

석사학위논문

고객 충성도 개선을 위한 PNA
기반 협업필터링 개인화 추천
시스템

2021년

한 성 대 학 교 대 학 원

산 업 경 영 공 학 과

산 업 경 영 공 학 전 공

김 승 형

석사학위논문
지도교수 김형수

고객 충성도 개선을 위한 PNA 기반 협업필터링 개인화 추천 시스템

A recommendation system based on PNA
collaborative filtering for reinforcing customer
loyalty

2020년 12월 일

한 성 대 학 교 대 학 원

산 업 경 영 공 학 과

산 업 경 영 공 학 전 공

김 승 형

석사학위논문
지도교수 김형수

고객 충성도 개선을 위한 PNA 기반 협업필터링 개인화 추천 시스템

A recommendation system based on PNA
collaborative filtering for reinforcing customer
loyalty

위 논문을 산업경영공학 석사학위 논문으로 제출함

2020년 12월 일

한 성 대 학 교 대 학 원

산 업 경 영 공 학 과

산 업 경 영 공 학 전 공

김 승 형

김승형의 산업경영공학 석사학위논문을 인준함

2020년 12월 일

심사위원장 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

국 문 초 록

고객 충성도 개선을 위한 PNA 기반 협업필터링 개인화 추천 시스템

한 성 대 학 교 대 학 원
산 업 경 영 공 학 과
산 업 경 영 공 학 전 공
김 승 형

서비스 업종뿐만 아니라 일반 소비재 산업에서도 빅 데이터 분석 기술을 활용한 개인화 추천 서비스와 같은 CRM 전략의 중요성이 높아지고 있다. 그러나, 상품 추천 모형에 대한 기존 연구는 대부분 추천 적중률을 높이기 위한 머신러닝 모형 개발과 개선에 집중되어 왔기 때문에 추천 활동 이후 고객과의 중장기적인 관계는 고려되지 못했다. 본 연구에서는 이러한 한계를 극복하기 위하여 소셜네트워크분석(Social Network Analysis) 분석기법을 상품 네트워크 분석 영역에 응용한 PNA (Product Network Analysis) 기법을 협업필터링 모형에 연계하여 단순한 추천 적중률뿐만 아니라 중장기적인 고객의 충성도를 개선할 수 있는 추천 시스템을 제안하고자 한다. CF/PNA (Collaborative Filtering based on Product Network Analysis) 개인화 추천 시스템은 거래 데이터를 기반으로 상품 네트워크를 구성한 후, 다양한 후보 추천상품 중에 PNA 분석 결과를 토대로 중장기적으로 고객 충성도가 향상될 수 있는 상품을 선정하여 최우선적으로 추천하는 시스템이다. 국내 대형 아웃

도어 기업과의 협력으로 실제 고객 데이터를 기반으로 모델링과 검증을 수행한 결과, 모형의 추천 적중률뿐만 아니라, 거래빈도, 재구매율, 교차구매지수 등 거래의 질적수준을 평가하는 고객로열티 지수도 함께 증가함을 확인하였다. 본 연구는 추천 적중률이 좋아도 상품 간의 연계성이 부족하여 지속적인 고객관계를 유지하기 어려운 경우 CRM 관점에서는 오히려 추천을 유보하는 것이 바람직할 수 있다는 전략적인 통찰력을 얻을 수 있다.

[주요어] CF/PNA 개인화 추천 시스템, 상품네트워크분석, PNA, 협업필터링,
CRM

목 차

I. 서 론	1
1.1 연구의 배경	1
II. 문헌 연구	3
2.1 추천 시스템과 협업필터링	3
2.1.1 기술 중심 연구	10
2.1.2 이용자 중심 연구	11
2.2 소셜네트워크분석	14
2.2.1 소셜네트워크분석의 분석수준	15
2.2.2 중심성 분석	16
2.2.3 고객 네트워크와 상품 네트워크	18
2.3 고객로열티	22
2.3.1 태도적 접근관점	24
2.3.2 행동적 접근관점	25
2.3.3 통합적 접근관점	25
III. PNA 기반 협업필터링 개인화 추천 시스템	27
3.1 추천 시스템 개요	27
3.1.1 아이템 유사도 분석	28
3.1.2 상품 네트워크 구성	30
3.1.3 중심성 분석	31
3.1.4 중심성 가중치 적용	32
3.1.5 추천 리스트 생성	32

IV. 실험 및 평가	34
4.1 데이터 수집 및 전처리	34
4.2 추천 시스템 검증방법	35
4.3 검증결과 및 비	36
V. 결 론	39
5.1 연구요약 및 의의	39
5.2 연구의 한계 및 향후 연구방향	40
참 고 문 헌	41
ABSTRACT	50

표 목 차

[표 2-1] 협업필터링 기반 기술 및 이용자 중심 연구	13
[표 2-2] 고객 네트워크와 상품 네트워크 연구	21
[표 2-3] 고객로열티의 개념	23
[표 3-1] 거래 데이터 예시	28
[표 3-2] 고객X상품 구매 행렬 예시	29
[표 3-3] 상품 간 유사도 예시	30
[표 3-4] 상품 네트워크 상의 매개중심성 예시	32
[표 3-5] 매개중심성을 적용한 구매 예측 선호도 예시	32
[표 3-6] 최종 구매 예측 선호도 예시	33
[표 3-7] 추천리스트 생성 예시	33
[표 4-1] 데이터 예시	34
[표 4-2] 아이템기반 협업필터링과 PNA기반 협업필터링의 성능비교	37
[표 4-3] 아이템 기반 협업필터링 대비 개선율	38

그림 목 차

[그림 2-1] 추천 시스템의 종류	5
[그림 2-2] 사용자 기반 협업필터링 알고리즘	8
[그림 2-3] 아이템 기반 협업필터링 알고리즘	8
[그림 2-4] 고객네트워크와 상품네트워크 예시 그림	19
[그림 3-1] PNA 기반 협업필터링 개인화 추천 시스템 프로세스	27
[그림 3-2] 상품 네트워크 다이어그램	31
[그림 4-1] 추천 시스템 성능 비교 프로세스	36

I. 서론

1.1 연구의 배경

빅 데이터와 머신러닝의 비즈니스적인 활용도가 점차 높아지고 있는 가운데 기업들은 고객 빅 데이터를 기반으로 한 비즈니스 예측에 많은 투자를 하고 있다(복경수, 유재수, 2017). 특히, CRM (Customer Relationship Management) 전략의 한 영역으로 간주되었던 고객이탈예측이나 개인화 상품 및 서비스 추천은 상거래 관점의 4차 산업혁명의 핵심 화두로 떠오르고 있으며(복경수, 유재수, 2017), 많은 학술적 연구와 더불어 다양한 아이디어로 산업계에 직접 적용하려고 하는 신기술 기반의 스타트업 역시 폭발적으로 증가하고 있는 추세이다(서일석, 2016).

특히, 개인화 추천 영역은 비교적 단기간에 모델을 개발할 수 있을 뿐만 아니라, 비즈니스 규칙의 결과가 비교적 명확하고, 활용도가 높아 가시적인 경영 성과를 도출할 수 있기 때문에 비즈니스 머신러닝 분야에서 가장 높은 관심을 받고 있는 주제 중 하나이다(김필성, 김기봉, 조영택, 신용태, 2020). 개인화 상품추천을 위해 제시된 다양한 방법론 중 협업필터링은 학술적, 실무적으로 가장 많이 응용되어왔던 머신러닝 기법이다(i.e., 손지은, 김성범, 김현중, 조성준, 2015). 그러나, 협업필터링을 포함한 대부분의 추천 시스템 방법론에 대한 연구들은 주로 알고리즘의 공학적, 시스템적 개선을 통해 추천 모형의 정확도를 개선하는데 초점이 맞추어져 왔다(i.e., 황용석, 김기태, 2019).

개인화 상품 추천은 전략적인 관점에서 고객과의 관계를 유지하고, 강화하기 위한 교차 및 상승 판매 전략의 일환으로 보는 것이 타당하기 때문에 추천 상품에 대한 적중률뿐만 아니라, 추천상품 구매 이후의 연쇄적인 반응을 고려하는 것이 바람직하다. Pu, Chen & Hu(2012)은 이러한 시스템적 접근방식의 추천 알고리즘의 정확도와 효율성의 향상이 반드시 추천 서비스 사용자들의 서비스 채택 여부나 만족도로 이어지지 않는다는 문제점을 제기하였다. 또한, 국내 한 기관의 연구보고서에 따르면 개인화 추천 시스템을 도입한 국내 B2C 기업의

약 70%는 추천 서비스를 도입한 직후 고객 매출이 증가세를 보이다가 최대 2년 이내에 원점으로 회귀하는 현상을 경험했다고 보고하고 있다(한국CRM협회, 2016).

이러한 현상은 추천 시스템을 위한 머신러닝 모형 개발 과정에서 모형의 평가 기준이 지지도(support)나 신뢰도(confidence)와 같은 확률적인 척도나 평균 제곱근오차(RMSE)와 같은 추천 점수에 대한 오차 기준 척도처럼 정확도에 관련된 기술적인 관점에서만 모형의 우수성을 평가하려는 데 기인한다. 즉, 대부분의 추천 모형은 단기적인 관점에서 적중률(hit ratio)이 가장 높은 상품을 추천할 뿐, 그 상품이 갖는 특성이나 다른 상품과의 연관성을 전혀 고려하지 않는 것이다. 그러나, 고객의 생애주기 관점에서 각 관계 진화 단계마다 적절한 상품이나 부적절한 상품이 존재할 수 있고, 심지어 구매 상품에 따라 고객과의 관계가 유지 또는 단절될 수도 있다는 연구 결과에 따라(i.e., 박석철, 김성연, 2009; 서문식, 이지은, 조상현, 2010), 개인화 추천 시스템을 통한 상품 추천은 단순히 구매 적중률만을 고려하는 것이 부적절할 수 있다는 점을 시사하고 있다.

이에 본 연구에서는 개인화 추천 시스템 역시 고객과의 관계를 유지, 강화하고자 하는 일련의 CRM 전략으로 인식하고, 추천 정확도 뿐만 아니라 고객의 충성도를 함께 고려할 수 있는 개인화 추천 시스템을 제안하고자 한다. 구체적으로는 소셜네트워크분석(SNA: Social Network Analysis) 분석기법을 상품 네트워크 분석에 적용한 PNA (Product Network Analysis) 기법을 협업필터링(Collaborative Filtering) 추천 모형에 결합하여 단순히 추천 적중률이 아닌 연쇄적인 구매 파급효과가 큰 상품을 추천하도록 하였다. 상품 네트워크 분석은 고객의 구매패턴을 네트워크로 형상화하여 매개의 중심이 되는 상품을 파악함으로써 구매 활동이 단일 상품의 범주를 넘어 다른 상품들과 연계해서 판매되는 흐름을 파악할 수 있는 방법이다(김혜경, 김재경, Yi, C. Q., 2009). 이러한 상품 네트워크 분석 기법을 개인화 추천 시스템에 응용하면 추천 대상의 상품이 더 이상 독립적인 단일 객체가 아니라, 연쇄적인 구매 행태 네트워크를 고려한 그룹 객체가 되기 때문에 구매 확률뿐만 아니라 구매 이후의 연쇄 파급효과를 높이게 됨으로써 궁극적으로 고객의 반복구매나 교차구매 행동과 같은 고객로열티의 수준이 상승할 것으로 기대된다.

Ⅱ. 문헌 연구

2.1 추천 시스템과 협업필터링

최근 인터넷의 발달로 전자상거래 시장이 성장하면서 기업들이 취급하는 상품의 수가 크게 늘어났으며, 소비자 또한 다양한 정보를 바탕으로 보다 양질의 상품 구매가 가능해졌다. 따라서 다양한 정보들 중 소비자가 선호하고 필요로 하는 것을 자동으로 추천해주는 시스템도 발전하게 되었고, 이러한 추천 시스템은 기업의 수익적 성과를 위해 필수역량으로 자리잡았다. 오늘날 모든 전자상거래 웹사이트는 암시적 또는 명시적으로 사용자의 관심 항목을 발견하고 추천하는 자체 추천 시스템을 보유하고 있다(Tewari & Barman, 2018). 아마존이나 넷플렉스, 스포티파이 등 글로벌 기업들의 전방위적인 활용을 통해 널리 알려지게 된 개인화 추천 시스템은 고객 개인별로 스타일과 선호도를 파악하여 다양한 선호도 데이터를 추출하고, 이를 학습해서 맞춤형 상품이나 서비스를 제공하는 일련의 방법론이다(손지은, 김성범, 김현중, 조성준, 2015). 기업들이 추천 시스템에 관심을 갖고 투자하는 이유는 크게 3가지로 요약된다(Ahn, 2014).

