



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

碩士學位論文

개인형 추천 서비스를 위한  
사용자 성향분석 엔진

濟州大學校 大學院

컴퓨터工學科

金旻翹

2013년 6월

# 개인형 추천 서비스를 위한 사용자 성향분석 엔진

指導教授 李 尙 俊


金 旻 翹

이 論文을 工學 碩士學位 論文으로 提出함

2013年 06月

金旻翹의 工學 碩士學位 論文을 認准함


審査委員長

홍 망 천 

委 員

변 상 용 

委 員

이 상 진 

濟州大學校 大學院

2013年 06月

# 목 차

그림목차 .....	i
표 목차 .....	ii
국문초록 .....	iii
영문초록 .....	iv
약어표 .....	v
I. 서론 .....	1
II. 관련연구 .....	3
1. 추천 알고리즘에 관한 연구 .....	3
1) 협업적 필터링 .....	4
2) 내용기반 필터링 .....	6
3) 하이브리드 기법 .....	7
2. 시스템 적용에 관한 연구 .....	8
1) 음악 추천서비스 .....	8
2) 음식 추천서비스 .....	9
3) 영화 추천서비스 .....	9
III. 사용자 성향분석 엔진 .....	10
1. 데이터 요소 정의 .....	10
2. 수집 .....	12
3. 가중치 설정 .....	14
4. 성향분석 .....	15
5. 추천 및 피드백 .....	18
6. 사용자 성향분석 엔진 모형 .....	21
IV. 실험 환경 및 결과 .....	22
1. 실험환경 .....	22
1) 하드웨어 환경목록 .....	23
2) 소프트웨어 환경목록 .....	23
3) 추천 음식목록 .....	23

4) 데이터 요소 정의 .....	24
5) 환경설정 .....	24
6) 가중치 설정 .....	25
2. 실험결과 .....	27
1) 환경 및 사용자 설정: A결과 .....	27
2) 환경 및 사용자 설정: B결과 .....	29
V. 결론 .....	30
참고문헌 .....	31

## 그림 목 차

<그림 1. 음식 추천서비스에서의 협업적 필터링 예시> .....	4
<그림 2. 음식추천 서비스에서의 내용기반 필터링 예시> .....	6
<그림 3. 일반 데이터 요소 예시> .....	10
<그림 4. 정보수집 절차> .....	12
<그림 5. 추천목록 예> .....	20
<그림 6. 사용자 분석엔진 모형> .....	21
<그림 7. 실험 환경 시스템 구조도> .....	22

## 표 목 차

<표 1. 추천 시스템 개념> .....	3
<표 2. 일반적 데이터 요소 정의> .....	11
<표 3. 수집 데이터 목록> .....	13
<표 4. 기온의 가중치와 범위설정 예시> .....	14
<표 5. 환경설정 가정> .....	15
<표 6. 가중치 설정 : 사용자 성향 미적용> .....	16
<표 7. 습도에 민감한 사용자 민감도 예시> .....	16
<표 8. 가중치 설정 : 사용자 성향 적용> .....	17
<표 9. 사용자 민감도 정보 예시> .....	18
<표 10. 추천목록별 가중치 예시> .....	18
<표 11. 사용자 추천목록 예시> .....	18
<표 12. 사용자 민감도 변경 예시> .....	19
<표 13. 하드웨어 환경목록> .....	23
<표 14. 소프트웨어 환경목록> .....	23
<표 15. 추천 음식목록> .....	23
<표 16. 수집 데이터 목록> .....	24
<표 17. 환경 및 사용자 설정> .....	24
<표 18. 온도에 따른 음식별 가중치> .....	25
<표 19. 습도에 따른 음식별 가중치> .....	25
<표 20. 계절에 따른 음식별 가중치> .....	25
<표 21. 나이에 따른 음식별 가중치> .....	26
<표 22. 사용자 민감도 초기값 설정> .....	26
<표 23. 추천서비스 결과 : 피드백 미적용> .....	27
<표 24. 사용자 민감도 변경 목록> .....	28
<표 25. 추천서비스 결과 : 피드백 진행 적용> .....	28
<표 26. 추천서비스 결과 : 피드백 미적용> .....	29
<표 27. 사용자 민감도 변경 목록> .....	29
<표 28. 추천서비스 결과 : 피드백 진행 적용> .....	29

국문초록

## 개인형 추천 서비스를 위한 사용자 성향분석 엔진

컴퓨터공학과 김민교

지도교수 이상준

본 논문에서는 사용자의 환경, 양방향 정보 등을 통해 개인형 추천 서비스를 제공하기 위한 사용자 성향 분석 엔진 모델을 제안한다. 사용자 성향 분석을 위해 분석할 사용자 데이터 요소를 정의하고 수집, 가중치, 성향분석, 추천, 피드백 과정을 통해 연관관계를 구성하고, 개인형 추천 서비스를 개선된 형태로 제공할 수 있도록 한다. 또한 제안된 성향 분석 엔진을 음식 추천 서비스에 적용하여 응용방안을 제시한다.



## ABSTRACT

# User inclination analysis engine for the personalized recommendation service

KIM, MIN-KYO

Department of Computer Engineering

Graduate School

Jeju National University

This paper propose the user inclination analysis engine model for the personalized recommendation service. The application area of this study is selecting food based on individual user's preferences. We define the data elements of the user's preference for the foods. By analyzing the user's preferences about the foods data, we compose the associated relations of those data. This process facilitation provide the enhanced personal recommendation service. For those purposes, we present the inclination analysis engine for the food selection plan service.

## 약어 표

CF	Collaborative Filtering
CBF	Content Based Filtering

## I. 서론

최근 인터넷 환경이 진화하면서 많은 디지털 정보가 생성, 제공되고 있으며 인류가 한해 만들어 내는 디지털 정보의 양은 1조 기가바이트 수준까지 증가하였다[1].

디지털 정보의 양이 증가함에 따라 사용자 각자가 원하는 정보를 쉽게 전달받는 개인화 서비스의 중요도는 증가하고 있으며 사용자 성향을 분석하고 개인이 원하는 콘텐츠를 시간, 장소, 상황 등에 맞게 추천해주는 서비스가 제공되고 있다. 개인형 추천 서비스의 대표적인 회사로는 아마존 닷컴과 넷플릭스가 있으며 아마존 닷컴은 1997년 고객의 도서 구매이력을 바탕으로 도서 추천하는 서비스를 제공하였으며 추천 시스템 고도화를 위해 매년 이익의 10%를 투자하고 있다. 또한 전세계 40개국에 3300만명의 가입자를 확보하고 있는 온라인 스트리밍 서비스 성공기업인 넷플릭스는 “시네매치”라는 자체 추천 알고리즘을 개발하여 서비스를 제공하고 있으며 영화, TV드라마 시청의 70%가 이 엔진에 의한 것으로 분석하고 있다[2][3].

콘텐츠 추천서비스 연구동향은 크게 두 가지로 구분되며 추천 알고리즘 및 방법론에 관한 연구와 추천 시스템의 응용을 위한 시스템 구성 및 적용에 관한 연구이다. 추천 알고리즘 및 방법론에 관한 연구는 협업적 필터링, 내용기반 필터링과 앞의 두 기법을 결합한 하이브리드 기법으로 구분된다. 추천 시스템의 응용에 관한 연구는 2000년대 초반까지 주로 영화, 음악분야였으며 현재 전자상거래, 인터넷 콘텐츠, 교육, 스마트폰 등으로 확산 적용되고 있다[3].

