

스마트폰 사용패턴 분석을 위한 원격 로그데이터 수집 시스템 구현

송현지, 이민경, 정희원, 유석종
숙명여자대학교 컴퓨터과학부

e-mail : sarahpina7@gmail.com, mklee317@gmail.com, hwj2681@gmail.com, sjyu@sm.ac.kr

Implementation of a remote log-data collecting system for the analysis on smartphone usage pattern

Hyun-Ji Song, Min-Kyung Lee, Hee-Won Chung, Seok-Jong Yu
Dept. of Computer Science, Sookmyung Women's University

요 약

다수 사용자의 스마트폰 사용패턴을 협업적인 방법으로 분석할 경우 모바일 기기에 대한 선호도 분석, 과몰입 정도 판단 등 다양한 관련 연구에 활용될 수 있다. 본 연구는 스마트폰의 사용패턴 분석을 통한 사용자 맞춤형 서비스 개발을 위하여 로그데이터를 추출하여 서버에 저장하는 시스템을 설계하고 구현하는 것을 목표로 한다. 사용자의 스마트폰 로그데이터를 수집하기 위하여 모바일앱을 개발하고 모바일앱을 통해서 추출된 로그데이터를 저장할 서버 DB 를 구축하고 유사성 분석을 위한 협업필터링 엔진을 개발하였다. 개발된 시스템의 성능 평가를 위하여 일부 사용자에 대한 사용패턴 데이터셋 구축 실험을 수행하였으며 후속 연구를 위한 실험 환경을 설계하였다.

1. 서론

최근 스마트폰 사용자의 수가 급격히 증가하고 있으며 사용하는 연령층도 다양해지고 있다. 한국인터넷진흥원에서 2013 년 만 12 세 이상의 모바일인터넷 이용자 6000 명을 대상으로 한 조사에 의하면 모바일 인터넷 이용자의 95.5%가 스마트폰을 보유하고 있다 [1]. 이러한 스마트폰 사용 증가는 많은 데이터 발생을 의미하며 사용 패턴의 분석을 통하여 다양한 응용 연구에 사용될 수 있다. 특히 모바일앱 별 사용시간 측정을 통하여 선호도 분석을 통한 맞춤형 서비스가 가능하며 스마트폰 중독 여부를 판단하는 연구에도 활용할 수 있다[2][3].

본 연구에서는 스마트폰에서 발생하는 로그데이터를 수집하여 사용자의 사용패턴을 분석하는 시스템을 개발하고자 한다. 특히 수집된 데이터에 협업필터링 알고리즘을 적용하여 유사한 사용패턴을 보이는 사용자 집단을 탐색하는 기능을 구현한다[4][5].

Google play 스토어에 등록된 로그데이터 수집 모바일앱 중 일부를 본 시스템과 비교한 결과는 <표 1>과 같다. 대부분의 기존 시스템은 모바일앱이 설치된 기기 내에서만 데이터를 수집하고 외부 서버에 저장하는 기능은 제공하지 않는다.

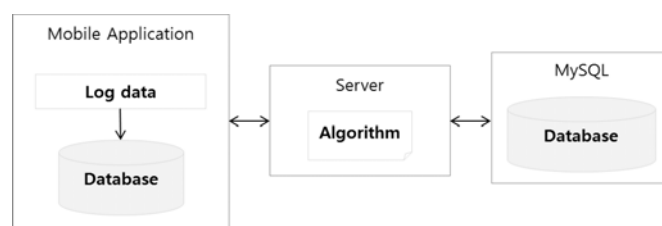
본 시스템은 모바일앱에서 서버로의 데이터 전송 기능, 협업필터링 분석기능을 제공하여 기존 로그데이터 수집 앱들과는 차별화된다.

<표 1> 로그데이터 수집 모바일앱의 특징 비교

모바일앱	개발자	로그데이터 수집	서버로 데이터 전송	협업필터링
본 시스템	본 연구팀	○	○	○
넌얼마나쓰니	RinaSoft	○	○	X
CatLog	Nolan Lawson	○	X	X
Log Collector	Xtralogic, Inc	○	X	X
MX 로그 컬렉터	J2 Interactive	○	X	X

2. 스마트폰 로그데이터 수집 분석 시스템

본 연구의 전체 시스템의 수행과정은 (그림 1)과 같으며, 시스템은 모바일앱 부분과 서버 부분으로 나눌 수 있다.



