



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

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

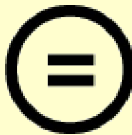
다음과 같은 조건을 따라야 합니다:



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.

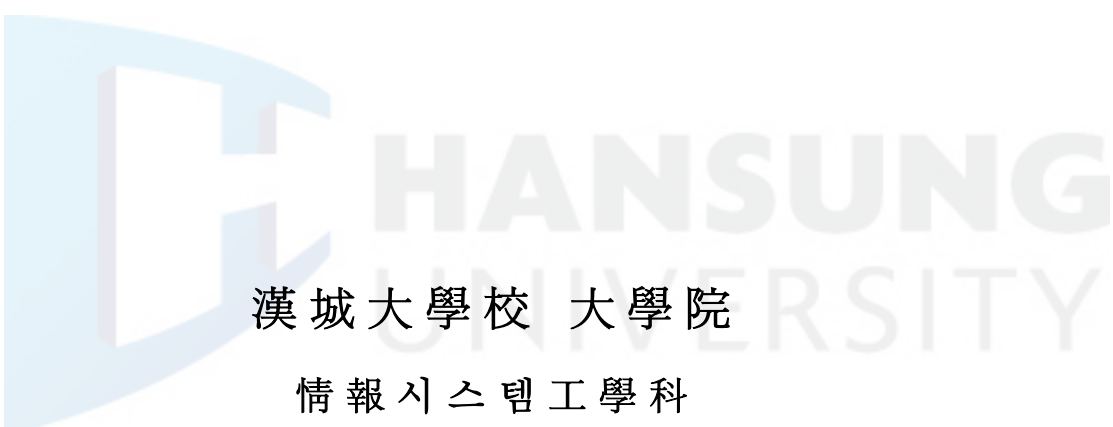
이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

碩士學位論文

모바일 환경에서 상황정보를 이용한
하이브리드 필터링 추천시스템 설계

2011年

The logo of Hansung University is a large, light blue stylized 'H' that incorporates a white square in the center. To the right of the logo, the words 'HANSUNG UNIVERSITY' are written in a large, light blue, sans-serif font.

漢城大學校 大學院

情報시스템工學科

情報시스템工學專攻

高 貞 敏

碩士學位論文
指導教授 南斗熙

모바일 환경에서 상황정보를 이용한
하이브리드 필터링 추천시스템 설계

Development of Hybrid Filtering Recommendation System
using Context-Information in Mobile Environments

2010年 12月 日

漢城大學校 大學院

情報시스템工學科

情報시스템工學專攻

高 貞 敏

碩士學位論文
指導教授 南斗熙

모바일 환경에서 상황정보를 이용한
하이브리드 필터링 추천시스템 설계

Development of Hybrid Filtering Recommendation System
using Context-Information in Mobile Environments

위 論文을 工學 碩士學位 論文으로 提出함

2010年 12月 日

漢城大學校 大學院

情報시스템工學科

情報시스템工學專攻

高 貞 敏

高貞敏의 工學 碩士學位 論文을 認准함

2010年 12月 日

審査委員長 _____ 印

審査委員 _____ 印

審査委員 _____ 印

목 차

제 1 장 연구의 개요	1
제 1 절 연구의 배경 및 필요성	1
제 2 절 연구의 목적	4
제 3 절 연구의 구성 및 흐름	6
제 2 장 이론적 배경	8
제 1 절 집단지성(Collective intelligence)	8
1. 개요	8
2. 기본 개념	9
3. 집단지성의 예	13
제 2 절 상황인식(Context awareness)	16
1. 개요	16
2. 기본 개념	17
3. 모바일 환경에서 상황정보 활용	20
제 3 절 정보추천 관련 기법	22
1. 개요	22
2. 정보 필터링 기법 분류	23
3. 협업 필터링(Collaborative filtering)	30
제 4 절 시사점 및 본 연구의 위치	35

제 3 장 상황정보를 이용한 하이브리드 필터링 37

제 1 절 상황정보를 이용한 하이브리드 필터링 알고리즘	37
1. 상황추출	38
2. 사용자 기반 협업 필터링	39
3. 아이템 기반 협업 필터링	47
제 2 절 추천 서비스 시나리오 예	51
1. 상황추출 단계	51
2. 사용자 기반 협업 필터링 단계	52
3. 아이템 기반 협업 필터링 단계	54

제 4 장 추천시스템 설계 및 구현 60

제 1 절 추천시스템 설계	60
1. 설계 개요	60
2. 요구사항 분석	61
3. 서비스 도출	62
4. 추천 시스템 논리적 설계	63
제 2 절 소프트웨어 아키텍처	69
1. 소프트웨어 아키텍처 개요	69
2. 추천시스템의 모듈 뷰	71
3. 추천시스템의 컴포넌트와 커넥터 뷰	72
4. 추천시스템의 할당 뷰	74
제 3 절 상황정보를 이용한 하이브리드 필터링 추천 시스템	76
1. 개발 환경	76
2. 추천 시스템 구성 및 실행과정	78

제 5 장	상황정보를 이용한 하이브리드 추천시스템 평가	87
제 1 절	실험 배경 및 절차	87
1. 실험	배경	87
2. 실험	절차	88
제 2 절	실험 데이터 및 실험군 구성	89
1. 실험	데이터 구성	89
2. 실험군	구성	90
제 3 절	실험 및 실험 결과	91
1. 실험		91
2. 실험	결과	95
제 4 절	하이브리드 필터링 검증	99
1. 정확도와	재현율	99
2. 검증	결과	100
제 6 장	결론	102
참고문헌		104
ABSTRACT		107

【 표 목 차 】

[표 2-1] 필터링 기법 분류	24
[표 2-2] 협업 필터링의 한계점	32
[표 3-1] 유사성 척도의 단점	41
[표 3-2] 선호 Matrix Table	44
[표 3-3] 각 고객의 유사도 산출 결과	45
[표 3-4] 사용자 A와 이웃집단의 선호 Matrix Table	46
[표 3-5] 선호 Matrix Table	49
[표 3-6] 각 아이템 간의 유사도 산출 결과	50
[표 3-7] 추천대상자 A와 이웃집단의 선호 Matrix Table	51
[표 3-8] 선호 Matrix Table	53
[표 3-9] 사용자 A에 대한 각 사용자의 유사도	53
[표 3-10] 사용자 A의 이웃사용자 집단 선호 Matrix Table	54
[표 3-11] 선호 Matrix Table	55
[표 3-12] 아이템 간의 상관관계분석결과	56
[표 3-13] 쇼핑물2의 이웃 아이템 집단 선호 Matrix Table	56
[표 3-14] 쇼핑물5의 이웃 아이템 집단 선호 Matrix Table	57
[표 3-15] 쇼핑물6의 이웃 아이템 집단 선호 Matrix Table	57
[표 3-16] 쇼핑물2의 이웃 아이템 집단 선호 Matrix Table	58
[표 3-17] 쇼핑물5의 이웃 아이템 집단 선호 Matrix Table	58
[표 3-18] 쇼핑물6의 이웃 아이템 집단 선호 Matrix Table	59
[표 4-1] 요구사항명세 테이블	61
[표 4-2] 추천시스템 제공 서비스	63
[표 4-3] 상황정보를 이용한 하이브리드 추천시스템 액터	64
[표 4-4] 추천시스템 유스케이스	65
[표 4-5] 하이브리드 필터링 추천 유스케이스 명세	68
[표 4-6] 시뮬레이션 테스트 환경	83
[표 4-7] 10명의 가상사용자에 대한 5개 음식점의 선호(평점)매트릭스	84

[표 5-1] 실험 데이터 구성(지면 한계 상 10여개의 데이터만을 표시하였음)	89
[표 5-2] 실험군 구성	90
[표 5-3] 1단계 실험 결과	95
[표 5-4] 2단계 실험 결과	95
[표 5-5] 3단계 실험 결과	96
[표 5-6] 4단계 실험 결과	97
[표 5-7] 상황정보 활용 유용성 평가 결과	98
[표 5-8] 검색 시스템에서 정확도와 재현율의 정의	99
[표 5-9] 추천 시스템에서 정확도와 재현율의 정의	100
[표 5-10] 상황정보를 이용한 하이브리드 추천 시스템 정확도 및 재현율	100



【 그림 목 차 】

<그림 1-1> 연구의 배경 및 필요성	2
<그림 1-2> 연구의 목표	5
<그림 1-3> 연구의 흐름	7
<그림 2-1> 집단지성의 개념도	8
<그림 2-2> 집단지성의 예	15
<그림 2-3> 상황인식 컴퓨팅 개념도	16
<그림 2-4> 내용기반 필터링 예-1	25
<그림 2-5> 내용기반 필터링 예-2	26
<그림 2-6> 협업 필터링 예(사용자01과 유사한 사용자_02 판별)	26
<그림 2-7> 협업 필터링 예(사용자01에는 존재하지 않는 ITEM_04에 대한 선호 예측) ...	27
<그림 2-8> 인구통계학적 필터링 예(사용자01, 02에게 레드 색상 스카프 추천) ..	28
<그림 2-9> 지식기반 필터링 예(명시한 상황/서비스/컨텐츠에서 적절한 ITEM_03 추천) ..	29
<그림 3-1> 상황정보를 이용한 하이브리드 필터링 추천 시스템	37
<그림 3-2> 상황추출 절차	38
<그림 3-3> 사용자 기반 협업 필터링 절차	39
<그림 4-1> 상황정보를 이용한 하이브리드 추천시스템 유스케이스 다이어그램 ..	66
<그림 4-2> 추천시스템 서버 모듈류 아키텍처	71
<그림 4-3> 추천시스템 서버 컴포넌트 & 커넥터 류 아키텍처	73
<그림 4-4> 추천시스템 서버 할당류 아키텍처	75
<그림 4-5> 추천시스템 목적행위 및 목적지/현위치 선택화면	78
<그림 4-6> 추천시스템 목적행위 및 목적지/현위치 선택화면	79
<그림 4-7> 추천시스템 아이템 표출 화면(왼쪽: 리스팅, 오른쪽: 지도)	80
<그림 4-8> 추천시스템 정보센터(서버)의 xml 데이터 아이템 표출 화면	81
<그림 4-9> 아이템 세부 정보 및 평점 입력 화면	83
<그림 4-10> 가상 사용자 C의 전체 아이템 리스트 그림 29	84
<그림 4-11> 가상 사용자 C의 임의 추천 리스트	85
<그림 4-12> 가상 사용자 C의 베스트 아이템 추천 리스트	85

<그림 4-13> 가상 사용자 C의 사용자 협업 필터링 추천 리스트	85
<그림 4-14> 가상 사용자 C의 하이브리드 필터링 추천 리스트	86
<그림 4-15> 가상 사용자 C의 내 과거 베스트 추천 리스트	86
<그림 5-1> 아이템 세부 정보 및 평점 입력 화면	88
<그림 5-2> 임의 아이템 추천화면	91
<그림 5-3> 베스트 아이템 추천화면	92
<그림 5-4> 사용자 협업 필터링 추천화면	93
<그림 5-5> 하이브리드 필터링 추천화면	94



제 1 장 연구의 개요

제 1 절 연구의 배경 및 필요성

오늘날 정보통신 기술의 급속한 성장 및 발전에 따라 유비쿼터스 네트워크 컴퓨팅에 대한 관심이 증폭되고 있다. 이와 더불어 인터넷, 전자 상거래, 통신 등 일반인의 생활 및 활동과 관련된 각기 다양한 분야에서 이용자의 성향이나 관심에 적합한 서비스 혹은 정보를 추천하는 이용자 맞춤 서비스에 대한 관심 역시 커지고 있다. 이러한 이용자 맞춤 서비스는 이용자의 기호나 선호에 따른 적합한 서비스 및 정보를 예측하고 추천함으로써 합리적인 의사결정에 도움을 줄 수 있다. 뿐만 아니라 정보의 범람이라 할 수 있을 정도로 과다한 정보들 사이에서 이용자의 관심에 적합하지 않은 정보의 검색을 사전에 차단함으로써 이용자의 정보 탐색시간 및 비용을 줄여줄 수 있다. 즉 예측을 통한 추천서비스의 제공은 인간의 정보인지 측면에서의 정보수용 및 활용능력을 함께 증진 시킬 수 있다.

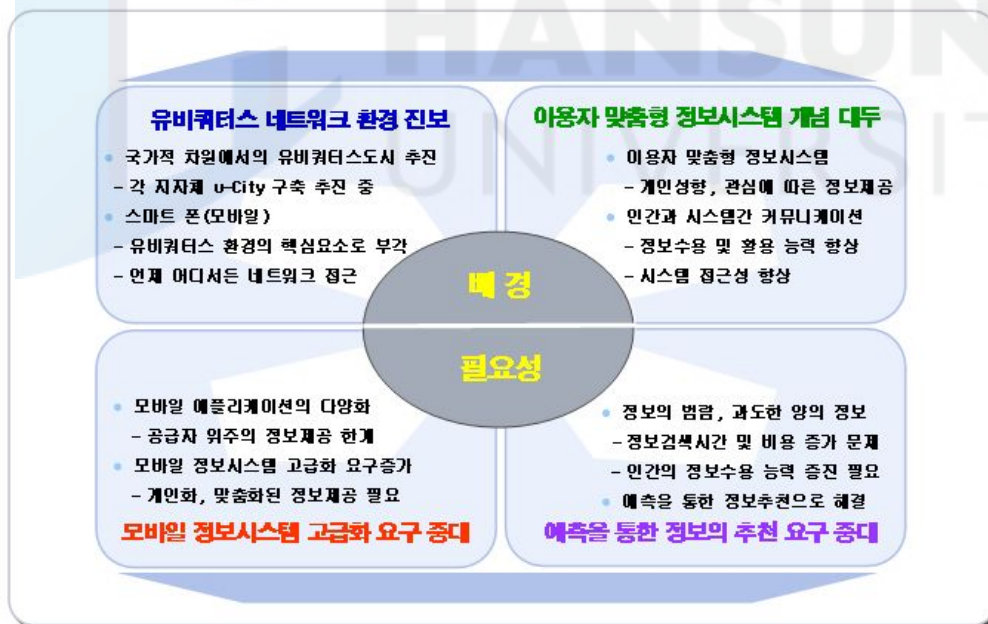
한편 최근 스마트폰(Smartphone)을 매개체로 모바일 관련 기술들이 급속도로 발전하며 큰 각광을 받고 있다. 스마트폰은 PC와 같은 기능과 더불어 고급 기능을 제공하는 휴대 전화이다. 스마트폰은 응용 프로그램 개발자를 위한 표준화된 인터페이스와 플랫폼을 제공하는 완전한 운영 체제 소프트웨어를 실행하는 전화로 볼 수 있으며 뿐만 아니라 전자 우편, 인터넷, 전자책 읽기 기능, 내장형 키보드나 외장 USB 키보드, VGA 단자를 갖춘 고급 기능이 있는 전화로도 볼 수 있다. 이러한 모바일기술의 급속한 발전은 언제, 어디서든 원하는 정보 및 서비스를 제공받을 수 있는 유비쿼터스 네트워크 환경이 도래했음을 의미한다.

이러한 환경 및 인프라의 발전에 따라 최근 모바일 환경에서 각 종 정보 및 서비스를 제공하는 다양한 모바일기반 응용소프트웨어들이 출시되고 있는 추세이다. 그러나 그 대부분이 공급자 위주의 정보시스템으로 단순히 다량의 정보들을 불특정 다수의 이용자들에게 제공하는데 목적을 두고 있어 앞서 설명한 이용자 개개인에 대해 맞춤화 혹은 개인화된 정보 및 서비스의 제공은 거

의 이루어지지 않고 있다. 이러한 한계를 탈피하여 이용자와의 상호작용을 통해 개인화 혹은 맞춤형 정보 및 서비스를 제공할 필요가 있다. 다시 말하면 급격히 발전하고 있는 유비쿼터스 컴퓨팅 및 정보통신기술과 함께 발맞춰 보다 진보된 형태의 사용자 중심 정보제공시스템을 구성하여 제공할 필요성이 대두되고 있는 것이다.

상기 배경 및 필요성에 따라 본 연구에서는 모바일환경에서 개인화 및 맞춤화를 위한 추천시스템을 설계 및 구현하고자 한다. 추천 시스템의 핵심 요소라 할 수 있는 정보의 필터링은 최근 웹 및 전자상거래 등의 추천시스템에서 가장 효과적인 기법으로 알려져 있는 협업필터링(Collaborative Filtering)을 이용한다. 이 때 보다 정확한 추천결과를 얻기 위해 단순 협업필터링 적용만이 아닌 각 종 정보필터링 기법의 장점만을 결합한 하이브리드 필터링(Hybrid Filtering)을 이용하여 추천 시스템을 구성한다. 또한 추천의 질을 향상시키기 위해 정보의 필터링이 실행되기 전 모바일 디바이스에서 활용 가능한 사용자의 목적행위, 위치 등의 상황정보(Context-information)를 이용하여 이용자 상황에 따른 효과적인 정보의 추천을 가능하도록 한다.

이를 정리하면, 모바일 환경에서 상황정보를 이용한 하이브리드 필터링 추



<그림 1-1> 연구의 배경 및 필요성

천시스템을 설계하고 구현하는 것이 본 연구의 목적이라 할 수 있다.

최근 유비쿼터스 네트워크 환경의 핵심요소로 모바일 환경이 큰 주목을 받고 있다. 본 연구를 통해 모바일 환경에서 이용자 상황정보 및 정보필터링 기법을 이용한 정보추천시스템을 구현함으로써 차후 모바일 단말기를 포함한 각종 커뮤니케이션 기기를 기반으로 이용자에게 효율적이고 편리한 정보획득 및 서비스 제공을 가능하게 할 수 있을 것이다. 또한 본 연구의 결과는 궁극적으로 유비쿼터스 네트워크 컴퓨팅 기술과 맞물려 개인화 및 맞춤형시스템에 관한 연구의 발전방향을 제시할 수 있을 것으로 기대된다.