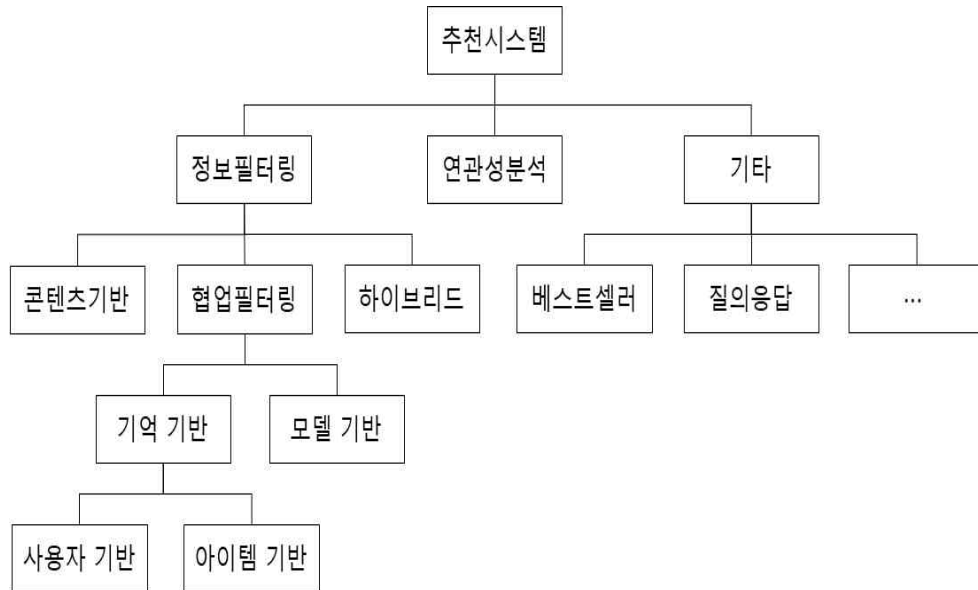
첫 번째, 정보 과부하(Information Overload) 문제를 해결할 수 있다(Schafer, Konstan, & Riedl, 2001). 최근에는 정보의 과부하로 인해 소비자들의 구매의사결정이 쉽지 않다. 많은 양의 데이터 속에서 소비자들은 이를 적절하게 선별하여 이용해야 하지만 정보 과부하로 인해 구매의사결정에 많은 시간이 소요된다. 추천 시스템은 개인의 취향을 고려해 적합한 상품을 추천해 줌으로써, 정보 과부하 문제에 효과적인 대안이 될 수 있다.

두 번째, 대중 맞춤(Mass Customization)을 실현한다(Schafer, Konstan, & Riedl, 2001). 대중 맞춤은 “수많은 고객들 개개인의 욕구를 충족시킬 수

있는 맞춤 상품 혹은 서비스를 제공하는 것”으로 Pine II이 1993년에 제안한 개념이다(Pine II, 1993).

세 번째, 추천 시스템은 판매자와 구매자 모두에게 가치를 제공한다. 판매자 입장에서 추천 시스템은 다음과 같은 3가지 방식으로 전자상거래를 촉진시킨다(Ahn, 2014; Schafer, Konstan, & Riedl, 2001). 첫째는 구매자의 구매 활동을 유도하여, 판매자의 매출을 확대한다는 것이다. 추천 시스템은 고객에게 그들이 가장 선호하는 상품이나 서비스를 적절히 노출 시킴으로써, 구매를 촉진한다. 둘째, 교차판매를 증대시킨다는 것이다. 추천 시스템은 고객에게 추가적으로 다른 적합한 상품을 제안함으로써, 구매를 촉진한다. 셋째, 높은 수준의 고객 충성도를 형성하는데 도움을 줄 수 있다. 추천 시스템은 고객들로 하여금 편리함과 더불어 가치 있는 맞춤 정보를 제공함으로써, 고객의 신뢰를 높일 수 있다. 결과적으로 해당 고객의 충성도를 더욱 증가하는데 도움을 줄 수 있다. 한편 구매자 입장에서는 추천 시스템과 같은 맞춤형 개인화 마케팅 도구는 상품 탐색비용 절감 및 만족 증대와 같은 실질적인 가치를 제공한다(Häubl & Trifts, 2000). 따라서 기업은 고객에게 고객 맞춤형 상품 추천을 하기 위한 추천 알고리즘 개발의 중요성을 인지해야 한다.

추천 시스템의 종류는 그 알고리즘 구현방식에 따라 베스트셀러 기반 추천 시스템, 연관성 분석 기반 추천 시스템, 최소질의대상 상품결정 추천 시스템, 정보필터링 추천 시스템 등으로 구분된다. 다양한 알고리즘의 추천 시스템이 있지만 정보필터링 추천 시스템은 다른 추천기법에 비해 정확도가 높아 가장 많이 활용된다. 정보필터링 추천 시스템은 다시 콘텐츠 기반 추천 시스템과 협업필터링 추천 시스템 그리고 하이브리드 추천 시스템으로 나뉘어 진다.



[그림 2-1] 추천 시스템의 종류

자료원 : 손지은, 김성범, 김현중, 조성준(2015) 인용

콘텐츠 기반(Content-based) 추천 시스템은 과거 고객의 구매 상품을 분석하여 상품들 간의 유사한 특성을 파악하고 고객이 선호하는 상품의 특성을 가진 상품을 추천하는 기법이다. 이 기법의 장점은 상품 자체에의 속성을 이용하여 모델링하는 기법으로 직접적이고 단순하다는 점이다(Wu & Chen, 2000). 그러나 이 기법은 추천을 위해 상품 자체의 속성을 추출해야 되는데, 다양한 상품이 계속적으로 출시되는 상황에서 다양한 종류의 상품 분석이 쉽지 않고 사실상 분석이 효과적으로 이루어지기 어렵다. 그리고 고객의 과거 상품구매이력이나 평가에 의존도가 높아서 고객의 과거형태 지향적인 한계점도 있고 이전에 좋게 평가한 상품과 비슷한 상품 군을 찾는 방식으로 이루어지기 때문에, 과도하게 특정 부분에 치우치게 된다는 점등의 한계점이 있다(Balabanović & Shoham, 1997).

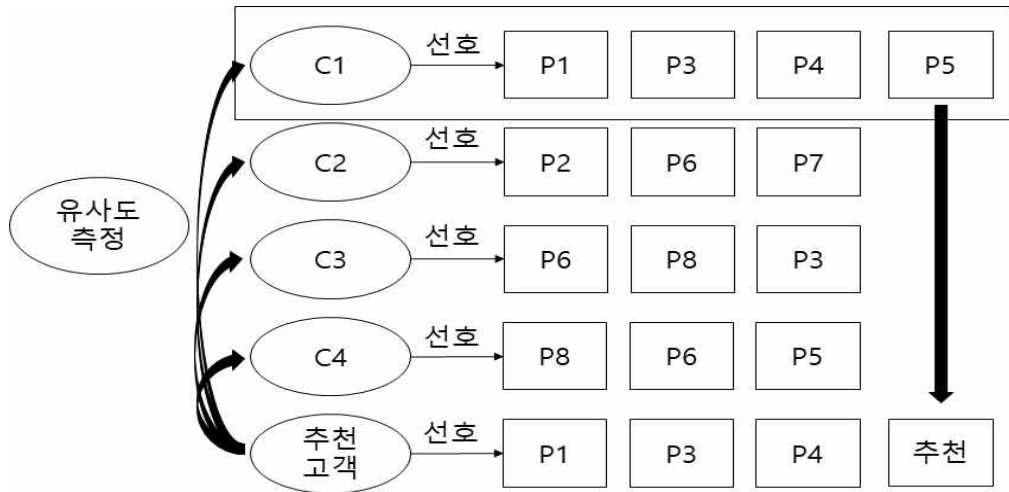
협업필터링(Collaborative Filtering) 추천 시스템은 지금까지 제시된 개인화 추천 시스템 기술 중 가장 많은 학술적 관심을 받아왔으며(i.e., 박

종학, 조운호, 김재경, 2009), 현재까지 가장 우수한 추천기법으로 알려져 있다. 특히, 아마존, 넷플렉스, CDNow 등 수많은 기업들이 적용함에 따라 실무적인 타당성도 상당 부분 증명되고 있다. 1992년 Goldberg에 의해서 처음으로 소개된 협업필터링(Goldberg, Nichols, Oki, & Terry, 1992)은 특정 상품에 대한 유사도가 비슷한 고객들은 다른 상품에 대해서도 비슷한 유사도를 보일 것이라는 가정을 바탕으로 사용자 또는 아이템 간 유사도를 기반으로 선호도를 예측하는 기법이다. 협업필터링은 고객들이 동일한 평가를 내린 상품 군에 대해 상대적으로 높은 예측력을 보이는데, 데이터가 충분할 경우에는 높은 예측력을 보인다는 장점을 가지고 있다(Konstan, Miller, Maltz, Herlocker, Gordon, & Riedl, 1997; Pazzani, 1999). 협업필터링은 사용자가 아이템을 평가한 정보를 이용해 유사도를 예측한다는 점에서 콘텐츠 기반 추천 시스템과 큰 차이점을 보인다. 즉, 추천대상 고객에게 그와 비슷한 취향의 고객을 선정하여 비슷한 취향 고객의 구매 아이템을 기반으로 추천하기 때문에 추천의 다양성이 콘텐츠 기반 추천 시스템보다 높다. 또한 협업필터링 추천 시스템은 결과가 직관적이며 상품의 구체적인 내용을 분석할 필요가 없다는 장점도 있다. 예를 들어, 상품A와 상품B를 구매했다고 하면 상품A와 상품B가 어떤 상품인지 서로 같이 구매되어야하는 관계인지 추가로 분석할 필요가 없다. 단지 사용자가 상품A와 상품B를 같이 구매했다는 거래 데이터를 바탕으로 추천대상 고객에게 추천한다.

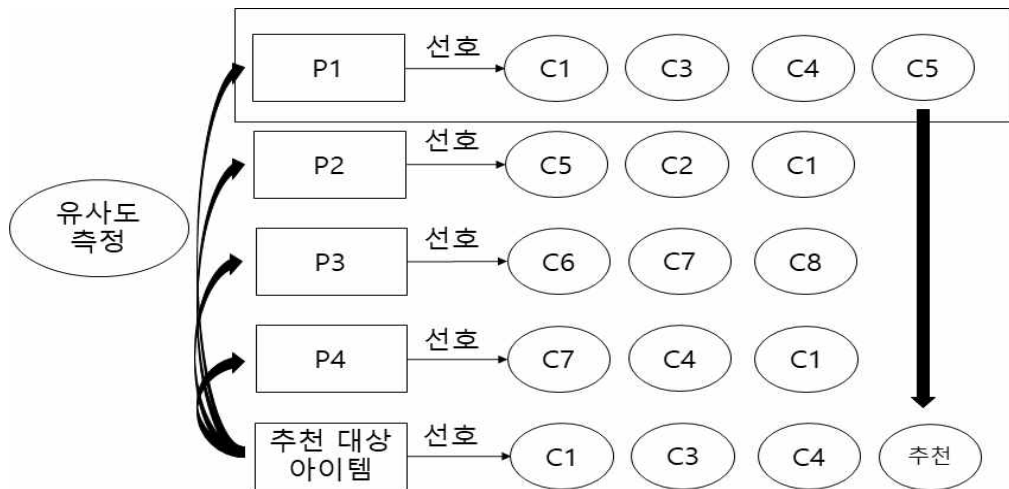
협업필터링은 다양한 상품추천 관련 연구에서 이용되고 있지만, 데이터가 충분히 축적되어 있지 않은 상황에서 작동하기 어려운 콜드 스타트(Cold-start) 문제, 혹은 데이터가 너무 과도하게 많아 연산시간이 과도하게 소요되는 확장성(Scalability) 문제 등의 한계점도 있다(Ahn, 2014). 이에 협업필터링 기반 추천 시스템의 한계점을 극복하기 위해 새로운 정보의 분석결과를 추천 알고리즘에 결합하는 하이브리드 접근법에 대한 연구들

이 다양하게 시도되고 있다(Ahn, 2014).

협업필터링 기법은 크게 기억 기반, 모델 기반으로 나눌 수 있다(손지은, 김성범, 김현중, 조성준, 2015). 기억 기반 추천 시스템은 다시 사용자 기반 협업필터링과 아이템 기반 협업필터링으로 구분된다. 사용자 기반 협업필터링은 사용자가 입력한 선호도 정보를 이용하여 고객의 비슷한 성향을 갖는 이웃 사용자를 선택하고, 선택된 이웃들이 공통적으로 선호하는 아이템을 고객에게 추천해주는 기법이다(손지은, 김성범, 김현중, 조성준, 2015). 아이템 기반 협업필터링은 특정 아이템을 기준으로 하여, 사용자들에 의해 평가된 점수가 비슷한 아이템을 이웃 아이템으로 선택한 후, 이웃 아이템을 평가한 점수를 기반으로 추천 대상 고객이 특정 아이템에 대해서 갖게 될 유사도를 예측한다(Sarwar, Karypis, Konstan, & Riedl, 2001). 예를 들어, [그림 2-2]는 사용자 기반 협업필터링의 기본적인 개념을 나타내고 있다. 추천 고객이 선정되면, 거래 데이터를 바탕으로 추천 고객과 나머지 사용자들 간의 유사도를 측정한다. 비슷한 구매 아이템이 많을수록 높은 유사도를 갖게 되며, 모든 유사도를 비교했을 때 유사도가 가장 높은 사용자를 이웃으로 선택한다. [그림 2-2]에서 나타나듯 추천 고객에게 가장 유사도가 높은 사용자는 'P1', 'P3', 'P4'를 구매한 'C1'이다. 여기서 추천 고객에게는 'C1'이 구매하였으나 추천 고객은 아직 구매하지 않은 'P5'를 최종적으로 선택하여 추천 고객에게 추천한다. [그림 2-3]은 아이템 기반 협업필터링의 기본적인 개념을 나타내고 있다. 추천 대상 아이템을 기준으로 유사한 아이템을 선정한 뒤, 선정된 'P1'은 구매하였으나 추천 대상 아이템은 구매하지 않은 'C5'에게 추천 대상 아이템을 최종적으로 추천하게 된다.



