

시니어 인지반응과 온라인 활동 이력을 활용한 개인화 추천 시스템 설계 *

윤유동, 지혜성, 임희석*

*고려대학교 컴퓨터학과

e-mail:2015010492@korea.ac.kr

Personalized Recommendation System Design Using Senior Recognition Response and Online Activity History

You-Dong Yun, Hye-Sung Ji, Heui-Seok Lim*

*Dept. of Computer Science and Engineering, Korea University

요 약

최근 통신 기술의 발달로 온라인을 통한 대규모 콘텐츠의 유통이 가능해졌으나, 사용자들은 수많은 콘텐츠 사이에서 원하는 정보를 찾는 시간이 단축되는 것을 원했다. 이로 인해 다양한 분야에서 개인화된 콘텐츠를 추천해주는 추천 시스템(recommendation system)에 대한 요구가 점차 높아졌다. 그럼에도 불구하고 시니어를 위한 추천 시스템에 대한 연구는 매우 부족하다. 또한, 시니어 세대의 변화에 따라 시니어 관련 콘텐츠 연구도 다양하게 진행되고 있으나, 스마트 기기 및 서비스가 젊은 층에 친화적으로 개발됨으로써 시니어 층의 접근성을 감소시키고 있다. 이에 본 연구에서는 다양한 신체적 변화를 겪는 시니어 세대 위해 추천 시스템에서 인지반응 데이터를 이용하여 콘텐츠를 시청하기 적합한 환경을 제공함과 동시에 활동 이력을 중심으로 개인화 추천 시스템을 설계하여 시니어 사용자들의 개념 변화(concept drift) 문제로 사용자가 원하지 않는 콘텐츠를 추천받을 가능성을 줄일 수 있도록 한다.

1. 서론

최근 통신 기술의 발달로 웹페이지를 통해 원하는 콘텐츠를 언제 어디서든 쉽게 소비할 수 있는 환경이 조성되면서 온라인을 통해 대규모 콘텐츠의 유통이 가능해졌다 [1]. 그러나 수많은 콘텐츠 사이에서 원하는 정보를 쉽게 찾을 수 없는 현상이 발생되었고, 사용자들은 이러한 과정에 소비되는 시간이 단축되기를 원했다 [2]. 이로 인해 다양한 분야에서 다양한 사용자에게 적합한 콘텐츠를 선별해주는 추천 시스템(recommendation system)에 대한 요구는 점점 더 높아지고 있다 [3,20]. 그럼에도 불구하고 시니어를 위한 추천 시스템에 대한 연구는 매우 부족한 상황이다.

통계청에 따르면, 우리나라의 2015년 고령화 지수는 13.1%로 이미 고령 사회로 진입한 것으로 나타났다. 스마트시대와 함께 고령 인구의 비율이 상승하게 되면서 시니어 세대 역시 스마트한 노년기를 보내는 것에 대한 관심이 높아지고 있다 [4]. 특히 최근 시니어 세대는 의료기술의 발달과 함께 막대한 정보 소비에 빠른 속도로 적응하여 과거와는 다르게 건강한 신체와 생각을 바탕으로 활동적인

사회와의 교류 행태를 보이고 있다 [5]. 이와 같은 변화에 따라 온라인에서 시니어에 관련된 실버콘텐츠도 다양한 분야에서 등장하고 있다. 그러나 스마트 기기 및 서비스가 현재 주요 소비층인 젊은 층에 친화적으로 개발됨으로써 시니어 층의 접근성을 감소시키고 있다 [6]. 이에 따라 지속적으로 증가하는 스마트 시니어 세대에게 차별 없고 원활한 콘텐츠 제공을 위한 시스템에 대한 요구는 점점 더 높아지고 있는 상황이다.

온라인에서의 추천 시스템은 사용자가 구매한 콘텐츠의 특성이나 사용자의 환경으로부터 만들어진 모델을 기반으로 사용자가 아직 고려하지 않은 콘텐츠나 선호도를 예측하여 사용자에게 콘텐츠를 제공해주는 일종의 정보 필터링 시스템이다 [7]. 이처럼 기존의 추천 시스템은 사용자의 선호도를 기반으로 콘텐츠를 제공해주지만, 시니어 세대는 고령화에 따라 다양한 신체적 변화를 겪게 된다. 시니어 세대는 인지 및 감각적 능력의 감소로 인해 스마트 기기의 사용을 꺼려하는 경향이 있으며, 감각적 능력의 감소로 인해 스마트 기기의 화면 인식이 어렵기 때문에 스마트 기기에서 제공하는 정보를 받아들이기 힘들다는 연구결과가 있다 [18,19]. 그렇기 때문에 시니어 세대를 위한 추천 시스템은 개인의 선호도뿐만 아니라 인지반응까지 고려하여 시청 환경까지 맞춤형 콘텐츠 제공이 필요하다 [8].