제 2 절 연구의 목적

협업 필터링이란 많은 사용자들로부터 얻은 기호정보(taste information)에 따라 사용자들의 관심사들을 자동적으로 예측하게 해주는 방법이다. 이는 최근 웹 2.0 등의 대두를 통해 많은 각광을 받고 있는 집단지성을 구현하는 하나의 방법론이라 할 수 있다.

협업 필터링 접근법의 근본적인 가정은 사용자들의 과거의 경향이 미래에서도 그대로 유지될 것이라는 전제에 있다. 예를 들어, 음악에 관한 협업 필터링 혹은 추천시스템은 사용자들의 기호(좋음, 싫음)에 대한 부분적인 목록을 이용하여 그 사용자의 음악에 대한 기호를 예측하게 된다. 이는 특정 사용자의 정보에만 국한 된 것이 아니라 많은 사용자들로부터 수집한 정보를 사용한다는 것이 특징으로 단순히 투표를 한 수를 기반으로 각 아이템의 관심사에 대한 평균적인 평가로 처리하는 방법과 차별화 된 것이다. 즉 협업필터링은 고객들의 선호도와 관심 표현을 바탕으로 비슷한 패턴을 가진 고객들을 식별해 내는 기법인 것이다. 비슷한 취향을 가진 고객들에게 서로 아직 구매하지 않은 아이템들을 교차 추천하거나 분류된 고객의 취향이나 생활 형태에 따라 관련 아이템을 추천하는 형태의 서비스를 제공하기 위해 사용된다. 이는 다수의 개체들이 서로 협력 혹은 경쟁을 통하여 얻게 되는 지적 능력에 의한 결과로 얻어진 집단적 능력인 집단지성의 활성화 형태라 할 수 있다.

협업 필터링은 크게 사용자 기반 협업 필터링과 아이템 기반 협업 필터링으로 구분되는데 각 필터링은 고유의 장단점을 가지고 있으며 최근에는 다양한 필터링 기법들의 유용한 부분만을 결합한 하이브리드 형태의 필터링이 주목받고 있다.

한편 상황정보(Context)란 일반적으로 실세계에 존재하는 개체의 상태를 특징화하여 정보를 의미한다. 이러한 상황정보를 활용하여 사용자가 처한 상황을 인지하고 사용자가 원하는 정보를 원하는 형태로 처해진 상황에 맞게 획득하고 이용할 수 있는 상태를 상황인식(Context-Awareness)이라 하며 최근 유비쿼터스 네트워크 컴퓨팅 환경을 구현하기 위한 방법론으로 큰 관심을 모으고 있다.

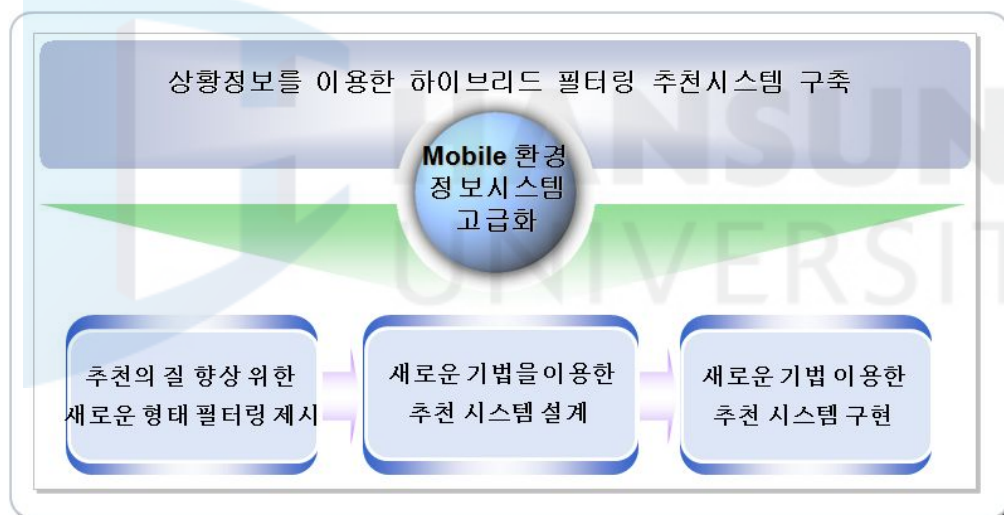
본 연구의 목적은 모바일 환경에서 상황정보를 이용한 하이브리드 필터링 추천시스템을 설계하고 구현하는 것이다. 이 추천시스템은 정보 추천의 질을 높이기 위해 모바일 환경에서 획득할 수 있는 사용자의 목적, 위치 등의 상황 정보를 활용하여 추천정보의 카테고리 및 범위를 결정한 후 사용자기반 협업 필터링, 아이템 기반 협업 필터링을 결합한 하이브리드 필터링을 이용하여 개인이 선호할 것이라 판단되는 정보를 예측하고 추천한다.

상기 최종목적을 실현하기 위해 본 연구에서 다루는 세부적인 목표를 다음과 같이 세 가지로 구분할 수 있다.

첫째, 추천 시스템의 예측능력 및 추천의 질을 향상시키기 위한 목적으로 새로운 형태의 정보 필터링 기법을 제시한다.

둘째, 새롭게 제시한 기법인 상황정보를 이용한 하이브리드 필터링 기반의 추천시스템을 설계한다.

셋째, 새롭게 제시한 기법인 상황정보를 이용한 하이브리드 필터링 기반의 추천시스템을 구현한다.



<그림 1-2> 연구의 목표

제 3 절 연구의 구성 및 흐름

앞서 제시한 연구의 목적을 달성하기 위해 먼저 본 연구에서 다뤄지는 이론적 배경들에 대한 문헌검토, 기술수준 분석 등을 통해 연구개발방향을 정립한다. 다음으로 효과적인 정보의 예측 및 추천을 위한 새로운 방법론을 제시한다. 본 연구에서 제시하는 방법은 상황정보를 이용한 하이브리드 필터링을 기초로 이용자들의 개인화 및 맞춤화를 실현하는 것으로 이를 핵심으로 한 추천시스템의 설계 및 구현을 실행한다. 그 다음으로 실험 대상자를 선정하여 개발된 추천시스템에 대한 테스트를 실시함으로써 효용성 및 정확성을 평가한다.

본 논문은 총 6개의 장으로 구성되어 있으며 각 장에 대한 설명은 다음과 같다.

제 1 장 연구의 개요에서는 연구의 배경과 목적 그리고 연구방법 및 구성에 대해서 설명한다.

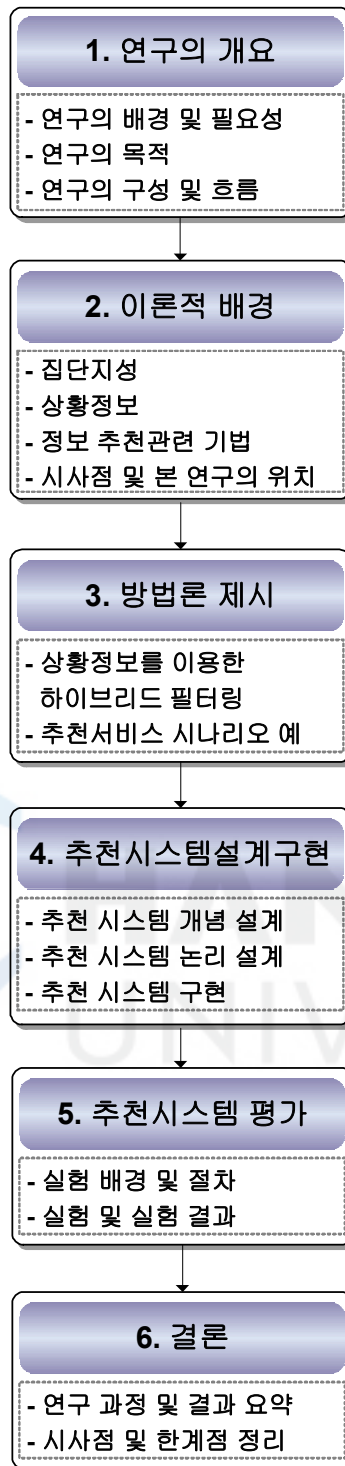
제 2 장에서는 집단지성, 상황인식 등 본 연구에서 고려되는 이론적 배경에 대해 설명하도록 한다. 또한 본 연구의 핵심요소라 할 수 있는 각 정보 필터링 기법에 대한 이론 및 선행 연구들에 대해 정리한다.

제 3 장에서는 본 연구에서 설계하고자 하는 상황정보를 활용한 하이브리드 필터링 기반의 추천시스템에 대한 구성 및 알고리즘을 제시한다.

제 4 장에서는 제 3 장에서 언급한 알고리즘과 기법을 이용하여 추천시스템을 설계 및 구현하고 추천 시스템의 구성 및 추천과정에 대해 설명한다.

제 5 장에서는 제 4 장에서 구현된 추천시스템 탑재 모바일 기기를 실제 이용자들에게 사용하게 한 뒤, 설문을 토대로 그 유용성을 평가한다. 또한 추천시스템의 평가 척도로 활용되는 평가기법을 이용하여 본 연구에서 설계된 추천시스템의 성능을 정량적으로 평가하도록 한다.

제 6 장 결론에서는 본 연구 과정 및 결과에 대해 요약하고 연구의 시사점 및 한계점에 대한 정리를 통해 향후 연구방향을 제시한다.



<그림 1-3> 연구의 흐름

제 2 장 이론적 배경

제 1 절 집단지성(Collective intelligence)

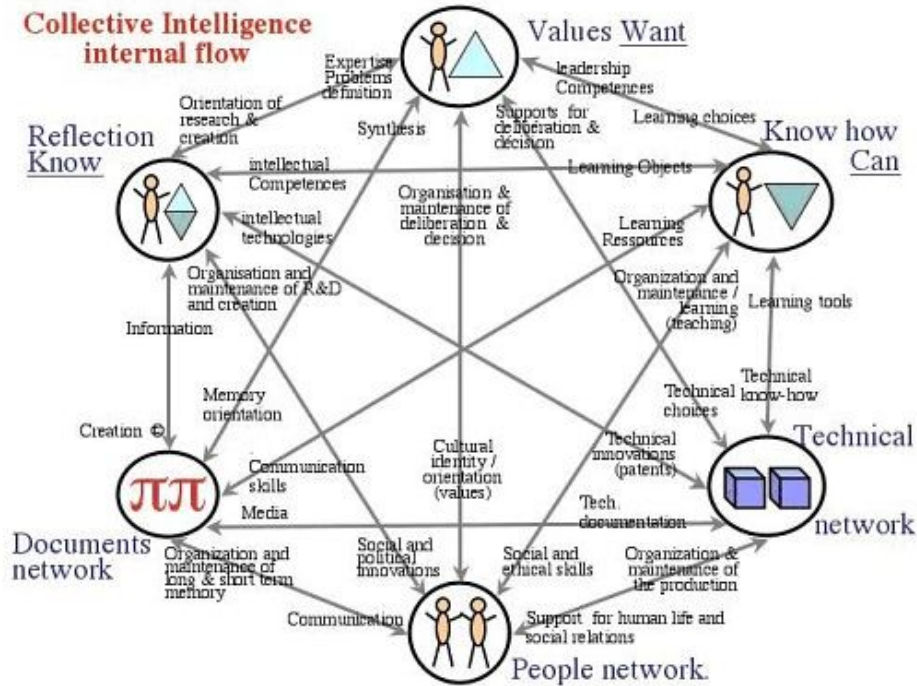
1. 개요

집단지성(集團知性, collective intelligence)이란 다수의 개체들이 서로 협력 혹은 경쟁을 통하여 얻게 되는 지적 능력에 의한 결과로 얻어진 집단적 능력을 말하며 중지(衆智, 대중의 지혜), 집단지능, 협업지성, 공생적 지능이라고도 한다.[Wikipedia, 2010] 집단지성은 박테리아, 동물, 인간 그리고 컴퓨터 네트워크 등 분야의 의사결정과정에서 다양한 형태로 나타나며 집단의 총의를 모으는 과정을 통한 결정 능력은 한 개체의 능력범위를 넘어선 힘을 발휘할 수도 있다고 주장한다. 집단지성은 사회학, 경영학, 컴퓨터 공학 등의 다양한 분야에서 연구되고 있다.

집단지성에 대한 개념은 Douglas Hofstadter(1979), Peter Russell(1983), Tom Atlee(1993), Pierre Lévy(1994), Howard Bloom(1995), Francis Heylighen(1995), Douglas Engelbart, Cliff Joslyn, Ron Dembo, Gottfried Mayer-Kress(2003) 등의 연구에서 등장하였으며, Norman L. Johnson에 의해 공생지능(Symbiotic intelligence)으로 소개되기도 하였다.

Tom Atlee 등의 일부 학자들은 집단지성의 주된 초점을 인간에게 맞추고 다양한 연구를 진행하였다. Atlee는 집단지성은 의사결정 과정에서 발생하는 집단사고(GroupThink)와 개개의 인지적 편향(individual cognitive bias)을 극복하여 향상된 의사결정을 달성할 수 있다고 생각했다.

집단지성은 최근의 정보통신 기술놀로지의 급속한 발전에 기인한 네트워크의 활성화, 즉 인터넷의 출현으로 대변될 수 있다. 인터넷의 발전에 기인한 웹 2.0은 데이터의 소유자나 독점자 없이 누구나 손쉽게 데이터를 생산하고 인터넷에서 공유할 수 있도록 한 사용자 참여 중심의 인터넷 환경이다. 인터넷상에서 정보를 모아 보여주지만 하는 웹 1.0에 비해 웹 2.0은 사용자의 직접적인 데이터 접근 및 조작을 위해 데이터를 제공하는 플랫폼이 정보를 더 쉽게 공



<그림 2-1> 집단지성의 개념도

자료 : jacksonmedeiros, 'Collective intelligence internal flow',
<http://www.flickr.com/photos/jacksonmedeiros/2301066413/>

유하고 서비스 받을 수 있도록 구축되어 있다. 이러한 사용자와 네트워크의 상호작용을 바탕으로 사용자로 하여 자신만의 콘텐츠를 직접 생성할 수 있도록 할 뿐만 아니라 개인이 아닌 다수에 의해 현존하는 정보 및 지식들의 집대성과 발전을 가능하게 하고 있다. 웹 2.0의 특성은 사용자의 참여, 정보의 개방, 공유에 있으며, 이를 반영해내는 가장 큰 요소가 바로 집단지성이라고 할 수 있다.

2. 기본 개념

Tom Atlee와 George Pór은 그룹이론 및 인공지능이 인간의 생활에 이점을 제공하는 분야라면 집단지성은 인간의 사고방식, 공유의지 그리고 지성 및 지능의 개방성의 활성화 형태로써 인간경영의 최우선으로 인식되어야 한다고 생각했다.[Atlee, T, 2008; George Por] Tom Atlee과 George Pór은 집단지성

을 통해 나타나는 의사결정의 강력한 힘에 대해 확신을 가졌으며 이는 모든 개별 개체의 각각의 능력보다 뛰어날 수 있을 것으로 보았다. Tom Atlee와 George Pór은 집단 지성의 극대화 그 집단의 모든 구성원으로부터 모든 잠재적인 유용한 제안, 즉 최적제안을 받아드리고 이를 발전시키고자 하는 조직의 능력 및 의지에 달려있다고 보았다. 한편 이들은 응집력이 높은 소규모 의사결정 집단에서 대안의 분석 및 이의 제기를 억제하고 합의를 쉽게 이루려고 하는 심리적 경향을 집단사고(GroupThink)라고 여겼는데 이러한 집단사고는 의사결정 과정에 나타나는 ‘집단착각 현상’으로 의사결정을 위한 자원의 투입을 제한하거나 최종 결정을 위해 개선 및 발전이 필요한 잠재적인 최적제안을 제거함으로써 집단지성에 지장을 초래할 수 있다고 보았다. 다시 말하면 집단사고는 집단구성원들이 대안에 대한 충분한 분석 및 토론이 없이 쉽게 합의하고 그 대안이 최선이라고 믿고 합리화하려고 하는 현상을 말한다. 집단사고에 빠지게 되면 조직구성원들은 새로운 정보나 변화에 민감하게 반응하지 못해 상황적응 능력이 떨어지게 된다. 집단사고에 의한 결정은 주어진 상황에서 임의의 또는 잘못된 정보에 의해 이루어질 가능성이 존재하며 집단사고에 의한 잘못된 의사결정 및 정보는 널리 퍼질 수 있기 때문에 Tom Atlee와 George Pór은 이를 경계하고자 했다.

다른 연구가들은 Tom Atlee와 George Pór의 견해와 다른 측면에서 집단지성을 접근했다. Pierre Lévy는 매스 커뮤니케이션의 관점에서 네트워크 정보통신 테크놀로지의 능력에 초점을 맞추고 사회의 지식의 바다를 한층 향상시키기 위해 집단지성을 연구했다. 그는 집단지성을 통해 인간으로 하여금 쉽고 빠른 상호작용, 공유 그리고 협동을 가능하도록 한 통신수단을 제안했다. Lévy의 생각과 같이 인터넷의 발전은 위키피디아(Wikipedia)와 같은 전에 없던 집단의 지식을 기반으로 하는 포럼의 탄생을 이끌었다.