[그림 2-2] 사용자 기반 협업필터링 알고리즘



[그림 2-3] 아이템 기반 협업필터링 알고리즘

모델 기반 협업필터링은 기억 기반 협업필터링 방식의 과정을 기본으로 하되, 군집화, 분류, 예측의 단계에서 기계학습 또는 데이터마이닝 기법을 활용하는 것이다(손지은, 김성범, 김현중, 조성준, 2015). 이는 기존의 기억 기반 협업필터링 적용 시 발생하는 콜드 스타트, 확장성 등 다양한 문제점을 보완하기 위해 제안된 기법으로 사용되는 모델은 나이브 베이즈, 군집화 등이

있다(Domingos & Pazzani, 1997; Sarwar, Karypis, Konstan, & Riedl, 2002). 이 기법은 기억 기반 협업필터링 기법보다 짧은 시간에 결과를 예측할 수 있고, 데이터의 양이 매우 많아져도 추천할 수 있다는 장점이 있다. 하지만, 사용자 선호 모델을 신속하게 업데이트해야 하는 환경에는 적합하지 않다. 최근 들어, 데이터에 내재되어 있는 패턴을 발견할 수 있는 적절한 모델을 적용하여 추천 정확도를 높이고자 하는 연구가 활발히 진행되고 있고, 실제 데이터에 적용했을 때 성능이 우수하다는 점에서 효용성을 인정받고 있다(Domingos & Pazzani, 1997; Sarwar, Karypis, Konstan, & Riedl, 2002). 일반적으로 사용자 선호도가 범주형 데이터일 경우 분류모델이 사용되며, 연속형 데이터일 경우 회귀모델이나 특이값 분해 기법이 사용된다.

하이브리드 추천 시스템은 콘텐츠 기반 추천 시스템과 협업필터링 추천 시스템에서 나오는 문제점을 해결하기 위해 두 가지 이상의 알고리즘을 결합한 추천 기법이다. 앞서 살펴본 바와 같이 콘텐츠 기반 추천 시스템과 협업필터링 추천 시스템은 각각의 장·단점을 가지고 있다. 최근에는 각 추천 시스템의 장점을 극대화하면서 단점을 보완하는 하이브리드 추천 시스템에 대한 연구가 활발히 이루어지고 있다(Al Mamunur Rashid, Karypis, & Riedl, 2006; Barragáns-Martínez, Costa-Montenegro, Burguillo, Rey-López, Mikic-Fonte, & Peleteiro, 2010; Burke, 2002). Adomavicius와 Tuzhilin(2005)는 하이브리드 추천 시스템을 모델의 형태에 따라 크게 네 가지로 분류하였다. 첫 번째는 독립된 추천 결과를 조합하는 것이다. 이는 서로 다른 기준의 여러 개의 알고리즘을 학습한 뒤, 각 알고리즘의 추천 결과를 결합하여 상품을 추천하는 방식이다. 두 번째는 콘텐츠 기반 정보를 협업필터링에 적용하는 것으로, 콘텐츠 기반 추천 시스템에 의해 생성되는 사용자 프로파일 정보를 협업필터링에 의해 만들어진 모델에 결합하여 추천하는 시스템이다(김병만, 이경, 김시관, 임은기, 김주연, 2004). 세 번째는 LSI, PLSI와 같은 알고리즘을 이용하여 협업필

터링의 정보를 콘텐츠 기반 추천 시스템에 융합하는 것이다. 네 번째는 협업필터링과 콘텐츠 기반 추천 시스템을 동시에 고려하는 단일 모델을 구축하는 방법이 있다. 단일 모델을 학습하기 위해서는 MCMC와 같은 추정 모델이나 베이지안 학습법이 이용되기도 한다(손지은, 김성범, 김현중, 조성준, 2015).

2.1.1 기술 중심 연구

협업필터링 기반의 추천 시스템에 대한 연구는 유통, 서비스, 소비재 등 다양한 업종에 걸쳐 적용되어 왔으나(i.e., Celma & Herrera, 2008; Zhang, Séaghdha, Quercia, & Jambor, 2012; 강호윤, 옥창수, 2015; 김민성, 임일, 2014), 모형의 기본적인 작동 기제가 기존 소비자에 대한 과거 구매이력을 바탕으로 학습되고, 더 자주 구매되는 상품에 대해 더 높은 추천 확률이 계산되는 것이기 때문에 전통적으로 구매이력이 적은 상품에 대한 데이터 희소성 문제나 독특한 취향의 고객이나 신규고객에 대한 추천 부족문제, 또는 신상품에 대해 추천이 제외되는 등의 이슈들이 제기되어 왔다(i.e., 김민성, 임일, 2014; 신창훈, 이지원, 양한나, 최일영, 2012; 조운호, 방정혜, 2009). 이에 따라 그동안의 협업필터링 기반 개인화 추천 시스템에 대한 연구는 주로 이러한 협업필터링의 한계점을 개선하여, 추천의 정확도를 높이하고자 하는 방향으로 진행되어 왔다(i.e., 황용석, 김기태, 2019).

이러한 협업필터링 프로세스나 알고리즘에 대한 공학적, 시스템적 접근 방식의 연구는 협업필터링이 갖는 근본적인 한계점을 개선하여 보다 정확한 예측을 제공할 수 있다는 방법론적인 차원에서 중요한 시사점을 제시한다. 그러나, 기술적인 타당성이 경영학적인 타당성을 모두 보장해주지 않듯, 이러한 기술적인 접근방식은 비즈니스 관점, 특히 고객과의 관계 유지와 강화를 중요시하는 CRM적인 관점에서는 일정한 한계를 갖는다(Knijnenburg, Willemsen, Gantner, Soncu, & Newell, 2012). 예를 들어 기술 중심의 협업필터링의 연구들이 산출하는 객관적 오차 지표가 실제 추천 서비스 이용자들의 이용경험 만족도를 충실히 반영하지 못한다(황용석, 김기태, 2019). 즉, 추

천 알고리즘의 추천 적중률이 높아진다고 해도 이용자들의 주관적인 추천 만족도가 비례해서 높아지지는 않는다(Cremonesi, Garzotto, & Turrin, 2012; Knijnenburg, Willemsen, Gantner, Soncu, & Newell, 2012; Pu, Chen, & Hu, 2012). 또한 이러한 기술 중심의 협업필터링의 연구들은 실제로 이용자가 추천에 만족하지 못한 경우, 기술 중심 연구만으로는 이용자가 왜 만족하지 못하였는지 그 이유를 알 수 없다는 점에서 연구한계를 가지고 있다(황용석, 김기태, 2019). Knijnenburg, Willemsen, Gantner, Soncu & Newell(2012)는 추천 만족도가 추천 적중률에는 관련이 없고, 추천 적중률 이외에 다양한 요소로 영향을 받는다고 한다. Cremonesi, Garzotto & Turrin(2012)는 추천 적중률이 협업필터링과 같은 머신러닝을 이용한 개인화 추천 시스템의 추천 만족도가 머신러닝을 이용하지 않은 베스트셀러 기반 추천 시스템 보다 높지 않았음을 보여주며, 추천 만족도는 이용자의 주관적인 인지된 품질이나 인지된 정확성 개념으로 측정한다고 주장한다. 이와 같이 일부 추천 시스템의 연구들은 추천 정확도를 높이는 기술 중심 연구가 아닌 다양한 관점에서 이용자들의 추천 만족도를 측정할 수 있는 이용자 경험 중심으로 연구 범위를 확대해 이용자 중심 관점에서 연구를 진행하고 있다.

2.1.2 이용자 중심 연구

이용자 중심 접근 연구들은 주로 실험실 연구나 서베이를 연구방법으로 사용한다. 구체적인 연구 방법은 동일한 추천 시스템을 사용 인터페이스 상의 몇 가지 요소를 변경하거나, 서로 다른 추천 시스템을 사용하게 한 후 추천 서비스 이용자들에게 실제 사용 경험을 서베이를 이용해 측정한다. 혹은 서베이 대신, 이용자들의 추천 서비스 이용 과정을 직접 관찰하거나 녹화를 통해 추천 성과를 측정하기도 하고, 인터뷰와 같은 질적 방법을 이용하기도 한다(황용석, 김기태, 2019).

이용자 중심 연구들의 몇 가지 사례를 살펴보자면, Bogdanov, Haro, Fuhrmann, Xambó, Gómez와 Herrera(2013)는 185명의 피실험자들에게 협업필터링 기반의 지니어스라는 음악 추천서비스를 사용하게 한 후, 만

족 혹은 불만족 여부를 서베이하였다. 그 결과 만족도는 추천 가수 리스트와 이용자가 선호하는 가수 리스트의 유사성보다는 선호 장르와 음악적 유사성에 좌우된다는 것을 알 수 있었다. Zhang, Séaghdha, Quercia & Jambor(2012)는 21명의 피실험자들에게 가상의 영화 추천사이트를 사용하게 한 후, 추천 정확성과 추천 의외성, 추천 만족도 세 변수의 관계를 살펴보았다. 그 결과 추천 정확성과 추천 만족도 사이는 정적인 관계를 보였으나, 추천 의외성이 클 경우 오히려 추천 정확성이 증가할수록 추천 만족도가 감소하는 경향을 발견하였다. 그러나 Celma와 Herrera(2008)는 288명의 피실험자들에게 협업필터링 기반 추천 시스템과 내용 기반 추천 시스템을 사용하게 한 후, 추천 정확성과 추천 의외성, 추천 만족도 세 변수의 관계를 살펴본 결과, 협업필터링 기반 추천 시스템의 추천 의외성이 내용 기반 추천 시스템에 비해 낮게 평가되었음에도 불구하고, 협업필터링 기반 추천 시스템이 내용 기반 추천 시스템보다 추천 만족도가 높게 나타났다.

이용자 중심 연구들이 예측 정확도 위주의 기술 중심 연구들에서 벗어나 이용자들의 만족도를 연구의 범위로 넓힌 점에서 중요한 역할을 했다(황용석, 김기태, 2019). 하지만 이용자 중심 연구들은 소비자에 대한 서베이 혹은 인터뷰 결과를 바탕으로 이루어져 있기 때문에 체계화된 방법론적 프레임워크가 제시되지 못하고, 연구자에 따라 상이한 측정 변수를 활용하고 있기 때문에 일관된 연구결과를 기대하기 어렵고 시스템적으로 개발되어 실무적으로 활용되기에는 많은 한계를 내포하고 있다(i.e., Huang, Zhao, & Lu, 2013; Knijnenburg, Willemsen, Gantner, Soncu, & Newell, 2012).

[표 2-1] 협업필터링 기반 기술 및 이용자 중심 연구

구분	연구	주요 내용
기술 중심 연구	조운호, 방정혜(2009)	상품 네트워크 중심성과 협업필터링을 결합 해 신상품 추천 문제 해결
	신창훈, 이지원, 양한나, 최일영, (2012)	고객 네트워크 중심성과 협업필터링을 결합 해 신규고객추천 문제 해결
	김민성, 임일(2014)	소셜네트워크 분석기법과 협업필터링을 결합 해 특이 취향 사용자 문제 해결
	강호윤, 옥창수(2015)	아이템 간 연관성을 고려한 협업필터링을 구 현해 데이터 희소성 해결
	정석봉(2017)	고객 구매 패턴을 파악하여 아이템 기반 협 업필터링 구현해 데이터 희소성 해결
이용자 중심 연구	Celma & Herrera(2008)	협업필터링과 내용 기반 추천 시스템의 사용 만족도 비교 연구
	Zhang, Séaghdha, Quercia, & Jambor (2012)	추천 시스템 알고리즘별 추천 정확성, 의외 성, 만족도 세 변수의 관계 비교 연구
	Bogdanov, Haro, Fuhrmann, Xambó, Gómez, & Herrera (2013)	협업필터링 기반 추천 시스템 지니어스의 사 용만족도 조사

2.2 소셜네트워크분석

사회연결망 또는 소셜네트워크는 네트워크 참여자들의 연결패턴으로 표현되는 네트워크 구조를 통해 사회적인 상호작용을 설명하려는 이론으로써 지난 10여년간 매우 다양한 사회과학 영역의 주요 연구주제로 다루어져 왔다(i.e., 강부식, 2010; 김병국, 정석봉, 권기석, 2013; 조윤희, 김인환, 2010). 소셜네트워크 분석의 목적은 소셜네트워크 형태의 특징을 도출하고, 관계성으로 체계를 설명하거나 체계를 구성하는 단위의 행위를 설명하는 것이라고 할 수 있다(김용학, 2014). 소셜네트워크는 대상 간 연결을 네트워크 형태로 표현하는데, 이러한 네트워크에서 연결대상은 노드(node)이며, 연결은 링크(link)가 된다. 여기서 노드는 상호 고유한 속성을 지닌 개별 개체를 의미하고, 링크는 노드들 간의 연결 관계를 나타낸다.

소셜네트워크 분석의 활용분야는 매우 다양하다. 예를 들어 네트워크를 추적해 핵심인물을 파악하는데 쓰일 수 있고(Zech & Gabbay, 2016) 대상들 간의 관계를 통해 어떤 것이 교류되고 있는지 탐색할 수도 있다. 또한 최근에는 트위터, 페이스북, 인스타그램 등 사회관계서비스(Social Network Service) 등의 소셜 미디어 발달과 함께 장기간 축적된 데이터로 정보의 흐름이 어떻게 흘러가고 있는지에 대한 분석도 가능하게 되었다(Kim, 2017). Christakis와 Fowler(2010)는 친구 네트워크를 통해 비만도 전염이 된다고 주장하였고, 이때 네트워크를 통해 감정도 전염될 수 있다고 하였다. 그는 본인 자신의 영향력이 3단계에 걸쳐 미친다고 주장하였고 이 단계가 6단계를 거치면 이 세상 모든 사람들이 연결될 수 있다고 주장하였다.