(그림 1) 본 시스템의 처리 과정

(1) 모바일 앱을 통한 로그데이터 실시간 추출

사용자가 로그데이터 수집 앱을 1회만 실행하면 사용자가 별도의 조작을 하지 않아도 자동으로 모바일 앱을 통해 로그데이터가 서버로 전송되도록 설계하였다. 로그데이터를 수집하는 모바일 앱은 스마트폰 사용 패턴 분석에 활용할 수 있는 핵심 데이터들을 24시간 단위로 수집한다. 수집되는 항목으로는 총 사용시간, 화면 켜짐 횟수, 데이터 사용량, 앱 카테고리 별 사용시간과 실행 횟수이다. 데이터 사용량은 WIFI 사용량과 3G/LTE 사용량으로 구분하여 추출하였고, 앱 카테고리 별 사용시간은 사람들이 주로 사용하는 카테고리 별로 각각 추출하였다. 선별한 카테고리는 <표 2>와 같다. Google Play에서 설치한 모바일 앱이 아닌 기존에 내장되어 있던 모바일 앱들은 성격에 맞게 직접 분류하였다.

<표 2> Google Play 제공 카테고리 분류 결과

모바일 앱 카테고리	Google Play 제공 카테고리
채팅	커뮤니케이션
SNS	소셜
게임	롤플레이, 보드, 스포츠, 시뮬레이션, 아케이드, 액션, 어드벤처, 자동차 경주, 전략, 카드, 카지노, 캐주얼 게임, 퀴즈, 퍼즐
음악 및 동영상	음악 및 오디오 미디어 및 동영상
인터넷 (뉴스 + 검색)	다음, 네이버, 네이버, Google 검색, 크롬은 인터넷 앱으로 분류
쇼핑	쇼핑
사진	사진

로그데이터 수집과정에서 사용자의 개입을 최소화하기 위해 <표 3>과 같이 전원이 꺼지는 경우와 작업관리자 실행에 대비하는 기능을 구현하였다.

<표 3> 사용자 개입을 최소화 하기 위한 부가 기능

제약 사항	구현 기능
전원	휴대폰 시작 시 자동으로 실행
	갑자기 전원이 꺼지는 경우를 대비하여 주기적으로 데이터 저장
작업관리자	강제 종료 방지

(2) 서버 DB에 로그데이터 분류 및 저장

수집된 로그 데이터는 사용자의 성별, 연령, 직업과 같은 기본 프로필 정보와 함께 서버로 전송된다. 전송 시, HTTP 통신 방식을 통해 데이터를 전송하여 DB에 저장될 수 있도록 하였다. DB에 데이터를 저장하는 주기는 로그 데이터 수집 후 바로 전송하여 저장하도록 하였기 때문에 로그 데이터 수집 주기와 같은 24시간이 된다.

전송은 기본적으로 WIFI를 사용 중일 때 자동 전송되도록 하고, 3G나 LTE를 통한 전송은 사용자의 동의를 얻은 경우에만 이루어지도록 한다. 만일 어떠한 네트워크도 연결되어 있지 않아 로그데이터를 전송할 수 없는 경우에는 아이콘이 표시되어 추후에 사용자가 직접 전송할 수 있다.

(3) 스마트폰 사용패턴 분석

새롭게 수집한 데이터와 이전에 수집되어 DB에 저장되어 있던 다른 사용자들의 스마트폰 사용패턴과의 유사성을 유클리디언 거리(Euclidean distance) 계산 또는 협업필터링(Collaborative Filtering)을 통해 계산한다. 측정 알고리즘의 구성은 로그 데이터의 점수화 작업과 협업필터링을 통한 사용자 간의 유사도 점수 도출로 이루어진다.

사용자의 사용패턴 비교가 완료되면, 사용자의 단말기에 비교 결과와 사용자의 스마트폰 사용패턴을 보여준다.

3. 시스템 구현 및 데이터셋 구축 실험

(1) 시스템 개발환경

본 시스템의 개발환경은 <표 4>와 같다.

<표 4> 시스템 개발 환경

	모바일 앱	서버	DB
언어	JAVA	JSP	SQL
환경	Eclipse	Tomcat 7.0	MySQL

(2) 데이터셋 구축 실험

<표 5>는 실제 사용자들의 스마트폰에 개발된 모바일 앱을 설치하여 추출된 사용패턴 데이터셋과 이를 점수로 표준화한 결과이다. 사용패턴 데이터셋에 구현된 협업필터링 알고리즘을 적용하여 사용패턴이 유사한 사용자 그룹을 탐색할 수 있고 이를 통하여 특정 사용자의 선호도 또는 사용수준, 중독 정도 등을 예측할 수 있다. 본 연구에서는 특히 스마트폰 과몰입도를 진단하는 비교실험에 초점을 맞추고 있으며, 이 부분은 개발된 모바일 앱을 배포한 후 충분한 데이터셋이 확보된 이후에 후속 연구로 진행하기로 한다.