본 논문에서는 개인을 그룹으로 성향을 구분 짓는 것이 아닌 개인의 사용자 환경, 양방향 정보 등을 통해 성향 분석 엔진 모델을 제안하고 내용기반 필터링 기법을 이용한 추천 서비스의 정확도를 높일 수 있는 응용모델을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 추천서비스와 관련된 기본 개념 및 기술을 살펴본다. 3장에서는 사용자 성향 분석을 위해 분석할 사용자 데이터 요소를 정의하고 수집, 가중치, 성향분석, 추천, 피드백 과정을 통해 연관관계를 구성하고 개인형 추천 서비스를 개선된 형태의 모델을 제안한다. 4장에서는 제안된

성향 분석 엔진을 음식 추천 서비스에 적용하여 응용방안을 제시하고 시험 및 분석을 한다. 끝으로 5장에서 결론을 맺는다.

## II. 관련연구

### 1. 추천 알고리즘에 관한 연구

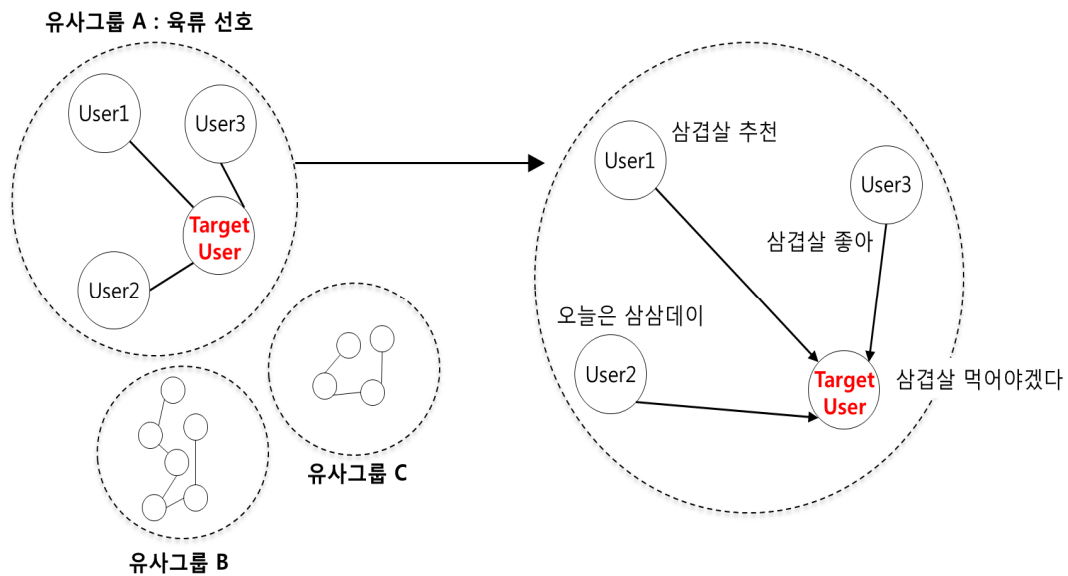
추천 알고리즘에 대한 연구는 협업적 필터링(collaborative filtering), 내용 기반 필터링(content based filtering)과 앞의 두 기법을 결합한 하이브리드(hybrid method)기법으로 구분된다. 추천 알고리즘을 분석하고 학습하는 내용은 크게 3가지로 구성되며 Contents, Neighbor, Context로 구성 된다. Contents는 선호하는 정도, Neighbor는 유사선호 그룹을 의미하며 Context는 환경정보를 의미한다. 추천 알고리즘을 통해 Contents, Neighbor, Context를 분석하고 학습하는 과정을 반복하여 사용자에게 추천 서비스를 제공해 준다[3].

<표 1. 추천 시스템 개념[3]>

목록	음식 추천서비스 예시
Contents	<ul style="list-style-type: none"> <li>- 어떤 종류의 음식을 좋아하는가?</li> <li>- 못 먹는 음식은 없는가?</li> <li>- 술은 좋아하는가?</li> </ul>
Neighbor	<ul style="list-style-type: none"> <li>- 유사 선호 그룹은 어디인가?</li> <li>- 유사 선호 그룹이 좋아하는 음식은 어떤 것인가?</li> </ul>
Context	<ul style="list-style-type: none"> <li>- 저녁시간에 어떤 음식을 좋아하는가?</li> <li>- 80% 습도에 파전을 좋아하는가?</li> <li>- 여름에 차가운 음식을 좋아하는가?</li> </ul>
<p><b>협업적/내용기반 필터링을 통해 Contents, Neighbor, Context를 분석, 학습하고 개인형 추천서비스 제공</b></p>	

## 1) 협업적 필터링

협업적 필터링은 각 사용자별로 그 사용자와 가장 선호정보가 유사한 사용자 이웃(neighbor)을 찾는 다음, 이웃들이 선호하는 상품을 그 사용자에게 추천하는 것이다. 사용자의 유사한 이웃을 찾을 때는 Pearson correlation coefficient, Jaccard coefficient, Cosine vector 등의 알고리즘이 활용된다. 이러한 방법을 활용하여 각 사용자와 성향이 유사한 사용자를 찾아낸 다음, 각각의 가중평균으로 사용자가 접하지 못한 목록에 대한 선호도를 예측하는 것이다. 각각의 유사도 측도에 따라 이웃이 다르게 형성되고 이에 따라 선호도 점수가 매우 다르게 나타나므로, 이웃을 형성하는 과정이 협업적 필터링에서 중요한 요소라 할 수 있다 [3].



<그림 1. 음식 추천서비스에서의 협업적 필터링 예시>

<그림 1>의 예시와 같이 협업적 필터링은 육류선호 그룹에 유사도 측도에 따라 이웃을 형성하고 이웃들이 Target User에게 음식 추천서비스를 제공해주는

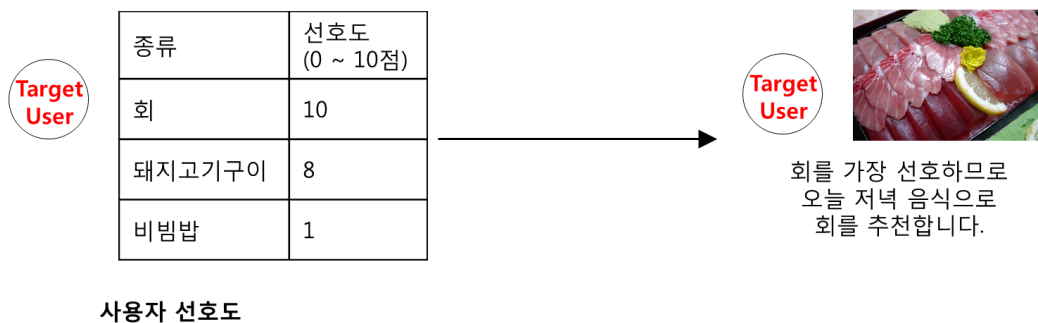
구조이다.

1990년대 중반 협업 필터링에 대한 논문이 처음 발표된 이래로 현재에도 많은 연구자들이 성능을 개선하기 위한 연구를 꾸준히 진행하고 있고, 온톨로지를 이용한 성능 개선에 대한 연구도 진행되고 있다[7]. 2000년대 중반까지 이루어진 협업적 필터링에 관한 연구는 주로 사용자가 접한 뉴스 또는 음악의 선호도를 사용자가 직접 평가한 데이터(explicit data) 또는 상품의 구매여부(binary data)와 같은 데이터를 바탕으로 협업적 필터링에 대한 연구가 이루어졌다. 그러나 2000년대 중후반부터는 전자상거래 사이트에서 사용자의 행동 및 탐색 패턴을 고려한 협업적 필터링 기법이 개발되었고 이 방법은 사용자가 직접 입력하지 않아도 자동으로 분석할 수 있다[3].

