을 친구로 잘못 추론한 경우가 종종 있었다. 그림 5는 사용자 1과 애인(lover)간의 관계 추론결과를 보여주는 것으로, 통화 로그가 많아질수록 정확하게 추론하는 것을 확인할 수 있다.

앞의 실험에서 추론한 관계는 통화패턴으로부터 획득한 사용자와 상대방 간의 상호작용 정보이기 때문에 사용자의 실제 관계와는 정확히 일치하지 않을 수 있다.

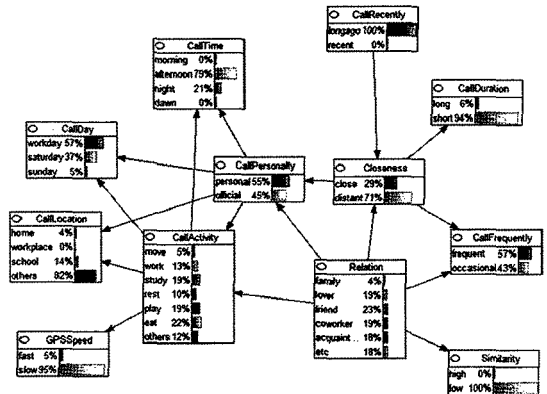
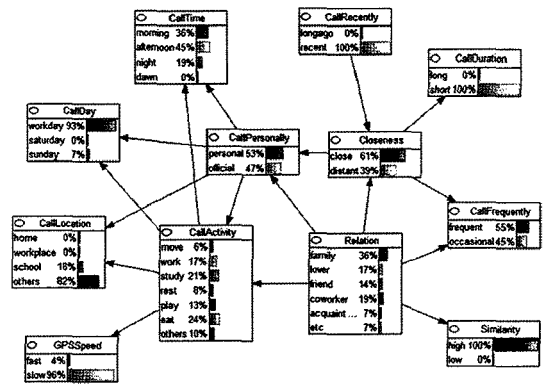


그림 4 사용자 1과 두 명(상/하)간의 소셜 컨텍스트 추론 결과 예

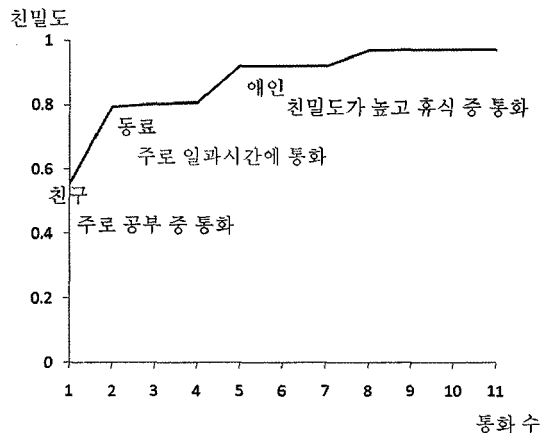


그림 5 로그 축적에 따른 사용자1과 애인간의 관계 추론 변화

하지만 소셜 네트워크 서비스 제공을 위한 중요한 정보로 활용될 수 있다. 예를 들면 추론 관계에 따라 그룹을 생성해주고, 같은 그룹에 속한 사용자들 간에 관련있는 컨텍스트를 공유한다거나, 통화상대를 추천하는 등의 서

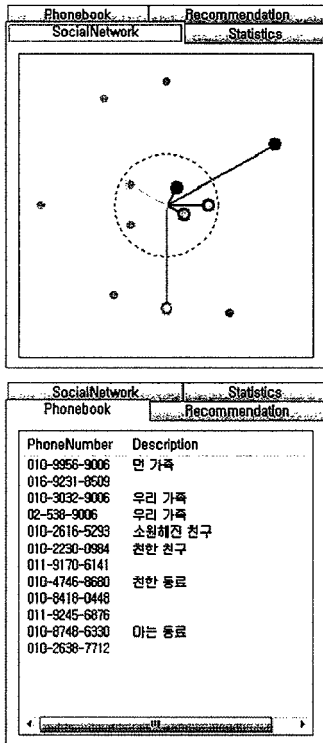


그림 6 추론한 소셜 컨텍스트의 활용 예. 사용자 관계그 래프(상)와 전화번호부(하)

비스가 가능하다. 그림 6은 본 논문에서 제안하는 소셜 컨텍스트 추론 결과를 이용한 간단한 어플리케이션의 예를 보여준다. 그림 6(상)은 사용자가 자신(중심)과 주변인들(노드)간의 관계를 한눈에 확인할 수 있는 소셜네트워크 그래프로, 친밀한 사람일수록 짙은 색의 선으로 나타내주며 친밀하지 않은 사람은 옅은 색으로 보여지거나 화면에 표시되지 않는다. 선의 길이는 최근 연락 여부를 나타내는 것으로, 일정기간동안 연락하지 않은 경우는 점점 선이 길어져서 사용자의 설정 값(점선으로 표시된 원)보다 멀어지면 알람으로 경고해줄 수 있다. 그림 6(하)는 추론한 관계와 친밀도를 전화번호부에 표시해준 것으로, 사용자가 별명을 붙이거나 그룹 등을 관리할 때 유용될 수 있다.

5. 결론

본 논문에서는 베이지안 네트워크를 이용하여 친밀도, 연락 중 행동, 관계 등의 모바일 소셜 컨텍스트를 추론하고, 이를 이용하여 소셜 네트워크를 모델링하는 방법을 제안하였다. 연락 패턴에 기반하여 추출한 사용자의 사회적 관계는 주소록에 그룹을 자동으로 생성해준다거나 다양한 소셜네트워크 기반 서비스를 제공해주는데

활용 가능하다. 하지만 수집된 로그의 양이 적을 경우 정확한 추출이 어렵고, 모델이 정적이기 때문에 다양한 사용자의 생활패턴을 효과적으로 모델링하지 못한다는 한계점이 있었다. 향후에는 정적 모델의 한계를 극복하기 위하여 실시간으로 모델의 구조와 파라미터를 학습하는 방법을 연구할 계획이다. 또한 다양한 직업 및 생활패턴을 가진 피험자로부터 상호작용 데이터를 수집하여 실제 네트워크형식의 사회관계를 추론하고, 이를 모바일 기반 추천 서비스에 적용하여 사용성 평가 등을 통해 유용성을 분석할 계획이다.

참고 문헌

- [1] Y. Chen, et al., "Adapting web pages for small-screen devices," *IEEE Internet Computing*, 9(1), pp.50-56, 2005.
- [2] N. Eagle, and A. Pentland, "Reality mining: Sensing complex social systems," *Personal and Ubiquitous Computing*, 10(4), pp.255-268, 2006.
- [3] D. Siewiorek, et al., "SenSay: A context-aware mobile phone," *7th IEEE International Symposium on Wearable Computers*, pp.248-249, 2003.
- [4] P. Pulli, et al., "CyPhone: Future personal tele-cooperation device," *IFIP World Computer Congress*, pp.529-544, 1998.
- [5] Y. Cao, et al., "Follow me, follow you-spatiotemporal community context modeling and adaptation for mobile information systems," *IEEE International Conference on Mobile Data Management*, pp. 108-115, 2008.
- [6] H. Lee, et al., "An agent-based context-aware system on handheld computers," *IEEE International Conference on Consumer Electronics*, pp.229-230, 2006.
- [7] K.-S. Hwang, and S.-B. Cho, "Modular Bayesian network for uncertainty handling on mobile device," *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp.402-408, 2008.
- [8] F.V. Jensen, *Introduction to Bayesian Networks*, Springer, NewYork, 1996.
- [9] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Mateo, CA, 1988.