소셜네트워크분석의 핵심적인 기능은 다음 4가지 이다. 첫 번째, 네트워크에서 가장 영향력 있는 대상을 선정하여 평가할 수 있다. 두 번째, 대상 간의 관계 경로를 통해 지름길이나 사회적 거리를 탐색한다. 세 번째, 네트워크 속에 응집된 사회적 그룹을 추출하여 커뮤니티를 파악한다. 네 번째, 중심적 연결자나 중재자처럼 중요한 사회적 역할을 추정한다.

2.2.1 소셜네트워크분석의 분석수준

소셜네트워크분석은 분석수준에 따라서 네트워크 수준의 분석, 노드 수준의 분석, 네트워크의 내재된 속성 분석, 중심성 분석, 하위집단 수준의 분석, 에고 네트워크 수준의 분석 등 6가지 형태로 구분할 수 있다(이수상, 2012).

첫 번째, 네트워크 수준의 분석은 거시 수준의 분석지표를 제시하는데, 밀도, 크기, 포괄성, 집중도 등이 해당하고 전체적인 네트워크의 특성을 파악하기 위한 분석방법이다(이수상, 2012). 두 번째, 노드 수준의 분석은 네트워크 수준의 분석과 다르게 미시 수준의 분석지표를 제시하는데, 각 노드의 특성, 각 노드들 간의 링크 관계에 대한 특성을 분석한다. 노드 수준의 분석 지표는 인접성 지표와 연결성 지표로 구분되며, 인접성 지표는 연결거리, 직경, 연결성 지표는 연결정도, 연결강도 등이 해당된다(이수상, 2013). 세 번째, 네트워크 내재된 속성 분석은 하이브리드 수준의 분석으로 특정 노드들이 밀접하게 구성되는 상호성, 이행성, 군집화 계수 등의 분석을 포함한다(이수상, 2012). 네 번째, 중심성 분석은 네트워크 내에서 각 노드들의 영향력을 나타내는 정도를 측정하는 지표로서 소셜네트워크분석에서 가장 많이 활용되는 지표이다(이수상 2012). 중심성 지표는 네트워크 유형에서 볼 경우, 전체 네트워크의 중심성에 해당하며, 네트워크 내에서 각 노드들의 영향력 크기를 순위화 하는데 사용이 가능하다(이수상, 2010). 다섯 번째, 하위집단 분석은 전체 네트워크의 하위집단을 분류하고 네트워크 구조를 분석하는 방법이다. 하위집단 분석지표는 응집도를 기반으로 분석하는 컴포넌트 분석과 파당분석 등이 있다. 여섯 번째, 에고 네트워크 분석은 한 특정한 개인 에고를 중심에 위치시키고 그 노드와 연결된 다른 노드와의 관계를 표현하는 분석 방법이다(손동원, 2002). 에고 네트워크 분석은 전체 네트워크에서 에고 노드와 연결된 노드들로 구성된 부분 네트워크의 특성을 파악하는데 유용하다.

2.2.2 중심성 분석

대부분의 소셜네트워크의 중심적 위치를 갖는 사람이나 기관은 구조적 위치로 인하여 보다 나은 정보 접근성과 정보 전달력을 갖는다(De Nooy, Mrvar, & Batagelj, 2018). 예를 들어 중심성이 높은 사람은 높은 지위를 가지거나, 인기가 많거나, 정보를 받거나 전달하는데 있어서 큰 영향력을 행사할 수 있는 사람이다(김용학, 김영진, 2016). 이처럼 중심성 분석은 네트워크의 중심성을 바탕으로 한 네트워크에서 핵심적인 역할을 하는 중심 노드가 무엇인지를 도출하는 유용한 방법이다(김혜경, 김재경, Yi, C. Q., 2009). 이러한 중심성을 측정하는 데는 다양한 기준이 있는데, 예를 들어 한 노드가 얼마나 다수의 노드들과 연결되어 있는지 측정할 수 있고, 다른 노드에 도달하기 위하여 특정 노드를 얼마나 많이 지나가는지를 계산하여 중심성을 측정할 수 있다. 또는 한 노드가 다른 노드로 가기 위하여 얼마나 많은 노드를 거쳐가야 하는지도 측정할 수 있다. 이렇게 중심성을 측정하는 방법은 다양하고 각각의 목적에 따라서 연구문제에 적합한 중심성 분석을 선택해야한다(Robins, 2015). 중심성은 연결 정도 중심성(Degree Centrality), 근접중심성(Closeness Centrality), 매개중심성(Betweenness Centrality), 고유벡터 중심성 (Eigenvector Centrality) 등으로 구분할 수 있다(곽기영, 2014).

첫 번째 연결 정도 중심성은 한 노드는 많은 연결을 가질수록 더욱 많은 권력을 가진다는 기본적인 믿음에 근거하여 많은 연결을 가진 노드는 선택의 폭이 넓기 때문에 더욱 많은 기회를 가진다고 본다(Hanneman & Riddle, 2005). 즉, 직접 연결된 이웃 노드가 많을수록 연결 정도 중심성이 높다고 할 수 있다(사이람, 2017).

연결 정도 중심성을 산출 방법은 해당노드가 가진 값을 최댓값으로 나눈다. 이 때의 최댓값은 자기 노드를 제외한 모든 노드와 연결되어 있을 때의 값이다. 따라서 n 을 네트워크를 구성하는 전체 노드의 수, d_i 를 i 번째 노드의 연결 수라고 하였을 때 연결 정도 중심성의 식은 다음과 같다.

$$\frac{d_i}{n-1}, i=1, \dots, n \quad (1)$$

두 번째 근접중심성은 특정 노드가 얼마나 네트워크 내의 중앙에 있는지를 측정하여 다른 노드들과 얼마나 가까운지를 보여준다. 근접중심성이 높은 노드는 확률적으로 가장 빠르게 다른 노드에 영향을 주거나 받을 수 있으므로 영향의 즉효성을 측정하는데 적합하다(사이람, 2017). 네트워크에서 모든 다른 노드로 가는 경로의 합이 가장 작은 노드가 근접중심성이 가장 높은 노드이다(Freeman, 1978).

근접중심성을 산출하는 방법은 거리가 짧을수록 중심성이 큰 것이기 때문에 결과값에 역수를 취해준다. 거리를 측정값으로 가진다. 따라서 n 을 네트워크를 구성하는 전체 노드 수, $d(i, j)$ 를 노드 i 번째 노드와 $j(\neq i)$ 번째 최단경로의 거리값이라고 하였을 때, 근접중심성의 식은 다음과 같다.

$$\frac{n-1}{\sum_{j \neq i}^n d(i, j)}, i=1, \dots, n$$

세 번째 매개중심성은 특정 노드가 다른 노드들 사이에 놓여있는 정도를 측정한다. 매개중심성은 한 노드가 네트워크 상 정보전달에 있어서 필수적인 위치에 존재할 때 높아진다. 즉, 한 노드를 거쳐 가는 연결이 많을수록 그 노드의 매개중심성은 높아진다(De Nooy, Mrvar, & Batagelj, 2018). 예를 들어, 노드 A와 노드 B가 오직 노드 C를 통해서만 관계를 맺을 수 있는 경우, 노드 C의 매개중심성은 높다고 할 수 있다. 여기서 노드 C는 정보가 흘러가는 것에 대한 통제력을 가지는 것이 특징이다. 이 노드가 사라질 경우 네트워크 전체 연결과 흐름에 큰 타격을 줄 수 있다(사이람, 2017).

매개중심성을 산출하는 방법은 해당 노드를 경유하는 최단경로와 그 노드를 경유할 확률의 곱의 총합을 네트워크 내 모든 경로의 수로 나눈 것과 같다. 따라서 $g_{jk}(n_i)$ 를 노드 j 와 k 를 연결하는 최단경로 중 노드 i 를 거치는

경로의 수, $\sum_{j < k} g_{jk}(n_j)/g_{jk}$ 를 각각의 노드 쌍의 최단경로 상 노드 i 가 포함될 확률의 누적 합, $[(g-1)(g-2)/2] = n_i$ 를 포함하지 않는 모든 노드 쌍의 수라고 하였을 때, 매개중심성의 식은 다음과 같다.

$$\frac{\sum_{j < k} g_{jk}(n_j)/g_{jk}}{[(g-1)(g-2)/2]}$$

네 번째 고유벡터 중심성은 Bonacich(1987)에 의해 최초로 소개되었는데 자신과 연결된 이웃노드들의 중요성을 자신의 중심성을 판단하는데 사용하는 개념이다. 노드 간의 연결이 되었을 때, 중요성이 높은 노드와 연결된 노드의 고유벡터 중심성이 중요성이 높지 않은 노드와 연결된 노드의 고유벡터 중심성보다 높다. 즉, 연결된 다른 노드의 영향력이 강할수록 고유벡터 중심성도 높아진다.

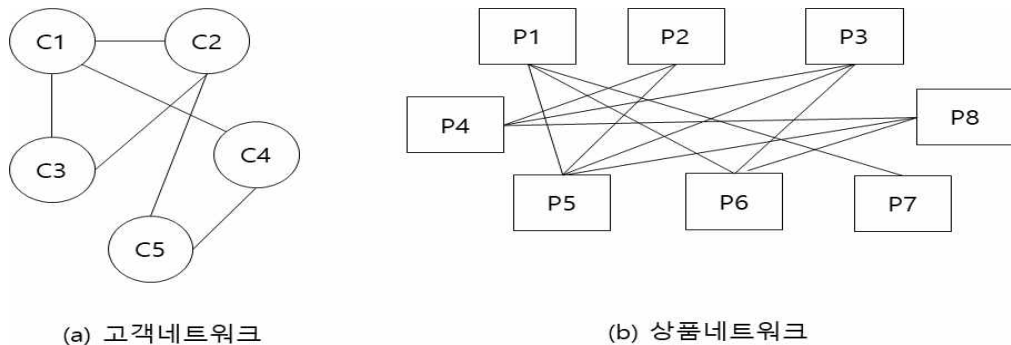
고유벡터 중심성을 산출하는 방법은 다음과 같다. 고유벡터 중심성은 인접 행렬을 요인 분석으로 차원 축소를 진행할 때 고유값이 가장 큰 벡터일 때의 각 노드의 값이 고유벡터 중심성 값이다. 즉, 네트워크의 연결된 값들을 어떤 노드가 가장 잘 설명하는지를 보여주는 값이다(김용학, 김영진, 2016). 고유벡터 중심성은 이웃 노드들의 점수를 합산한 값에 비례하여 부여된다. 따라서 λ 를 고유값, A 를 1-모드 네트워크(one mode network), n 을 전체 노드수라고 하였을 때, 고유벡터 중심성의 식은 다음과 같다.

$$\frac{1}{\lambda} \sum_{j=1}^n A_{ji} c_e(v_j)$$

2.2.3 고객 네트워크와 상품 네트워크

소셜네트워크는 분석 데이터의 형태와 주제에 따라 고객 네트워크나 상품 네트워크와 같은 다양한 형태의 소셜네트워크 개념으로 확장하여 사용할 수

있다(i.e., Kim, Kim, & Chen, 2012; 조운호, 김인환, 2010). 가령, 고객 데이터베이스와 거래 데이터베이스를 조인하여 고객별 구매 이력 데이터를 추출할 경우 고객이나 상품을 노드로 하는 하나의 소셜네트워크를 구성할 수 있다(i.e., 김병국, 정석봉, 권기석, 2013). 예를 들어 고객이 상품을 구매한 거래 데이터에서, 고객-상품의 행렬을 구성한다. 고객 i 가 상품 j 를 구매하면 셀(i, j)는 '1'로, 구매하지 않으면 '0'으로 표시한다. 고객-상품 행렬에서 상품을 구매한 사람들 사이에 직접적인 연결 관계가 없더라도, 관계를 인위적으로 설정하여 고객과 고객 또는, 상품과 상품 사이의 관계를 나타내는 1-모드 네트워크를 생성할 수 있다. 이를 그림으로 표현하면 [그림2-4]의 고객 네트워크(a)와 상품 네트워크(b)가 된다. 이때 고객을 노드로 하는 소셜네트워크를 고객 네트워크라고 할 수 있고(e.g., 박종학, 조운호, 김재경, 2009), 반대로 상품을 노드로 하는 소셜네트워크를 상품 네트워크라고 할 수 있다(김병국, 정석봉, 권기석, 2013). 여기서 고객 네트워크는 고객들이 서로 동일한 상품을 1개 이상 구매하였을 경우, 그 고객들은 서로 상호작용이 있다고 판단할 수 있고 상품 네트워크는 고객이 함께 구매한 상품들 간에는 서로 상호 작용 관계가 있다고 판단한 것이다.