한국정보화진흥원에서 개발된 S-척도[7]는 스마트폰중독 정도를 지수화한 것으로 설문조사를 통해 다수 항목을 질문하여 측정된 점수에 따라 스마트폰 사용자를 고위험군, 잠재적 위험군, 일반 사용자군의 3 단계로 구분하고 있다. 후속 연구에서는 본 연구에서 개발된 모바일 앱을 이용하여 다수사용자로부터 수집된 로그 데이터셋 분석결과와 일부 사용자의 대한 S 척도 측정 지수를 상호 연관하여 모바일 앱상에서 자신의 과몰입 정도를 미리 예측하는 기능을 제공하고

자 한다.

(3) 모바일앱별 사용패턴 측정결과

모바일앱 설치 후 발생하는 로그데이터는 스마트폰 내장 DB 에 저장되어 사용자에게 낱짜 별로 보여진다. 실행 화면은 (그림 2)와 같다. 사용한 모바일앱은 총 사용시간이 많은 순서대로 정렬되어 보여진다.



(그림 2) 낱짜 별 측정 결과 화면

<표 5> 측정된 로그데이터 및 표준점수화 결과

사용자	총 사용시간	화면켜짐 횟수	데이터 사용량	총 사용시간 점수	화면켜짐 점수	데이터 사용량 점수
1	3시간 1분 7초	128번	242.84MB	4	3	3
2	2시간 52분 12초	129번	28.31MB	3	3	1
3	4시간 31분 50초	192번	26.34MB	5	4	1
4	4시간 1분 46초	132번	18.40MB	5	3	1
5	4시간 6분 46초	102번	54.36MB	5	3	1
6	3시간 28분 54초	170번	86.01MB	4	4	1
7	5시간 47초	153번	65.19MB	6	4	1
8	3시간 6분 50초	158번	42.67MB	4	4	1
9	5시간 35분 21초	286번	313.81MB	6	6	4
10	2시간 33분 2초	131번	19.70MB	3	3	1

4. 결론

본 연구에서는 스마트폰에서 발생하는 로그데이터를 추출하는 모바일앱을 개발하고 서버 DB 에 저장하여 데이터셋을 구축하는 시스템을 개발하였다. 또한 협업필터링을 통해 스마트폰 사용자와 유사한 사용패턴을 가지는 사용자 집단을 찾는 알고리즘을 구현하

였다. 개발된 시스템을 통해 스마트폰 사용자의 사용패턴을 정량적으로 분석하는 것이 가능해졌으며 선호도 분석, 과몰입 진단, 맞춤형 사용자 서비스 등에 활용할 수 있다.

향후, 개발된 모바일앱을 Google play 스토어에 등록을 통해 배포하여 사용패턴 데이터셋을 확장하고 협업필터링 엔진을 기반으로 다수 사용자의 특성 분석 연구에 활용할 예정이다.

감사의 글

이 논문은 2013 년도 정부(교육부)의 재원으로 한국 과학창의재단(대학생 창의융합형 연구과제 지원사업)의 지원을 받아 수행된 연구임.

참고문헌

- [1] 한국인터넷진흥원 보고서, “2013년 모바일인터넷이용실태조사”, 2013.
- [2] 강준명 외, “이동통신망 관리를 위한 스마트폰 사용패턴 수집 및 분석”, KNOM Conference, 2011.
- [3] 강동주 외, “스마트폰의 애플리케이션 사용 경향 분석을 위한 프레임워크”, 정보과학회논문지: 컴퓨팅의 실제 및 레터 제 19 권 제 10 호, 2013.
- [4] 전종수 외, “2012년 인터넷중독 실태조사”, 한국정보화진흥원, 2012.
- [5] 오강탁, 이제은, “스마트 라이프 혁명의 실제와 스마트폰 중독”, Internet and Information Security 제3권 제4호(2012년 제4호) pp. 21~43, 2012.
- [6] 토비 세가란, “집단지성 프로그래밍”
- [7] 한국정보화진흥원, “스마트폰중독 진단척도 개발 연구”