Francis Heylighen, Valerie Turchin, 그리고 Gottfried Mayer-Kress는 컴퓨터 공학 및 인공두뇌학(Cybernetics)의 측면에서 집단지성을 바라봤다. Howard Bloom은 1억 년 전 시대의 박테리아 등 세균의 집단지성으로부터 현재까지 이어지는 집단지성의 진화에 관한 연구를 통해 다양한 종들이 탄생 이후 그들의 지능이 어떻게 작동하는지를 연구했으며 집단지성으로 지구상에 존

채하는 모든 구성요소들의 생물학적 적응 즉 학습적인 행태임을 강조했다.[Howard Bloom, 1995]

Peter Russell, Elisabet Sahtouris, Barbara Marx Hubbard 등은 세상의 정보 대뇌피질이라 할 수 있는 인간의 지적 활동(빠르게 변화하는 집단적 지능)의 가능성에 대해 연구했다. 정보의 대뇌피질과 인터넷은 유사점을 가지고 있다. 1995년 인터넷 사회라 정의된 세계적 정보 시스템은 통신 및 관련된 인프라를 바탕으로 정보의 제공, 접근. 공개 혹은 비공개, 높은 수준의 서비스 이용을 가능하게 하였다.[Leiner et al. 2003].

인터넷은 점진적으로 대뇌피질화 되어왔다고 할 수 있다. 인터넷이 출현한 1950년대 후반부터 WWW(World Wide Web)이 릴리즈 된 1991년, 그리고 2005년에 이르러 전 세계적으로 1,018,057,389의 인구가 인터넷을 사용하고 있다.[CIA, 2008]. 이러한 인터넷의 대뇌피질화는 웹에 접근하는 전 세계의 다양한 사용자들로 하여 지성간의 회합 또는 지식의 협업 활동을 가능하게 하고 있다.

집단지성은 의사결정을 위한 다수의 사용자들로 부터의 지성간의 회합, 지식의 협업 활동으로 볼 수 있으며 이러한 개념을 위해 다음의 4가지 원칙이 충족되어야 한다. 이는 개방성, 동등계층 생산, 공유성, 전역 활동성이며 세부적인 설명은 아래와 같다.[Tapscott, D, & Williams, A. D., 2008]

가. 개방성(Openness)

통신기술 초기, 아이디어, 지적재산, 정보 등의 자원은 기업 또는 인간 간의 경쟁우위를 확보할 수 있게 해줌에 따라 자원 보유자들의 자원에 대한 폐쇄적인 행태는 당연시되었다. 그러나 점진적인 아이디어의 개방을 통해 제품을 급진적인 개선을 얻는 사례가 발생하기 시작 했으며 뿐만 아니라 경쟁사 간의 개방으로 서로의 감시 및 경쟁을 통해 얻어지는 이점이 정보 및 아이디어의 폐쇄보다 더 큰 수익구조를 형성함에 따라 시간이 지나면서 이러한 자원 및 정보의 폐쇄 규제는 점차 완화되었다.

나. 동등계층 생산(Peering)

동등한 계층 생산은 가치를 추구하기 위한 수평조직의 형태라고 할 수 있으며 공동의 결과물을 생산하기 위해 자발적으로 모인 개인들이 자기 조직화된 평등한 커뮤니티에 전적으로 의존하여 제품과 서비스를 생산하는 방식이다. 한 예로 리눅스 프로그램을 들 수 있다. 리눅스는 프로그램 사용자에게 무료로 배포되며 개방적인 형태로서 프로그램의 수정이 가능할 뿐 아니라 이를 다른 사람에게 제공할 수 있다. 이것은 집단지성의 개념으로 동등성은 개인이 참여에 대한 동기를 가지고 서로 다른 결과를 달성하여 제품 및 서비스의 개선을 가능하게 한다.

다. 공유성(Sharing)

최근의 연구에 따르면 점차적으로 더 많은 기업들이 자신의 잠재력, 핵심 특허권 등 일부에 대한 공유를 시작하고 있다. 기업들이 자신의 지적 재산권의 공유에 대한 제한을 하는 것은 발전 및 도약을 위한 가능한 모든 기회들을 사전에 가로막는 것이나 마찬가지이기 때문이다. 공유는 기업의 측면에서 시장의 발전, 제품들의 빠른 출시 등을 가능하게 할 수 있으며 인간의 측면에서 지식 및 지능의 분배를 통한 집대성을 가능하게 할 수 있다.

라. 전역 활동성(Acting Globally)

통신 기술의 발전은 글로벌 기업 혹은 e-Commerce 등의 출현을 가속화시키고 있으며 거대 기업 뿐 아니라 저렴한 간접비용의 개인의 비즈니스에도 급속도로 영향을 미치고 있다. 크고 작은 범세계적으로 활동하는 기업들은 지리적인 한계를 뛰어넘어 매우 광범위하게 연결되어 새로운 시장, 아이디어, 기술에 대한 접근을 가능하게 하고 있다. 따라서 웹을 통한 전역적인 기업의 경쟁력 유지와 실시간 접근성에 대한 중요성이 부각되고 있다.

3. 집단지성의 예

웹 2.0에서 집단지성은 전 세계의 웹을 이용하는 다수의 사용자들의 협력과 공헌이 응집되어 하나의 커다란 지식창고, 즉 새로운 정보의 원천이 생성되는 것을 말한다. 집단 지성은 웹 2.0을 설명하는데 있어서 핵심이 되는 개념 가운데 하나로서 웹 2.0이 표방하는 "공유와 참여 및 개방"의 철학을 잘 내포하고 있다. 단 여기에는 다수 사용자의 참여가 어느 일정 수준을 넘어설 때 가치가 있는 정보 및 원천과 그렇지 않은 것은 것을 여과해 낼 수 있고 뿐만 아니라 새로운 가치를 만들어 낼 수 있다는 전제조건이 기반이 되어있다. 사용자의 참여와 협력을 바탕으로 새로운 가치를 창조해내는 것에 집단 지성의 의미가 있다.

가. 위키피디아(wikipedia)

인터넷과 웹 기술을 기반으로 새로이 등장한 위키피디아는 집단지성의 실현을 가능하도록 한 좋은 예로 볼 수 있다. 위키피디아는 웹을 통해 시공간상에서 서로 독립적으로 존재하는 다수의 지성적 존재들이 함께 서로의 지식과 경험을 응집하여 합쳐 백과사전을 만들어가는 곳이다. 이는 단순히 지적 결과물을 만들어낼 뿐만 아니라 서로 간의 상호작용을 바탕으로 한 지식의 확대와 확인 그리고 유기적 연결을 가능하게 한다. 위키피디아는 명확한 위계질서에 하에 조직되어 있지 않으며 서로 독립적이고 다채로운 배경을 가지고 있는 사용자들에 의해 구성되고 편집된 내용으로 만들어진다.

위키피디아의 가장 큰 특징은 어느 정도 비평가 상호평가에 의해 운영되지만 사전 내용의 편집과 관리에 대해 제한이 없다는 것이다. 하지만 이와 동시에 비교적 엄격하게 내용을 검사하고 상호비판을 통해 정리해 나가기도 하는데 이는 집단지성적 결과물의 공동창출이라는 가치와 사용자 자신의 만족의 가치가 동시에 적절히 균형을 이루고 있다는 점에 기반 한다. 또한 사용자들이 특별한 이익과 이해를 위해 일하지 않는 점에서 그 원인을 찾을 수 있다. 결과적으로 단시간에 기존의 백과사전을 대체하고 있으며 속도와 내용 면에서 비약적인 발전을 이루고 있다.

영문 위키피디아는 2001년에 시작되어 1000만 여개의 항목에 대한 내용이 만들어져 있으며 계속적으로 성장하고 있다. 현재 전 세계 200여 개 언어로 만들어 지고 있으며 국내의 경우 2002년 10월부터 도입되었다.

나. 아마존(amazon.com)

아마존닷컴은 인터넷을 통해 물건을 파는 최초의 주요 회사들 가운데 하나였다. 아마존닷컴은 생산자의 데이터(아이템설명 등)와 사용자의 참여(리뷰)를 이용하여 온라인 서점으로서의 입지를 확실히 다졌다. 아마존은 책을 구매한 사용자가 책에 대한 리뷰 및 평가를 웹 사이트에 올릴 수 있도록 하고 이를 통해 축적된 리뷰들을 다른 사용자들에게 제공함으로써 책 정보, 커버 이미지, 목차 등 동일한 정보만을 제공하는 다른 온라인 서점과 차별화 전략을 시행했다. 아마존닷컴은 사용자의 집단적인 참여, 즉 집단의 구매경향 및 평가를 통해 얻을 수 있는 유용한 정보 데이터베이스를 구축하고 이를 통해 더 나은 검색 결과, 추천 정보 등을 사용자에게 제공함으로써 동종 업계에서 큰 성공을 거둔 예라고 볼 수 있다.

다. 오픈소스

오픈 소스 소프트웨어 개발 프로젝트 역시 집단 지성의 가치를 매우 잘 보여주는 사례 중 하나이다. 많은 수의 소프트웨어가 오픈 소스 프로젝트로 개발되고 있는 가운데 웹의 인프라가 되는 여러 가지 소프트웨어 즉, 리눅스, 아파치 웹 서버, MySQL 데이터베이스, 펄(Perl), PHP 역시 다수의 개발자의 참여로 만들어지고 있다. 수많은 사용자들은 소스를 활용하여 자유로운 치운 소프트웨어를 만들고 이는 소프트웨어끼리의 경쟁으로 이뤄져 더 좋은 소프트웨어의 생존, 혹은 각기 장단점을 상호 보완된 새로운 소프트웨어의 탄생이 가능하도록 한다.

SourceForge.net에는 10만 개 이상의 프로젝트가 등록되어 있고 100만 명 이상의 사용자가 등록되어 있다. 누구든지 프로젝트를 추가하고 프로젝트에 참여해 코드를 작성하여 추가하거나, 코드를 다운 받고 이용할 수 있다.

라. 아고라

아고라란, 고대 그리스의 도시국가(폴리스)의 중심 시에 있는 광장이다. 정치적인 광장과 시장을 겸한 독특한 것으로 그 주변에는 관청과 신전 등 공공건물이 많이 세워져 있었다. 어원은 아고라조(모이다)로서, 사람들의 모임이나 모이는 장소를 의미하였다. 고대 그리스의 도시국가(폴리스)의 시민은 집 밖의 공공생활을 즐겼으며 하루의 대부분을 아고라에 모여 정치와 사상 등을 토론하는 등 아고라는 일상생활의 중심이었다. 현대의 아고라는 이 어원을 그대로 살린, 사람들이 토론 등을 하기 위해 모이는 온라인 광장이라고 할 수 있다. 토론이라는 다소 엄격한 주제를 일상생활에서 쉽게 접근 할 수 있도록 인터넷 기술을 바탕으로 그 역할을 잘 살린 모습이라고 할 수 있다. 특정 주제에 대한 해아릴 수 없을 만큼의 수많은 의견들이 쏟아지고 논쟁되어지며 이를 발전시킴으로써 사용자들 간의 판단에 따라 여론이 형성 되어 가는 것은 집단지성의 부분적 요소라고 할 수 있다. 아고라는 웹상의 온라인 광장에서 집단의 참여와 소통의 문화를 형성을 통하여 효과적인 의사결정을 수행하는 집단지성의 대표적인 형태로 자리 잡고 있다.



<그림 2-2> 집단지성의 예

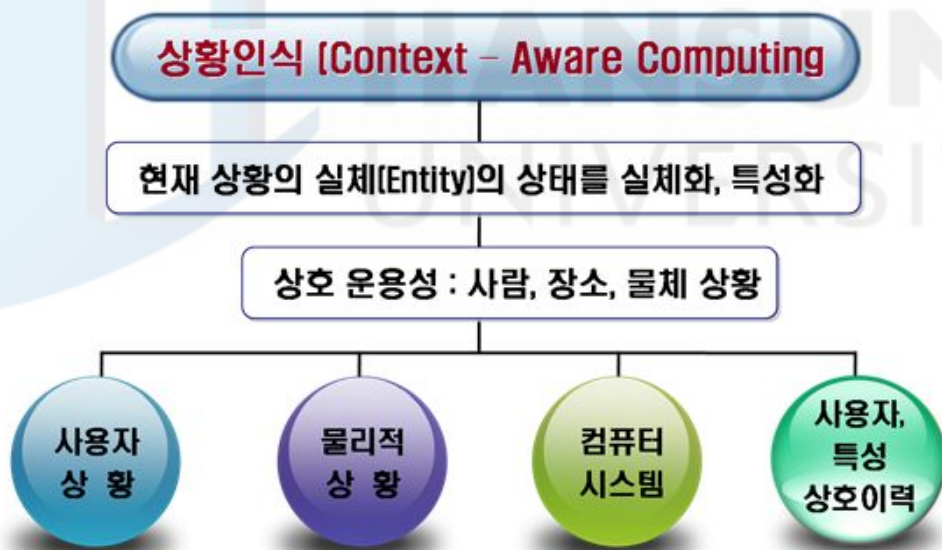
제 2 절 상황인식(Context awareness)

1. 개요

오늘날 정보통신 테크놀로지 기술이 발전함에 따라 인간과 컴퓨터간의 장벽 해소에 대한 관심이 날로 커지고 있다. 이를 위해 다양한 방법론들이 연구되고 있으며 그 중 하나로 상황인식 기법에 대한 연구들이 주목받고 있다.

상황인식(Context awareness)이란 일반적으로 실세계에 존재하는 개체의 상태를 특징화하여 정의한 정보, 즉 상황정보(Context)를 활용하여 사용자가 처한 상황을 인지하고 사용자가 원하는 정보를 원하는 형태로 처해진 상황에 맞게 획득하고 이용할 수 있는 상태를 의미한다.[이도경, 2010]

상황인식 컴퓨팅은 이러한 상황인식 기술을 활용하여 인간의 삶을 더욱 세분화시키고, 정교한 서비스를 구현하기 위해 제안된 컴퓨팅 환경이다. 상황인식 컴퓨팅은 현실공간과 가상공간을 연결하여 가상공간에서 현실의 상황을 정보화하고 이를 활용하여 사용자 중심의 지능화된 서비스를 제공하는 기술이다. 현실 세계의 모든 상황을 표현하는 기술적 수단을 제시하며, 이를 기반으



<그림 2-3> 상황인식 컴퓨팅 개념도

자료: 김택천, 김택준, “웹3.0 상황인식(Context-Aware computing)현재상황의 실체화, 특성화

로 상황 인식, 상황의 특징 추출, 학습, 추론 등의 지능화된 기법을 적용하여 인간 중심의 자율적인 서비스를 가능하게 한다.[정보통신용어사전, 2010]

언제 어디서나 사용자의 잠재적인 요구에 부합하여 서비스 및 콘텐츠를 제공할 수 있는 유비쿼터스 컴퓨팅 환경이 주목받고 있는 가운데 이를 구현하기 위한 하나의 방법론이라 할 수 있는 상황인식기술에 대한 관심이 커지고 있는 것이다. 상황인식기술을 바탕으로 현실세계에 대한 정보화를 통해 잠재적인 이용자의 요구사항을 파악할 수 있으며 이를 바탕으로 이용자 중심의 지능화 및 맞춤형 서비스 제공을 할 수 있다.

2. 기본 개념

가. 유비쿼터스 컴퓨팅

유비쿼터스 컴퓨팅은 라틴어에서 유래 되었으며 사전적 의미로 “언제 어디서든지 어떤 사물을 통해서도 컴퓨터를 활용할 수 있음을 이르는 말”로서 1988년, Xerox Palo Alto Research Center (PARC)의 수석 기술자 Mark Weiser에 의해 처음 사용되었다. Mark Weiser는 PARC의 또 다른 수석 과학자인 John Seely Brown과 함께 혹은 단독의 연구를 통해 최초로 유비쿼터스 컴퓨팅의 기본적인 철학 개념을 제안했으며 이는 다음과 같다.[Mark Weiser et al, 1991; 1996; 1999]

- Disappear computing

Disappear의 개념은 컴퓨터의 특성이 사라짐으로써 컴퓨터와 현실세계의 사물을 구분할 수 없는 것을 의미한다. 예를 들면, 머그컵이 기존의 것과는 달리 위치정보 알람 기능을 내장해 단말기로 머그컵의 위치 정보를 받거나, 온도에 따라 머그컵의 색상이 변할 수 있는 것이다.

- Invisible computing

Invisible의 개념은 다수의 컴퓨터를 현실세계의 물리적 환경에 배치함으로써 기존 컴퓨터의 능력을 향상시키고 사용자의 능률을 높이는 것을 의미한다. 컴퓨팅이 현실세계에서 보이지 않으려면 소형 컴퓨터 및 칩을 내장할

수 있는 기술이 바탕이 되어야 한다.

- Calm computing

Calm의 개념은 인간의 지각과 인지 능력에 대한 개념으로 어떠한 방식으로 바탕으로 인간과 컴퓨터의 정보 환경과 상호작용할 수 있을가에 대한 문제이다.

우리나라의 경우 2003년에 정보통신부 산하의 유비쿼터스 컴퓨팅 사업단, 유비쿼터스 컴퓨팅 연구센터, 유비쿼터스 컴퓨팅 연구소가 출범했으며 단장을 역임한 아주대학교 전자공학부 조위덕 교수는 다음과 같이 유비쿼터스 컴퓨팅의 특징을 기술하고 있다.

- 끊임 없는 연결 (Seamless Connectivity; HC Infra Network)

모든 사물들이 네트워크에 항상 연결되어 있어야 한다.

- 사용자 중심 인터페이스 (User Centered Interface)

사용자가 기기 사용에 있어서 어려움이 없으며 처음 접하는 사람을 포함해 누구나 쉽게 사용할 수 있는 인터페이스가 제공 되어야 한다.

