

Appingpot : 하둡 및 스파크를 활용한 어플리케이션 큐레이션 플랫폼

전상우, 심의석, 지정희
건국대학교 인터넷 미디어 공학과
email: {jsw1539, nasuk, jhchi}@konkuk.ac.kr

Appingpot : Application curation platform based on Hadoop and Spark

Sangwoo Jeon, Euseok Shim, Jeonghee Chi
Internet & Multimedia Engineering, Konkuk University

요 약

현재 해외뿐만 아니라 국내에서도 큐레이션 서비스가 활발히 운영중이다. 폭발적으로 증가한 어플리케이션 시장 에서 사용자들은 자신에게 맞는 앱을 찾고 설치하기 어려워지고 있다. 이에 대응하여 본 논문에서는 어플리케이션 큐레 이션 서비스인 Appingpot 시스템을 제안한다. Appingpot에서는 사용자들로부터 수집된 앱 로그데이터와 Facebook 친구 정보를 기반으로 Hadoop 과 Spark 를 통해 사용자들에게 적합한 앱을 추천하는 서비스를 제공한다.

1. 서론

모바일 앱 시장은 스마트폰, 스마트워치 등 다양한 모바일 기기와 함께 폭발적으로 성장하고 있다. 2016 년 6 월 기준으로 앱스토어는 약 200 만개, 구글플레이 역시 약 200 만개의 앱을 보유하며 빠르게 성장하고 있다[1]. 그러나 긍정적인 미래 전망만 있는 것은 아니다. 스마트폰 앱 시장의 사용자들은 일부 카테고 리에 한정된 이용 패턴을 보이고 있다. 이런 현상은 모바일 앱 시장에 너무 많은 앱이 있다 보니 사용자 들이 자신에게 맞는 앱을 찾아 설치하기 어려워 지기 때문이다[2,3].

구글 플레이와 애플스토어에서는 이런 문제를 해결 하고자, Android 팁 추천, 에디터 선정 추천 리스트, genius 추천 기술 등을 이용하여 사용자들에게 적합한 앱을 추천해 주고 있지만, 앱 시장의 추천은 상업적 목적의 추천이 많다는 문제가 있다. 실제로 구글 플 레이의 경우 일정한 돈을 지불하면 상위에 노출되게 할 수 있다. 이는 앱에 대한 공정한 평가를 통해 사용자들이 좋은 앱을 사용할 기회를 제공하고 있지 않 다는 것을 의미한다.

마켓 차원의 추천뿐만 아니라 추천 서비스를 제공 하는 앱들도 있다. 이러한 앱들은 사용자들의 앱 사 용패턴을 이용하여 개인화된 추천을 제공하지만 신뢰 도가 높지 않다.

이 논문에서는 상업적 추천이 아니고, 기존의 단순 히 사용자의 사용패턴만을 고려한 추천 방식에서 벗 어나 사용자의 사용패턴뿐만 아니라, 페이스북 친구

정보를 이용함으로써 추천 앱의 신뢰도와 정확도를 높일 수 있는 시스템을 제안한다.

2. 관련연구

소셜 큐레이션 서비스(Social Curation Service)는 원하 는 정보만 골라서 볼 수 있도록 하는 네트워크 서비 스를 말한다. 이러한 서비스에서는 ‘관심사’를 중심 으로 사용자들간의 네트워크가 형성되는 것이 특징이 다. 패션에 관심이 있다면 패션만, 야구에 관심이 있 다면 야구에 관한 정보만 집중해서 받아볼 수 있다.

해외에서는 이미 ‘핀터레스트’, ‘Netflix’[4] 등 의 큐레이션서비스가 활발히 운영 중이다. 국내에서 도 ‘Watcha’, ‘인테리어’ 등 다양한 큐레이션 서비 스들이 사용자들의 인기를 얻고있다.

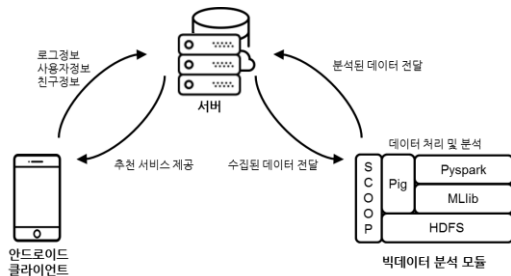
특히 사용자가 원하는 서비스를 제공하기 위해서는 추천알고리즘이 중요하다. ‘Netflix’, ‘Amazon’를 비롯하 여 많은 기업들이 협업필터링알고리즘을 사용하여 추 천서비스를 제공한다. 협업필터링알고리즘[5]이란 다 양한 상품에 대한 사용자의 평점과 상호작용이 혼합 하여 새로운 상품을 추천해주는 기술이다.

본 시스템에서는 앱을 콘텐츠로한 큐레이션 서비 스를 제공한다. 기존의 서비스인 구글플레이 Android 팁 추천, 아마존의 개인화 추천 서비스 등의 서비스와는 다르게 Facebook 친구 정보를 활용하여 이미 형성된 사용자들 간의 네트워크를 사용하여 사용자 확보와 편의성을 향상시킬 수 있다. 이러한 친구정보는 앱을 추천시 새로운 앱에 대한 진입장벽 낮추는데도 기여

할 수 있을 것으로 기대된다.