그러나 협업적 필터링의 경우 사용자 선호정보에 따라 분석할 수 있는 한계가 있다. 사용자 선호정보에 따라 유사 이웃을 찾고 이웃들이 선호하는 상품을 추천해주는 방식이기 때문이다. 선호정보가 없는 Target User가 협업적 필터링을 통해 추천 서비스를 받기 위해서는 많은 수의 사용자와의 유사도를 계산해야하며 이를 통한 추천서비스 또한 만족도가 떨어진다. 이러한 문제점을 보완하기 위한 연구도 진행되었으며 차원감소기법을 이용하여 연산속도를 개선하고 추천서비스의 만족도도 높일 수 있다[8].

## 2) 내용기반 필터링

내용기반 필터링은 각 사용자의 특정 콘텐츠 또는 상품의 속성을 학습하는 것을 기반으로 한다. 내용기반 필터링 기법은 정보검색 또는 정보추출 분야에서 발전된 것으로써, 콘텐츠 또는 상품의 추천을 위해 콘텐츠 또는 상품의 내용과 사용자가 요구하는 정보간의 유사도를 계산한 다음, 유사도가 높은 결과를 추천하는 것이다. 이러한 방법으로 내용을 기반으로 하여 추천하는 기법을 내용기반 필터링이라고 한다. 이를 구현하기 위해 사용되는 방법으로는 가중치 기법, 적합성 피드백, 확률검색 모형 등을 활용한다. 내용기반 필터링 기법의 단점으로는 사용자 정보와 콘텐츠 또는 상품을 비교하여 추천되므로 유사 콘텐츠 또는 상품이 지속적으로 추천되며 콘텐츠 또는 상품별로 일일이 정의해야하는 번거로운 점이 있다. 현재까지는 주로 음악 장르 또는 문서의 분류에 적용되고 있다[3].



<그림 2. 음식추천 서비스에서의 내용기반 필터링 예시>

<그림 2>의 예시와 같이 내용기반 필터링은 사용자가 가장 선호하는 음식종류에 대해 분석을 하고 이와 유사한 음식을 사용자에게 추천해주는 서비스를 제공해주는 구조이다.



### 3) 하이브리드 기법

하이브리드 기법은 협업적 필터링과 내용기반 필터링을 결합하여 서로의 단점을 보완하는 형식의 연구가 진행되고 있다. 협업적 필터링은 사용자 선호정보가 부족할 경우 추천서비스의 한계가 있으며 유사 이웃을 찾는데 연산이 많아지며 추천서비스의 만족도가 낮아진다. 내용기반 필터링은 사용자 정보와 콘텐츠 또는 상품을 비교하여 추천되므로 유사 콘텐츠 또는 상품이 지속적으로 추천되며 콘텐츠 또는 상품별로 일일이 정의해야하는 번거로운 점이 있다. 이와 같은 단점을 서로 보완하는 형식의 연구가 주를 이루고 있으며[3] 협업적 필터링과 내용기반 필터링을 적절하게 조합하여 추천서비스의 만족도를 높일 수 있도록 사용자 이용 빈도수, 패턴에 따라 필터링 방법을 다르게 적용하는 연구가 있다[9]. 또한 추천 대상이 되는 목표 고객과 가장 가까운 이웃집단을 발견하여 그들이 선호하는 상품을 발견하는 과정과 목표 고객이 평가한 상품과 유사한 선호도를 나타내는 상품을 발견하는 과정을 동시에 고려하는 하이브리드 추천 알고리즘을 제시하고 상품 사이의 유사도와 고객들 사이의 유사도를 계산하여 추천의 정확도를 높이는 방법에 연구도 진행되었다[10].

## 2. 시스템 적용에 관한 연구

1990년대 후반부터 2000년대 초반까지 사용자의 선호도가 비교적 분명히 드러나는 영화와 음악에 관한 추천서비스가 주를 이루었다. 연구 그룹의 데이터 공유로 인하여 연구자들이 해당 데이터를 쉽게 취득할 수 있었기 때문이다. 하지만 사용자에게서 데이터를 입력받아서 추천을 제공해주는 서비스는 한계가 있었다. 아마존닷컴과 넷플릭스가 상업적이고 시스템 적용과 활용을 위한 초기모델이라 할 수 있다[3].

2000년대 중반에는 전자상거래 사이트에서 고객 클릭 패턴을 대상으로 연구가 이루어졌으며[3] 협업적 필터링에 대한 연구도 수행되었다. 협업적 필터링에 관련된 다양한 이슈들(희박성, 확장성, 이웃집단의 형성방법, 추천 상품의 선정방법 등)에 대해 설명하고, 데이터 집단 간의 수치 실험을 진행하였다[10].

2000년대 후반에는 전자상거래, 인터넷 콘텐츠 사이트, 교육, 블로그 추천 등으로 활용범위가 확대되었으며 스마트 디바이스를 통한 추천서비스로도 적용되고 있다.

### 1) 음악 추천서비스

음악과 관련된 서비스는 제목, 가수, 작곡가 등의 메타데이터를 이용하여 검색하고 듣는 형태에서[11] 협업적 필터링과 내용기반 필터링을 활용하여 음악 및 아티스트를 추천하는 서비스를 제공한다[3]. 사용자가 좋아하는 음악의 대표 선율을 추출하고 내용기반 필터링을 통해 유사한 선율의 음악을 추천해주기도 하며[11] 각각의 사용자의 정취정보를 시간 개념단위로 나누고 협업적 필터링 이용하여 음악을 추천해주는 연구도 있다[7].

음원제공 사이트 멜론은 협업적 필터링과 내용기반 필터링을 활용하여 음악 및 아티스트를 추천하는 서비스와 추천 근거 또한 제시하였다[3].

## 2) 음식 추천서비스

음식추천 서비스에 관련된 연구로서 트위터로부터 자동 생성한 사용자 정보를 이용하여 내용기반 필터링과 협업적 필터링을 진행하여 음식 추천서비스를 제공할 때 사용자 정보가 부족하여 적절한 추천이 못한다는 문제점을 해결하려 했으며[12], 모바일 단말을 통해 현재 사용자 위치정보를 파악하고 사용자 상황정보와 개인 선호 정보 등을 이용하여 맛집을 추천해주는 서비스 적용 방법에 대한 연구도 있다[13].

스마트폰 보급과 더불어 상황정보(context)를 쉽게 얻을 수 있으며 이를 통하여 음식 추천서비스에 제공한다. 그 예로 2004년 창립한 미국 지역 정보 사이트인 Yelp는 사용자 성향을 그룹화 한 후 레스토랑 및 각종 소매점을 추천 서비스를 제공하였으며 2011년 8,300만 달러의 매출을 올렸다[3].