[그림 2-4] 고객네트워크와 상품네트워크 예시 그림

고객 네트워크는 본질적인 소셜네트워크의 형태를 그대로 적용하는 것이기 때문에 개인화 추천 시스템 등 고객기반 마케팅 자동화 시스템 연구에 자주 사용되어 왔다(e.g., 강부식, 2010; 박종학, 조운호, 김재경, 2009; 정석봉, 2014; 조운호, 김인환, 2010). 박종학, 조운호, 김재경(2009)은 고객 간의 구매패턴의 유사성을 기준으로 고객 네트워크를 구성한 후 중심성 개념을 협업필터링에 적용하여 백화점 신규고객을 대상으로 상품을 추천하는 방법을 제안하였다. 이때 유사도가 특정 임계치 이상일 경우에만 해당 고객들을 링크로 연결하는 방법을 사용하였다. 강부식(2010)은 고객 네트워크를 구성한 후 구조적 공백 개념을 협업필터링에 적용하여 신규고객을 대상으로 추천하는 방법을 제안하였다. 조운호와 김인환(2010)은 고객 네트워크를 구성한 후 인공신경망 모형을 이용하여 협업필터링 추천 시스템의 성능을 예측하였다. 정석봉(2014)은 인터넷 쇼핑물의 효과적인 고객분류와 선호상품 추출을 위해 고객 네트워크의 중심성 개념을 활용하였다. 상품 네트워크 역시 다양한 비즈니스 연구에서 활용되어 왔다. 김병국, 정석봉, 권기석(2013)은 상품 네트워크를 구성하기 위하여, 고객에 의해 동시구매가 발생한 상품들을 링크로 연결하였다. 그 후 중심성 개념을 적용하여 구매 품목간의 다양한 관계 정보를 분석하였다. 최인호, 김병국, 정석봉(2014)은 상품 네트워크를 구성한 후 중심성 개념을 활용하여 새로운 상품분류기준을 제시하였다. 김혜경, 김재경, Yi, C. Q.(2009)은 백화점 거래 데이터를 활용하여 상품 네트워크를 구성한 후 중심성과 집중도의 개념을 활용하여 고객 구매패턴의 중심이 되는 상품을 파악하였으며, 다른 상품들과 직간접적으로 연계되어 판매되는 상품 관계를 파악하였다. Kim, Kim & Chen(2012)은 장바구니 분석과 상품 네트워크 분석의 중심성 개념을 활용해 교차 판매, 상향 판매 등 새로운 마케팅 방안을 도출하였다. 상품 네트워크의 경우 분석을 통해 새로운 마케팅 전략을 도출하는데 연구가 주로 이루어져 있고 고객 네트워크에 비해 상품 네트워크 분석을 연계한 심도 있는 응용연구가 진행되고 있지는 못하는 실정이다.

[표 2-2] 고객 네트워크와 상품 네트워크 연구

구분	연구	주요 내용
고객 네트워크	정석봉(2014)	고객 네트워크의 중심성을 활용한 추천 시스템 제안
	강부식(2010)	고객 네트워크의 구조적 공백을 활용한 협업필터링 추천 시스템 제안
	박종학, 조운호, 김재경(2009)	고객 네트워크 중심성을 활용한 협업필터링 추천 시스템 제안
	조운호, 김인환(2010)	고객 네트워크를 구성, 인공신경망 모델을 이용하여 협업필터링 추천 시스템의 성능을 예측
상품 네트워크	김병국, 정석봉, 권기석(2013)	상품 네트워크를 구성하여 구매 품목 간의 관계를 정의한 후 마케팅전략 도출
	최인호, 김병국, 정석봉(2014)	상품 네트워크를 구성하여 중심성 지표를 이용한 새로운 상품분류기준 제시
	김혜경, 김재경, Yi, C. Q.(2009)	상품 네트워크를 구성하여 다른 상품들과 직·간접적으로 연계되는 상품을 파악하여 마케팅 전략 도출
	Kim, Kim & Chen(2012)	장바구니 분석과 상품 네트워크 분석을 결합해 새로운 마케팅방안 도출

2.3 고객로열티

기업이 추진하는 모든 경영전략은 결국 고객의 로열티(Loyalty)를 증대시킬 수 있는 방향으로 모색되어야 한다(Oliver, 1999). 마케팅 분야 학자들이나 참가자들은 모든 서비스산업에서 전략적인 목표로서 고객로열티의 중요성을 인식하고 있다(Gustafsson, Johnson, & Roos, 2005).

고객로열티는 선호하는 상품이나 서비스를 미래에도 지속적으로 구매하려는 의도나 행동을 말한다. Dick과 Basu(1994)는 고객로열티를 고객이 특정 기업, 브랜드 및 상품 등에 대해 가지고 있는 충성스러운 결속의 신념이라 하였다. Gremler와 Brown(1996)는 고객로열티는 서비스 제공자로부터 재구매하는 행동, 긍정적인 태도, 서비스에 대한 니즈가 생겼을 때 오직 해당 서비스 제공자만 이용하려고 고려하는 정도를 말한다고 정의한다. Oliver(1999)는 고객로열티는 제품이나 서비스를 미래에도 지속적으로 재구매하고 애호하는 깊은 몰입의 상태로 정의한다. Walsh, Dinnie & Wiedmann(2006)은 고객로열티는 미래에 선호하는 상품이나 서비스를 일관되게 재구매하거나 재이용 하려는 깊이가 있는 일종의 몰입으로 정의한다. 그리고 Johnson, Herrmann & Huber(2006)은 고객이 이탈하지 않고 재구매 확률과 현재에 이용하는 브랜드를 추천하게 될 가능성에 영향을 주는 것이라 정의한다. 최웅, 이영우, 조승현(2012)은 고객로열티를 고객의 행동적 반응으로서 특정 제품이나 서비스를 구매하는 횟수나 구매의 일관성 및 서비스에 대한 호의적 태도라고 하였고, 장나영과 이규민(2016)은 고객로열티를 특정 기업의 사람, 서비스 및 제품에 대한 애착과 애정이라고 하였다. 이준혁과 박헌진(2019)은 고객로열티는 지속적인 이용에 대한 강한 심리적 애착이라고 하였다. [표 2-3]은 국내·외 고객로열티 선행연구에서 제시하고 있는 고객로열티의 개념을 정리한 것이다.

[표 2-3] 고객로열티의 개념

연구	개념
Dick & Basu(1994)	고객이 특정 기업, 브랜드 및 상품 등에 대해 가지고 있는 충성스러운 결속의 신념
Gremler & Brown(1996)	서비스 제공자로부터 재구매하는 행동, 긍정적인 태도, 서비스에 대한 니즈가 생겼을 때 오직 해당 서비스 제공자만 이용하려고 고려하는 정도
Oliver(1999)	제품이나 서비스를 미래에도 지속적으로 재구매하고 애호하는 깊은 몰입의 상태
Walsh, Dinnie & Wiedmann(2006)	미래에 선호하는 상품이나 서비스를 일관되게 재구매하거나 재이용하려는 깊이가 있는 일종의 몰입
Johnson, Herrmann & Huber(2006)	고객이 이탈하지 않고 재구매 확률과 현재에 이용하는 브랜드를 추천하게 될 가능성에 영향을 주는 것
최웅, 이영우, 조승현(2012)	고객의 행동적 반응으로서 특정 제품이나 서비스를 구매하는 횟수나 구매일관성 및 서비스에 대한 호의적 태도
장나영, 이규민(2016)	특정 기업의 사람, 서비스 및 제품에 대한 애착과 애정
이준혁, 박헌진(2019)	지속적인 이용에 대한 강한 심리의 애착

기업이 고객에 기대에 맞는 가치를 제공하여 고객만족이 이루어지면 자연스럽게 해당 기업의 로열티가 상승할 것이고, 이는 재구매로 이어질 가능성이 높을 것이다. 결국 높은 수준의 고객로열티가 형성되어있다면 기업이 얻는 수익은 높아질 것이다. 이는 기업입장에서 이탈회원감소, 고객유지비용 감소, 교차판매, 상승판매 등 다양한 부분에서 긍정적인 영향을 미칠 것이다. 따라서 실질적인 고객 행동으로 이어지지 않는 경우가 많은 고객만족도에 비해 구체적인 고객행태로의 전환 가능성이 크기 때문에 기업은 고객과의 우호적인 관계 구축을 통해 고객로열티 기반의 유기적 성장을 달성하려고 노력하고 있다(강양, 곽용섭, 2012). 이러한 고객로열티에 대한 연구는 크게 태도적, 행동

적, 통합적 관점에서 이루어져 왔다(Dekimpe, Steenkamp, Mellens, & Abeele, 1997; Dick & Basu, 1994).

2.3.1 태도적 접근관점

태도적 접근관점에서는 고객선호 또는 심리적 몰입의 측면에서 특정 서비스 제공자에 대해 갖게 되는 호의적인 태도와 구매가능성을 다루었다(Mathieu & Zajac, 1990). 즉, 태도적 접근관점에서의 로열티는 브랜드를 좋아하는 선호 또는 애착과 비슷한 개념으로 파악된다. 로열티에 내재하는 감정과 심리적 애착을 반영하기 위해서는 태도 관련 자료를 이용한다. 태도 관련 자료는 감정과 관련된 것으로 호의적인 태도를 가지고 반복적으로 같은 상품 또는 서비스를 구매하는 것을 의미한다. 태도적 로열티는 경제적인 이유 혹은 상품이나 서비스를 이용하기 어려운 환경 등으로 인해 반복 구매를 할 수는 없으나, 이러한 부분을 개선하고 해결하면 반복구매를 할 수 있는 잠재적인 로열티를 의미한다. 따라서 태도적 로열티는 잠재적 고객을 찾기 위한 측정도구로 활용되기도 한다. 태도적 로열티를 측정하는 방법은 실제 구매 행동보다는 고객의 호의적 태도를 중심으로 상품이나 서비스에 대한 선호도 조사를 중심으로 많이 측정되어왔다.

Oliver(1997)는 태도적 로열티를 경쟁적 고립(Competitive insulation)이라고 정의하였는데, 이는 고객이 특정한 상점이나 점포에 일관된 충성도를 보이는 것을 뜻한다. 즉, 태도적 로열티가 높은 고객은 타 경쟁사에 전환 유혹이 발생하더라도 동일 브랜드를 재구매하는 특징을 갖는다고 주장하였다. 하지만 태도적 로열티 측정은 극히 개인적인 상황과 주관적 의견이 반영되기 때문에 현실적으로 측정이 어렵다는 의견도 있다(안성식, 박연옥, 강병옥, 2011). Oh(1995)는 태도적 로열티 측정만으로는 실제 구매행동으로 이어지는지 정확한 통계를 낼 수 없고, 경쟁 혹은 변화된 시장상황에 대해 구체적인 근거를 제시할 수 없다는 단점을 지적하였다. 또한 태도적인 관점에서 측정한 고객로열티는 학술적으로 검증된 서베이 기반의 측정으로 개념적인 타당성을 보장하고, 고객로열티의 선행적인 요인과의 인과관계 등을 검증하는 실증연구에서

의미가 있지만, 객관적으로 정량화된 지표가 아니기 때문에 기업이 상시적으로 측정하고, 관리하는 데에는 일정한 한계가 존재한다.

2.3.2 행동적 접근관점

행동적 접근관점에서의 고객로열티는 일정기간 동안 고객이 상품 또는 서비스에 대하여 반복구매를 측정하기 때문에 실제 구매행동과 관련이 많다. 행동적 로열티의 측정방법으로는 행동적인 측면에서의 반복구매 행동, 재구매의사, 추천의사, 구매빈도, 관계지속여부, 구매비율, 구매순서 등이 측정지표로 이용되었다(Baloglu, 2002; Jones & Sasser, 1995; Sharp & Sharp, 1997). 이러한 행동적 로열티 측정방법의 문제점은 반복구매가 반드시 브랜드를 향한 심리적 몰입을 뜻하는 것은 아니라는 것이다. 즉, 행동적 로열티가 높은 고객이라도 그들이 모두 높은 로열티를 지녔다고 할 수 없다. 기업의 성과와 수익에 직접적인 영향을 미칠 수 있는 행동적 로열티를 높이기 위해서는 기업은 기업입장에서의 이익이나 관점만을 반영하는 것이 아닌 객관적인 서비스를 제공해야하며, 고객 맞춤형 서비스를 제공하는데 노력을 기울여야할 것이다. 또한 행동적 로열티 측정은 정량적인 데이터를 기초로 측정하기 때문에 객관적이고, 신뢰성 있는 분석방법이지만, 엄연히 로열티의 개념을 대리하여 측정하는 일종의 프록시(Proxy) 지표라는 점 때문에 측정 타당성(Validity)의 한계를 내포하고 있다.

2.3.3 통합적 접근관점

통합적 접근관점은 태도적 측정방법과 행동적 측정방법을 모두 고려한 것으로 소비자의 호의적인 태도와 구매가능성, 재구매의사, 구매빈도, 반복 구매 행동 등을 포괄적으로 측정하는 방법이다. 이는 로열티의 경우 한 측면만으로 측정하는 경우 측정 불가능한 부분이 발생할 수 있다는 의견에서 출발했다. 즉 통합적 접근관점에서는 앞서 제시한 행동과 태도라는 두 가지 변수를 통합하여 하나의 개념으로 보는 것이다(Dick & Basu, 1994). Prus와

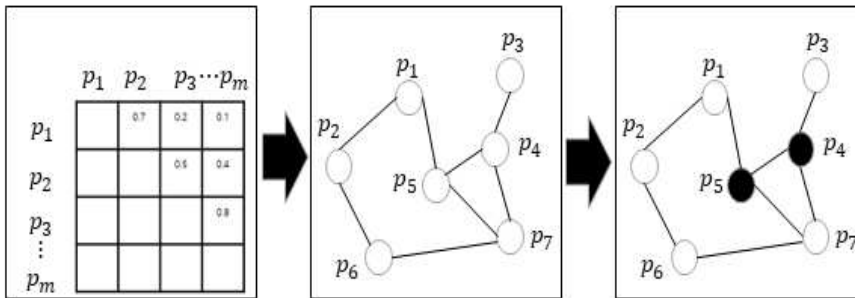
Brandt(1995)는 고객로열티가 태도와 행동의 결합이라 주장하면서 태도는 재구매 의도, 공급자의 권유 의지, 경쟁브랜드 전환 거부 등의 개념을 포함하였으며, 행동은 구매브랜드로부터 다른 상품이나 서비스를 더 많이 재구매하거나 주변 지인들에게 추천한다는 개념을 포함하였다. Ganesh, Arnold & Reynolds(2000)는 통합적 접근관점에서의 고객로열티 측정에 있어 긍정적 구전과 추가구매 행동을 능동적 충성도로 서비스 유지를 수동적 충성도로 구분하여 측정할 것을 제안하기도 하였다.