- 컴퓨팅 기능이 탑재된 사물 (Smart Things)

가상공간이 아닌 현실 세계의 어디서나 컴퓨터의 사용이 가능해야 한다.

- 의미론적 상황인식 동작 (Semantic Context awareness)

사용자의 상황(장소, ID, 장치, 시간, 온도, 날씨 등)에 따라 서비스가 변해야 한다.

나. 상황인식 컴퓨팅

유비쿼터스 컴퓨팅의 구현을 위한 방법론 중 하나라고 할 수 있는 상황인식 (context-awareness) 기술은 1994년 Schilit에 의해 소개되었다. 컴퓨터 공학에서 상황인식은 컴퓨터를 기반으로 현재 상황에 관련된 인지와 그에 대한 반응을 가능하게 하는 것을 의미한다. 상황인식 시스템은 실세계에서 특정 기준이

나 운영을 위한 환경, 혹은 지능적인 기법을 통해서 현재 상황의 상태를 특징화하여 정의된 정보를 파악하고 그에 대해 적절한 반응을 실행하도록 한다. 이를 바탕으로 사용자의 현재 상태에 대한 분석 뿐 아니라 사용자의 미래 행동에 대한 추측 및 가정을 도출할 수도 있다. Dey(2001)는 상황(Context)에 대해 “실세계에 존재하는 개체의 상황에 특징을 부여함으로써 정의한 정보”라고 하였다.[Dey, Anind K., 2001]

상황인식에 관한 Dey의 연구초기와 컴퓨터 공학 분야에서 처음 상황(Context)에 대한 개념을 접했을 때 이는 사용자의 위치에 대한 문제라고 여기는 경향이 대다수였다. 이에 따라 초기의 상황인식 관련 연구에서는 위치기반 및 인간의 행동인식에 대해 깊게 초점이 맞춰져왔다. 그러나 최근의 연구들을 통해 상황인식은 단순히 위치 정보뿐만 아니라 사용자가 포함된 다양한 프로세스의 부분, 예를 들면 적응형 인터페이스, 응용 프로그램의 데이터 설정, 정보 검색의 정확도 향상, 서비스의 발견, 사용자와의 상호작용, 스마트 환경 등에 대한 고려가 되고 있다. 또한 상황인식의 응용을 위해 정교하며 일반적인 상황모델이 제안되고 있다.[Cristiana Bolchini et al, 2007]

상황인식 시스템은 크게 상황정보의 획득(인식센서 사용을 통한 상황의 인지), 상황에 대한 추상화 및 이해(센서로 얻어진 정보와 정의된 상황과의 매칭), 인식된 상황에 바탕으로 한 애플리케이션의 동작(상황에 따른 작업 실행) 세 부분으로 구성된다고 볼 수 있다.[Schmidt, Albrecht 2003]

상황인식은 유비쿼터스 컴퓨팅 시스템을 지원하기 위한 기술이라고 볼 수 있다. 상황인식은 혁신적인 사용자 인터페이스를 디자인하는데 사용되기도 하며, 유비쿼터스 및 웨어러블 컴퓨팅(Wearable computing) 환경에서 사용되기도 한다. 뿐만 아니라 인터넷과 하이브리드 검색 엔진의 출현의 시작과도 관련 있다고 할 수 있다.

3. 모바일 환경에서 상황정보 활용

가. 상황정보(Context)

상황인식이 연구된 초기에 Dey는 현실 세계의 개체들의 특징을 구분하기 위한 가장 일반적이고 명확한 기준을 Location, Identity, Activity, Time 4가지로 보았다. 그러나 상황인식에 대한 연구가 계속되며 더 복잡하고 상세한 상황에 대한 개념이 정립되기 시작했다. 이에 따라 상황은 유비쿼터스 컴퓨팅과 관련하여 사용자와 다른 사용자, 시스템, 혹은 디바이스의 애플리케이션 간 상호 작용에 영향을 미치는 사람, 장소, 사물, 개체, 시간 등 상황(situation)의 특징을 규정하는 정보. 좀 더 구체적으로는 네트워크 연결 상태, 통신 대역폭, 그리고 프린터·디스플레이·워크스테이션과 같은 컴퓨팅 상황(Computing context), 사용자의 프로필·위치·주변의 사람들을 비롯한 사용자 상황(User context), 조명·소음 레벨·교통 상태·온도 등 물리적 상황(Physical context), 시간·주·달·계절 등 시간적 상황(Time context)등으로 그 개념이 확대되었다.

이런 상황 정보들은 시스템으로 하여금 다양한 감지 장치(Sensing Device)와 애플리케이션을 통해 파악되고 모아져 다양한 응용 서비스 제공에 이용되거나, 다른 상황 정보와 묶여 제3의 결론을 얻는 추론에 사용되기도 한다.[정보통신용어사전, 2010]

나. 모바일 환경에서 상황정보 활용

국외 모바일 기술 전문 블로그 C. Enrique Ortiz(2007)에서는 모바일 상황정보(Mobile Context)는 인간-중심 컴퓨팅 구현과 매우 밀접한 관련 있다고 평가하고 있다. 모바일 환경의 급속한 발전으로 모바일의 특성인 이동성을 이용한 각 종 서비스 등이 가능해지고 있으며 이들은 사용자들 중심으로 사회적 네트워크 서비스(Social Network Service) 즉, 더 좋은 의사결정을 위한 정보의 교류 및 공유, 주변 환경에 대한 학습, 정보의 구매 등을 포함하고 있다.

다시 말하면 모바일 환경의 발전은 사용자 개인을 중심으로 하여 지속적으로 발전하고 있으며 이를 구성하는 주요요소는 사용자의 상황정보(Context)라 할 수 있다. 상황정보는 일반적인 정보와 달리 누가(Who), 언제(Where), 어디

서(When), 어떤 것(What)들과의 사용자 상호작용을 의미한다.

국내 모바일 전문 블로그 mobizen(2009)에서는 모바일 상황정보(Mobile Context)에 관하여 다음과 같이 설명하고 있다.

모바일 상황정보는 크게 사용자가 속해있는 현재의 상황(Situation)과 통신(Communication)으로 구성된다. 첫 번째 사용자의 현재 상황은 시간, 위치, 장소, 관심사 뿐 아니라, 어떠한 것이든 사용자가 필요로 하는 것을 이야기 할 수 있다. 두 번째 통신은 목소리, 텍스트, SMS, MMS, E-mail 등 다양한 형태의 대화 혹은 정보로써 이는 모바일 상황정보의 또 다른 핵심 요소라 할 수 있다. 이는 사용자의 현재 상황 즉 Situation과 달리 모바일 디바이스(Device)와 관련되어 있으며 디바이스의 성능, 지원 서비스 영역, 통신 속도에 따라 상이한 통신 방법을 선택하여 개인의 사회적 네트워크를 유지하거나 확장 할 수 있다.

모바일의 상황정보 활용에는 대표적인 예로 개인화 서비스, Push형 정보 전달, Social의 확장이 있으며 이에 대한 간략한 설명은 다음과 같다.

- 개인화 서비스

모바일은 철저하게 개인화된 디바이스로 사용자 상황에 맞는 맞춤형 서비스를 제공해야 한다. 위치 정보 서비스, Mobile 프로파일링, 개인화 서비스 등을 활용한 서비스가 필요하다. 이중 Web에 비해 Mobile Web이 우위를 갖는 LBS는 가장 핵심이 될 것이다.

- Push형 정보 전달

기존에는 '정보'를 얻기 위해서는 사용자가 PC에서 브라우저를 실행시킨 후 웹주소를 입력해야만 했다. 모바일은 기존의 Pull을 통한 정보 인식을 Push형으로도 제공할 수 있다.

- Social의 확장

현재의 상황을 인식하여 사용자의 Presence를 얻어내고, 이를 개인의 Social Network와 확장할 수 있는 서비스가 모바일에서의 주요 쟁점이다.

제 3 절 정보추천 관련 기법

1. 개요

오늘날 인터넷의 급속한 발전에 따라 사용자들이 웹상에서 접할 수 있는 아이템과 정보들의 수가 기하급수적으로 늘고 있다. 이러한 현상은 아이템 및 정보의 다양화라는 긍정적인 영향을 불러오기도 하였지만 사용자들의 효과적인 의사결정에 악영향을 미친다는 평가를 받고 있기도 하다. 이에 따라 추천 시스템(Recommendation System)에 대한 관심이 점점 확대되어 왔다.

추천 시스템은 사용자에게 적절한 아이템 및 정보를 추천해주는 시스템으로 현재 뉴스, 영화, 도서, 음악 등의 추천에 적용되어 널리 이용되고 있다.

인터넷의 등장과 발전에 따라 발생한 대표적인 정보추천 관련 기법인 정보 검색과 정보필터링의 개념에 대해 간략히 정리하면 다음과 같다.

- 정보 검색(Information Retrieval)

정보를 수집, 분류, 저장하였다가 필요한 경우 추출하는 일련의 과정을 말한다. 기존 도서관이나 서점 등에서 자료를 찾기 위해 이용되었던 정보 검색은 인터넷이 확산됨에 따라 그 개념이 웹 사이트나 웹문서를 찾기 위한 검색포털로 발전되었다. 주로 디렉토리 검색, 기사검색 웹페이지 검색 등에 이용되어 왔으며 폭발적인 인터넷 정보의 증가에 따라 정확하고 핵심적인 정보를 제공하는 것이 주요 목적이라 할 수 있다. 사용자의 정보판단을 위해 문서 혹은 자료의 제목을 제시하며 또는 문서의 내용 중에 검색 키워드가 존재할 경우 그 앞 뒤 일부분을 제시한다. 그러나 사용자가 원하는 정보를 정확하게 나타내지는 못하기 때문에 사용자들은 검색 결과의 적합성을 판단하는데 일정 시간 및 노력을 기울여야 한다.[박철제, 전종근, 2002]

- 정보 필터링(Information Filtering)

웹을 통해 들어오는 정보들을 브라우저로 보내기 전에, 불필요한 정보를 삭제하거나 추가적인 내용을 덧붙이거나하는 작업을 정보 필터링이라고 한

다. 일반적인 정보 필터링의 운영 과정으로는 사용자의 관심도를 기반으로 프로파일 모델링 단계 후 검색된 문서를 표현하기 위한 특정 문서추출 단계를 거친다. 이때 사용자 프로파일과 문서의 유사성을 측정하여 검색된 문서가 사용자의 관심문서인지를 판단한다.[송미란 외 1999 : 1]

정보 필터링의 동적인 특성을 갖는 정보원천을 이용한다는 점이 정보검색과 가장 큰 차이점이라 할 수 있다. 이에 따라 사용자는 제공된 정보의 적합성을 재판단하지 않아도 된다. 이는 사용자의 관심을 처리하기 위한 프로파일 구축을 바탕으로 실현된다. 정보필터링의 단점으로 지적되는 사항은 사용자가 새로운 정보 분야를 요구할 경우 지능적으로 대처하지 못할 수 있다는 점이며 이는 사용자의 선호도가 충분히 구축되어 있지 않는 경우 발생한다. 또한 사용자의 관심이 너무 한쪽으로 치우칠 경우 새로운 정보에 대한 접근 기회가 사전에 차단될 수 있는 문제점도 있다.[김진 외, 2000: 324] 정보 필터링의 세부 요소로는 협업 필터링, 내용기반 필터링 등이 있다.

정보 검색은 고정된 문서 집합에서 관련 문서를 검색하기 때문에 검색내용은 줄지만 내용을 파악하기 위한 시간과 노력이 요구되는 단점이 존재한다. 이러한 단점을 보완하기 위해 정보 필터링기법이 대두되었다. 정보 필터링은 대량의 온라인 정보를 줄임으로써 정보과적을 줄이고자 하는 것은 정보검색과 그 개념이 동일하나 사용자 프로파일을 이용하여 정보검색의 단점을 보완하고자 하는 기법이다.

정보필터링은 협업, 내용기반, 인구통계학적, 유용성기반, 지식기반으로 총 5가지 유형의 기법으로 구분되며 이 중 협업 필터링은 개인자체의 프로파일뿐만 아니라 유사한 선호도를 가진 집단의 성향을 기준으로 정보를 추천하기 때문에 적절한 정보제공을 가능하게 한다.[손창환, 2006: 12-15]

2. 정보 필터링 기법 분류

정보 필터링기법은 고객과의 상호작용, 인터페이스의 유형에 따라 분류되는 것이 아니라 사용자가 시스템 기반 데이터와 그 데이터를 이용하는 방법에 따

라 분류될 수 있다. 일반적으로 추천과정이 시작되기 전에 시스템 내에 존재하는 Background Data, 추천을 하기 위해 고객에 의해 입력되어지는 Input Data, 그리고 추천하기 위해 Background Data와 Input Data를 조합하는 알고리즘 등으로 형성되어있으며 이때 다음과 같이 크게 5가지 유형의 기법으로 분류된다.[Burke, 2002]

[표 2-1] 필터링 기법 분류

기법	Background	Input	Process
협업	각 항목에 대해 점수 부여	각 항목에 대해 점수 부여	특정고객이 부여한 점수를 토대로 유사고객 파악
내용기반	각 항목의 특징	각 항목에 대해 점수 부여	각 항목에 대한 사용도와 점수에 분류자 도출
인구통계학적	고객의 인구통계학 정보 및 각 항목에 대해 점수 부여	고객에 대한 인구통계학 정보	유사 고객 파악 시 인구통계학적 요인 구분에 의해 유사 고객 파악
지식기반	고객욕구를 충족시킬 수 있는 항목에 대한 지식	고객의 필요와 흥미에 대한 설명	고객의 욕구를 충족시키는 항목 추정
유용성기반	각 항목의 특징	고객이 선호하는 각 항목의 유용성	유용성에 따라 각 항목을 분류 후 항목선택

자료 : Burke(2002), Hybrid Recommender Systems : Survey and Experiment, User Modeling and User-Adapted Interaction

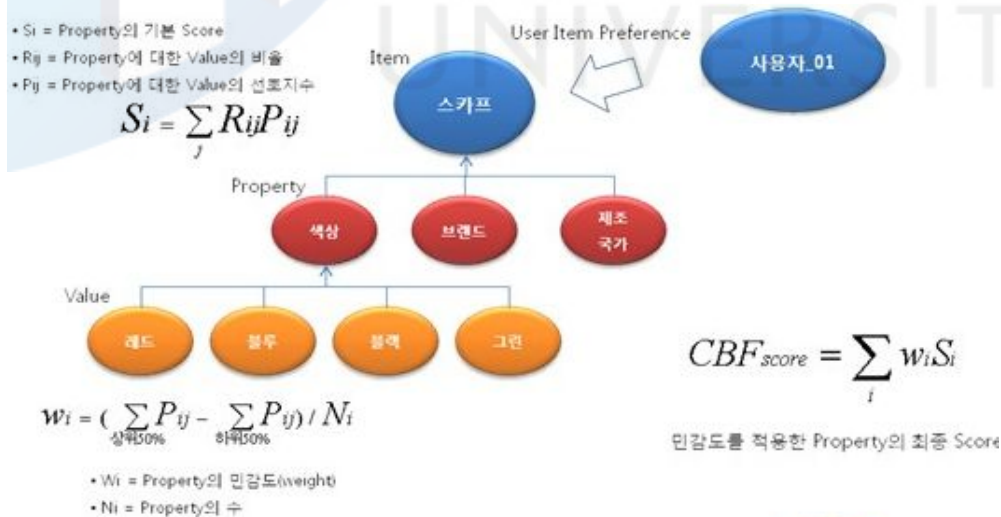
정보마이닝 및 시맨틱 기술 관련기업인 솔트룩스(Saltlux)는 자사 포털사이트의 IT wiki부분에서 정보추천기술에 대해 다음과 같이 설명하고 있다.