이에 본 논문에서는 다양한 신체적 변화를 겪는 시니어

* **Acknowledgement.** 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2016년도 문화기술 연구개발 지원 사업으로 수행되었음. [2016. 스마트 시니어세대의 문화향유를 위한 인지반응 맞춤형 UI/UX기술 개발]

세대를 위하여 추천 시스템에서 인지반응 데이터를 이용하여 콘텐츠를 시청하기 적합한 환경을 만들어준다. 그리고 사용자의 활동 이력을 중심으로 개인화 추천 시스템을 설계하여 시간이 지남에 따라 변화하는 사용자들의 선호도 문제, 즉 개념 변화(concept drift) 문제로 사용자가 원하지 않는 콘텐츠를 추천받을 가능성을 줄일 수 있도록 한다 [21]. 본 연구를 통해 시니어 사용자의 활동 성향을 파악하고, 활동 성향에 따라 적합한 콘텐츠를 제공해 줄 수 있을 뿐만 아니라, 인지반응에 따라 콘텐츠를 시청하기에 가장 적합한 환경까지 제공해 줄 수 있다.

2. 관련 연구

2.1. 시니어 인지반응

시니어 세대는 나이가 들면서 인체의 생리기능이 쇠퇴기에 들어가는 시기의 세대이다. 이러한 생리기능의 쇠퇴는 시니어 세대가 온라인 환경에서 콘텐츠를 사용함에 있어서 어려움을 유발할 수 있다 [6]. 그렇기 때문에 시니어 세대의 인지반응을 면밀히 파악하고 이에 따라 콘텐츠 시청에 적합한 환경을 제공해 주는 것은 시니어 문화 향유에 중대한 영향을 끼칠 것으로 보인다.

대표적인 생리기능의 감퇴로는 시각기능의 감퇴, 청각기능의 감퇴, 반응의 둔화 등이 있을 수 있다. 첫 번째로 시각기능의 감퇴는 시선과 시야가 좁아지고, 원근감 파악이 어려워지고, 색채감각이 감소하여 색의 차이 구별이 어려워지고, 가시거리가 감소하는 등의 특징이 있다 [8,9,10]. 두 번째로 청각기능의 감퇴는 저음보다 고음이 듣기 어려워지고, 소리가 선명하게 들리지 않는 등의 특징이 있다 [8,9,10]. 마지막으로 반응의 둔화는 생리기능, 특히 신경계 기능저하에 의한 속도가 요구되는 기본적 처리 능력이 저하되고, 정신적인 반응의 저하, 새로운 사물에 대한 적응에 시간이 걸리는 등의 특징이 있다 [8,10].

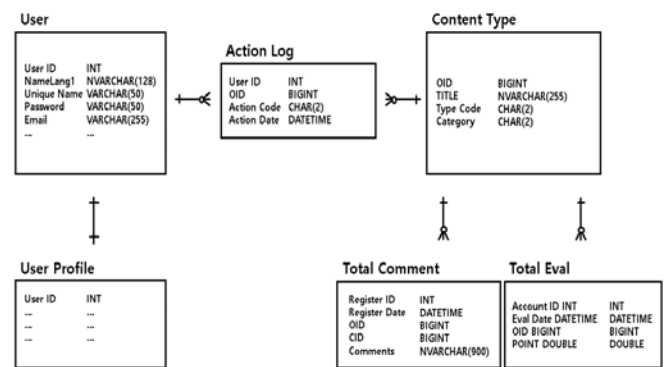
2.2. 온라인 활동 이력

온라인 활동 이력은 웹 로그(web log)의 한 부분으로서, 웹페이지에서 사용자들의 콘텐츠를 구매하거나 댓글을 작성하거나 게시물을 게시하는 등의 활동을 통해 발생하는 데이터이다. 웹사이트에서 정보 획득을 목표로 활동을 하는 사용자들은 웹 로그의 형태로 흔적을 남기게 되는데, 이러한 데이터를 기반으로 사용자 모델링(modeling)을 실행할 수 있다 [11].