Ⅲ. PNA 기반 협업필터링 개인화 추천 시스템

3.1 추천 시스템 개요

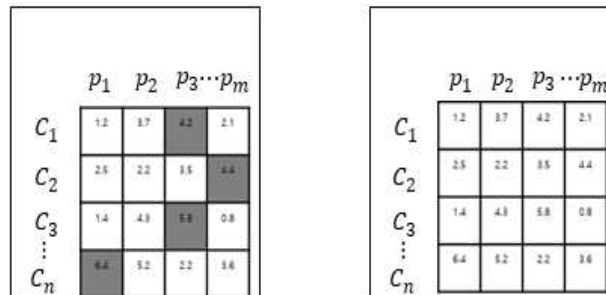
본 연구에서 제안하는 개인화 추천 시스템은 기존 아이템 기반 협업필터링 추천 과정에 상품 네트워크 분석(Product Network Analysis) 기법을 결합한 고객충성도 중심의 개인화 추천 프로세스이다 [그림 3-1 참조]. 각 단계에 대해서 상세한 설명은 3.1.1부터 3.1.5절에 걸쳐 제시하며, 실제 연구용 분석에 사용한 데이터는 결과 표가 방대하여 간단히 8개의 상품과 4명의 고객에 대한 예시를 들어 이해를 도모하고자 하였다.

1단계 : 아이템 유사도 분석 2단계 : 상품 네트워크 구성 3단계 : 중심성 분석



5단계 : 추천 리스트 생성

4단계 : 중심성 가중치 적용



[그림 3-1] PNA 기반 협업필터링 개인화 추천 시스템 프로세스

3.1.1 아이템 유사도 분석

아이템 유사도 분석 단계에서는 기존 상품들에 대한 고객들의 구매성향을 분석하고, 상품 간 유사도를 도출하는 과정으로써 세부적으로는 고객들의 구매성향을 나타내는 고객-상품 구매 행렬 도출 과정과 이를 바탕으로 상품 간의 유사도를 계산하는 과정으로 구분될 수 있다.

구매 이력 데이터셋이 [표 3-1]와 같은 구조를 가지고 있다고 할 때, 고객-상품 구매 행렬은 구매 이력 데이터셋을 바탕으로 특정 고객(C)이 특정 상품(P)을 몇 개 구매 했는지를 나타내는 행렬로써 행렬 $A = (A_{ij})$ 는 같이 표현할 수 있다.

$$A_{ij} = \begin{cases} n : \text{고객 } C \text{가 상품 } P \text{를 구매한 횟수} \\ 0 : \text{고객 } C \text{가 상품 } P \text{를 비구매} \end{cases}$$

[표 3-1] 거래 데이터 예시

고객ID	상품ID	구매일자	구매금액
$C1$	$P5$	2018-02-22	23,000
$C2$	$P1$	2018-05-25	44,000
$C1$	$P5$	2018-07-11	23,000
$C3$	$P1$	2018-08-11	44,000

[표 3-2]의 예시에서 첫 번째 행은 고객 $C1$ 이 상품 $P5$, $P7$, $P8$ 을 각각 2개, 4개, 6개 구매한 것을 의미한다.

[표 3-2] 고객X상품 구매 행렬 예시

	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$
$C1$	0	0	0	0	2	0	4	6
$C2$	4	2	1	0	0	0	0	5
$C3$	5	0	5	6	0	2	1	0
$C4$	0	5	3	4	1	5	0	0

그리고, 상품 간의 유사도는 코사인 유사도를 사용하여 식(1)와 같이 두 상품 i 와 j 의 아이템 유사도 $\cos(P_i, P_j)$ 를 통해 산출할 수 있다.

$$\cos(P_i, P_j) = \frac{\overrightarrow{P_i} \cdot \overrightarrow{P_j}}{\|\overrightarrow{P_i}\| \cdot \|\overrightarrow{P_j}\|} \quad (1)$$

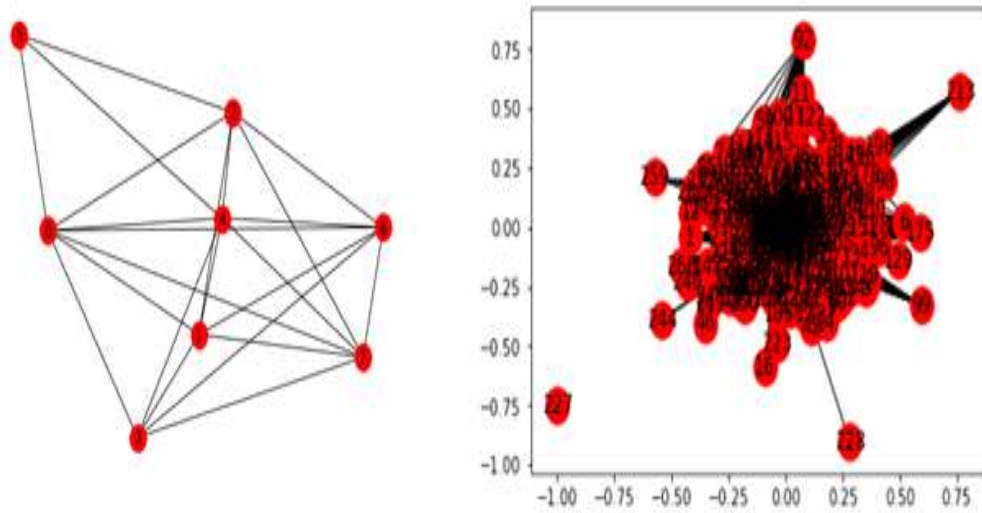
상품 간의 유사도를 측정하는 방법으로는 피어슨 상관계수, 코사인 계수, 자카드 계수 등이 있으나(김형수 2020), 본 연구에서는 아이템 기반 협업필터링에서 성능이 우수하다고 알려진 코사인 계수를 사용하였다(강호운, 옥창수, 2015). 계산된 상품 간 유사도는 0부터 1사이의 값을 가지며, 1에 가까울수록 아이템 간 유사도가 높다고 판단할 수 있다. [표 3-3]는 [표 3-2]를 바탕으로 코사인 유사도를 계산한 결과이다.

[표 3-3] 상품 간 유사도 예시

	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>	<i>P7</i>	<i>P8</i>
<i>P1</i>	1.0000	0.6659	0.5547	0.5547	0.4003	0.5883	0.3805	0.1009
<i>P2</i>	0.6659	1.0000	0.3430	0.8575	0.5941	0.0000	0.6471	0.6655
<i>P3</i>	0.5547	0.3430	1.0000	0.0000	0.8660	0.7071	0.6860	0.3638
<i>P4</i>	0.5547	0.8575	0.0000	1.0000	0.2887	0.0000	0.1715	0.6063
<i>P5</i>	0.4003	0.5941	0.8660	0.2887	1.0000	0.4082	0.7921	0.7702
<i>P6</i>	0.5883	0.0000	0.7071	0.0000	0.4082	1.0000	0.0000	0.0000
<i>P7</i>	0.3805	0.6471	0.6860	0.1715	0.7921	0.0000	1.0000	0.4991
<i>P8</i>	0.1009	0.6655	0.3638	0.6063	0.7702	0.0000	0.4991	1.0000

3.1.2 상품 네트워크 구성

상품 네트워크 구성 단계는 아이템 유사도 분석 단계를 통해 도출된 구매 상품 간의 유사도를 바탕으로 상품을 노드로 하는 소셜네트워크, 즉 상품 네트워크를 도출하는 단계이다. 상품 네트워크에서는 상품을 노드로, 상품 간 유사도를 링크로 하는 일종의 소셜네트워크이다. 즉, 상품 네트워크에서는 모든 상품에 대해서 임의의 두 상품 간 구매 유사도를 링크로 연결한다. 가령, 상품 P1과 P2의 유사도가 0.6659이면, 이 두 상품을 연결하는 링크값은 0.6659가 된다. [그림 3-2]의 왼쪽 그림은 8개 상품 예시에 대한 상품 네트워크 개념도이고, 오른쪽 그림은 본 연구에서 사용한 266개의 상품 네트워크 구성도이다.



[그림 3-2] 상품 네트워크 다이어그램

3.1.3 중심성 분석

중심성 분석 단계에서는 도출된 상품 네트워크에서 각 노드(상품) 중심성을 측정하는 단계이다. 본 연구에서는 소셜네트워크 이론의 다양한 중심성 개념 중에서 네트워크 내에서 한 노드가 담당하는 매개자 혹은 중개자 역할의 정도를 나타내는 매개중심성(Betweenness Centrality)을 활용하였다. 만약, 상품 j와 상품 k가 오직 상품 i를 통해서만 연계된다면 상품 i는 높은 매개중심성을 갖는다고 할 수 있다. 기존 연구에서도 매개중심성이 높은 상품들이 교차판매나 연쇄판매, 가격할인 등의 마케팅 전략 차원에서 유의미한 효과를 가져올 수 있다고 하였다(김병국, 정석봉, 권기석, 2013; 김혜경, 김재경, Yi, C. Q., 2009). 아래 [표 3-4]는 8개 상품 예시에 대한 매개중심성 도출 결과 테이블이다.

[표 3-4] 상품 네트워크 상의 매개중심성 예시

	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$
매개	0.026	0.026	0.034	0.015	0.026	0.015	0.026	0.015
중심성	984	984	921	873	984	873	984	873

3.1.4 중심성 가중치 적용

중심성 가중치 적용 단계는 각 노드의 매개중심성을 기존 협업필터링의 결과인 구매 선호도 값에 가중치로 적용하여 새로운 구매 선호도 값을 생성하는 단계이다. 기존 협업필터링을 통한 구매 선호도 값에 매개중심성을 가중치로 적용하게 되면 각 고객에게 매개중심성이 높은 상품은 더 강하게, 매개중심성이 낮은 상품은 더 약하게 추천하는 효과를 갖게 된다. [표 3-5]는 상기 8개 상품 예시에 대해 매개중심성을 반영하여 새롭게 구한 고객별 구매 선호도 예시를 나타내고 있다.

[표 3-5] 매개중심성을 적용한 구매 예측 선호도 예시

	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$
$C1$	0.085 191	0.060 903	0.067 143	0.020 688	0.258 875	0.018 901	0.275 431	0.164 371
$C2$	0.195 073	0.125 487	0.198 016	0.073 229	0.121 235	0.058 236	0.126 527	0.114 022
$C3$	0.364 164	0.233 206	0.576 508	0.254 301	0.116 585	0.196 951	0.117 726	0.052 158
$C4$	0.202 532	0.361 222	0.480 378	0.220 586	0.184 164	0.244 061	0.073 943	0.034 929

3.1.5 추천 리스트 생성

마지막으로 추천 리스트 생성 단계에서는 구매 예측값이 높은 상품 중 지정된 개수만큼 최상위 추천 상품 리스트를 추출하는 단계이다. 이미 구매한 상품을 다시 추천하지 않도록 하기 위해 매개중심성 기반 구매 예측 선호도

분석 결과에서 기존 구매 이력이 있을 경우 구매 선호도를 강제로 0으로 치환하였다. [표 3-6]는 기존 구매 이력이 있는 구매 선호도를 0으로 치환한 결과 테이블이다.

[표 3-6] 최종 구매 예측 선호도 예시

	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>	<i>P7</i>	<i>P8</i>
<i>C1</i>	0.085 1910	0.060 9030	0.067 1430	0.020 6880	0.000 0000	0.018 9010	0.000 0000	0.000 0000
<i>C2</i>	0.000 0000	0.000 0000	0.000 0000	0.073 2290	0.121 2350	0.058 2360	0.126 5270	0.000 0000
<i>C3</i>	0.000 0000	0.233 2060	0.000 0000	0.000 0000	0.116 5850	0.000 0000	0.000 0000	0.052 1580
<i>C4</i>	0.202 5320	0.000 0000	0.000 0000	0.000 0000	0.000 0000	0.000 0000	0.073 9430	0.034 9290

고객별 조정된 구매 선호도를 바탕으로 높은 순서대로 정렬한 후 지정된 개수만큼 추천 목록을 도출한다. 상기 예시에서 고객별 상위 2개의 상품을 출력한 결과는 [표 3-7]과 같다.