추천 결과를 제공하기 위한 방법으로 사용되는 정보필터링은 크게 사용자가 선택한 콘텐츠의 속성 정보를 바탕으로 한 내용기반 필터링과 사용자의 사회적 환경 및 관계 정보를 바탕으로 한 협업필터링, 인구 통계학적 기법을 근간으로 하여 사용자를 대표하는 군집을 추출하고 이들 공통 패턴과 속성을 뽑아 이를 사용자의 선호정보로 이용하는 인구통계학적 필터링, 사용자의 명시적인

입력에 의하여 추천하는 지식기반 필터링 등으로 나뉜다. 각각에 대한 설명은 다음과 같다.[솔트룩스(Saltlux), 2010]

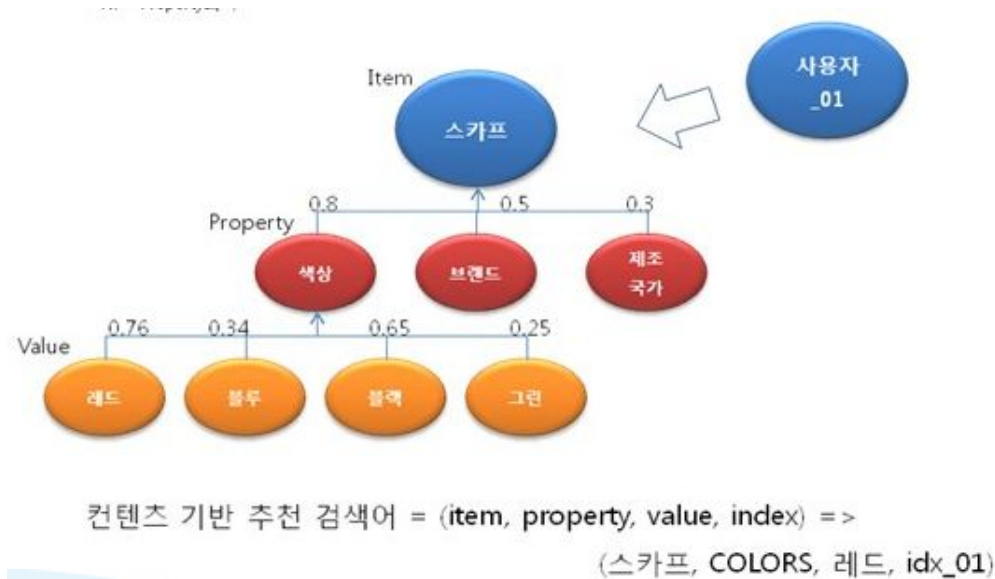
가. 내용기반 필터링(Content-based filtering)

내용기반 필터링은 사용자의 선호 정보를 제공받아 그와 가장 유사한 아이템을 검색해 주는 기법이다. 정보검색 기술에 기반을 두고 있는 내용기반 필터링 기법은 공학적인 관점에서 볼 때 매우 상식적인 접근이라 할 만하나, 몇 가지 단점을 가지고 있다. 추천의 범위를 사용자들 자신이 입력한 좁은 프로파일 안에 가두게 될 염려가 있으며 기법 자체가 추천 아이템의 도메인에 지나치게 종속적인 점이 문제이다. 또한 사용자는 되도록 적게 입력하되 양질의 추천을 얻으려는 성향이 있어 사용자의 온전한 선호 정보 프로파일을 얻는 것이 어려워진다. 내용기반 필터링의 알고리즘은 대상 도메인의 아이템에 대하여 속성과 값을 추출하고 아이템을 구성하는 속성의 중요도를 계산하고 각 속성에 대한 사용자의 민감도를 계산한다. 민감도는 사용자가 색상, 브랜드, 제조국가에 영향 받는 정도를 의미하며, 내용기반 필터링 점수를 계산하는데 가중치로 사용된다. 민감도는 상위 50%의 하위분류 값들의 평균과 하위 50% 하위분류 값들의 평균의 차이를 통해 계산된다. 즉, 사용자의 평가점수의 편차가 큰자는소에 대해서는 사용자의 선호도의 차이가 크다고 가정한 것이다. 이



<그림 2-4> 내용기반 필터링 예-1

는 도메인에 따라서 다르게 적용이 가능하다. 민감도의 값들의 합이 1이 되도록 보정한 후, 앞에서 구한 요소들의 점수에 가중치로 사용하여 내용기반 필터링 스코어를 계산한다.

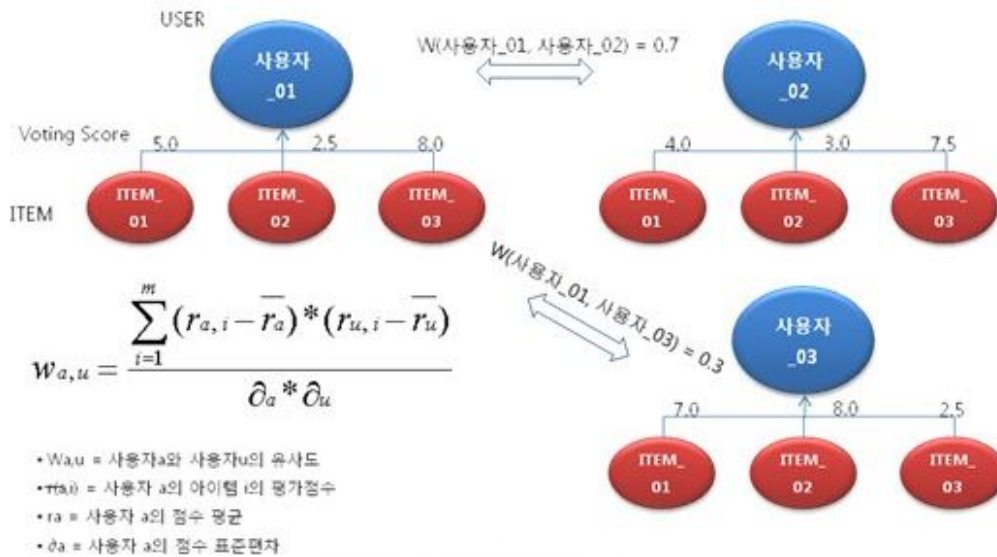


<그림 2-5> 내용기반 필터링 예-2

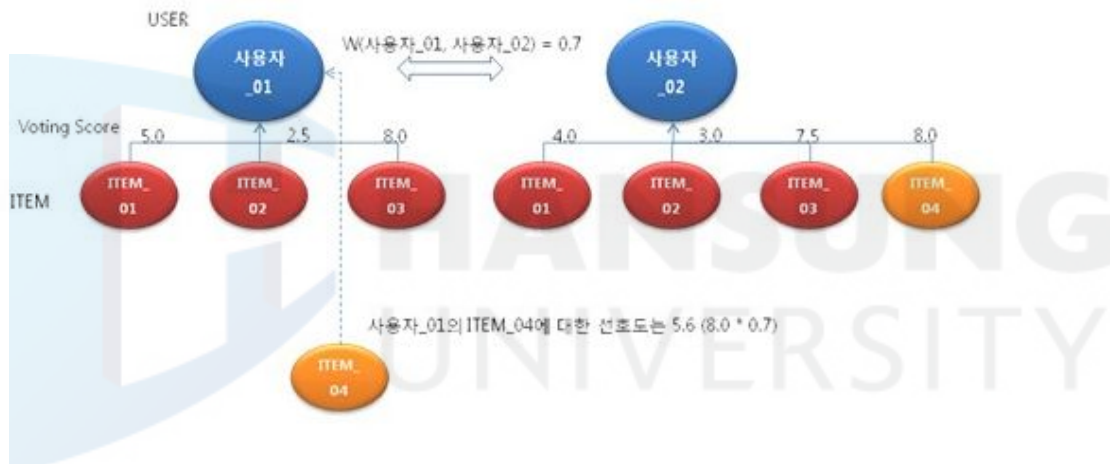
나. 협업필터링(Collaborative filtering)

사회적인 관점에서 볼 때 내용기반필터링 기법은 사용자 개인의 내부에서 기인하는 추천이고, 협업필터링 기법은 사용자의 외부의 다른 사람으로부터 오는 추천이라고 할 수 있다. 외부의 다른 사람들 소리에 귀를 기울이고 따를 것인가의 문제는 각 개인에게 있어 중요한 아이템 결정의 요소가 된다.

협업필터링 기법은 대상 아이템이 상세한 설명 정보를 갖고 있지 않거나 컴퓨터가 프로세싱하기 어려운 분야의 것이라고 해도 적용이 가능하다. 사용자로부터의 상세한 선호 정보가 없어도 적용이 가능하며 아이템과 사용자의 관계를 분석하기에 아이템에 대한 상세한 선호 정보가 없어도 가능하다. 단점으로는 아이템에 대한 사용자들의 평가가 필요하다는 점과 시스템에 새로운 아이템이 추가되어도 평가 전에는 추천될 경로가 봉쇄된다는 점이다. 이는 새로운 사용자가 추가되어도 협업자를 도출하기 전까지는 마찬가지이다,



<그림 2-6> 협업 필터링 예(사용자01과 유사한 사용자_02 판별)

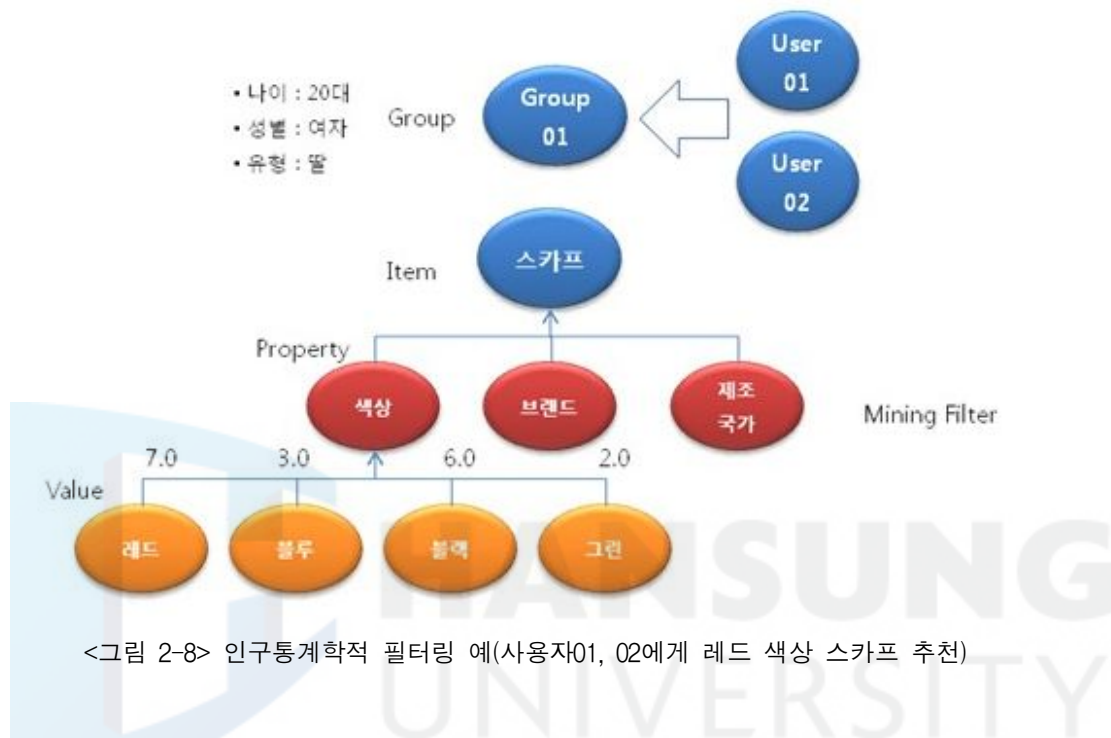


<그림 2-7> 협업 필터링 예(사용자01에는 존재하지 않는 ITEM_04에 대한 선호 예측)

다. 인구통계학적 필터링(Demographic Filtering)

인구 통계학적 기법을 근간으로 하여, 사용자를 대표하는 군집을 추출하고 이들 공통 패턴과 속성을 뽑아 이를 사용자의 선호정보로 이용하는 기법이다. 마케팅 자료로 많이 이용되어 왔으며, 공학적인 추천 기법으로는 외면되어 왔다. 대표적으로 같은 성별로부터 온 추천, 전문가 그룹에서 온 추천, 영화의

박스오피스와 같은 실제 유효성이 존재하는 추천 방식으로 구현이 가능하므로, 다중 필터 중 하나로 인구통계학적 필터링을 구현하여 추천에 복합적으로 사용할 수 있다. 데이터마이닝의 같은 축적된 자료로부터 공통의 그룹을 형성하고 패턴을 추출하는 기법을 적용하여 인구통계학적 필터링 기법의 근간이 되는 통계학적 자료를 마련하고 이를 통해 추천의 정확성과 해석을 기대할 수 있다.



라. 지식기반 필터링(Knowledge-based Filtering)

지식기반 필터링 기법은 사용자의 명시적인 입력에 의해 상황과 서비스, 아이템 따른 추천 콘텐츠를 입력하게 하여, 사용자에게 제공하는 기법으로 가장 단순하지만, 효과가 좋은 기법이라고 할 수 있다. 역으로 사용자가 제외하고 싶은 아이템에 대해서도 적용할 수 있다. 하지만 사용자에게 명시적으로 다양한 서비스, 상황에 따른 추천 받을 아이템을 입력 받는 것이 쉽지만은 않으므로, 내용기반 필터링 기법과 마찬가지로 암묵적으로 이러한 사용자의 선호를 추출하는 기법에 대한 연구가 이뤄져야 한다.

Profbuilder[Wasfi, 1999]는 내용 기반 필터링과 협업 필터링 모두를 사용하여 웹페이지를 추천하였다. 사용자들은 웹사이트로부터 두 개의 추천 목록을 받게 되는데 하나는 내용 기반 필터링에 의해 얻어진 결과이며 나머지는 협업 필터링에 의해 얻어진 결과이다.

그러나 내용기반 필터링과 협업 필터링의 결합에서 내용기반 필터링의 한계점이 그대로 문제점으로 지적되어 왔는데 이는 내용기반 필터링을 적용하고자 할 때 아이템은 컴퓨터가 이해 할 수 있는 형식(예: 텍스트)으로 되어 있어야 하고 프로파일에 존재하는 단어와 항목의 내용에 나타는 단어가 정확하게 일치 하지 않을 경우 항목을 선택할 수 없는 문제점을 가지고 있다는 것이다. [김병만 외, 2004] 이러한 한계점에 따라 국내에서는 내용 기반 필터링 기법을 사용하지 않는 새로운 형태의 하이브리드 필터링에 대한 연구가 진행되어 왔다.

박지선(2002)은 협업 필터링의 세부 기법이라 할 수 있는 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 비교분석하고 이를 결합하고자 하였다. 추천 시스템의 예측 능력을 향상시키기 위하여 유사한 선호도를 가지는 고객들의 평가에 근거하여 아이템들 간의 유사도를 구해 특정 아이템에 대한 고객의 선호도를 예측하고자 하였다.

손창환(2006) 역시 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 결합하고자 하였다. 사용자 기반 협업 필터링을 먼저 적용 한 후 아이템 기반 협업 필터링을 적용하는 것이 예측의 정확성을 높일 수 있음을 주장하였으며 이를 기반으로 웹상에서 도서추천시스템을 구축하였다.

다음 단락에서는 본 연구에서 제시하고자 하는 하이브리드 필터링의 근간인 협업 필터링에 대해 자세히 설명하도록 한다.

3. 협업 필터링(Collaborative filtering)

가. 협업 필터링 개요

오늘날 협업 필터링은 추천의 정확도가 다른 필터링에 비해 높다고 인정을 받아 다양하게 이용되고 있으며 뿐만 아니라 가장 성숙한 기법으로 평가되고

있다. 일반적으로 협업 필터링은 아이템에 대한 평가를 종합하고, 이 평가 점수를 토대로 고객 간의 유사성을 인지하여 새로운 아이템을 추천하는 기법이라 할 수 있다.[Sarwar et al, 2001]

협업필터링은 메일 시스템인 Tapestry에서 유래 되었으며 잦은 전자 메일 사용으로 발생하는 사용자들의 부담감을 해결하기 위해 연구되었다. 많은 양의 메일로 인한 업무효율 저하를 방지하기 위해 사용자의 관심도를 바탕으로 메일을 필터링하는 시스템을 구축했다. 그러나 워크그룹과 같은 공동체 구성원들의 의견으로만 메일을 정제했다는 한계점을 가지고 있다.[Goldberg et al.,1992]

이후 개인화 추천서비스의 제공을 위해 사용자선호도 기반의 자동 추천 시스템 등이 개발되었는데 GroupLens Research System은 사용자의 직접적인 평가 및 사용자의 직접적인 평가가 없는 경우 다른 사용자 프로파일 기반으로 추천시스템을 구현하였다. 이는 평가가 없을 시 같은 관심을 가진 집단의 평가를 바탕으로 문제를 해결하고자 한 방법으로 대표적인 협업 필터링 기반의 추천시스템으로 인정받았다.[Sawar, 2000]

GroupLens Research System은 Filterbot이라는 가상의 사용자를 이용하는데 이들은 새로운 기사 혹은 정보가 출현하자마자 그들에 대한 평가 점수를 부여하도록 되어있다. 협업 필터링의 단점 중 하나인 희소성의 문제를 해결하긴 하지만 가상의 사용자들의 평가를 기반으로 하기에 현실세계의 사용자 성향과 차이가 발생할 수 있다.[김병만 외, 2004]

협업 필터링은 내용기반 필터링이 가지고 있는 몇몇 문제점을 해결한다. 내용기반 필터링에서 자동으로 분석이 어려웠던 아이템의 속성들이 사용자의 평점으로 환원되기 때문에 내용기반 필터링이 다루지 못했던 아이템들에 대한 추천이 가능하다. 또한 사용자의 취향이나 아이템의 질에 기반을 둔 추천을 가능하게 하며 무엇보다 다른 사용자의 경험을 바탕으로 하기 때문에 사용자가 기존에 선호해 왔던 아이템 외에도 사용자가 높이 평가 할 수 있는 아이템에 대한 추천도 가능하다.[안신현, 2007]

나. 협업 필터링의 한계점

협업 필터링 기법은 일반적으로 확장성(Scalability) 및 희박성(Sparsity) 측면에서 한계를 가지고 있다고 알려져 있다.[김재경, 2005]

희박성의 측면에서 볼 때 시스템에 새로운 아이템이 추가되었을 때, 이에 대한 사용자의 평점이 쌓이기 전까지는 이 아이템을 추천할 수 있는 방법이 없다. 아이템 수에 비해 사용자 수가 적을 경우에는 평점이 존재하지 않는 아이템이 많을 수 있기 때문에 추천대상이 될 만한 아이템의 범위가 좁아진다. 확장성의 측면에서 볼 때 사용자 및 아이템의 수가 증가한다면 사용자와 이웃 집단을 판별하기 위한 연산량이 기하급수적으로 늘어날 수 있다. 협업 필터링에 대한 기본적인 한계점을 정리하면 다음과 같다.

[표 2-2] 협업 필터링의 한계점

확장성(Scalability)	희박성(Sparsity)
- 고객과 아이템의 수가 증가함에 따라 이웃 고객군을 찾기 위한 연산량이 기하급수적으로 늘어날 수 있다. - 실시간 추천이 목적인 추천시스템에서 이는 심각한 확장성 문제일 수 있다.	- 아이템 혹은 정보의 수가 기하급수적으로 늘어남에 따라 고객의 선호도가 입력되지 않은 콘텐츠의 수가 상대적으로 증가한다. - 이웃 고객군 형성과정에서 적은 수의 평가 데이터만을 사용하게 됨으로 유사도 측정에 신뢰성이 떨어질 위험이 있다.