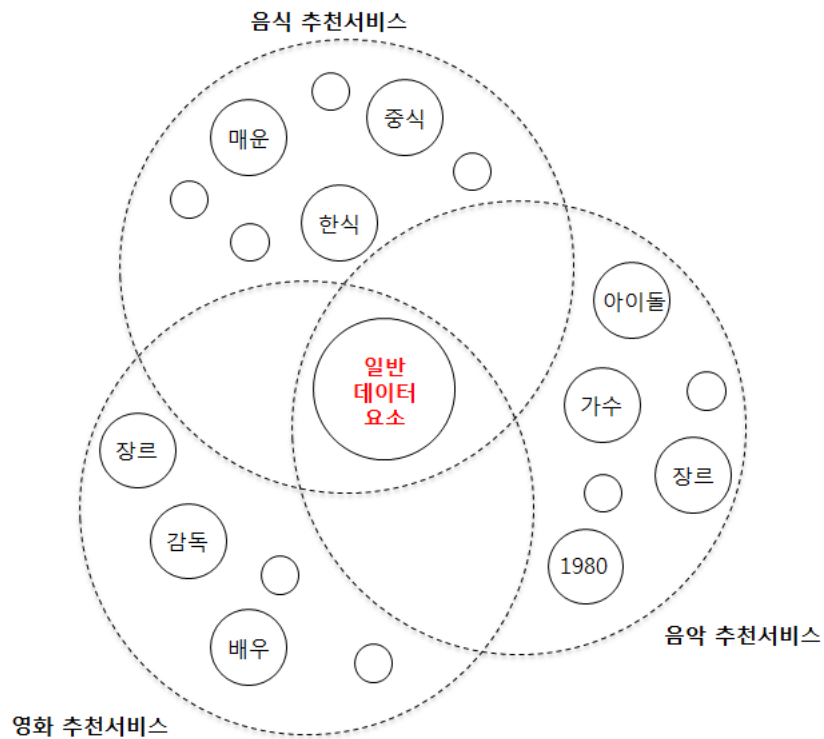
## 3) 영화 추천서비스

영화 추천서비스에 관련된 연구로서 협업적 필터링과 퍼지시스템을 이용하여 사용자가 영화를 본 후 평가점수를 예측하는 시스템을 제안하였으며 퍼지시스템을 통해 협업적 필터링의 단점을 해결하려 하였고[5] 협업적 필터링의 단점을 콘텐츠 정보를 결합하여 해결하려고 한다[13].

### III. 사용자 성향분석 엔진

#### 1. 데이터 요소 정의

여러 추천 서비스에 따라 다양한 데이터 요소가 존재하겠지만 일반적 데이터 요소로 정의를 하고 각각의 추천 서비스에 따라 데이터 요소를 추가하여 정의하기로 한다. <그림 3>의 일반 데이터 요소 예시와 같이 음식, 영화, 음악 추천서비스에서 사용되는 일반 데이터 요소를 정의하고 추가되는 추천 서비스의 성격에 따라 데이터 요소를 추가하는 것이다.



<그림 3. 일반 데이터 요소 예시>

사용자 성향분석을 위해 수집할 일반적 데이터는 환경, 사용자 특성, 양방향 데이터 요소로 크게 구분된다. 사용자들의 환경 데이터 요소는 위치, 기온, 습도,

계절, 시간 등으로 정의하며 특성 데이터 요소는 성별, 나이, 선호정보 등으로 정의한다. 양방향 데이터 요소는 환경에 따라 추천 받은 결과를 평가하여 성향 분석 요소로 정의한다. 데이터 요소를 일반적으로 정의한 내용은 아래의 <표 4>와 같다.

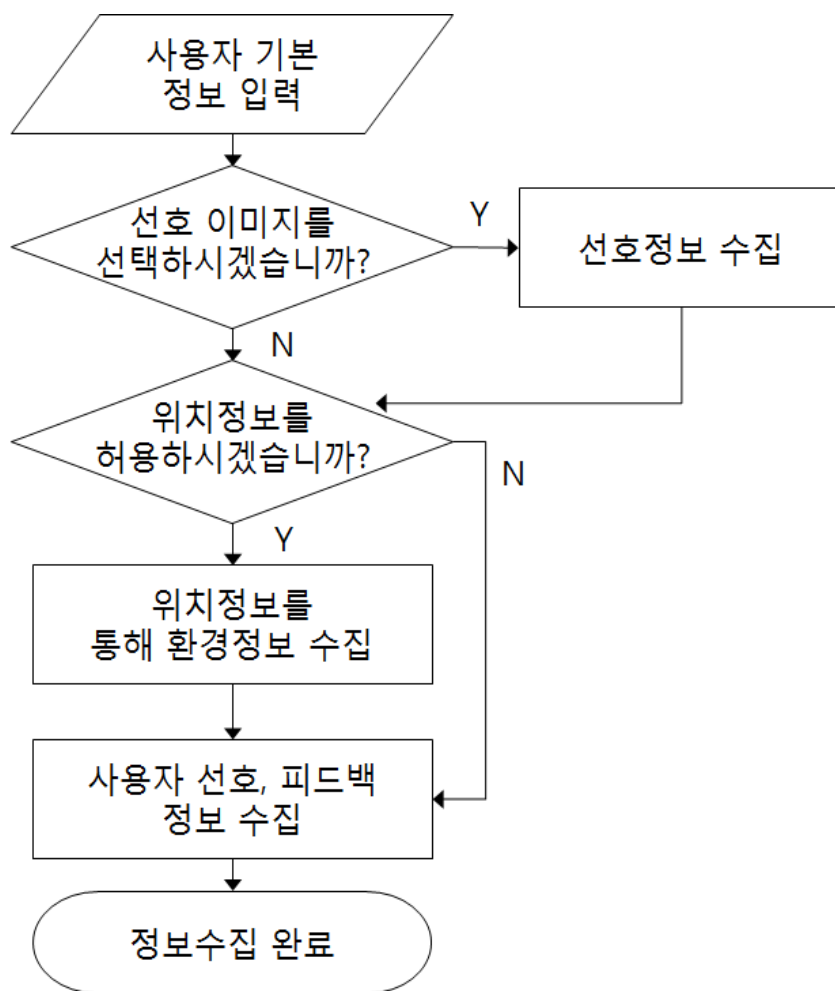
<표 2. 일반적 데이터 요소 정의>

분류	데이터 요소
환경	위치, 기온, 습도, 시간
사용자	성별, 나이, 선호 정보
양방향	추천정보, 평가점수
기타	행동정보, 전문가정보 등

<표 4>와 같이 데이터 요소 정의를 하였다. 요소별 가중치를 통해 사용자 성향을 분석하고 추천 및 피드백 과정을 진행할 것이다.

## 2. 수집

정의한 데이터 요소를 수집하기 위해 다음의 절차를 진행한다. 첫째, 사용자 정보를 입력받고 추천 이미지 정보를 보여주고 선택하게 하여 기본 선호정보를 수집한다. 둘째, 사용자가 추천정보를 요청할 때 마다 위치정보를 이용하여 기온, 습도를 수집하고 계절을 수집한다. 셋째, 과거에 추천 받았고 사용자가 선택한 정보가 있으면 추천정보와 그 정보에 따른 가중치 정보를 수집한다.



<그림 4. 정보수집 절차>

수집되는 개인별 정보 데이터 요소는 환경, 사용자, 피드백 정보이며 데이터 목록은 <표 3>와 같으며 구분에 따라 데이터베이스에 저장된다.

<표 3. 수집 데이터 목록>

분류	요소	구분
환경정보	위치	위도, 경도
	온도(℃)	-10 ~ 40
	습도(%)	0 ~ 100
	계절	봄, 여름, 가을, 겨울
사용자정보	나이(세)	-
	선호정보	-
	피드백정보	-

위 <표 3>에 저장된 데이터 요소별로 개인 성향별 가중치가 주어지며 가중치 값에 따라 성향분석을 진행한다.

### 3. 가중치 설정

온도, 습도, 계절 등에 대한 개인의 민감한 정도를 분석하기 위해 내용기반 필터링 기법을 이용한다. 개인과 이웃 간의 유사도를 측정하여 추천을 제공하는 협업적 필터링은 개인의 성향분석에 적당하지 않으며, 내용기반 필터링 기법을 이용하여 데이터 요소별로 분석이 가능하다. 성향분석을 위해 추천 서비스 성격에 따라 각 데이터 요소별 가중치와 범위를 추천 서비스 내용에 따라 설정한다. 데이터 요소 중 기온을 예를 들면 <표 4>과 같이 가중치와 범위를 설정한다. 추천 서비스 성격에 따라 중요한 데이터 요소의 경우 범위와 가중치를 세분화 시킨다. <표 4>의 한치물회의 경우 흑돼지구이보다 온도에 영향을 많이 받으므로 가중치를 크게 설정하였다. 요소별 가중치는 해당 분야 전문가가 값을 정의하고 값의 범위는 7점 척도이다.

