

MBTI 기반 협업 추천 시스템 : 웹툰 콘텐츠 사례 연구

이명연[○], 이오준^{*}, 홍민성^{*}, 정재은^{*}

[○]중앙대학교 컴퓨터공학부, 중앙대학교 컴퓨터공학과

e-mail : {dalaetm[○],concerto34^{*}, minsung87^{*}, j3jung^{*}}@cau.ac.kr

MBTI-based Collaborative Recommendation System : A Case Study of Webtoon Contents

Myeong-Yeon Yi[○], O-Joun Lee^{*}, Min-sung Hong^{*}, Jason J.Jung^{*}

[○]Dept. of Computer Science and Engineering, Chung-Ang University, School of Computer Science and Engineering, Chung-Ang University

● Abstract ●

웹툰의 양은 방대하여 사용자가 원하는 웹툰을 찾는 데 어려움이 있기 때문에 체계적인 추천 시스템이 필요하다. 하지만 기존의 추천 시스템은 조회수가 많은 인기 웹툰을 추천하는 방식과 사용자와 비슷한 연령대, 성별의 사용자들이 조회한 콘텐츠를 추천해주는 인구 통계학적 추천(demographic filtering) 방식, 그리고 비슷한 사용자를 분석하여 추천해주는 협업적 추천(collaborative filtering) 방식에 국한되어 있어, 개인의 성향을 반영하여 추천하고 있다고 보기 어렵다. 따라서 사용자 개인의 성향을 분석하는 방식에 대한 시도가 필요하다. 본 연구에서는 이러한 한계를 극복하기 위해서 개인의 성향을 분석하는 지표로 MBTI(Myers-Briggs Type Indicator) 유형을 이용하고, 같은 MBTI 유형의 사용자간의 협업적 필터링 추천 방식을 제안하였다. 또, 협업적 필터링 방식에서 발생하는 콜드 스타트 문제와 초기 평가자 문제를 해결하는 방안을 제시하였다.

키워드: 웹툰(webtoon), 추천(recommendation), MBTI(Myers-Briggs Type Indicator), 협업 필터링(collaborative filtering)

I. Introduction

야후코리아의 ‘카툰 세상’은 2002년에, 다음의 ‘만화 속 세상’은 2003년에 정식 출시한 데에 반해, 비교적 늦은 2004년에 출시한 네이버 웹툰도 2014년에 누적 연재된 작품 수는 520편을 기록하였다 [1]. 웹툰 플랫폼 기업 티파스미디어는 서비스 1년만에 누적 작품 편수가 2만1500개에 이른다 [2]. 이처럼 공급되는 웹툰의 양은 방대하여 사용자가 원하는 웹툰을 찾는 데 어려움이 있기 때문에 체계적인 추천 시스템이 필요하다.

하지만 기존의 추천 시스템은 조회수가 많은 인기 웹툰을 추천하는 방식과 사용자와 비슷한 연령대와 성별을 갖는 다른 사용자들이 조회한 콘텐츠를 추천해주는 인구 통계학적 추천(demographic filtering) 방식, 그리고 유사한 사용자를 분석하여 추천해주는 협업적 추천(collaborative filtering) 방식에 국한되어 있다.

단순 조회수와 평점, 인구 통계학적인 정보를 기반으로 하는 기존 추천방식은 개인의 성향을 반영하여 추천하고 있다고 보기 어렵다. 협업적 추천 방식은 사용자가 어떤 웹툰을 봤을 때 그 웹툰을 본 다른 사용자들이 많이 본 웹툰들을 추천하는 정도에 그치기 때문에, 이 또한 개인의 성향을 반영하고 있지 않다.

따라서 웹툰 사용자의 개인적 성향을 분석하여 추천에 반영 할

수 있는 연구가 필요하다.

II. 관련 연구

1. 기존 웹툰 추천 사례

Fig. 1은 Naver에서 운영하는 웹툰 서비스의 추천 플랫폼이다. 조회수 또는 평점이 높은 인기 웹툰을 추천하는 방식과 사용자와 비슷한 연령대, 성별의 사용자들이 조회한 콘텐츠를 추천해주는 인구 통계학적 추천방식, 그리고 비슷한 사용자를 분석하여 추천해주는 협업적 추천 방식에 국한되어 있다. 이 방식들은 사용자 개인의 성향과 취향을 반영하지 못하는 한계점을 갖는다.

이러한 한계점을 해결할 문헌적 연구 분석 사례는 없다. 또한 웹툰 콘텐츠 자체의 분석 정보를 활용한 추천 연구 사례는 없다.