이와 같은 단점에도 불구하고 다양한 연구에서 협업 필터링 기법은 다른 기법보다 예측력이 우수한 것으로 인정받고 있다. 이는 사람들의 취향이 무작위로 분포된 것이 아니라, 그룹이나 사람들의 취향 사이에는 일반적인 트렌드와 패턴이 있다는 가정에서 비롯된 것이다. 실제로 사람들은 주변의 지인이나 인터넷을 통한 다른 사용자들과의 커뮤니케이션 등을 통해 아이템이나 경험에 대한 추천을 받고, 이를 선택에 참조한다.[안신현, 2007]

협업 필터링은 이러한 입소문(Word of Mouth)의 개념을 자동화 한 것이라고 할 수 있다.[Shardanand, Maes, 1995]

다. 협업 필터링의 구분

협업 필터링은 크게 사용자 기반 협업필터링(User-Based Collaborative filtering)과 아이템 기반 협업필터링(Item-Based Collaborative filtering)으로 구분된다.

▪ 사용자 기반 협업필터링

사용자 기반 협업필터링은 유사한 선호도를 가지는 다른 고객들, 즉 비슷한 트렌드나 패턴을 보이는 이웃집단군의 아이템 평가를 바탕으로 해당 고객이 선호할만한 아이템을 예측하여 추천을 하는 기법이다.

즉 사용자들의 아이템에 대해 과거의 구매결정이나 선호를 표시한 데이터를 토대로 사용자 프로파일을 구성하고 각 사용자들의 프로파일을 비교하여 유사도 산출을 통해 이웃집단군을 형성한 후 새로운 아이템에 대해 미리 경험한 이웃집단의 선호도에 근거하여 사용자의 새로운 아이템에 대한 선호도를 예측하여 아이템을 추천한다.

사용자 협업 필터링은 두 고객의 선호도에 대한 상관관계는 두 고객 모두 선호도를 표시한 아이템에 대해서만 계산되므로 고객 간 유사도를 측정할 때 선호도가 표시되지 않은 아이템은 배제된다는 단점이 있다. 또한 비록 두 이용자가 선호도에 따른 상관관계가 높지 않더라도 다른 고객의 선호도 예측에 좋은 자료가 될 수 있으나 상관관계가 높지 않다는 이유로 정보로 활용되지 못하는 한계점을 가지고 있다. 끝으로 오직 두 사용자 사이에서만 상관관계가 산출된다는 것이다. 예를 들어 사용자 A와 B가 아주 높은 상관관계에 있고, B와 C도 그렇다고 가정하면 A와 C도 상관관계가 높다고 할 수 있다. 그러나 A와 C가 공통된 아이템 어느 것에도 선호도를 표현하지 않았을 시 A와 C의 상관관계를 산출 할 수 없다.

▪ 아이템 기반 협업필터링

아이템 기반 협업필터링은 고객이 기존에 평가한 각각의 아이템들과 특정 고객의 선호도를 예측하고자 하는 아이템이 얼마나 비슷한가를 계산하여 K

개의 가장 비슷한 아이템을 선택하는 것으로 사용자 기반 협업 필터링과 반대로 사용자 간의 유사도가 아닌 아이템간의 유사도를 측정하여 아이템을 추천하는 것이다.

아이템 기반 협업필터링은 아이템들 간의 유사도를 산출하기 위하여 이 두 아이템에 모두 선호도를 입력한 사용자들의 선호도를 이용하는데 이 때 사용자들 간의 유사도가 전혀 고려되지 않는 한계점이 있다. 이에 따라 만약 특정 고객과 전혀 선호도가 비슷하지 않은 사용자들의 평가를 기반으로 예측을 수행한다면 아이템들 간의 유사도의 정확도가 떨어지고 아울러 추천 시스템의 예측 능력과 추천 능력이 저하 될 수 있는 문제점을 가지고 있다. 또한 아이템들 간의 유사도를 계산하는 과정에서 고객이 선호도를 표현한 아이템에 대한 고려가 되지 않는 경우도 발생할 수 있다. 현재 고객이 가장 선호하는 제품이 추천과정에서 누락되어짐에 따라 추천의 정확도가 떨어질 수 있다는 것이다.

아이템 기반 협업필터링은 고객이 선호도를 입력한 기존의 아이템들과 예측하고자 하는 아이템의 상관관계를 산출하여 선호도를 예측한다. 이 기법은 사용자들 간의 유사도가 전혀 고려되지 않기 때문에 만약 특정 고객과 전혀 선호도가 비슷하지 않은 사용자들의 평가를 기반으로 한다면, 아이템들 간의 유사도가 정확하지 않고 아울러 추천시스템의 예측 능력과 추천능력이 저하 될 수 있다. 이러한 문제점을 보완하기 위해 유사한 선호도를 가지는 사용자들의 평가에 근거하여 아이템간의 유사도를 구할 수 있다. 즉 사용자 기반 협업 필터링으로 유사이웃집단을 형성 한 후 아이템 기반 협업 필터링을 적용하여 아이템들 간의 유사도를 구하는 것이다.[박지선 외, 2002]

제 4 절 시사점 및 본 연구의 위치

본 장에서는 본 연구에서 활용되고 다루지는 집단지성, 상황인식, 정보추천에 관한 이론적 배경에 대해 정리하였다. 각각에 대한 시사점은 다음과 같다.

본 연구에서 설계하고 구축하는 추천시스템의 근간이 되는 협업필터링 기반의 정보 추천은 사용자 외부의 집단지성으로부터 오는 추천이라고 할 수 있다. 외부의 다른 사람들 소리에 귀를 기울이고 따를 것인가의 문제는 각 개인에게 있어 중요한 아이템 결정의 요소가 된다. 집단지성은 위키피디아, 다음아고라 등의 시발점이 된 개념으로 집단의 총의를 모아 최적의 의사결정을 하고자 하는 목적을 가지고 있다. 협업 필터링은 집단지성을 실현하기 위한 매개체로 유사 선호도를 가지고 있는 이웃집단의 입소문(Word of Mouth)을 전산화를 통해 구성하며 이를 의사결정을 위해 사용한다. 이는 집단지성을 구현하기 위한 방법론의 한 가지 형태라고 할 수 있다.

상황인식이란 일반적으로 실세계에 존재하는 개체의 상태를 특징화하여 정의한 정보, 즉 상황정보(Context)를 활용하여 사용자가 처한 상황을 인지하고 사용자가 원하는 정보를 원하는 형태로 처해진 상황에 맞게 획득하고 이용할 수 있는 상태를 의미한다. 상황정보를 좀 더 구체적으로 컴퓨터상황(네트워크 연결 상태, 통신대역폭, 프린터, 디스플레이 등), 사용자 상황(사용자프로파일, 위치, 선호도, 목적, 주변사람), 물리적 상황(계절, 조명, 소음레벨, 교통), 시간적 상황(시간, 주, 달)으로 구분 할 수 있다.

한편 사용자가 개인 단말기를 이용하는 모바일 환경에서는 사용자의 상황정보를 활용하여 개인화 서비스를 제공할 수 있다. 정보필터링의 전 단계로 사용자 상황정보를 이용하여 추천정보의 카테고리 및 범위를 결정한다면 추천의 질 또한 높일 수 있을 것이다. 따라서 본 연구에서는 기존의 추천기법과 달리 사용자의 상황정보를 이용한 하이브리드 필터링을 수행하여 추천을 실행하도록 한다.

정보필터링은 방식에 따라 내용기반, 협업, 인구통계학적, 유용성기반, 지식기반 등의 필터링으로 구분되는데 이 중 내용기반, 협업 필터링이 가장 효과적이라 알려져 있다. 한편 하이브리드 필터링은 단점이 없는 둘 또는 그 이상

의 필터링 기법을 결합하여 좀 더 정확한 예측을 하고자 하는 기법으로 관련 연구들에 따르면 내용기반 및 협업 필터링을 결합한 방식에 대한 연구가 주를 이루어 왔으나 내용기반 필터링의 한계점에 따라 최근에는 협업 필터링의 세부 기법인 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 결합에 대한 다양한 연구들이 시도되고 있다.

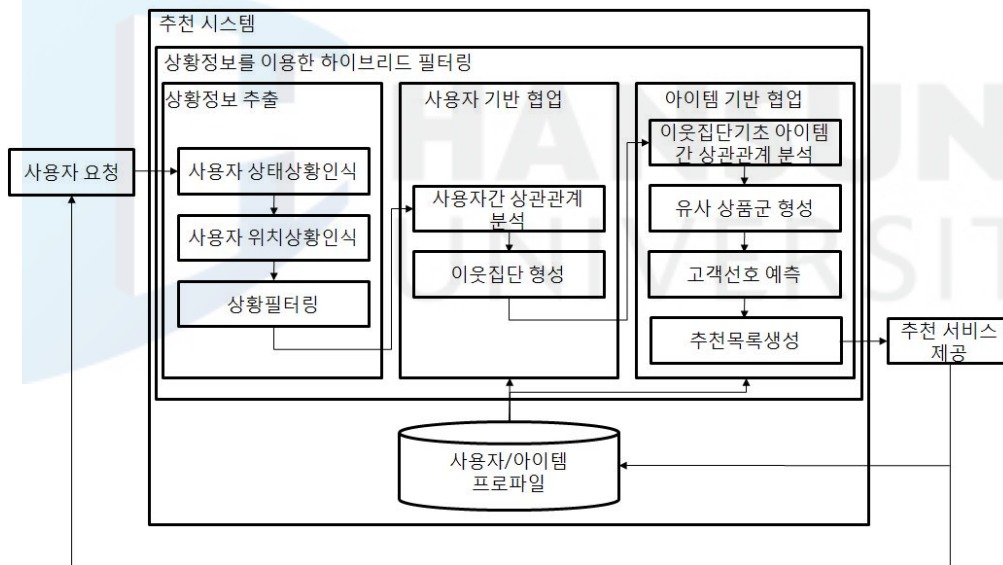
상기 내용을 요약하여 정리하면 본 연구에서는 집단지성을 구현하기 위한 방법론 중 하나라 할 수 있는 협업필터링을 기반으로 하여 하이브리드 필터링을 형성한다. 협업필터링은 사용자 기반 협업 필터링과 아이템 기반 협업 필터링으로 구분되는데 추천의 질을 높이기 위해 각 필터링의 이점만을 결합하는 것이다. 또한 모바일 이용자의 상황정보를 활용하여 이용자의 목적, 위치에 대한 정보를 파악한 후 하이브리드 필터링 적용함으로써 추천 서비스의 질을 향상시키고자 한다.



제 3 장 상황정보를 이용한 하이브리드 필터링

제 1 절 상황정보를 이용한 하이브리드 필터링 알고리즘

본 연구에서는 새로운 형태의 추천시스템 구축을 위해 상황정보를 이용한 하이브리드 필터링 기법을 이용한 정보의 추천을 제시한다. 상황정보를 이용한 하이브리드 필터링은 사용자가 선호할 만한 아이템을 예측하기 위해 먼저 사용자의 현재 상황에 대한 정보를 파악한다. 이는 정보 필터링의 전 단계로 사용자의 목적, 위치에 따라 사용자에게 제공되어야 할 정보 혹은 아이템의 범위를 결정한다고 볼 수 있다. 사용자의 목적, 위치에 따라 추천하고자 하는 서비스가 결정되면 사용자 기반 협업 필터링의 적용으로 추천 목록이 형성되며 이 추천 목록은 다시 아이템 기반 협업 필터링의 단계를 거쳐 최종 추천 목록이 생성된다. 상황정보를 이용한 하이브리드 필터링 알고리즘의 논리적 개념도는 다음과 같다.



<그림 3-1> 상황정보를 이용한 하이브리드 필터링 추천 시스템

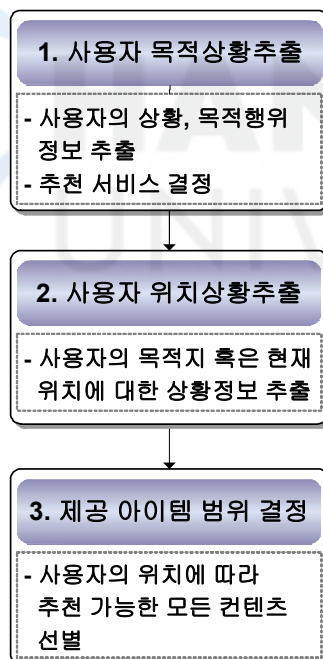
다음 세 개의 단락에서 상황정보를 이용한 하이브리드 필터링을 구성하는 상황추출, 사용자 기반 협업 필터링, 아이템 기반 협업 필터링의 세부적인 메

커니즘을 설명하도록 한다.

1. 상황추출

상황인식이란 실세계에 존재하는 개체의 상태를 특징화하여 정의한 정보, 즉 상황정보(Context)를 활용하여 사용자가 처한 상황을 인지하고 사용자가 원하는 정보를 원하는 형태로 처리된 상황에 맞게 획득하고 이용할 수 있는 상태를 의미한다.

본 연구에서의 상황추출은 사용자에게 추천 대상이 되는 서비스를 선별하는 과정으로 사용자의 목적행위, 위치 등의 상황정보를 추출하고 이를 고려하여 추천 대상이 되는 서비스를 결정한다. 예를 들면 사용자의 상황 혹은 선호 서비스에 대한 인식을 수행함으로써 추천하고자 하는 서비스를 결정한다. 다음으로 사용자의 목적지 혹은 현재 위치를 인식하고 해당 위치의 일정 범위 내에 존재하는 추천 정보 및 아이টে를 선별할 수 있다. 본 연구에서 제시하는 상황인식 절차를 순서도로 표현하면 다음과 같다.



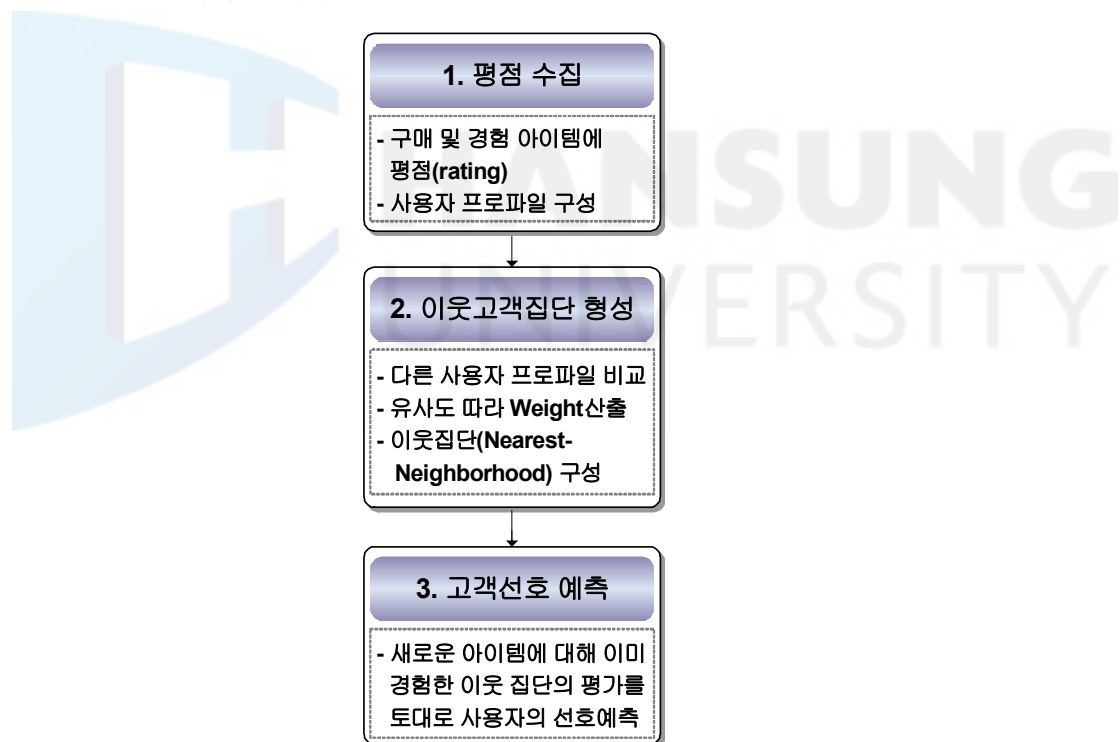
<그림 3-2> 상황추출 절차

사용자가 시스템을 구동하면 시스템은 사용자의 상태상황인식으로 사용자의 목적행위를 판별한다. 이용자가 자신의 쇼핑을 원한다고 가정하면 다음 단계에서 시스템은 사용자의 위치정보 추출로 사용자의 목적지 혹은 현재 위치를 판별한다. 이에 따라 해당 사용자의 상황에 추천 대상이 되는 일정 범위 내의 쇼핑 매장을 선별 할 수 있다. 그 다음 단계로 모든 쇼핑 매장 아이템에 대해 사용자들의 과거 경험이나 선호를 토대로 이웃집단을 선별하는 사용자 기반 협업 필터링이 실행된다.

2. 사용자 기반 협업 필터링

사용자 기반 협업필터링은 고객이 좋아할 만한 아이템을 예측하기 위해 비슷한 선호도를 가지는 다른 고객들, 즉 비슷한 트렌드나 패턴을 보이는 유사 집단군의 아이템 평가를 바탕으로 하여 추천하는 기법이다.[박지선 외, 2002]

사용자 기반 협업필터링의 일반적인 절차는 다음과 같다.