이병주(2014)는 온라인 활동 이력으로 접속, 콘텐츠 시청 등의 활동 로그를 활용한 연구를 실시했다. 이한진(2011)은 콘텐츠 시청, 콘텐츠 평가, 콘텐츠 게시, 콘텐츠 시청 등의 활동 로그를 활용한 연구를 실시했다. 장상호(2010)는 콘텐츠 시청, 방문 시간 등의 활동 로그를 활용한 연구를 실시했다. 이동철(2011)은 콘텐츠 구매 등의 활동 로그를 활용한 연구를 실시했다. 최성이(2014)는 콘텐츠 카테고리를 주요 분석 요인으로 사용했다.

3. 온라인 활동 데이터 설계

다음 (그림 1)에 나타나 있는 본 연구에서 설계된 온라인 활동 데이터에는 사용자(User), 콘텐츠 유형(Content Type)의 2가지 개체(entity)가 있다. 사용자 개체에서는 사용자 ID(User ID), 비밀번호(Password) 등의 개인정보 관련 속성(attribute)으로 구성되어있으며, 사용자 프로파일(User Profile)과 일대일(1:1) 관계를 가지고 있다. 콘텐츠 유형은 콘텐츠 ID(OID), 콘텐츠 이름(Title), 콘텐츠 유형 코드(Type Code), 콘텐츠 카테고리(Category) 속성으로 구성되어있으며, 사용자와 콘텐츠 유형 개체는 다대다(N:M) 관계를 가지고 있다. 두 개체 사이에 있는 관계(Relationship)인 활동 로그(Action Log)는 사용자와 콘텐츠 유형의 기본 키(primary key)인 사용자 ID, 콘텐츠 ID와 활동 코드(Action Code), 활동 일시(Action Date) 속성으로 구성되어있다.



(그림 1) 온라인 활동 데이터 설계도

콘텐츠 유형에서 하위 개체로 구분되는 댓글(Total Comment)에서는 콘텐츠 유형의 기본 키인 콘텐츠 ID와 댓글 ID(CID), 작성자 ID(Register ID), 작성 일시(Register Date), 댓글 내용(Comments)로 구성되어있으며, 평가(Total Eval) 개체에서 역시 콘텐츠 유형의 기본 키인 콘텐츠 ID와 계정 ID(Account ID), 평가 일시(Eval Date), 평가 점수(POINT)로 구성되어있다.

Action Log				Total Comment				
User ID	OID	Action Code	Action Date	Register ID	OID	CID	Register Date	Comment
1	1	BU(Buy)	2016.08.18	1	32	1	2016.08.18	Good Content!
2	2	CO(Comment)	2016.08.26	2	27	7	2016.08.26	Amazing!
3	3	EV(Evaluation)	2016.08.31	3	45	4	2016.08.31	What is this?...
...

Content Type				Total Eval			
OID	Title	Type Code	Category	Account ID	OID	Eval	Eval Date
1	Economy	EB(E-Book)	EC(Economy)	1	1	4	2016.08.18
2	Easy Society	GA(Game)	SO(Society)	2	2	5	2016.08.26
3	Baseball Today	MU(Music)	SP(Sports)	3	3	2	2016.08.31
...