기존 웹툰 추천 시스템

2. MBTI 유형

MBTI(Myers-Briggs Type Indicator)는 성격에 따른 선호를 측정하는 대표적인 심리검사이다. MBTI는 C.G.Jung의 인간의 행동의 차이는 개인이 인식(Perception)하고 판단(Judgement)하는 특징이 다르기 때문이라는 심리유형론을 근거로 한다.

MBTI는 인간의 선호를 E (Extroversion) & I (Introversion)와 S (Sensing) & N (Intuition), T (Thinking) & F (Feeling), J (Judging) & P (Perceiving) 4개지 척도로 측정한다. 이를 다시 조합으로써 16가지의 성격 유형이 가능하다 (ISFJ, ISFP, ISTJ, ISTP, INFJ, INFP, INTJ, INTP, ESFJ, ESFP, ESTJ, ESTP, ENFJ, ENFP, ENTJ, ENTP) [3].

3. MBTI 유형을 이용한 추천 사례

3.1 영화 추천 연구 사례

송미나 등은 MBTI 유형을 이용하여 영화 추천 시스템에 관련된 연구를 진행하였다 [4]. 이 연구에서는 MBTI 유형이 같은 사람이 영화 취향이 비슷하지 알아보기 위해, 피실험자에게 영화를 보여준 후 감정을 반영하는 단어에 대한 선택과 평점을 주도록 유도하였고, 결과를 피실험자의 MBTI 유형에 따라 분류하였다.

결과, MBTI 유형 별로 피실험자들은 비슷한 단어를 선택하였다. 또한 굉장히 큰 인기를 얻은 영화가 아니라면 유형별로 영화에 대한 선호도가 비슷하다는 것을 발견하였다. 이를 토대로 같은 MBTI 유형의 사용자는 한 영화를 보고 비슷한 감정과 만족감을 보인다는 것을 제시하였고, MBTI 유형을 영화 추천 지표로 설정함이 합당하다는 결론을 이끌어 냈다.

3.2 음악, 책, 영화의 장르 추천 사례

사용자에게 음악, 책, 영화의 장르를 추천하는데 있어서도 MBTI 유형을 적용하려는 연구가 있었다. Animesh Pandey는 MBTI 유형에 따라 많이 접하는 음악, 책, 영화의 장르를 확인하기 위해

“personalitycafe.com”, “typologycentral.com”, 페이스북 그룹과 “reddit.com”을 통해 온라인 설문을 진행하였다. 그리고 그 결과를 바탕으로 사용자의 MBTI 유형에 따라 장르를 추천해주는 방식을 제안하였다 [5].

4. 웹 상에서 콘텐츠 추천 연구 사례

웹 상에서 콘텐츠를 추천해 주는 시스템은 사용하는 정보의 특성에 따라 내용 기반 추천(contents-based filtering), 인구 통계학적 추천(demographical filtering), 협업적 추천(collaborative filtering)으로 나눌 수 있다.

내용 기반 추천은 사용자에게 있어 이전에 선호한 항목과 비슷한 특성을 갖는 항목을 사용자 취향에 맞을 가능성이 높다고 보고 선호된 항목들의 속성을 이용하여 추천하는 방식이다. 인구 통계학적 추천은 사용자의 나이, 성별, 생활수준 등과 같은 인구 통계학적 정보를 바탕으로 항목의 선호도를 추정하는 방법이다. 협업적 추천은 특정 사용자가 선호하는 항목들과 다른 사용자들이 선호하는 항목들을 바탕으로 항목의 선호도를 추정하는 방법이다 [6].

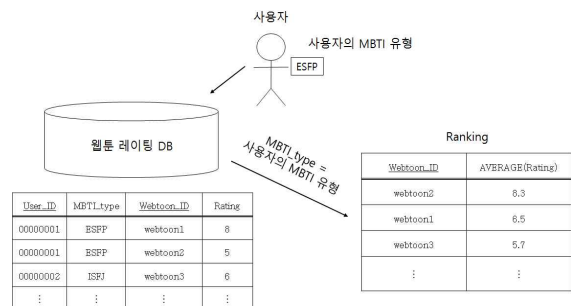
본 연구에서는 MBTI를 기반으로 유사한 성향의 사용자를 뽑고 이들이 선호한 콘텐츠를 추천하는 협업적 추천 기법을 제안한다.

III. MBTI 유형을 이용한 협업 필터링 시스템

제안한 MBTI 기반의 협업 필터링 시스템은 사용자의 MBTI 유형을 입력하면 같은 MBTI 유형의 사용자들이 선호한 웹툰을 추천해 준다.

Fig. 2에서 웹툰 레이팅 데이터베이스는 사용자의 식별자와 MBTI 유형, 레이팅을 매긴 웹툰의 식별자, 그리고 선호도를 나타내는 레이팅을 속성으로 갖는다. 즉, 설계된 데이터베이스의 각 레코드는 MBTI를 갖는 한 사용자의 한 웹툰에 대한 선호도를 의미한다.