[표 3-7] 추천리스트 생성 예시

	1순위	2순위
<i>C1</i>	<i>P1</i>	<i>P3</i>
<i>C2</i>	<i>P7</i>	<i>P5</i>
<i>C3</i>	<i>P2</i>	<i>P5</i>
<i>C4</i>	<i>P1</i>	<i>P7</i>

IV. 실험 및 평가

4.1 데이터 수집 및 전처리

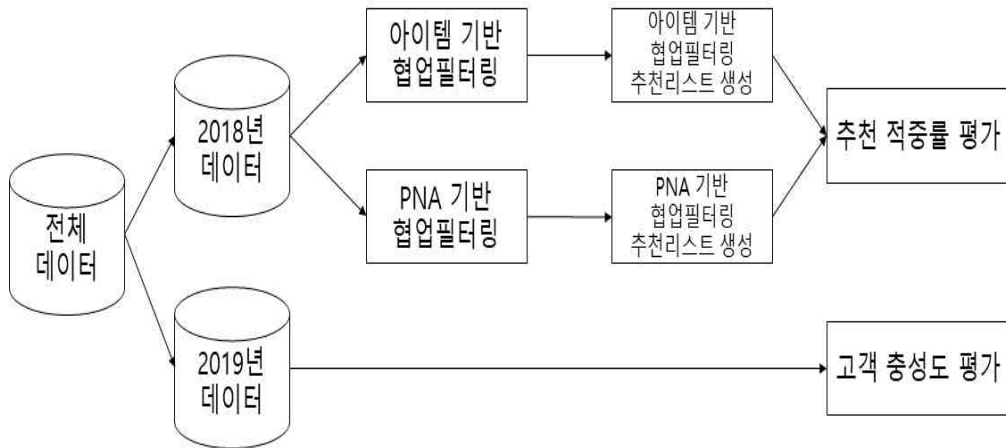
본 연구에서 제안하는 PNA 기반 협업필터링 개인화 추천 시스템의 검증을 위해 아웃도어 A사의 2018년 1월 1일부터 2019년 12월 31일까지 총 2년치의 고객 거래 데이터 1,566,925건을 분석용 데이터마트로 구성하여 사용하였다. 이중 2018년 데이터는 모델링 및 적중률 검증을 위해 사용하고, 2019년은 추천이 이루어진 이후 증장기적인 고객들의 충성도 변화를 추적하기 위해서 사용하였다. 한편, 개인화 추천을 위한 데이터셋은 거래이력이 최소 2건 이상인 경우만 해당하므로 전체 고객에서 한번만 구매한 고객은 제외하였고, 상품 종류가 매우 많아 모델링 시 데이터의 희소성 문제가 발생할 수 있기 때문에 상세 상품 품목을 중분류 단위로 카테고리화하여 분석 대상의 상품으로 취급하였다. 모델링에 사용된 2018년 거래데이터는 322,651명의 고객, 688,468건의 거래이력, 그리고 266개의 상품품목으로 구성되었다. [표 4-1]은 데이터 예시이다.

[표 4-1] 데이터 예시

고객 ID	상품 ID	구매금액	상품카테고리	구매일자
<i>C1</i>	<i>P3</i>	33,000	3S95	2018-01-12
<i>C2</i>	<i>P5</i>	23,000	1J1D	2018-07-20
<i>C3</i>	<i>P7</i>	27,000	1J2Y	2018-10-15
⋮				
<i>C4</i>	<i>P3</i>	33,000	3S95	2019-12-31

4.2 추천 시스템 검증방법

[그림 4-1]과 같이 기존의 아이템 기반 협업필터링을 경쟁모델로 하여 제안 시스템의 타당성을 검증하고자 하였다. 2018년 거래 데이터 기준 고객 322,651명을 7:3 비율로 학습용 및 평가용 데이터로 분리하고 각각 아이템 기반 협업필터링과 PNA 기반 협업필터링을 적용하여 고객별 추천 상품을 3개씩 도출하였다. 추천알고리즘의 추천 정확도를 평가하기 위해서는 평균 MAE가 이용되기도 하지만 실제 추천 대상 고객에게는 예측 선호도를 이용하여 추천 대상 고객이 가장 선호하리라 생각되는 N개의 상품 목록을 추천하는 방법이 적용되고 있다. 추천 시스템의 다양성을 고려한 성능 비교 연구를 위해서는 각 고객에게 복수 개의 상품을 추천하는 방법이 일반적이다(김형수, 박예린, 이정민, 2019; 손지은, 김성범, 김현중, 조성준, 2015). 예측 결과로써 최종 구매 상품을 모형이 얼마나 잘 맞추었는지 확인하기 위해 고객들의 구매 이력 중 최종 구매 상품은 모델링 과정에 포함하지 않았다. 모형이 예측한 고객별 추천상품 3개 중 고객들의 실제 최종 구매 상품이 차지하는 비율을 개인화 추천 모형의 적중률로 평가하였다. 이 지표는 실제 산업 현장에서 각 사용자별로 몇 개의 최상위 추천대상 상품을 추려내 추천이 이루어진다는 점을 고려할 때, 현실세계의 개인화를 반영하는 실용적인 지표라는 점에서 의미를 갖는다(Adomavicius & Kwon, 2007). 한편, 본 연구의 주요 목적이 단기적인 추천 적중률뿐만 아니라, 추천 이후 중장기적인 고객들의 충성도의 개선에 있으므로 모든 고객들의 2019년 거래 이력을 분석하여 고객들의 구매금액, 거래빈도, 구매량, 교차구매지수, 재구매율 등 고객 충성도를 정량적, 객관적으로 평가할 수 있는 지표를 도출하여 두 모형을 비교하였다.



[그림 4-1] 추천 시스템 성능 비교 프로세스

4.3 검증결과 및 비교

기존 아이템 기반 협업필터링과 PNA 기반 협업필터링의 성능 비교 결과는 [표 4-2]과 같다. 두 가지 경쟁 모형 모두 총 7번의 반복 실험을 하였고, 아래 표는 7번 측정 결과이다.

분석 결과, 모든 성과지표 관점에서 PNA 기반 협업필터링이 기존 아이템 기반 협업필터링보다 우수하게 나타났다. [표 4-3]은 7번 반복 측정에 대한 아이템 협업필터링 대비 개선율이다. 제안 모형의 평균 개선율은 추천 적중률의 경우 7.94%, 평균 구매금액은 21.43%, 평균 거래빈도는 14.06%, 평균 구매량은 19.23%, 평균 교차구매지수는 6.44% 개선되어 확연한 차이를 볼 수 있었으며, 재구매율의 경우 미약하게나마 0.35% 개선되었다는 알 수 있다. 특히, RFM 관점의 충성도 지수인 거래빈도나 구매량 및 구매금액의 경우 매우 큰 개선효과를 보인다는 것을 알 수 있다. 즉, 상품 네트워크 상의 노드별 매개중심성을 개인화 추천에 응용하면 상품 간의 연계성이 높아져 전통적인 협업필터링의 문제였던 데이터 희소성을 완화할 수 있어 단기적인 추천 적중률도 개선할 수 있을 뿐만 아니라, 추천 이후 연쇄구매효과가 높은 상품을 우선적으로 추천하기 때문에 중장기적인 구매행태 기반의 고객로열티를 개선할 수 있다고 해석할 수 있다.

[표 4-2] 아이템기반 협업필터링과 PNA기반 협업필터링의 성능비교

구분	실험횟수	추천 적중률	평균 구매금액	평균 거래빈도	평균 구매량	평균 교차구매지수	재구매율
아이템 기반 협업필터링	1회	20.76%	409,094	2.34	4.29	3.02	26.40%
	2회	20.71%	439,907	2.34	4.55	3.01	27.03%
	3회	20.76%	412,507	2.39	4.49	3.01	27.56%
	4회	20.78%	426,627	2.45	4.52	3.04	27.14%
	5회	20.72%	449,188	2.46	4.65	3.12	26.85%
	6회	20.84%	415,439	2.36	4.42	3.01	27.23%
	7회	20.66%	479,536	2.63	4.92	3.11	27.75%
	평균	20.75%	433,185	2.43	4.55	3.05	27.14%
PNA 기반 협업필터링	1회	22.63%	553,830	2.78	5.64	3.27	26.42%
	2회	22.42%	552,767	3.02	5.73	3.29	27.23%
	3회	22.48%	468,093	2.53	4.92	3.16	27.44%
	4회	22.38%	507,996	2.66	5.33	3.21	27.53%
	5회	22.21%	503,993	2.63	5.18	3.22	26.69%
	6회	22.37%	526,002	2.71	5.28	3.22	27.58%
	7회	22.27%	564,061	3.02	5.84	3.32	27.74%
	평균	22.40%	525,249	2.76	5.42	3.24	27.23%

[표 4-3] 아이템 기반 협업필터링 대비 개선율

구분	실험횟수	추천 적중률	평균 구매금액	평균 거래빈도	평균 구매량	평균 교차구매지수	재구매율
개선율	1회	9.00%	35.38%	18.88%	31.57%	8.45%	0.07%
	2회	8.26%	25.66%	28.90%	25.76%	9.17%	0.74%
	3회	8.29%	13.48%	5.93%	9.69%	5.03%	-0.44%
	4회	7.71%	19.07%	8.78%	17.76%	5.74%	1.45%
	5회	7.23%	12.20%	6.62%	11.40%	3.06%	-0.61%
	6회	7.32%	26.61%	14.87%	19.69%	6.90%	1.28%
	7회	7.78%	17.63%	14.48%	18.73%	6.72%	-0.04%
	평균	7.94%	21.43%	14.06%	19.23%	6.44%	0.35%

V. 결 론

5.1 연구요약 및 의의

본 연구에서는 소셜네트워크의 노드간 관계 및 구조적 특성을 분석하기 위해 활용되고 있는 중심성 개념을 이용하여 상품 네트워크상에서 중심성이 높은 상품을 찾고, 이를 기반으로 아이템 기반 협업필터링을 적용하는 새로운 상품 추천 시스템을 제안하였다. 제안 모형은 아이템 유사도 분석, 상품네트워크 구성, 중심성 분석, 중심성 가중치 적용, 그리고 추천리스트 생성 등의 5가지 프로세스에 의해 운용되고, 모형의 타당성을 검증하기 위해 국내 아웃도어 브랜드 A사의 협력으로 실제 고객 및 구매 데이터를 활용하여 검증하였다. 검증 결과, 추천 모형의 단기적인 성과지표인 추천 적중률뿐만 아니라, 고객의 구매 행태적 충성도를 나타내는 다양한 정량적 성과지표에서도 상대적인 우수한 성능을 보였다.

본 연구의 학술적 의의로는 첫째, 기존의 협업필터링 연구에서는 고려하지 않았던 추천 이후의 고객과의 관계를 고려하여 고객 충성도에 관련된 성과지표를 함께 고려하였다는 점이다. 교차판매전략이나 개인화 추천 전략이 고객관계를 기반으로 하는 CRM 전략 임에도 불구하고, 기존의 개인화 추천 연구들은 모형의 단기적인 정확도 개선에 집중되어 왔다. 앞으로 다양한 비즈니스 머신러닝 연구분야에서도 기술적인 개선 노력뿐만 아니라, 마케팅 이론과 전략이 함께 고려되는 전략적인 접근이 더욱 중요해질 것이다.

둘째, 소셜네트워크 이론을 개인화 추천 연구 분야에 접목했다는 점이다. 상품 간의 관계를 분석하거나, 고객 네트워크 분석을 기반으로 마케팅 전략을 제시했던 연구는 많았으나, 상품 네트워크의 특정 지표를 개인화 추천 알고리즘의 주요 파라미터로 실제로 적용했던 연구는 상대적으로 일천했던 것이 사실이다.

셋째, 실험을 위해 가공된 데이터나 대중에게 공개된 공공 데이터가 아

나라, 현재 운영 중인 주요 기업의 실제 고객 빅 데이터를 활용하여 모형을 검증하였다는 점이다. 실제 데이터 확보에 어려움이 많은 관계로 머신러닝 모형 개발 연구의 상당 부분은 가공된 데이터나 오픈 데이터를 기반으로 진행되는 경우가 많은데, 이러한 모형들이 실제 기업에 적용될 경우 실험 데이터에는 존재하지 않았던 실제 데이터의 수많은 노이즈로 인해 예상했던 결과와 큰 차이를 보이는 경우가 많다.

5.2 연구의 한계 및 향후 연구방향

한편 본 연구의 한계점으로는 첫째, 의류/패션이라는 하나의 업종만을 대상으로 진행한 연구결과이므로 일반화된 이론으로 주장하기에는 한계가 있다는 점이다. 주요 산업별 특성이 반영될 수 있도록 제조, 유통, 서비스, 금융 등 복수 업종의 데이터를 적용하여 업종별 차이를 보는 연구가 진행된다면 매우 의미 있을 것이다. 둘째, CRM 이론상 고객충성도는 관계진화단계에 따라 달라질 수 있음에도 불구하고(김형수, 김영걸, 2014), 모든 고객을 단일 집단으로 가정하고 연구를 진행하였다는 점이다. 총 2년치의 규모가 큰 데이터셋을 사용했음에도 불구하고, 비교적 구매 주기가 긴 아웃도어 상품의 특성을 고려하여 고객의 관계진화단계를 구분하고자 한다면 최소 5년 이상의 관측 데이터를 필요로 한다. 셋째, 데이터 희소성 문제를 해결하기 위해 상세 상품을 중분류 품목 단계로 일반화시켜 추천이 이루어진 것이므로 이것을 그대로 실무에 적용하기는 어렵다. 본 시스템을 실제 기업에서 운용하기 위해서는 추천 목록 추출 단계 이후 실제 상세 상품으로 다시 매핑시키는 단계가 더 필요하다.