(그림 2) 온라인 활동 데이터 수집 예시

상단의 (그림 2)에서 온라인 활동 데이터가 어떤 방식으로 수집되는지 예시를 통해 살펴볼 수 있다. 활동 코드

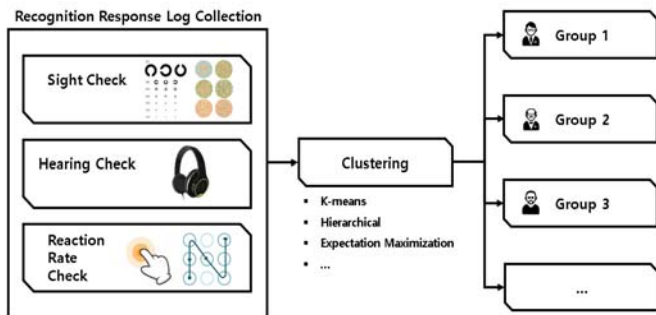
는 웹페이지에서 사용자가 취할 수 있는 활동을 의미하며, 사용자는 구매, 평가, 댓글, 게시/등록 등의 활동을 취할 수 있다. 유형 코드는 콘텐츠의 유형을 의미하며, 콘텐츠 유형으로는 전자책, 게임, 음악, 웹, 뉴스, 사진, 동영상 등이 있다. 콘텐츠 카테고리는 콘텐츠가 속해있는 범주를 의미하며, 콘텐츠 카테고리의 종류로는 정보통신, 사회/정치, 과학, 스포츠, 경제, 연예/문화 등이 있다.

4. 개인화 추천 시스템 설계

4.1. 인지반응 데이터 모델링

인지반응 데이터 모델링은 온라인 활동에 앞서 설문지, 게임과 같은 다양한 인지반응 측정 콘텐츠를 이용하여 수집된 시각, 청각, 반응 속도 등의 인지반응 데이터를 클러스터링(clustering) 알고리즘을 사용하여 인지반응 유형에 따라 군집을 분류하는 것을 의미한다. 인지반응 데이터 모델링 과정은 인지반응 유형에 따라 가장 적절한 콘텐츠 시청 환경 제공을 위해 수행되는 과정이다. 다음 (그림 3)에서 대략적인 모델링 과정이 나타나 있다.

만약 시각기능에 이상이 있는 유형의 군집에게는 글씨 크기 확대, 글씨 색 변경 등의 맞춤형 콘텐츠 시청 환경을 제공할 수 있으며, 청각기능에 이상이 있는 유형의 군집에게는 기본 소리 크기 조절 등의 맞춤형 콘텐츠 시청 환경을 제공할 수 있다.



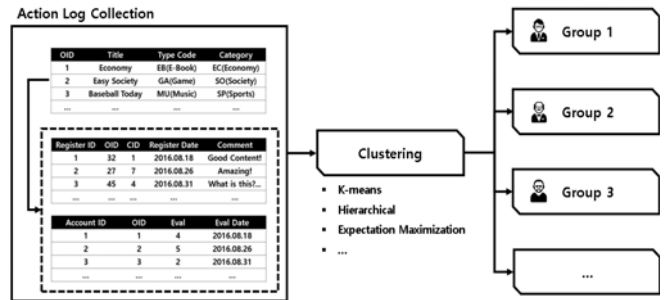
(그림 3) 인지반응 데이터 모델링 프로세스

4.2. 활동 데이터 모델링

활동 데이터 모델링은 설계된 온라인 활동 데이터에 따라 수집된 활동 로그들을 기반으로 다양한 클러스터링 알고리즘을 통해 모델링 과정을 거쳐 활동 유형에 따라 군집을 분류하는 것을 의미한다. 온라인 활동 기반의 사용자 모델링 과정은 활동 유형이 같은 사용자 집단을 통해 목표 사용자에게 콘텐츠를 추천하기 위해 수행되는 과정이다. 다음 (그림 4)에서 대략적인 모델링 과정이 나타나 있다.

데이터 모델링을 위해서 사용하는 클러스터링 알고리즘으로는 K-means, Hierarchical, Expectation Maximization 등의 클러스터링 알고리즘을 사용한다. 알고리즘마다 방법과 특징이 다르기 때문에 다양한 알고리즘을 반영하여 신뢰도 있는 모델링을 수행할 수 있도록 한다.

사용자 유형 모델링 결과는 어떤 활동을 많이 하는가에 따라 최적화된 k개의 군집으로 분류되어 나타나게 된다. 모델링 결과 분류된 군집들은 각각의 군집별로 다른 특징을 가지고 있기 때문에 모델링 결과를 바탕으로 사용자마다 적합한 콘텐츠 시청 환경 또는 콘텐츠를 제공하는 기반이 마련될 수 있다.