사용자에게 웹툰을 추천하기 위해, 먼저 웹툰 레이팅 데이터베이스에서 사용자와 같은 MBTI 유형의 행들을 모두 찾는다. 그리고 그 행들의 ‘Rating’값을 추출한다. (웹툰 레이팅 데이터베이스를 수집하는 방법은 아래 III.1, III.3.1에서 논한다.) 그 후 각 웹툰의 ‘Rating’값의 평균을 내서 정렬하고, 상위 N개를 추천하는 Top-N 방식을 이용하여 추천한다.



웹툰 레이팅 데이터베이스를 이용하여 추천하는 방법

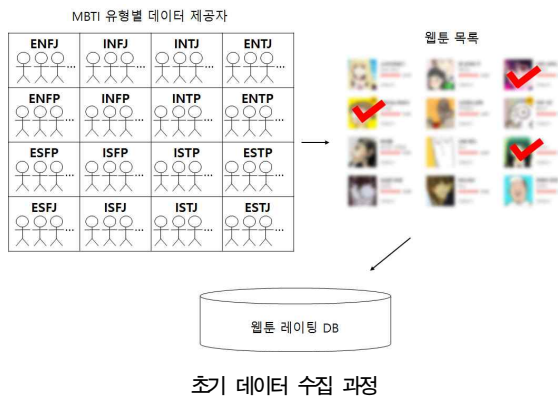
협업 필터링 방식은 두 가지 문제가 있다. 첫 번째는 처음이기

때문에 어떤 아이템도 사용자에게서 평가받지 못한 상태여서 발생하는 콜드 스타트 문제(Cold-Start Problem)와 두 번째는 새로운 아이템이 나타나 이 아이템은 어떤 사용자에게서도 평가받지 못하여서 발생하는 초기 평가자 문제(First-Rater Problem)가 있다. 이를 보완하기 위한 방법은 아래 III.1, III.2에서 논할 것이다.

1. 초기 데이터 수집

콜드 스타트 문제는 Fig. 3과 같은 방법의 초기 데이터를 수집을 통해 완화시킬 수 있다.

우선 각 MBTI 유형 별로 다수의 데이터 제공자에게 여러 웹툰이 있는 웹툰 목록을 보여주고 이전에 봤던 웹툰인지 여부를 떠나 선호하는 웹툰을 복수로 고르게 한다. 이렇게 골라진 웹툰은 ‘check’라는 레이팅 값을 갖고 데이터에 누적된다. 여기서 데이터 누적 형태는 III.3.2에서 다룬다. 시스템 관리자는 데이터 제공자수를 높게 잡거나 ‘check’값을 바꿔가며 이 추천 시스템의 유효성을 증가시킬 수 있다. 이렇게 초기 데이터를 수집한 후, 사용자에게 웹툰 추천 시스템을 제공함으로써 협업적 추천 시스템의 주된 문제인 콜드 스타트 문제를 어느 정도 완화할 수 있다.



2. 팝업 선호도 조사

신규 웹툰의 경우 이전에 평가된 바가 없어 추천이 되지 않고, 이에 따라 평가에서 배제되어 발생하는 초기 평가자 문제는 다음과 같은 방식을 통해 해결할 수 있다.

웹툰이 새로 나오면 모든/일부 사용자에게 이 웹툰을 선호하는지 아닌지를 팝업 선호도 조사를 통해 물어본다. 여기서 선호는 위 III.1에서와 같이 이전에 봤던 웹툰인지 여부와 상관없이 체크 받는다. 선호한다면 Rating값은 ‘check’가 되어 웹툰 레이팅 데이터베이스에 추가가 되고, 선호하지 않는다면 추가하지 않는다.

3. 데이터베이스 관리

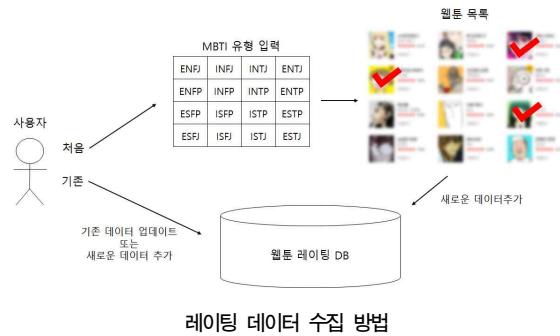
3.1 사용자의 레이팅 수집

초기 데이터 수집 이후에는 사용자가 추천 시스템을 이용할 수 있다. 여기서 사용자는 Fig. 4.에서처럼 크게 처음 사용자와 기존 사용자 둘로 나뉜다.