3. Appingpot 시스템

[그림 1]은 본 논문에서 제안하는 시스템 구성도를 나타내고 있다. 본 시스템은 사용자에게 앱을 추천 및 정보를 제공하는 클라이언트 모듈과 로그데이터를 수집하는 서버모듈, 수집된 데이터를 가공하고, 분석하는 빅데이터모듈로 구성되어 있다.



[그림 1] Appingpot 시스템 구성도

3.1 Appingpot 서버모듈

서버는 Nodejs 로 구성되었으며, 안드로이드 클라이언트로부터 전달되는 각종 로그데이터를 수집하여 Mysql 데이터베이스에 저장하여 빅데이터 모듈에서 이를 분석할 수 있도록 한다. 또한 빅데이터 모듈로부터 분석된 추천 결과를 클라이언트에 전송하여 추천 서비스를 제공한다.

3.2 Appingpot 빅데이터 분석 모듈

Appingpot 빅데이터 모듈은 Apache Sqoop 을 사용하여 HDFS(Hadoop Distributed File System)에 서버 모듈의 앱로그 데이터를 로드한다. 로드가 완료되면 Pig 를 사용하여 분석에 필요한 전처리 작업을 진행한다. PySpark 와 머신러닝 라이브러리인 MLlib 을 활용하여 사용자에게 적합한 앱을 추천한다. 추천된 결과는 Apache Sqoop 을 사용하여 서버모듈의 Database 에 저장한다. 추천 알고리즘으로는 협업필터링 방식을 채택하였다. 이를 적용하기 위하여 먼저, 수집한 앱 사용량을 정규화과정을 거쳐 평점을 계산하였으며 이를 통해 UserId, PackageName, Rating 필드로 이루어진 데이터셋을 생성하였다. 이를 기반으로 ALS(Alternating Least Squares) 추천 알고리즘을 이용하기 위하여 실험을 통하여 RMSE(평균제곱근편차) 값이 낮은 파라미터를 ALS 의파라미터로 선택하여 학습시켰다. 유저가 추가되어 새로운 데이터가 들어오면 새로운 데이터를 학습시키고 기존 데이터에서 평가하지 않은 항목들 내에서 추천데이터를 추출하여 서비스 DB 에 저장한다.

3.3 Appingpot 클라이언트모듈

안드로이드 클라이언트는 첫 설치 시 스마트폰에 설치된 앱 정보 리스트를 저장하고, 페이스북 로그인 시 친구 목록을 저장하여 서버에 전송한다. 이후 지속적으로 사용자의 앱 사용 로그를 저장하여, 스마트

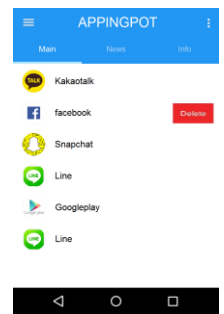
폰에 저장하고, 주기적으로 서버에 전송한다. 서버로부터 추천 정보를 전달받아 사용자에게 유용한 앱을 추천한다.

4. 구현 결과

사용자가 사용할 추천받은 결과를 활용하는 안드로이드 화면과 앱관리화면을 다음과 같이 설계하였다.



[그림 2] 추천결과화면



[그림 3] 앱관리화면

[그림 2] Appingpot 앱의 추천받은 결과화면을 나타내고 있으며, 추천하는 앱, 이 앱을 사용하는 Facebook 친구 프로필이미지, 패턴일치율을 제공한다. [그림 3]은 앱관리화면이다. 설치된 앱을 확인할 수 있고, 삭제하고자하는 앱을 드래그하면 쉽게 삭제할 수 있는 기능을 제공한다.

5. 결론

본 논문에서는 사용자에게 앱을 추천하기 위해서, 사용자의 로그 데이터를 수집하고 수집한 데이터를 가공하여 ALS 알고리즘을 통해 사용자에게 어플리케이션을 추천하는 시스템을 제안하였다.

사용자가 직접 어플리케이션의 평점을 직접 입력하지 않고 앱 로그 데이터 기반으로 평점을 계산하기 때문에 사용자의 만족도를 높일 수 있으며, Facebook 친구정보를 활용하여 앱 설치에 대한 진입장벽을 낮출 수 있을 것이라 예상된다. 향후 연구로는 관심사 태그 정보, 페이스북 좋아요 페이지 등의 다양한 정보를 추가한다면 추천 정확도와 만족도를 더 높일 수 있을 것이라 예상된다.

6. 참고문헌

- [1] "Number of Apps Available in Leading App Stores as of July 2016", Statista.com (2016a)
- [2] "오선주. 스마트폰 앱 선택요인과 구매요인의 차이에 대한 탐색적 연구", 한국전자거래학회지, 2013
- [3] "스마트폰 애플리케이션 시장의 성장 한계요인 연구", 박희운, 고혜영, Journal of Korea Design Knowledge, Vol.22, pp.127-138, 2012.6
- [4] "<https://www.netflix.com/>", Netflix, 2015. 10
- [5] "Amazon. com recommendations: Item-to-item collaborative filtering", Linden, Greg, Brent Smith, and Jeremy York, Internet Computing, IEEE 7.1: 76-80. 2003