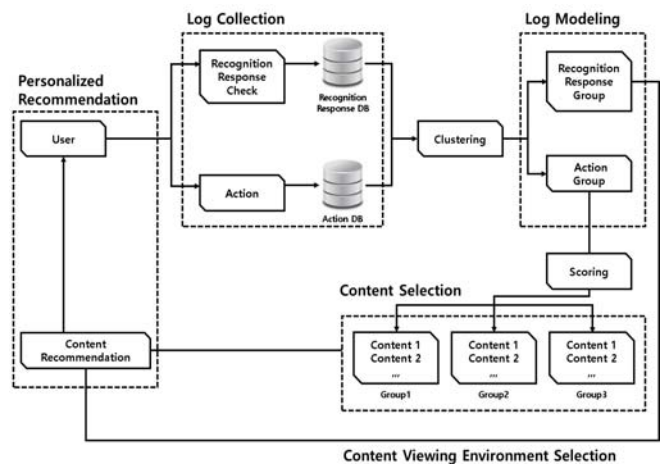


(그림 4) 활동 데이터 모델링 프로세스

4.2. 개인화 콘텐츠 추천

개인화 콘텐츠 시청 환경 추천은 인지반응 모델링 과정을 거쳐 유사한 인지반응 유형을 보이는 사용자들을 군집화하고, 목표 사용자의 인지반응에 따라 적합한 콘텐츠 시청 환경을 제공해준다.

개인화 콘텐츠 추천은 수집된 로그(Log Collection) 데이터를 기반으로 인지반응, 활동 데이터 별로 모델링 과정을 거쳐 유사한 활동 패턴을 보이는 사용자들을 군집화하게 된다(Log Modeling). 이렇게 구축된 인지반응 데이터 군집을 통해 시니어 사용자들의 맞춤형 콘텐츠 시청 환경을 제공해 주며(Content Viewing Environment Selection), 활동 데이터 군집들을 이용하여 같은 군집의 사용자들이 시청하거나 평가했던 콘텐츠의 점수를 계산한다(Content Selection). 점수가 높은 상위 N개의 콘텐츠를 선정 이후 선정된 장르의 콘텐츠를 목표 사용자에게 제공해준다(Personalized Recommendation) [17]. 다음 (그림 4)에서 본 연구의 개인화 추천 과정이 자세하게 나타나 있다.



(그림 4) 개인화 콘텐츠 추천 프로세스

5. 결론

최근 통신 기술의 발달로 콘텐츠로의 접근성이 향상되어 다양한 분야에서 개인화된 콘텐츠를 추천해주는 추천 시스템에 대한 요구가 점차 높아졌다. 개인화 추천 시스템은 사용자에게 맞춰진 콘텐츠를 제공해야하기 때문에 사용자가 남긴 웹 로그에서 선호도 파악을 어떻게 하는지가 매우 중요하다. 그런데 기존의 개인화 추천 연구에서는 대부분 사용자로부터 직접 입력받은 정보나 사용자의 콘텐츠 선택 정보를 기반으로 연구가 진행되었다. 그러나 사용자들의 선호도는 개념 변화 문제로 인해 유사한 사용자들이 선택한 콘텐츠나 그 콘텐츠에 대한 평가를 기반으로 콘텐츠 추천 시, 사용자가 원하지 않는 콘텐츠일 수 있다. 이러한 문제 해결을 위해 본 연구에서는 온라인에서 사용자의 활동 이력을 기반으로 활동 성향을 파악하고, 사용자에게 맞춰진 콘텐츠를 제공하는 방안을 제안하였다.

본 연구에서 제안하는 온라인 활동 이력을 개인화 추천 방안은 웹페이지 사용자들의 활동 로그를 기반으로 사용자들을 모델링하고, 모델링 결과에 따라 개인화된 콘텐츠를 추천해주는 구조이다. 제안된 개인화 추천 방안은 사용자의 활동 로그를 기반으로 콘텐츠를 추천해주기 때문에 기존의 개인화 추천 방안에서 제기되었던 사용자의 선호도 변화 문제 해결을 기대할 수 있다.

현재의 연구는 초기 상태로 시스템을 구성하는 데이터 베이스를 설계하고, 데이터베이스의 활동 로그 데이터를 기반으로 사용자 유형 모델링을 수행하고 있는 단계이다. 향후 수집한 사용자들의 온라인 활동 로그 데이터를 기반으로 모델링을 수행하고, 모델링 결과를 기반으로 사용자에게 맞춰진 콘텐츠를 추천해준다면 개념 변화 문제에 대한 해결책을 마련할 수 있을 것으로 기대한다.