<표 4. 기온의 가중치와 범위설정 예시>

음식 추천품목	기온범위(°C)		가중치 (0 ~ 7)
	최소	최대	
한치물회	-10	0	2
	0	10	2
	10	20	5
	20	30	7
	30	40	7
흑돼지구이	-10	0	5
	0	10	5
	10	20	5
	20	30	3
	30	40	3

내용기반 필터링 기법을 사용하기 위하여 각 추천 품목별로 환경정보 요소에 맞추어 가중치를 설정해 주어야한다. 또한 일반적인 사용자 정보 요소에 맞는 가중치도 설정해 주어야한다. 설정 후 성향분석과 피드백을 통해 사용자 성향에 맞게 가중치를 변화시켜갈 것이다.



#### 4. 성향분석

성향분석 계산식(1)은  $i$ 는 정보요소 구분,  $w_i$ 는 요소별 가중치,  $x_i$ 는 요소별 사용자 민감도를 의미한다.

$$\sum_{i=0}^n w_i x_i \quad (1)$$

추천목록 개별의 가중치를 식(1)로 계산할 수 있으며 가중치 값을 통해 사용자의 성향을 분석하고자 한다. 예시로 성향분석 계산을 위해 <표 5>와 같이 환경설정을 가정하고 <표 6>과 같이 가중치를 설정한다.

<표 5. 환경설정 가정>

분류	요소	정보
환경정보	위치	제주 제주시 아라동
	온도(℃)	34
	습도(%)	70
	계절	여름
사용자정보	나이(세)	33
	선호정보	-
	피드백정보	습도에 민감함

<표 5>와 같이 가정하고 추천 목록별로 성향분석 계산식(1)으로 가중치를 계산하며 계산식(1)을 나열하면 계산식(2)와 같다.  $w_t$ 는 온도 가중치,  $w_h$ 는 습도 가중치,  $w_s$ 는 계절가중치,  $w_a$ 는 나이 가중치를 의미하며,  $x_t$ 는 온도 민감도,  $x_h$ 는 습도 민감도,  $x_s$ 는 계절 민감도,  $x_a$ 는 나이 민감도를 의미한다.

$$w_t x_t + w_h x_h + w_s x_s + w_a x_a \quad (2)$$

<표 6>은 <표 5>의 환경설정이라고 가정 하에 가중치를 설정한 것이다.

<표 6. 가중치 설정 : 사용자 성향 미적용>

음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
한치물회	7	1	7	6	21
흑돼지구이	3	3	3	6	15
고등어조림	3	3	3	5	14
갈치구이	3	3	3	6	15
삼계탕	7	3	7	6	23
방어회	3	1	2	6	12

환경설정에 맞추어 가중치를 설정하고 음식 추천서비스 제공 항목은 삼계탕, 한치물회, 흑돼지구이, 갈치구이 등의 순이다. 하지만 사용자 성향이 습도에 민감한 사용자로 가정하였을 경우 환경정보, 사용자 정보의 요소별 사용자 성향 가중치는 <표 7>과 같으며 적용결과는 <표 8>과 같다.

<표 7. 습도에 민감한 사용자 민감도 예시>

	온도 민감도	습도 민감도	계절 민감도	나이 민감도
사용자 민감도	1	7	1	1

<표 8. 가중치 설정 : 사용자 성향 적용>

음식 추천목록	온도	습도	계절	나이	합계
사용자 민감도	1	7	1	1	-
한치물회	7	7	7	6	27
흑돼지구이	3	21	3	6	33
고등어조림	3	21	3	5	32
갈치구이	3	21	3	6	33
삼계탕	7	21	7	6	41
방어회	3	7	2	6	18

<표 8>은 사용자 성향 가중치를 적용한 결과이며 음식 추천서비스 제공 항목은 삼계탕, 흑돼지구이, 갈치구이, 고등어조림, 한치물회 등의 순이다. 습도에 민감한 음식인 한치물회가 사용자 성향에 따라 추천 순위가 아래로 변경되었다.

## 5. 추천 및 피드백

환경·사용자 정보를 수집하고 성향분석 계산식(1)을 이용하여 추천목록별 합계 점수를 계산한다. 합계 점수가 높은 목록들을 사용자에게 추천하고 사용자가 선택한 항목을 피드백을 통해 사용자 민감도 정보를 변경한다. 예를 들어 <표 9>와 <표 10>과 같은 상황을 가정하였을 경우 사용자에게 추천되는 항목은 <표 11>과 같다.

<표 9. 사용자 민감도 정보 예시>

	온도	습도	계절	나이	-
사용자 민감도	1	1	1	1	-

<표 10. 추천목록별 가중치 예시>

음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
한치물회	7	1	7	6	21
흑돼지구이	3	3	3	6	15
고등어조림	3	3	3	5	14
갈치구이	3	7	3	6	19
삼계탕	7	7	7	6	27
방어회	3	1	2	6	12

<표 11. 사용자 추천목록 예시>

	1순위	2순위	3순위	4순위	5순위
추천목록	삼계탕	한치물회	갈치구이	흑돼지구이	고등어조림
합계점수	27	21	19	15	14

<표 11>의 추천 목록 중에서 사용자가 호감도가 높은 항목을 선택하면 민감도 계산식(3)을 통해 사용자 민감도 정보를 수정한다.

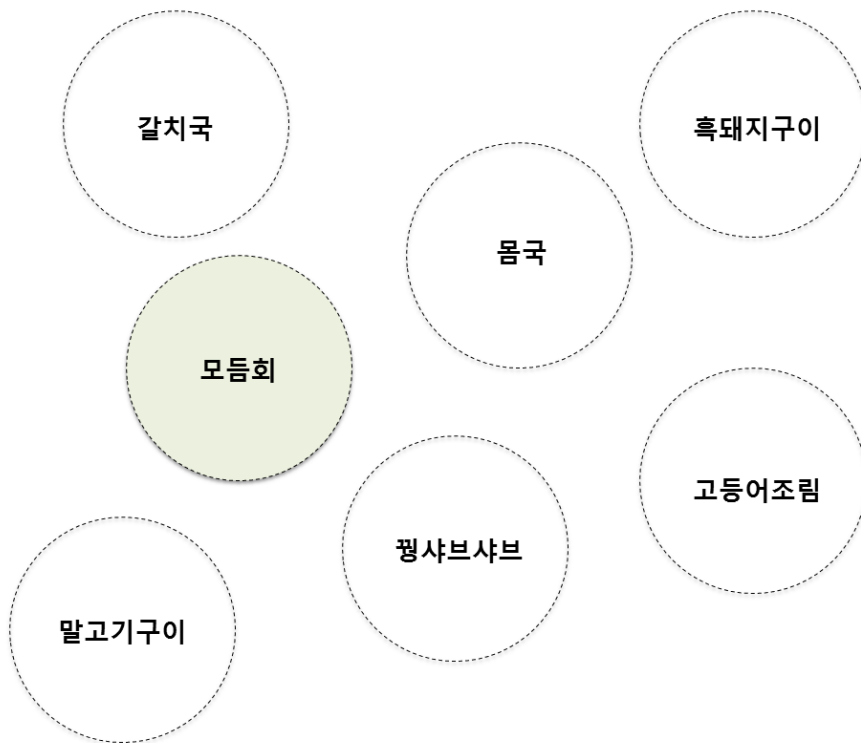
$$(R_i)_{new} = (R_i)_{old} - ((w_i)_b - (w_i)_s) \times R_c \quad (3)$$

민감도 계산식(3)의  $i$ 는 정보요소 구분,  $R_i$ 는 민감도,  $w_i$ 는 요소별 가중치,  $R_c$ 는 민감도 계수를 의미한다. 요소별로  $(R_i)_{new}$ 는 수정 민감도,  $(R_i)_{old}$ 는 기존 민감도,  $(w_i)_b$ 는 최적 추천항목의 가중치,  $(w_i)_s$ 는 사용자 선택 추천항목의 가중치를 의미한다. 민감도 계수는 사용자가 선택하는 항목에 따라 민감도가 변경되는 강도의 크기를 나타내는 계수이다. <표 12>는 <표 11>에서 사용자가 선택하는 항목에 따라 사용자 민감도가 변경된 예시이며 민감도 계수는 0.1로 설정하였다.