<그림 3-3> 사용자 기반 협업 필터링 절차

가. 사용자 평점(Rating) 수집

협업 필터링은 사용자와 유사한 선호를 가지는 이웃집단의 아이টে에 대한 선호도를 토대로 추천을 실행하기 때문에 무엇보다 사용자들의 평점을 수집하는 것이 중요하다. 단순히 사용자/아이템/평점과 같은 형태로 나타낼 수도 있다. 좀 더 양질의 추천을 제공하기 위해서 사용자들은 자신이 선택하거나 경험했던 아이টে에 대한 평점을 시스템에 제공해야 한다. 예를 들어 영화, 음반, 서적 등 카테고리 안의 각 아이টে에 대해서 사용자는 자신이 구매하거나 관심이 있는 아이টে에 대해 5점 또는 7점 등의 점수로 평점을 매긴다. 최근 인터넷의 발달로 대부분의 추천시스템이 웹을 바탕으로 이뤄짐에 따라 이러한 현상은 쉽게 찾아볼 수 있다.

나. 이웃 고객집단 형성

다음으로 사용자의 평점, 즉 고객의 과거 구매 기록이나 아이টে에 대한 선호도를 토대로 고객 간의 상관관계를 파악하여 유사이웃집단을 구성한다. 고객 별 유사도를 측정하기 위한 다양한 방법론이 존재하는데 피어슨 상관관계수(Pearson Correlation Coefficient)가 가장 많이 사용된다.[박지선 외, 2002]

- 피어슨 상관관계수(Pearson Correlation Coefficient)

$$r_{ij} = \frac{\sum_k (S_{ik} - \bar{S}_i) \times (S_{jk} - \bar{S}_j)}{\sqrt{\sum_k (S_{ik} - \bar{S}_i)^2 \times \sum_k (S_{jk} - \bar{S}_j)^2}}, -1 \leq r_{ij} \leq 1 \quad (1)$$

r_{ij} : 사용자 i , 이웃 사용자 j 간의 유사도

S_{ik} : 고객 i 가 아이টে k 에 대해 평가한 선호도

$\bar{S}_i$: 고객 i 의 선호도 평균

S_{jk} : 고객 j 가 아이টে k 에 대해 평가한 선호도

$\bar{S}_j$: 고객 j 의 선호도 평균

피어슨 상관계수는 -1에서 1사이의 값을 가지는데 1에 가까울수록 두 사용자간의 유사도가 높은 것이고 -1에 가까울수록 상반된 선호도를 가지고 있음을 나타낸다. 또한 0에 가까운 경우 선호도간 관계가 성립하지 않음을 의미한다. 사용자간 유사도는 아이템의 예측에 있어서 가중치로 이용되며 피어슨 상관계수의 경우 사용자와 선호도가 비슷한 이웃 사용자의 평가 뿐 아니라 선호도가 다른 사용자의 평가도 반영 할 수 있다.

피어슨 상관계수가 많이 이용되는 이유는 대용량의 데이터일 경우, 가장 적합하다고 평가되기 때문이다.[Breese, 1998]

이외에도 추천을 위한 유사성 산출의 척도로서 코사인 유사도, 베이지안 분류, 유클리디안 거리 계산 등이 있다. 다음은 피어슨 상관계수외의 유사성 척도들의 단점을 정리한 내용이다.

[표 3-1] 유사성 척도의 단점

유사성 척도	유사성 척도의 단점
코사인 유사도	- 군집 영역이 원점에 대하여 떨어져 있어야 한다 - 군집 영역들도 충분히 떨어져 있어야 한다.
베이지안 분류	- 속성 값들이 범주형 데이터일 때 적용 가능하다 - “하나의 레코드에 각 속성들은 서로 독립적이다”라는 가정의 제한이 존재한다.
유클리디안 거리	- 수치값이 아닌 데이터에 적용이 힘들다 - 데이터의 수가 많아질수록 상당히 많은 양의 계산과정이 필요하며 처리 속도의 저하를 가져오기도 한다.

자료 : 손창환(2006), “web 상에서 개인화된 아이템추천을 위한 Hybrid 추천시스템에 관한 연구”에서 인용

다. 고객선호예측

사용자 기반 협업 필터링 기법은 이웃 사용자 간의 유사도를 바탕으로 특정 아이템에 대한 사용자의 선호도를 예측한다. 그러나 본 연구에서 새롭게 제시하는 기법에서는 이웃집단을 형성 한 후 아이템 기반 협업 필터링을 적용하여 아이템 간의 유사도를 바탕으로 이웃 아이템 군을 형성하기 때문에 고객 선호 예측 과정은 아이템 기반 협업 필터링이 적용된 후 실행된다.

고객선호 예측의 척도로는 Weighted Average, Bias-From-Mean Average가 주로 이용된다. Weighted Average는 이웃 사용자와의 유사도를 Weight로 하여 추천대상 사용자의 아이템에 대한 평점의 평균을 내는 방법이다. 계산 방법이 간단하고 기대할 만큼의 성능을 내기 때문에 가장 많이 사용된다. 그러나 각 사용자의 평점의 평균을 반영하지 않는다는 단점을 갖고 있다. Bias-From-Mean Average는 사용자마다의 평점 평균을 고려하여 평균값을 낸다. 이는 Weighted Average 보다 좋은 성능을 보이는 것으로 알려져 있다. 다음은 각각 Weighted Average, Bias-From-Mean Average의 정의이다

▪ Weighted Average

$$P_{ik} = \frac{\sum_{j=1}^n r_{ij} \times S_{jk}}{\sum_{j=1}^n r_{ij}}$$

i : 사용자, j : 이웃 사용자, n : 이웃 사용자의 수

P_{ik} : 아이템 k 에 대한 사용자 i 의 예상 평점

r_{ij} : 사용자 i 와 사용자 j 의 유사도

S_{jk} : 아이템 k 에 대한 사용자 j 의 평점

▪ Bias-From-Mean Average

$$P_{ik} = \bar{S}_i + \frac{\sum_{j=1}^n r_{ij} \times (S_{jk} - \bar{S}_j)}{\sum_{j=1}^n r_{ij}} \quad (3)$$

i : 사용자, j : 이웃 사용자, n : 이웃 사용자의 수

P_{ik} : 아이템 k 에 대한 사용자 i 의 예상 평점

r_{ij} : 사용자 i 와 사용자 j 의 유사도

S_{jk} : 아이템 k 에 대한 사용자 j 의 평점

$\overline{s}_i$: 사용자 i 의 전체 평점 평균

$\overline{s}_j$: 사용자 j 의 전체 평점 평균

고객 i 의 아이템 k 에 대한 선호도 점수를 예측하기 위해서 아이템 k 를 평가한 고객 j 의 평가점수와 이들 고객들과 고객 i 의 상관계수를 이용하여 계산한다. 이때, 아이템 k 를 평가한 모든 고객들에 대해 계산한다.

상기 내용을 요약하면 과거의 구매 기록이나 고객이 표현한 선호도를 식 (1), 즉 피어슨 상관계수를 적용하여 이웃 집단을 추출하여 형성하고, 그 다음으로 이웃 집단의 과거 구매 기록을 토대로 식 (2) 혹은 (3), 즉 Weighted Average, Bias-From-Mean Average, Bias-From-Mean Average를 이용하여 특정 고객이 선호할 만한 K 개의 제품을 예측한다. 이와 같이 특정고객에게 예측한 아이템을 추천하는 것이 사용자 기반 협업 필터링이다.

본 연구에서 새롭게 제시하는 기법은 사용자의 상황에 대한 인식 이후에 이웃사용자 집단을 형성하고 이를 토대로 아이템 기반 협업 필터링을 적용하기 때문에 고객선호예측과정은 아이템 기반 협업 필터링이 적용된 후 실행된다.

라. 사용자 기반 협업 필터링의 이웃집단 형성 예

사용자 기반 협업 필터링을 기반으로 이웃집단 형성의 예를 들기 위해 고객들의 영화 평가 내역을 가정한 후에 그 데이터 셋을 이용하여 사용자 기반 협업 필터링 알고리즘을 적용해보도록 한다.

▪ 평점 수집 및 선호 Matrix Table

다음은 가상의 과거 평점 데이터를 바탕으로 선호 Matrix Table을 형성한 것이다.

[표 3-2] 선호 Matrix Table

	A	B	C	D	E
심야의 FM	3	3	5	1	0
검우강호	1	1	0	3	1
무적자	5	3	2	1	5
평점평균	3	2.33	2.33	1.666	2

▪ 유사도 산출을 통한 이웃집단 형성

만일 A 고객과 B 고객의 유사도를 파악하고자 한다면 A 및 B의 평점 기록을 다음과 같이 식 (1), 피어슨상관계수에 대입하여 유사도를 구하도록 한다.

$$\begin{aligned}
 r_{ij} &= \frac{\sum_k (S_{ik} - \bar{S}_i) \times (S_{jk} - \bar{S}_j)}{\sqrt{\sum_k (S_{ik} - \bar{S}_i)^2 \times \sum_k (S_{jk} - \bar{S}_j)^2}}, -1 \leq r_{ij} \leq 1 \quad (1) \\
 &= \frac{\sum_k (S_{A, \text{심야의 FM}} - \bar{S}_{A\text{평점평균}}) \times (S_{B, \text{검우강호}} - \bar{S}_{B\text{평점평균}}) + (\dots)}{\sqrt{\left[\sum_k (S_{A, \text{심야의 FM}} - \bar{S}_{A\text{평점평균}})^2 + (\dots)^2 \right] \times \left[\sum_k (S_{B, \text{검우강호}} - \bar{S}_{B\text{평점평균}})^2 + (\dots)^2 \right] + [(\dots)]}} \\
 &= \frac{0 + 2.66 + 1.33}{\sqrt{(0 + 4 + 4) \times (0.44 + 1.77 + 2.66)}} = 0.866
 \end{aligned}$$

이 때 r_{ij} 는 고객 i 와 j 의 상관관계를 나타내는데 [표 3-5의] 선호 데이터 테이블을 식 (1), 즉 피어슨 상관계수에 대입하여 산출한 각 고객의 유사도는 다음과 같다.

[표 3-3] 각 고객의 유사도 산출 결과

	A	B	C	D	E
A	$r_{A,A}=1$	-	-	-	-
B	$r_{B,A}=0.866$	$r_{B,B}=1$	-	-	-
C	$r_{C,A}=0.397$	$r_{C,B}=0.802$	$r_{C,C}=1$	-	-
D	$r_{D,A}=-0.866$	$r_{D,B}=-1$	$r_{D,C}=-0.803$	$r_{D,D}=1$	-
E	$r_{E,A}=0.755$	$r_{E,B}=0.327$	$r_{E,C}=-0.300$	$r_{E,D}=-0.327$	$r_{E,E}=1$

피어슨 상관계수를 바탕으로 산출된 각 고객의 유사도인 r_{ij} 는 두 고객의 선호도가 유사한 경우에는 1에 가까운 값을 가지게 되고, 상반된 선호도를 갖는 경우 -1에 가까운 값을 가지게 된다.

예를 들어 "A" 고객에 대해 사용자 기반 협업 필터링의 적용으로 판정된 유사한 이웃집단은 "B" 고객(0.866)과 "E" 고객(0.755)으로 볼 수 있으며, "C" 고객(0.397) 역시 이웃집단이긴 하지만 유사도가 높지 않은 것을 확인할 수 있다. "D" 고객(-0.866)의 경우 "A" 고객과 상반된 선호도를 가진 것으로 나타났다.

▪ 고객선호 예측

[표 3-4]에서 사용자 "E"는 "심야의 FM"에 대한 선호도는 "0"이다. 이를 사용자 "E"가 "심야의 FM"에 대해 경험이 없는 상태라 가정하면 이웃집단의 선호정보를 통해 사용자 "E"의 "심야의 FM"에 대한 선호도를 예측할 수 있다.

먼저 [표 3-5]에서 산출된 각 고객의 유사도를 바탕으로 사용자 "E"와 이웃집단군을 정리하면 다음과 같다. 마지막 행의 유사도는 사용자 E와 각 사용자들 간의 유사도를 나타낸다.

[표 3-4] 사용자 A와 이웃집단의 선호 Matrix Table

	E	A	B
심야의 FM	0	3	3
검우강호	1	1	1
무적자	5	5	3
평점평균	2	2.33	2.33
유사도	-	0.755	0.327

다음으로 사용자 “E”와 이웃사용자들의 유사도와 각 이웃 사용자들의 아이
템에 대한 선호를 식(2), 즉 Weighted Average에 대입하면 다음과 같다.

$$P_{ik} = \frac{\sum_{j=1}^n r_{ij} \times S_{jk}}{\sum_{j=1}^n r_{ij}} \quad (2)$$

$$P_{ik} = \frac{\sum_{j=1}^n (r_{i \text{와 이웃사용자 } j \text{의 유사도}}) \times (S_{\text{상품 } k \text{에 대한 사용자 } j \text{의 선호도}}) + (\dots)}{\sum_{j=1}^n r_{i \text{와 이웃사용자 } j \text{의 유사도}} + (\dots)}$$

$$P_{ik} = \frac{(0.755 \times 3) + (0.327 \times 3)}{(0.755 + 0.327)} = 3$$

Weighted Average에 따라 사용자 “E”에 대한 ”심야의 FM” 선호도 예측은
“3”으로 산출되었다.

한편 사용자 “E”와 이웃사용자들의 유사도와 각 이웃 사용자들의 아이
템에 대한 선호를 식(3), 즉 Bias-From-Mean Average에 대입하면 다음과 같다.

$$P_{ik} = \bar{S}_i + \frac{\sum_{j=1}^n r_{ij} \times (S_{jk} - \bar{S}_j)}{\sum_{j=1}^n r_{ij}} \quad (3)$$

$$P_{ik} = 2 + \frac{(0.755 \times (3 - 3)) + (0.327 \times (3 - 2.33))}{(0.755 + 0.327)} = 2.20$$

Bias-From-Mean Average는 사용자마다 평균 평점(선호도)을 고려하여 평균값을 내기 때문에 Weighted Average보다 좋은 성능을 보인다고 알려져 있다. 위의 예에서 확인할 수 있듯이 각각 Bias-From-Mean Average, Weighted Average로 산출된 값은 차이가 존재한다. 그러나 아이টে을 예측하는 과정에서 아이টে에 대해 산출되는 값이 틀려질 뿐이지 산출되는 값을 비교하여 추천아이টে에 대한 순위를 결정할 때는 각 기법에 대한 결과가 거의 유사한 성능을 보인다고 판단되었다. 따라서 본 연구에서는 계산 방법이 간단하고 상당한 수준의 성능을 보이는 Weighted Average를 사용하도록 한다.

3. 아이টে 기반 협업 필터링

가. 아이টে 기반 협업 필터링

고객이 기존에 평가한 각각의 아이টে들과 특정 고객의 선호도를 예측하고자 하는 아이টে이 얼마나 비슷한가를 계산하여 K개의 가장 비슷한 아이টে를 선택하는 것이 아이টে 기반 협업 필터링이다.[박지선 외, 2002]

아이টে 기반 협업 필터링은 앞서 설명된 사용자 간의 유사도가 아닌 아이টে 간의 유사도를 바탕으로 한다.

먼저 아이টে i 와 j 에 대해서 과거 기록을 토대로 두 아이টে 모두를 평가한 사용자들을 찾아 집합 U 라 놓을 수 있다. 집합 U 에 속한 각 고객들이 아이টে i 와 j 에 입력한 평가를 기반으로 한 i, j 아이টে간의 유사도는 사용자기반 협업필터링과 마찬가지로 피어슨 상관계수에 의해 계산될 수 있다.

- 피어슨 상관계수(Pearson Correlation Coefficient)

$$s_{ij} = \frac{\sum_u (R_{ui} - \bar{R}_i) \times (R_{uj} - \bar{R}_j)}{\sqrt{\sum_u (R_{ui} - \bar{R}_i)^2 \times \sum_u (R_{uj} - \bar{R}_j)^2}}, -1 \leq s_{ij} \leq 1 \quad (1)$$

s_{ij} : 아이템 i , 이웃 아이템 j 간의 유사도

R_{ui} : 고객 u 가 아이템 i 에 대해 평가한 선호도

$\bar{R}_i$: 아이템 i 에 대해 각 사용자가 평가한 선호도 평균

R_{uj} : 고객 u 가 아이템 j 에 대해 평가한 선호도

$\bar{R}_j$: 아이템 j 에 대해 각 사용자가 평가한 선호도 평균

아이템들 간의 상관관계 분석으로 이웃 아이템 집단이 형성되면 사용자 기반 협업 필터링과 마찬가지로 고객선호예측을 위해 Bias-From-Mean Average, 혹은 Weighted Average 등을 이용하여 특정 아이템과 유사한 아이템에 대해 고객이 평가한 선호도를 바탕으로 그 고객이 특정 아이템에 대해 가지는 선호도를 예측할 수 있다.

나. 아이템 기반 협업 필터링의 이웃 아이템군 형성 예

아이템 기반 협업 필터링을 기반으로 이웃 아이템군 형성의 예를 들기 위해 고객들의 영화 평가 내역을 가정한 후에 그 데이터 셋을 이용하여 아이템 기반 협업 필터링 알고리즘을 적용해보도록 한다.

▪ 평점 수집 및 선호 데이터 테이블 형성

다음은 가상의 과거 평가 데이터를 바탕으로 선호 Matrix Table을 형성한 것이다.