처음 사용자는 위 III.1의 방법과 같이 사용자의 MBTI 유형을

입력한다. 그리고 웹툰 목록에서 이전에 봤는지 여부를 떠나 선호하는 웹툰을 복수로 고르게 한다. 선택된 웹툰은 ‘check’값을 레이팅으로 하여 웹툰 레이팅 데이터베이스에 저장된다.

기존 사용자는 처음 시스템을 이용했을 때 선호한다고 선택한 웹툰 중에, 이전에 본 웹툰이 있다면 그것의 레이팅을 업데이트 할 수 있다. 또한 이후 웹툰을 본 후 새로운 레이팅을 추가할 수 있다.



3.2 데이터베이스 저장 형태

Table 1.은 웹툰 레이팅 데이터베이스의 생성 쿼리 및 예시이다. 한 개의 레코드는 한 사용자가 한 웹툰에 대해 평가한 평점에 대한 정보를 나타낸다.

데이터베이스 저장 형태

```

CREATE TABLE webtoonRatingDB(
    User_ID CHAR(8) NOT NULL,
    MBTI_type CHAR(4),
    Webtoon_ID CHAR(8) NOT NULL,
    Rating INT,
    PRIMARY KEY(User_ID, Webtoon_ID),
    CHECK (Rating >=0 AND Rating<=10)
);
    
```

User_ID	MBTI_type	Webtoon_ID	Rating
00000001	ESFP	webtoon1	8
00000001	ESFP	webtoon2	5
00000002	ISFJ	webtoon3	6
00000001	ESFP	webtoon3	check
⋮	⋮	⋮	⋮

따라서 사용자의 고유 식별자인 User_ID와 웹툰의 고유 식별자인 Webtoon_ID의 합성키가 기본키가 된다. MBTI_type에는 사용자의 MBTI 유형을 저장한다. Rating은 0~10까지의 정수 값을 갖는다. 처음 시스템을 이용하였을 때 선호한다고 답했던 웹툰에 관해서는 Rating을 ‘check’값으로 저장한다.

IV. Conclusions

기존에 웹툰을 추천하는 방식은 조회수 또는 평점이 높은 인기 웹툰을 추천하는 방식과 사용자와 비슷한 연령대, 성별의 사용자들이 조회한 콘텐츠를 추천해주는 인구 통계학적 추천방식, 그리고 비슷한 사용자를 분석하여 추천해주는 협업적 추천방식에 국한되어 있다. 이 방식들은 사용자 개인의 성향과 취향을 반영하지 못한다는 한계점을 갖고 있다.

이러한 한계점을 극복하기 위해, 본 논문에서는 사용자들을 개인의 성향을 분류하는 지표로 MBTI 유형을 이용하고, 이를 기반으로 유사한 사용자들의 웹툰리스트를 이용하는 협업적 추천 방식을 제안하였다. 또한 협업적 추천 방식을 이용함으로써 발생하는 콜드 스타트 문제와 초기 평가자 문제를 해결하는 방안을 제시하였다.

하지만 이 방법은 협업적 필터링 방식을 이용했기 때문에 웹툰 콘텐츠 자체에 대한 분석이 없고 단순히 기호 가치들(preference values)에만 의존한다는 점에서 한계가 있다.

이러한 한계를 개선하기 위해, 향후에는 웹툰 콘텐츠 자체에 대한 분석 연구가 필요하다. 또, 본 논문에 제안된 시스템의 유효성을 측정하고 개선할 연구가 필요하다.

사사

이 논문은 2015년 미래창조부 및 정보통신기술진흥센터의 재원으로 서울이코드활성화사업의 지원을 받아 수행된 연구임(R0613-15-1 205). 또한, 이 논문은 2014년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2014R1A2A2A050 07154).

References

- [1] Naver campaign, <http://campaign.naver.com/webtoon/>
- [2] Daum ruliweb news, <http://m.ruliweb.daum.net/news/536/view/MD20140105181709558.daum>
- [3] P.H. Harasym, E.J. Leong, B.B. Juschka, G.E. Lucier, and F.L. Lorcheider, "Myers-Briggs Psychological Type and Achievement in Anatomy and Physiology," *The American Journal of Physiology*, pp. 61-65, 1995.
- [4] M.N. Song, H. Namgoong, H.G. Kim, and J.H. Eune, "A Proposed Movie Recommendation Method Using Emotional Word Selection," *The Journal of Online Communities and Social Computing*, Vol. 5621, pp. 525-534, 2009.
- [5] Animesh Pandey, "Idea of a new Personality-Type based Recommendation Engine," 2013.
- [6] J.P. Michael, "A Framework for collaborative, content-based and Demographic Filtering," *Artificial Intelligence Review*, Vol. 13, No. 5-6, pp. 393-408, 1999.