참 고 문 헌

1. 국내문헌

- 강부식. (2010). 사회연결망의 구조적 공백을 활용한 신규고객 웹 상품추천방법. 『산업경제연구』, 23(5), 2371-2385.
- 강양, 곽용섭. (2012). 고객기반의 호텔 브랜드자산에 관한연구-브랜드 확장의 관점으로. 『Tourism Research』, (34), 59-81.
- 강호운, 옥창수. (2015). 아이템 연관성을 고려한 협력적 필터링 기반 추천시스템 개발. 『Entrue Journal of Information Technology』, 14(1), 135-144.
- 곽기영. (2014). 『소셜네트워크분석=Social network analysis』. 서울: 청람.
- 김민성, 임일. (2014). 소셜 네트워크 분석 기법을 활용한 협업필터링의 특이 취향 사용자 (Gray Sheep) 문제 해결. 『지능정보연구』, 20(2), 137-148.
- 김병국, 정석봉, 권기석. (2013). 사회연결망분석에 의한 온라인 쇼핑몰의구매 품목 관계 분석에 대한 연구. 『디지털융복합연구』, 11(11), 209-217.
- 김병만, 이경, 김시관, 임은기, 김주연. (2004). 추천시스템을 위한 내용기반 필터링과 협력필터링의 새로운 결합 기법. 『정보과학회논문지: 소프트웨어 및 응용』, 31(3), 332-342.
- 김용학. (2014). 『사회연결망분석』. 서울: 박영사.
- 김용학, 김영진. (2016). 『사회연결망 분석 (4 판)』. 서울: 박영사.
- 김필성, 김기봉, 조영택, 신용태. (2020). 개인화 상품 추천이 구매의도에 미치는 영향에 관한 연구: Matrix Factorization 알고리즘을 중심으로. 『한국 IT 정책경영학회 논문지』, 12(4), 1847-1852.
- 김형수. (2020). 『Step by Step 비즈니스 머신러닝 인 파이썬』. 서울: 프레딕스.
- 김형수, 김영걸. (2014). 『CRM 고객관계관리 전략 원리와 응용』. 서울: 영출판사.

- 김형수, 박예린, 이정민. (2019). 머신러닝을 이용한 공연문화예술 개인화 장르 추천 시스템. 『Information Systems Review』, 21(4), 31-45.
- 김혜경, 김재경, Yi, C. Q. (2009). 백화점 거래 데이터를 이용한 상품네트워크 연구. 『지능정보연구』, 15(4), 159-180.
- 박석철, 김성연. (2009). 인터넷 쇼핑몰 이용에 따른 불만족 귀인과 구매 후 행동의 관계. 『사이버커뮤니케이션학보』, 26(2), 85-116.
- 박종학, 조운호, 김재경. (2009). 사회연결망: 신규고객 추천문제의 새로운 접근법. 『지능정보연구』, 15(1), 123-140.
- 복경수, 유재수. (2017). “4차 산업혁명에서 빅데이터”. 『정보과학회지』, 35(6), 22-39.
- 사이람. (2017). 『NetMiner를 이용한 소셜 네트워크 분석: 기본교육』. 서울: 사이람.
- 서문식, 이지은, 조상현. (2010). 온라인 브랜드 커뮤니티에서의 관계단절 행동. 『한국콘텐츠학회논문지』, 10(12), 352-365.
- 서일석. (2016). 빅데이터와 핀테크 스타트업의 기회 및 동향. 『정보보호학회지』, 26(2), 20-24.
- 손동원. (2002). 『사회 네트워크 분석』. 서울: 경문사.
- 손지은, 김성범, 김현중, 조성준. (2015). 추천 시스템 기법 연구동향 분석. 『대한산업공학회지』, 41(2), 185-208.
- 신창훈, 이지원, 양한나, 최일영. (2012). 협력필터링과 사회연결망을 이용한 신규고객 추천방법에 대한 연구. 『지능정보연구』, 18(4), 19-42.
- 안성식, 박연옥, 강병오. (2011). 레스토랑 유형에 따라 서비스 품질이 태도적, 행동적 충성도에 미치는 영향력 차이에 관한 연구. 『Culinary Science & Hospitality Research』, 17(1), 26-43.
- 이수상. (2010). 공저 네트워크 분석에 관한 기초연구: 문헌정보학 분야 4개 학술지를 중심으로. 『한국도서관·정보학회지』, 41(2), 297-315.
- 이수상. (2012). 『네트워크 분석 방법론』. 서울: 논형.
- 이수상. (2013). 연구자 네트워크의 중심성과 연구성과의 연관성 분석: 국내 기록관리학 분야 학술논문을 중심으로. 『한국도서관·정보학회지』, 44(3),

405-428.

- 이준혁, 박헌진. (2019). 산업체 급식서비스 요인이 환경친화성, 고객만족 및 고객충성도에 미치는 영향-밀레니얼세대의 인식을 중심으로. 『지역산업연구』, 42(1), 113-131.
- 장나영, 이규민. (2016). 한식 뷔페 레스토랑 선택속성이 지각된 가치와 고객만족도 및 충성도에 미치는 영향. 『외식경영연구』, 19, 93-114.
- 정석봉. (2014). 인터넷 쇼핑몰의 고객분류 및 선호상품 추출을 위한 사회연결망분석 기반의 새로운 접근. 『Entrue Journal of Information Technology』, 13(1), 57-68.
- 정석봉. (2017). 오프라인 쇼핑몰에서 고객의 과거 구매 패턴을 활용한 아이템 기반 협업필터링 성능 개선에 관한 연구. 『Journal of Information Technology Applications & Management』, 24(4), 1-12.
- 조운호, 김인환. (2010). 사회연결망분석과 인공지능망을 이용한 추천시스템 성능 예측. 『한국지능정보시스템학회학회지』, 16(4), 159-172.
- 조운호, 방정혜. (2009). 신상품 추천을 위한 사회연결망분석의 활용. 『한국지능정보시스템학회학술지』, 15(4), 183-200.
- 최웅, 이영우, 조승현. (2012). 워터파크 브랜드 자산가치가 고객만족 및 소비자 충성도에 미치는 영향. 『호텔리조트연구』, 11(2), 23-43.
- 최인호, 김병국, 정석봉. (2014). 상품네트워크 분석을 통한 새로운 상품분류 기준에 관한 연구. 『Entrue Journal of Information Technology』, 13(2), 93-104.
- 한국CRM협회. (2016). 『국내 상장기업들의 CRM 운영실태 평가 보고서』. 한국CRM협회.
- 황용석, 김기태. (2019). 개인화 서비스 진전에 따른 자동추천 시스템 연구 동향과 방법론적 특성 연구. 『사이버커뮤니케이션학보』, 36(2), 221-253.

2. 국외문헌

- Adomavicius, G., & Kwon, Y. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems*, 22(3), 48–55.
- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6), 734–749.
- Ahn, H. (2014). Improvement of a Context-aware Recommender System through User's Emotional State Prediction. *Journal of Information Technology Applications and Management*, 21(4), 203–223.
- Al Mamunur Rashid, S. K. L., Karypis, G., & Riedl, J. (2006). ClustKNN: a highly scalable hybrid model-& memory-based CF algorithm. *Proceeding of webKDD*.
- Balabanović, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3), 66–72.
- Baloglu, S. (2002). Dimensions of customer loyalty: Separating friends from well wishers. *Cornell Hotel and Restaurant Administration Quarterly*, 43(1), 47–59.
- Barragáns-Martínez, A. B., Costa-Montenegro, E., Burguillo, J. C., Rey-López, M., Mikic-Fonte, F. A., & Peleteiro, A. (2010). A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition. *Information Sciences*, 180(22), 4290–4311.
- Bogdanov, D., Haro, M., Fuhrmann, F., Xambó, A., Gómez, E., & Herrera, P. (2013). Semantic audio content-based music recommendation and visualization based on user preference examples. *Information Processing & Management*, 49(1), 13–33.
- Bonacich, P. (1987). Power and centrality: A family of

- measures. *American journal of sociology*, 92(5), 1170–1182.
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4), 331–370.
- Celma, Ò., & Herrera, P. (2008). A new approach to evaluating novel recommendations. *In Proceedings of the 2008 ACM conference on Recommender systems* (pp. 179–186).
- Christakis, N. A., & Fowler, J. H. (2010). *Connected: The amazing power of social networks and how they shape our lives*. HarperCollins UK.
- Cremonesi, P., Garzotto, F., & Turrin, R. (2012). Investigating the persuasion potential of recommender systems from a quality perspective: An empirical study. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(2), 1–41.
- Dekimpe, M. G., Steenkamp, J. B. E., Mellens, M., & Abeele, P. V. (1997). Decline and variability in brand loyalty. *international Journal of Research in Marketing*, 14(5), 405–420.
- Dick, A. S., & Basu, K. (1994). Customer loyalty: toward an integrated conceptual framework. *Journal of the academy of marketing science*, 22(2), 99–113.
- Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine learning*, 29(2–3), 103–130.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215–239.
- Ganesh, J., Arnold, M. J., & Reynolds, K. E. (2000). Understanding the customer base of service providers: an examination of the differences between switchers and stayers. *Journal of marketing*, 64(3), 65–87.
- Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). Using

- collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 61–70.
- Gremler, D. D., & Brown, S. W. (1996). Service loyalty: its nature, importance, and implications. *Advancing service quality: A global perspective*, 5(1), 171–181.
- Gustafsson, A., Johnson, M. D., & Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of marketing*, 69(4), 210–218.
- Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing science*, 19(1), 4–21.
- Huang, G. Q., Zhao, W. J., & Lu, Q. Q. (2013). Bat algorithm with global convergence for solving large-scale optimization problem. *Application Research of Computers*, 30(3), 1–10.
- Johnson, M. D., Herrmann, A., & Huber, F. (2006). The evolution of loyalty intentions. *Journal of marketing*, 70(2), 122–132.
- Jones, T. O., & Sasser, W. E. (1995). Why satisfied customers defect. *Harvard business review*, 73(6), 88.
- Kim, H. K., Kim, J. K., & Chen, Q. Y. (2012). A product network analysis for extending the market basket analysis. *Expert Systems with Applications*, 39(8), 7403–7410.
- Kim, J. (2017). How Did the Information Flow in the# AlphaGo Hashtag Network? A Social Network Analysis of the Large-Scale Information Network on Twitter. *Cyberpsychology, Behavior, and Social Networking*, 20(12), 746–752.
- Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4–5), 441–504.

- Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM*, 40(3), 77–87.
- Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Psychological bulletin*, 108(2), 171.
- De Nooy, W., Mrvar, A., & Batagelj, V. (2018). *Exploratory social network analysis with Pajek: Revised and expanded edition for updated software (Vol. 46)*. Cambridge University Press.
- Oh, H. C. (1995). *An empirical study of the relationship between restaurant image and customer loyalty* (Doctoral dissertation, Virginia Tech).
- Oliver R, L. (1997). *Satisfaction: A behavioral perspective on the consumer*. New York ' NY: Irwin-McGraw-Hill.
- Oliver, R. L. (1999). Whence consumer loyalty?. *Journal of marketing*, 63(4_suppl1), 33–44.
- Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, 13(5–6), 393–408.
- Pine II, B. J. (1993). *Mass Customization*, Harvard Business School Press. Boston: MA.
- Prus, A., & Brandt, D. R. (1995). Understanding your customers. *Marketing tools*, 10–13.
- Pu, P., Chen, L., & Hu, R. (2012). Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction*, 22(4–5), 317–355.
- Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*.
- Robins, G. (2015). *Doing social network research: Network-based research design for social scientists*. Sage.

- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. *In Proceedings of the 10th international conference on World Wide Web* (pp. 285–295).
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2002). Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering. *In Proceedings of the fifth international conference on computer and information technology* (Vol. 1, pp. 291–324).
- Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data mining and knowledge discovery*, 5(1–2), 115–153.
- Sharp, B., & Sharp, A. (1997). Loyalty programs and their impact on repeat-purchase loyalty patterns. *International journal of Research in Marketing*, 14(5), 473–486.
- Tewari, A. S., & Barman, A. G. (2018). Sequencing of items in personalized recommendations using multiple recommendation techniques. *Expert Systems with Applications*, 97, 70–82.
- Walsh, G., Dinnie, K., & Wiedmann, K. P. (2006). How do corporate reputation and customer satisfaction impact customer defection? A study of private energy customers in Germany. *Journal of Services Marketing*.
- Wu, Y. H., & Chen, A. L. (2000). Index structures of user profiles for efficient web page filtering services. *In Proceedings 20th IEEE International Conference on Distributed Computing Systems* (pp. 644–651).
- Zech, S. T., & Gabbay, M. (2016). Social network analysis in the study of terrorism and insurgency: From organization to politics. *International Studies Review*, 18(2), 214–243.

Zhang, Y. C., Séaghdha, D. Ó., Quercia, D., & Jambor, T. (2012). Auralist: introducing serendipity into music recommendation. *In Proceedings of the fifth ACM international conference on Web search and data mining* (pp. 13–22).

ABSTRACT

A recommendation system based on PNA
collaborative filtering for reinforcing customer
loyalty

Kim, Seung-Hyeong

Major in Industrial & Management Engineering

Dept. of Industrial & Management Engineering

The Graduate School

Hansung University

In the many industry, the importance of CRM strategies using big data analysis technology is increasing. However, most of the research on the existing recommendation models, so the mid- to long-term relationship with customers after the recommendation activity has not been considered. In this study, the PNA technique, which applied the social network analysis method to the product network analysis, is linked to the collaborative filtering model, we would like to propose a recommendation system that can improve loyalty. The CF/PNA recommendation system configures a product network based on transaction data, and then selects products that can improve customer loyalty in the mid to long term based on PNA analysis results among candidate recommended products. As a result of modeling and verification based on actual customer data in cooperation with outdoor companies, it

was confirmed that customer loyalty index such as buying frequency, repurchase rate, and cross-purchase ratio evaluating the qualitative level of purchasing as well as the hit ratio have been increased. This study provides strategic insights that it may be desirable to withhold recommendations from a CRM viewpoint when specific product has not sure to maintain a continuous customer relationship due to lack of linkage between other products.

[Keywords] CF/PNA recommendation system, Product Network Analysis, Collaborative Filtering, CRM