<표 12. 사용자 민감도 변경 예시>

	선택항목	온도	습도	계절	나이
기 존 사 용 자 민 감 도	-	1	1	1	1
변 경 된 사 용 자 민 감 도	한치물회	1	0.8	1	1
	흑돼지구이	0.6	1	0.6	1
	고등어조림	0.6	1	0.6	0.9
	갈치구이	0.6	1	0.6	1
	삼계탕	1	1	1	1
	방어회	0.6	0.8	0.5	1

**온도는 35도이고 습도 80%로 불쾌한 날씨**

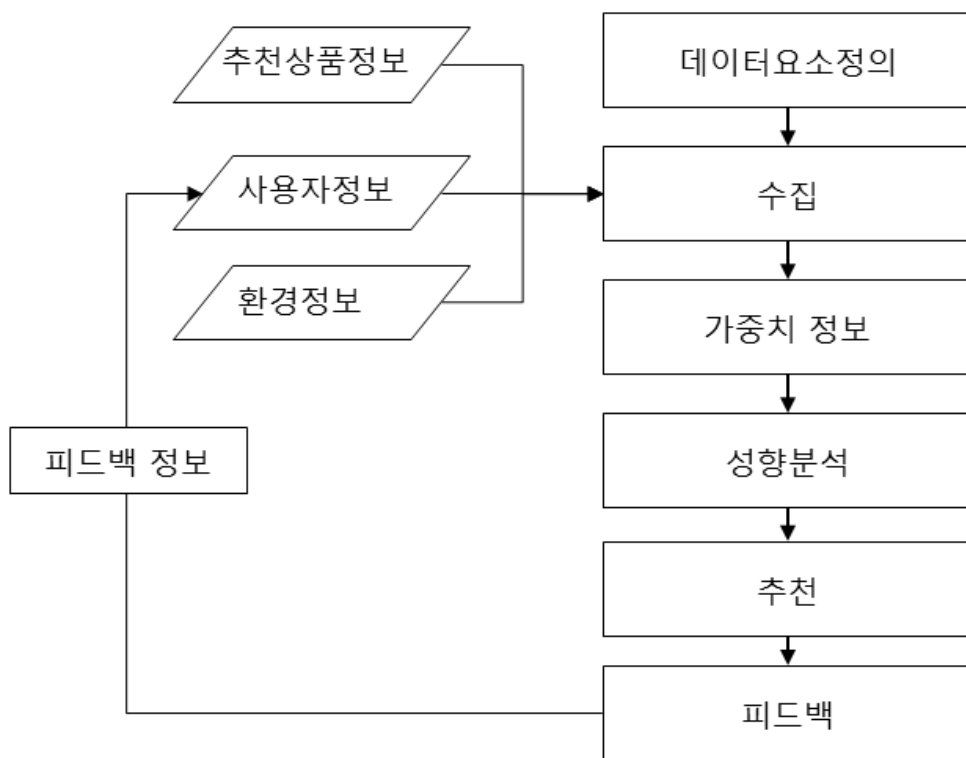


<그림 5. 추천 목록 예>

사용자 민감도를 피드백을 통해 수정함으로써 온도는 35도이며 습도는 80%인 불쾌한 날씨인 경우에도 <그림 5>와 같이 사용자는 모듬회를 추천받을 수 있다. 피드백 정보를 통해 사용자 성향분석을 하고 사용자가 어떠한 요소 또는 추천서비스 제품, 콘텐츠 등에 민감하고 둔감한지 알 수 있다.

## 6. 사용자 성향분석 엔진 모형

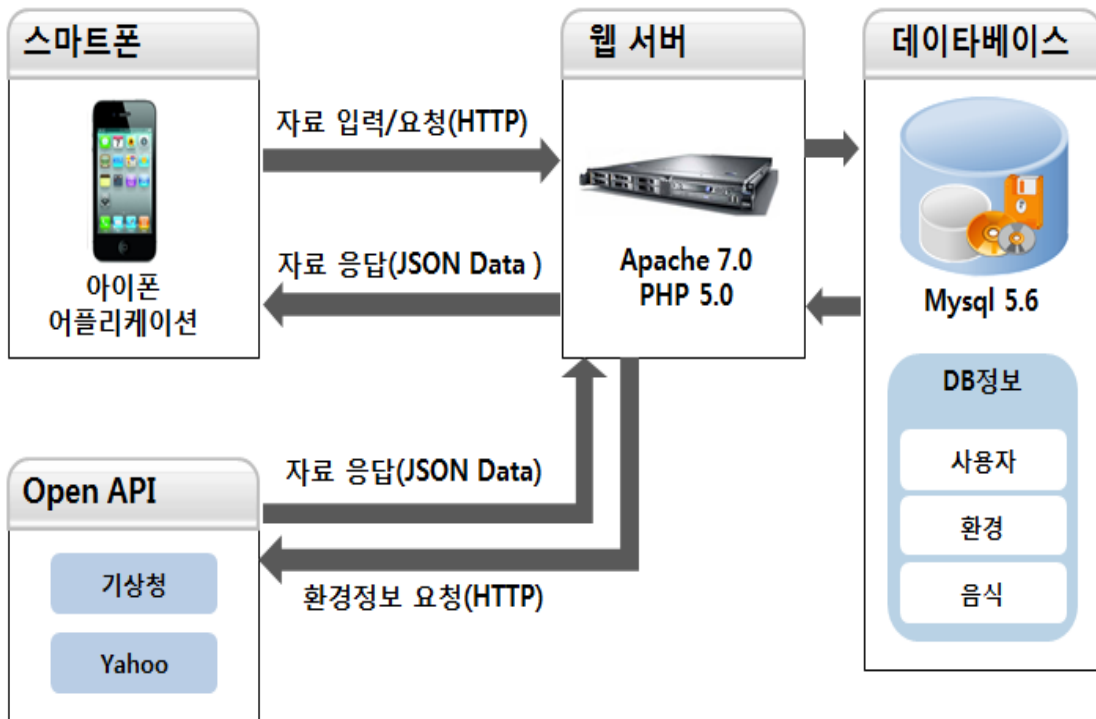
사용자 성향분석 엔진 모형은 추천 서비스의 성격에 맞게 데이터 요소를 정의하고 데이터 요소별로 가중치를 설정하고 성향분석, 추천, 피드백 과정을 진행한다. 추천 결과목록을 표출하여 주고 그 중 하나의 항목을 사용자가 선택한다. 선택한 항목 정보를 통해 피드백을 진행하고 사용자 성향 정보를 수정한다.



<그림 6. 사용자 분석엔진 모형>

## IV. 실험 환경 및 결과

### 1. 실험환경



<그림 7. 실험 환경 시스템 구조도>

실험환경은 사용자가 스마트폰을 이용하여 음식 추천서비스를 제공 받는다. 사용자가 사용자 기본정보를 입력하고 상황정보(Context) 중 하나인 위치정보와 함께 웹 서버로 음식 추천자료를 요청한다. 웹 서버는 사용자 위치정보를 통해 환경정보인 기온, 습도 등의 정보를 수신받는다. 이러한 정보를 데이터베이스에 저장하고 사용자 성향분석을 통해 사용자 스마트폰에 음식 추천정보를 제공한다. 제공받은 음식 추천정보를 선택하여 피드백 정보를 웹 서버에 전달한다.