참고문헌

- [1] 김민정 · 박두순 외 (2015), “상황기반과 협업필터링 기법을 이용한 개인화 영화 추천 시스템”, 한국정보처리학회논문지, Vol.4, No.9, pp.289-296.
- [2] 김현모 · 김민용 · 박재홍 (2014), “음원 추천시스템이 온라인 디지털 음원차트에 미치는 과급효과에 대한 연구”, 한국경영정보학회, Vol.16, No.3, pp.49-68.
- [3] 이상민 (2015), “협업적 태그를 이용한 추천시스템 개인화”, 서울대학교 융합과학기술대학원 석사학위논문.
- [4] 이보경 · 김성훈 (2015), “액티브시니어를 위한 스마트 헬스케어기반의 앱 콘텐츠 UX/UI 연구”, 한국디자인문화학회지, Vol.21, No.4, pp.433-445.
- [5] 이동민 · 김정곤 (2014), “노년층의 라이프 스타일 변화에 따른 새로운 주거공간에 대한 UX 디자인 방향 연구”, 디자인지식저널, Vol.30, No.-, pp.135-144.
- [6] 고유미 (2013), “50대 뉴시니어세대를 위한 스마트 앱 TV UI 개선 연구 :태블릿PC용의 메뉴구조와 내비게이션을 중심으로”, 이화여자대학교 대학원 석사학위논문.
- [7] 강호윤 · 옥창수 (2015), “아이템 연관성을 고려한 협력적 필터링 기반 추천시스템 개발”, 엔트루 저널, Vol.14, No.1, pp.135-144.
- [8] 이형준 (2012), “장노년층의 정보접근성 개선을 위한 스마트TV 인터넷 텐저블 UI/UX에 관한 연구”, 광운대학교 대학원 석사학위논문.
- [9] 루오양 (2016), “노인을 위한 텔레비전 UX디자인 연구”, 경희대학교 대학원 석사학위논문.
- [10] 노영서 · 김보연 (2013), “베이비 붐 세대의 고령화를 대비한 UI디자인 사용성 연구”, 디지털디자인학연구, Vol.13, No.1, pp.151-160.
- [11] 서동렬 · 김두진 외 (2015), “웹 사이트 이용 고객의 행동 정보를 기반으로 한 고객 선호지수 산출 방법”, 한국IT서비스학회 학술대회, pp.243-248.
- [12] 이병주 · 권정숙 외 (2014), “효과적인 웹 사용자의 패턴 분석을 위한 하둡 시스템의 웹 로그 분석 방안”, 한국IT서비스학회지, Vol.13, No.4, pp.231-243.
- [13] 이한진 (2011), “디지털 사진 커뮤니티 회원들의 커뮤니케이션 행태 분석 : 네이버 포토갤러리 서비스를 중심으로”, 연세대학교 대학원 석사학위논문.
- [14] 장상호 (2010), “데이터마이닝을 활용한 웹 사이트 분석 : 웹 로그 분석을 중심으로”, 고려대학교 대학원 석사학위논문.
- [15] 이동철 · 이은주 외 (2011), “인터넷 쇼핑의 구매의도 영향요인”, 경영정보연구, Vol.30, No.1, pp.211-226.
- [16] 최성이 · 김남규 (2014), “토픽 분석을 활용한 웹 카테고리별 방문자 관심 이슈 식별 방안”, Journal of information technology applications & management, Vol.21, No.4, pp.415-429.
- [17] 최자현 · 하인에 외 (2011), “협업적 여과 시스템의 성능 향상을 위한 장르 패턴 기반 사용자 클러스터링”, 한국컴퓨터정보학회논문지, Vol.16, No11, pp.17-24.
- [18] Massimi, M., Ronald, M., Michael W. (2007), “Using participatory activities with seniors to critique, build, and evaluate mobile phones.” Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, pp.155-162.
- [19] Melenhorst, A.S., Rogers, W.A., Bouwhuis, D.G. (2006), Older adults’ motivated choice for technological innovation: evidence for benefit-driven selectivity, Psychology and aging, 21(1), 190-195.
- [20] Hung, J. C, Weng, J. D, Chen, Y. H (2016), “A recommendation system based on mining human portfolio for museum navigation”, Evolving Systems, Vol.7, No.2. pp.145-158.
- [21] Yehuda Koren (2009), “Collaborative Filtering with Temporal Dynamics”, KDD '09 Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp.116-142.