### 1) 하드웨어 환경목록

<표 13. 하드웨어 환경목록>

하드웨어 목록	사양정보
아이폰	운영체제 : iOS 6.1.4 모델 : 아이폰 5 32GB/MD297KHA
웹 서버	CPU : Intel Core i7 M620 2.67GHz 운영체제 : 윈도우 7 32bits 메모리 : 8GB
데이터베이스 서버	CPU : Intel Core i7 M620 2.67GHz 운영체제 : 윈도우 7 32bits 메모리 : 8GB

### 2) 소프트웨어 환경목록

<표 14. 소프트웨어 환경목록>

소프트웨어 목록	
iOS 어플리케이션(Objective-C)	Mysql 5.6
Apache 7.0	PHP 5.0
기상청 OpenAPI	Yahoo OpenAPI

### 3) 추천 음식목록

실험을 위해 추천 음식서비스를 제공하는 음식목록을 6개로 설정하였다.

<표 15. 추천 음식목록>

한치물회	흑돼지구이
고등어조림	갈치구이
삼계탕	방어회

#### 4) 데이터 요소 정의

추천 음식서비스를 제공하기 위한 데이터 요소를 <표 16> 같이 정의한다.

<표 16. 수집 데이터 목록>

분류	요소	구분
환경정보	위치	위도, 경도
	온도(℃)	-10 ~ 40
	습도(%)	0 ~ 100
	계절	봄, 여름, 가을, 겨울
사용자정보	나이(세)	-
	선호정보	-
	피드백정보	-

#### 5) 환경 및 사용자 설정

실험을 위해 환경 및 사용자 설정을 <표 17>과 같이 A와 B로 설정한다.

<표 17. 환경 및 사용자 설정>

설정분류	요소	정보
A	위치	제주 제주시 아라동
	기온(℃)	35
	습도(%)	80
	계절	여름
	나이	45
B	위치	제주 서귀포시 천지동
	기온(℃)	3
	습도(%)	10
	계절	겨울
	나이	22

## 6) 가중치 설정

성향분석은 내용기반 필터링 기법을 이용하며 추천 서비스 성격에 따라 각 데이터 요소별 가중치와 범위를 추천 서비스 내용에 따라 설정한다. 요소별 가중치는 해당 분야 전문가가 값을 정의하고 값의 범위는 7점 척도이다. 실험을 위해 <표 18 ~ 21>과 같이 요소별 음식목록에 따른 가중치 값을 설정한다.

<표 18. 온도에 따른 음식별 가중치>

온도 (°C)	-10~0	0~10	10~20	20~30	30~40
음식목록					
한치물회	2	2	5	7	7
흑돼지구이	5	5	5	3	3
고등어조림	6	6	4	3	3
갈치구이	6	6	3	3	3
삼계탕	4	4	5	7	7
방어회	6	6	4	3	3

<표 19. 습도에 따른 음식별 가중치>

습도 (%)	0~30	30~60	60~100
음식목록			
한치물회	3	6	2
흑돼지구이	3	3	3
고등어조림	5	5	3
갈치구이	3	3	3
삼계탕	3	5	3
방어회	5	5	1

<표 20. 계절에 따른 음식별 가중치>

계절	봄	여름	가을	겨울
음식목록				
한치물회	3	7	4	2
흑돼지구이	5	3	4	6
고등어조림	5	3	4	5
갈치구이	6	3	3	7
삼계탕	3	7	3	3
방어회	7	2	2	7

<표 21. 나이에 따른 음식별 가중치>

나이 (세) 음식목록	0~10	10~20	20~30	30~40	40~50	50~60	60~
한치물회	3	4	5	6	4	6	5
흑돼지구이	6	6	6	6	6	4	4
고등어조림	3	3	4	5	5	5	5
갈치구이	5	5	6	6	6	6	6
삼계탕	6	6	6	6	6	5	5
방어회	4	4	4	6	6	4	4

<표 22>와 같이 사용자 민감도 초기값은 요소별로 1로 설정하였으며 민감도 계수는 0.3, 0.2로 설정하였다.

<표 22. 사용자 민감도 초기값 설정>

설정분류	온도 민감도	습도 민감도	계절 민감도	나이 민감도	민감도 계수
A	1	1	1	1	0.3
B	1	1	1	1	0.2

## 2. 실험결과

성향분석 계산식(4)를 이용하여 사용자의 추천 항목을 계산하였으며 식(4)의  $i$ 는 정보요소 구분,  $w_i$ 는 요소별 가중치,  $x_i$ 는 요소별 사용자 민감도를 의미한다.

$$\sum_{i=0}^n w_i x_i \quad (4)$$

민감도 계산식(5)을 통해 사용자가 선택한 항목에 따라 사용자 민감도 정보를 변경하였다. 사용자 민감도 정보가 변경됨에 따라 사용자 추천목록 순위도 변경이 되었다.

$$(R_i)_{new} = (R_i)_{old} - ((w_i)_b - (w_i)_s) \times R_c \quad (5)$$

민감도 계산식(5)의  $i$ 는 정보요소 구분,  $R_i$ 는 민감도,  $w_i$ 는 요소별 가중치,  $R_c$ 는 민감도 계수를 의미한다. 요소별로  $(R_i)_{new}$ 는 수정 민감도,  $(R_i)_{old}$ 는 기존 민감도,  $(w_i)_b$ 는 최적 추천항목의 가중치,  $(w_i)_s$ 는 사용자 선택 추천항목의 가중치를 의미한다. 민감도 계수는 사용자가 선택하는 항목에 따라 민감도가 변경되는 강도의 크기를 나타내는 계수이다.

### 1) 환경 및 사용자 설정: A결과

환경 · 사용자정보를 통해 나온 가중치 추천결과는 <표 23>과 같으며 삼계탕, 한치물회, 갈치구이, 흑돼지구이, 고등어조림, 방어회 순이다.

<표 23. 추천서비스 결과 : 피드백 미적용>

음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
한치물회	7	2	7	4	20
<b>흑돼지구이</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>6</b>	<b>15</b>
고등어조림	3	3	3	5	14
갈치구이	3	3	3	6	15
<b>삼계탕</b>	<b>7</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>23</b>
방어회	3	1	2	6	12

사용자가 흑돼지구이를 선택하고 피드백을 진행하였다. 피드백 과정을 진행할 때마다 사용자 민감도가 <표 24>와 같이 변경되었으며 추천서비스결과도 <표25>와 같이 변경되었다.

<표 24. 사용자 민감도 변경 목록>

피드백 횟수	온도 민감도	습도 민감도	계절 민감도	나이 민감도	민감도 계수
초기	1	1	1	1	0.3
1차	-0.2	1	-0.2	1	
2차	-1.4	1	-1.4	1	
3차	-1.4	1	-1.4	1	

<표 25. 추천서비스 결과 : 피드백 진행 적용>

피드백 횟수	음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
1차	한치물회	7	2	7	4	3.2
	<b>흑돼지구이</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>6</b>	<b>7.8</b>
	고등어조림	3	3	3	5	6.8
	갈치구이	3	3	3	6	7.8
	<b>삼계탕</b>	<b>7</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>6.2</b>
	방어회	3	1	2	6	6
2차	한치물회	7	2	7	4	-13.6
	<b>흑돼지구이</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>6</b>	<b>0.6</b>
	고등어조림	3	3	3	5	-0.4
	갈치구이	3	3	3	6	0.6
	<b>삼계탕</b>	<b>7</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>-10.6</b>
	방어회	3	1	2	6	0
3차	한치물회	7	2	7	4	-13.6
	<b>흑돼지구이</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>6</b>	<b>0.6</b>
	고등어조림	3	3	3	5	-0.4
	갈치구이	3	3	3	6	0.6
	<b>삼계탕</b>	<b>7</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>-10.6</b>
	방어회	3	1	2	6	0

최적의 추천항목이 삼계탕에서 흑돼지구이, 갈치구이 등의 순으로 변경되었다. 민감도 계수를 0.3으로 설정하여 민감도 변화 강도를 크게 설정하였으므로 1회의 피드백 결과만으로 최적의 추천항목이 변경되었다.

2) 환경 및 사용자 설정: B결과

환경 · 사용자정보를 통해 나온 가중치 추천결과는 <표 26>과 같으며 방어회, 갈치구이, 흑돼지구이, 고등어조림, 삼계탕, 한치물회 순이다.

<표 26. 추천서비스 결과 : 피드백 미적용>

음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
<b>한치물회</b>	<b>2</b>	<b>3</b>	<b>2</b>	<b>5</b>	<b>12</b>
흑돼지구이	5	3	6	6	20
고등어조림	6	5	5	4	20
갈치구이	6	3	7	6	22
삼계탕	4	3	3	6	16
<b>방어회</b>	<b>6</b>	<b>5</b>	<b>7</b>	<b>4</b>	<b>22</b>

사용자가 한치물회를 선택하고 피드백을 진행한다고 설정하였다. 피드백 과정을 진행할 때 마다 사용자 민감도가 <표 27>와 같이 변경되었다. 추천서비스결과도 3번의 피드백 과정을 진행한 후 <표28>과 같이 변경되었다.

<표 27. 사용자 민감도 변경 목록>

피드백 횟수	온도 민감도	습도 민감도	계절 민감도	나이 민감도	민감도 계수
초기	1	1	1	1	0.2
1차	0.2	1	0	0.8	
2차	0	1.4	-0.4	0.4	
<b>3차</b>	<b>-0.8</b>	<b>1</b>	<b>-1</b>	<b>0.6</b>	
<b>4차</b>	<b>-0.8</b>	<b>1</b>	<b>-1</b>	<b>0.6</b>	

<표 28. 추천서비스 결과 : 피드백 진행 적용>

피드백 횟수	음식 추천목록	온도 가중치	습도 가중치	계절 가중치	나이 가중치	합계
3차	<b>한치물회</b>	<b>2</b>	<b>3</b>	<b>2</b>	<b>5</b>	<b>2.4</b>
	흑돼지구이	5	3	6	6	-3.4
	고등어조림	6	5	5	4	-2.4
	갈치구이	6	3	7	6	-5.2
	삼계탕	4	3	3	6	0.4
	<b>방어회</b>	<b>6</b>	<b>5</b>	<b>7</b>	<b>4</b>	<b>-4.4</b>

최적의 추천항목이 방어회에서 한치물회, 삼계탕 등의 순으로 변경되었다. 민감도 계수를 실험 A보다 작은 0.2로 설정하였다. 실험 A보다 민감도 변화가 작으므로 3회의 피드백 결과 후에 추천항목이 변경되었다.

## V. 결론

본 논문에서는 사용자의 환경, 양방향 정보 등을 통해 개인형 추천 서비스를 제공하기 위한 내용기반 필터링 기반의 사용자 성향 분석 엔진 모델을 제안했다. 사용자 성향 분석을 위해 분석할 사용자 데이터 요소를 정의하고 수집, 가중치, 성향분석, 추천, 피드백 과정을 통해 연관관계를 구성하고 개인형 추천 서비스를 개선된 형태로 제공했다. 피드백 정보를 통하여 엔진의 가중치를 변경할 수 있도록 하여 추천 서비스의 정확도를 높일 수 있었다. 개인을 그룹으로 성향을 구분 짓지 않고 사용자 개인이 습도, 기온, 계절 등에 어떠한 부분에 민감한지 성향을 분석할 수 있다.

내용기반 필터링 기법의 단점으로 상품별로 정의해야하는 번거로운 점이 있었으며 본 연구에서 제안하는 사용자 성향분석 엔진도 내용기반 필터링 기법 기반의 추천서비스를 제공하기 때문에 이러한 단점이 존재한다. 하지만 이러한 단점을 전문가, 행동 정보와 사용자, 환경정보에 따른 가중치 값의 입력을 통해 장점이 된다. 예를 들면 사용자 성향분석 엔진 모형은 환경정보와 상품별 전문가 정보가 수집되기 때문에 계절별, 제철별 음식 추천서비스, 전문가 음식 추천서비스 등을 제공할 수 있다. 내용기반 필터링 기법의 다른 단점인 사용자 정보와 상품간의 비교를 통해 유사 상품이 지속적으로 추천되는 단점이 있었다. 하지만 상황정보와 내용기반 필터링, 피드백 등을 통해 단점을 보완했다.

추후 사용자 성향분석 엔진을 통해 사용자에게 적절하게 추천되고 있는지에 대한 연구와 공통 데이터 요소, 엔진을 통해 다수의 추천 서비스 적용에 대한 연구가 필요하다.



## 참고문헌

- [1] 위키피디아, <http://www.wikipedia.org/>
- [2] 손재권기자의 점선잇기, [http://jackay21c.blogspot.kr/2013/03/blog-post\\_4169.html](http://jackay21c.blogspot.kr/2013/03/blog-post_4169.html)
- [3] 김용수, “개인화 서비스를 위한 추천 시스템의 연구동향”, ie매거진(2012년 봄호) 제19권 제1호(통권 55호), pp. 37 ~42
- [4] 류종민,홍창표,강경보,강동현,양두영,좌정우,“모바일 상황인식 추천맛집 서비스 개발”, 한국콘텐츠 학회논문지 '07 Vol. 7 No.5
- [5] 이수진,전태룡,백경동,김성신,“협업적 필터링 및 퍼지시스템 기반 사용자 성향 분석에 의한 영화평가 예측 시스템”, 한국지능시스템 학회 논문지 2009. Vol. 19. No. 2, pp. 243~247
- [6] 서효석,권일경,이상용, “모바일 환경에서 사용자 행동 추론에 기반을 둔 개인화 추천 서비스”, Proceedings of KIIS Fall Conference 2011 Vol. 21, No. 2, pp. 27~28
- [7] 이동주,이상근,이상구, “시간 상황 정보를 고려한 협업 필터링을 이용한 음악 추천”, 2009 한국컴퓨터종합학술대회 논문집 Vol.36, No.1(C), pp.123~128
- [8] 김용수,염봉진,Noramn kim, “차원 감소 기법을 이용한 전자 상거래 추천 시스템”. Journal of the Korean Institute of Industrial Engineers Vol. 36, No. 3, pp. 193~202, September 2010.
- [9] 김용수, “사용자의 이용패턴을 고려한 하이브리드 추천시스템 설계”, 2010년 대한산업공학회/한국경영과학회 춘계공동학술대회, pp. 269~273.
- [10] 김기수, “개인화된 추천을 위한 하이브리드 협업필터링 알고리즘 및 인터넷 전자상거래에서의 활용”, 인터넷전자상거래연구 제8권 제4호 (2008.12.31).
- [11] 원재용,구경이,김유성, “음악 추천 시스템에서 대표 선율을 이용한 내용 기반 필터링 기법”, 2004년도 한국정보과학회 가을 학술발표논문집 Vol.31, No.2, pp.229~231.
- [12] 정주석,강신재, “자동 생성된 사용자 프로파일을 이용한 하이브리드 음식 추천 시스템”, 한국지능시스템학회 논문지 2011, Vol.21, No.5, pp. 609~617.
- [13] 김상화,오병화,김문중,양지훈, “협력적 필터링과 콘텐츠 정보를 결합한 영화 추천알고리즘”, 정보과학회논문지 : 소프트웨어 및 응용 제39권 제4호(2012.4), pp. 261~268.