[표 3-5] 선호 Matrix Table

	심야의FM	검우강호	무적자	슈퍼베드	적인걸
A	3	1	7	2	0
B	5	0	6	1	2
C	0	1	7	6	1
평점평균	2.666	0.666	6.666	3	1

- 유사도 산출을 통한 이웃 아이템군 형성

만일 “심야의 FM”과 “적인걸”의 유사도를 파악하고자 한다면 “심야의 FM” 및 “적인걸”의 평점 기록을 다음과 같이 식 (1)에 대입하여 유사도를 구하도록 한다.

$$\begin{aligned}
 s_{ij} &= \frac{\sum_u (R_{ui} - \bar{R}_i) \times (R_{uj} - \bar{R}_j)}{\sqrt{\sum_u (R_{ui} - \bar{R}_i)^2 \times \sum_u (R_{uj} - \bar{R}_j)^2}}, -1 \leq s_{ij} \leq 1 \quad (1) \\
 &= \frac{\sum_k (S_{A, \text{심야의 FM}} - \bar{S}_{\text{심야의 FM 평점 평균}}) \times (S_{A, \text{적인걸}} - \bar{S}_{\text{적인걸 평점 평균}}) + (\dots)}{\sqrt{\left[\sum_k (S_{A, \text{심야의 FM}} - \bar{S}_{\text{심야의 FM 평점 평균}})^2 + (\dots)^2 \right] \times \left[\sum_k (S_{A, \text{적인걸}} - \bar{S}_{\text{적인걸 평점 평균}})^2 + (\dots)^2 \right] + [(\dots)]}} \\
 &= \frac{-0.334 + 2.66}{\sqrt{(1 + 1 + 0) \times (0.111 + 5.444 + 7.111)}} = 0.397
 \end{aligned}$$

이 때 s_{ij} 는 고객 i 와 j 의 상관관계를 나타내는데 [표 3-6의] 선호 데이터 테이블을 식 (1), 즉 피어슨 상관계수에 대입하여 산출한 각 아이템 간의 유사도는 다음과 같다.

[표 3-6] 각 아이템 간의 유사도 산출 결과

	심야의FM	검우강호	무적자	슈퍼베드	적인걸
심야의FM	$s_{\text{심,심}} = 1$	-	-	-	-
검우강호	$s_{\text{검,심}} = -0.802$	$s_{\text{검,검}} = 1$	-	-	-
무적자	$s_{\text{무,심}} = -0.802$	$s_{\text{무,검}} = 1$	$s_{\text{무,무}} = 1$	-	-
슈퍼베드	$s_{\text{슈,심}} = -0.976$	$s_{\text{슈,검}} = 0.654$	$s_{\text{슈,무}} = 0.654$	$s_{\text{슈,슈}} = 1$	-
적인걸	$s_{\text{적,심}} = 0.397$	$s_{\text{적,검}} = -0.866$	$s_{\text{적,무}} = -0.866$	$s_{\text{적,슈}} = -0.188$	$s_{\text{적,적}} = 1$

아이템간의 유사도를 파악한 결과에서 예를 들어 “심야의 FM”과 유사하다고 판정되는 아이템은 “적인걸”(0.397)이며, “검우강호”(-0.802), “무적자”(-0.802), “슈퍼베드”(-0.976)의 경우 상반된 유사도를 가진 아이템으로 분류되었음을 확인 할 수 있다.

이웃 아이템 집단을 형성한 후 특정 제품에 대한 고객의 선호를 예측하기 위해선 사용자 기반 협업 필터링과 동일하게 Weighted Average 혹은 Bias-From-Mean Average를 이용할 수 있다.

본 절에서는 본 연구에서 제시하는 상황정보를 이용한 하이브리드 필터링을 구성하는 세부요소인 상황정보 추출, 사용자 기반 협업 필터링, 아이템 기반 협업 필터링에 대해 자세히 살펴보았다. 사용자의 목적행위 및 위치 상황정보를 이용하여 사용자에게 추천할 정보의 범위가 결정되며 사용자 기반 협업 필터링의 과정으로 이웃 사용자 집단을 형성 한 후 아이템 기반 협업 필터링으로 아이템간의 상관관계 분석 및 특정 상품에 대한 선호도 예측을 수행한다.

다음 절에서는 본 논문에서 제시하는 기법을 기반으로 사용자에게 추천 서비스를 제공하는 과정을 시나리오로서 상세히 설명하도록 한다.

제 2 절 추천 서비스 시나리오 예

1. 상황추출 단계

본 절에서는 가상의 시나리오를 토대로 본 연구에서 제시하는 기법을 통해 정보의 추천이 이뤄지는 일련의 과정을 설명하고자 한다. 이는 크게 상황추출 단계, 사용자 기반 협업 필터링 단계, 아이템 기반 협업 필터링 단계로 구성된다. 이는 모바일 환경에서 구동되며, 모바일 환경의 특성으로 활용가능 한 사용자의 목적, 위치 등 상황정보 추출을 통해 제공 정보의 범위를 결정하는 것으로 시작한다. 다음 단계로는 각 사용자 기반 협업 필터링 및 아이템 기반 협업 필터링을 적용하여 사용자에게 추천하고자 하는 아이템에 대한 예측 선호를 산출하고 추천목록을 제공함으로써 추천과정이 종료된다.

[표 3-8]은 본 연구에서 제시하는 기법을 기초로 하여 추천서비스를 제공하기 위한 시나리오를 정리한 것이다.

[표 3-7] 추천대상자 A와 이웃집단의 선호 Matrix Table

사용자	상황정보: 사용자 목적	상황정보: 현재 위치 혹은 목적지	제공정보
A	쇼핑	현재 “가” 지하철역	- 사용자 “A” 가 선호할 것이라 예상되는 쇼핑물 - 해당 쇼핑물로 이동하기위한 교통 및 경로 정보

추천대상자 “A”는 모바일 환경에서 제공되는 추천시스템과의 상호작용을 통해 자신의 목적 또는 요구사항을 시스템에게 제공한다. 예를 들어 쇼핑, 여가, 식사, 병원 등의 다양한 목적 및 행위가 존재 할 수 있으며 본 시나리오의 경우, 추천대상자 “A”는 쇼핑을 선택하였다고 가정한다. 다음 단계로 사용자는 현재 위치 혹은 목적지에 대한 위치정보를 시스템에 제공함으로써 시스템이 위치기반에 따른 수행을 가능하도록 한다.

상기 두 과정을 통해 시스템은 추천대상자 “A”의 목적, 위치에 대한 상황정보 추출을 통해 제공해야 할 추천정보의 범위를 필터링 할 수 있다. 시나리오

의 측면에서 살펴보면 시스템이 저장하고 있는 수천, 수만의 POI(Point of Interest)의 정보가 사용자가 현재 위치한 "가" 지하철역에서의 일정 기준범위 안의 쇼핑물 정보만으로 필터링이 되는 것이다.

앞의 상황정보추출 단계가 종료되면 추천시스템은 다음 단계의 사용자 기반 협업 필터링을 통해 이웃 사용자 집단을 형성한 후 형성된 내용을 토대로 아이템 기반 협업 필터링 과정을 수행하여 추천대상자 "A"가 선호할 것이라 예상되는 일정 수의 쇼핑물 정보를 도출하고 예상선호에 대한 순위를 결정하여 사용자에게 제공할 수 있다. 또한 부가적으로 추천목록에서 사용자가 선택한 쇼핑물, 즉 해당지점으로 이동하기 위한 교통 및 경로 정보를 제공을 함으로써 본 연구에서 제안하는 추천과정이 완료된다.

2. 사용자 기반 협업 필터링 단계

상황정보 추출 단계가 종료되면 사용자 기반 협업 필터링 단계가 시작되며 과거 사용자 프로파일을 통해 사용자 간의 상관관계 분석 및 이웃 사용자 집단이 형성된다. 사용자간 유사도를 산출하기 위해 피어슨상관계수가 이용되며, 본 과정에서 형성된 이웃 사용자 집단은 다음 단계인 아이템 기반 협업 필터링 단계에서 이웃 아이템 군을 형성하기 위한 기초자료로 사용된다.

가. 사용자 평점(Rating) 수집 및 사용자 간 상관관계 분석

상황정보 추출 단계의 프로세스를 통해 사용자 "A"의 목적이 "쇼핑"이며 현재 위치는 "가" 지하철역 부근"이라는 것을 시스템이 인지하였으며 이에 따라 시스템은 시스템에 저장되어 있는 사용자 "A"의 현재위치에서 일정범위 안에 있는 모든 쇼핑물들을 추출한 후 그 다음으로 과거 사용자들의 각 쇼핑물에 대한 평점 혹은 선호도를 바탕으로 사용자간 상관관계를 분석하게 된다. [표 3-9]는 추천대상자 "A"의 현재 위치에 인접한 모든 쇼핑물(1~10)에 대해 사용자들이 과거에 사용 혹은 방문으로 만족도를 0~5로 평가한 선호 Matrix Table을 나타낸다.

[표 3-8] 선호 Matrix Table

	A	B	C	D	E	F	G	H	I	J
쇼핑몰1	3	3	5	1	2	1	4	5	1	5
쇼핑몰2	0	2	0	4	3	1	2	4	5	3
쇼핑몰3	4	5	1	0	4	1	5	5	5	5
쇼핑몰4	5	4	2	2	5	0	3	1	0	3
쇼핑몰5	0	1	5	1	3	2	2	4	3	3
쇼핑몰6	0	0	5	1	3	4	1	1	4	1
쇼핑몰7	2	5	3	3	2	0	1	4	3	5
쇼핑몰8	1	5	5	3	3	4	2	4	1	5
쇼핑몰9	3	2	1	3	4	5	0	4	1	3
쇼핑몰10	3	4	3	1	3	4	4	3	2	5
평균	2.1	3.1	3	1.9	3.2	2.2	2.4	3.5	2.5	3.8

선호 Matrix Table에 따르면 추천대상자 “A”는 쇼핑몰 2, 5, 6에 대한 경험이 없다고 가정할 수 있다. 따라서 본 추천과정은 “A”에게 쇼핑몰 2, 5, 6에 대한 선호도를 예측하는 과정으로 진행된다.

사용자간 유사도를 바탕으로 한 상관관계 분석은 앞의 절에서 다뤄진 내용에 따라 피어슨상관계수를 통해 산출된다. 피어슨 상관계수는 두 사용자의 과거 선호에 따라 산출되며 -1부터 1까지의 값을 가진다. 1에 가까울수록 두 사용자 간의 유사도가 높음을 나타내며 -1에 가까울수록 그 반대이다.

추천대상자 “A”에 대해 모든 사용자들의 유사도를 산출한 결과는 [표 3-10]와 같으며 “A”와 유사성이 높다고 판단된 사용자의 경우 음영표시를 하였다.

[표 3-9] 사용자 A에 대한 각 사용자의 유사도

	B	C	D	E	F	G	H	I	J
A	0.585	-0.351	-0.284	0.526	-0.271	0.495	-0.022	-0.470	0.408

나. 이웃 사용자 집단 형성

사용자간 상관관계 분석을 통해 기존 전체 사용자의 선호 Matrix Table은 [표 3-11]과 같이 추천대상자 “A”의 이웃사용자 집단 선호 Matrix Table로 정리된다.

[표 3-10] 사용자 A의 이웃사용자 집단 선호 Matrix Table

	A	B	E	G	J
쇼핑물1	3	3	2	4	5
쇼핑물2	0	2	3	2	3
쇼핑물3	4	5	4	5	5
쇼핑물4	5	4	5	3	3
쇼핑물5	0	1	3	2	3
쇼핑물6	0	0	3	1	1
쇼핑물7	2	5	2	1	5
쇼핑물8	1	5	3	2	5
쇼핑물9	3	2	4	0	3
쇼핑물10	3	4	3	4	5
평균	2.1	3.1	3.2	2.4	3.8
유사도	-	0.585	0.526	0.495	0.408

형성된 이웃사용자 집단의 선호 Matrix Table은 다음 단계인 아이템 기반 협업 필터링 단계에서 이웃 아이템 군을 형성하기 위한 기초자료로 사용된다.

3. 아이템 기반 협업 필터링 단계

사용자 기반 협업 필터링 단계가 종료되면 이웃 사용자 집단이 형성되고 이는 아이템 기반 협업 필터링 단계에서 이웃 아이템 군을 형성하기 위한 자료로 사용된다. 유사한 성향을 가진 사용자들의 데이터를 이용함으로써 고객들 간의 유사도가 고려되지 않는 아이템 기반 협업 필터링의 근본적인 문제점을 해결할 수 있다. 형성된 유사아이템 집단을 통해 사용자에게 추천하고자 하는

아이템의 예상 선호도를 정량적으로 예측할 수 있다.

가. 아이템 간 상관관계 분석

기본적으로 사용자 기반 협업 필터링과 아이템 기반 협업 필터링 기법은 동일한 메커니즘을 가지고 있으며 사용자간의 유사도를 산출하느냐 아이템간의 유사도를 산출하느냐에 따라 달라진다.

추천대상자 “A”의 사용자 기반 협업 필터링을 통해 형성된 이웃 사용자 집단의 선호 Matrix Table을 행과 열의 기준을 바꿔 [표 3-12]로 정리하였다.

[표 3-11] 선호 Matrix Table

	쇼핑물1	쇼핑물2	쇼핑물3	쇼핑물4	쇼핑물5	쇼핑물6	쇼핑물7	쇼핑물8	쇼핑물9	쇼핑물10
A	3	0	4	5	0	0	2	1	3	3
B	3	2	5	4	1	0	5	5	2	4
E	2	3	4	5	3	3	2	3	4	3
G	4	2	5	3	2	1	1	2	0	4
J	5	3	5	3	3	1	5	5	3	5
평균	3.4	2	4.6	4	1.8	1	3	3.2	2.4	3.8

앞서 사용자 기반 협업 필터링 단계에서는 추천대상자 “A”에 대해 다른 모든 사용자들의 유사도를 산출하였다. 그러나 아이템 기반 협업 필터링 단계에서는 추천대상자 ”A“에게 쇼핑물 2, 5, 6에 대한 선호도를 예측하여 추천을 결정하기 위해 각 각 쇼핑물 2, 5, 6에 대해 다른 모든 아이템에 대한 유사도가 산출 되어야 한다. 각 아이템들의 상관관계를 분석하기 위해 각 아이템에 대한 선호도를 피어슨상관계수에 대입하여 모든 아이템 간의 유사도를 산출한다. 그 결과는 [표 3-13]과 같다. 본 연구에서는 유사도가 0.2인 경우 이웃집단이라 분류하였으며 [표 3-13]에서는 음영표시를 하였다. 이웃집단으로 판정하기 위한 적절한 임계값을 산출에 관한 추가 연구가 이루어져야 한다.

[표 3-12] 아이템 간의 상관관계분석결과

	쇼핑몰1	쇼핑몰2	쇼핑몰3	쇼핑몰4	쇼핑몰5	쇼핑몰6	쇼핑몰7	쇼핑몰8	쇼핑몰9	쇼핑몰10
쇼핑몰1	1.000	-	-	-	-	-	-	-	-	-
쇼핑몰2	0.179	1.000	-	-	-	-	-	-	-	-
쇼핑몰3	0.721	0.373	1.000	-	-	-	-	-	-	-
쇼핑몰4	-0.877	-0.408	-0.913	1.000	-	-	-	-	-	-
쇼핑몰5	0.235	0.939	0.210	-0.383	1.000	-	-	-	-	-
쇼핑몰6	-0.358	0.667	-0.373	0.204	0.783	1.000	-	-	-	-
쇼핑몰7	0.352	0.327	0.488	-0.267	0.102	-0.327	1.000	-	-	-
쇼핑몰8	0.319	0.685	0.612	-0.419	0.450	0.000	0.896	1.000	-	-
쇼핑몰9	-0.405	0.135	-0.662	0.659	0.177	0.404	0.264	0.147	1.000	-
쇼핑몰10	0.891	0.488	0.873	-0.896	0.413	-0.129	0.639	0.702	-0.315	1.000

나. 이웃 아이템 집단 형성

추천대상자 “A”의 쇼핑몰 2, 5, 6에 선호도를 예측하기 위해 2, 5, 6 각 아이템에 대해 이웃 아이템 집단 선호 Matrix Table을 형성하였으며 이는 [표 3-14]~[표 3-16]와 같다. 각 테이블에서 맨 마지막 행의 유사도는 기준 쇼핑몰과 이웃 쇼핑몰간의 유사도라고 할 수 있다.

[표 3-13] 쇼핑몰2의 이웃 아이템 집단 선호 Matrix Table

	쇼핑몰2	쇼핑몰3	쇼핑몰5	쇼핑몰6	쇼핑몰7	쇼핑몰8	쇼핑몰10
A	0	4	0	0	2	1	3
B	2	5	1	0	5	5	4
E	3	4	3	3	2	3	3
G	2	5	2	1	1	2	4
J	3	5	3	1	5	5	5
평균	2	4.6	1.8	1	3	3.2	3.8
유사도	-	0.373	0.939	0.667	0.327	0.685	0.488

[표 3-14] 쇼핑물5의 이웃 아이템 집단 선호 Matrix Table

	쇼핑물5	쇼핑물1	쇼핑물2	쇼핑물3	쇼핑물6	쇼핑물8	쇼핑물10
A	0	3	0	4	0	1	3
B	1	3	2	5	0	5	4
E	3	2	3	4	3	3	3
G	2	4	2	5	1	2	4
J	3	5	3	5	1	5	5
평균	1.8	3.4	2	4.6	1	3.2	3.8
유사도	-	0.235	0.939	0.210	0.783	0.450	0.413

[표 3-15] 쇼핑물6의 이웃 아이템 집단 선호 Matrix Table

	쇼핑물6	쇼핑물2	쇼핑물4	쇼핑물5	쇼핑물9
A	0	0	5	0	3
B	0	0	0	0	2
E	3	3	5	3	4
G	1	2	3	2	0
J	1	3	3	3	3
평균	1	1.6	3.2	1.6	2.4
유사도	-	0.667	0.204	0.783	0.404

다. 고객선호 예측 및 추천서비스 제공

이웃 아이템 집단 형성 완료되면 특정 아이템과 유사한 아이템에 대해 사용자가 평가한 선호도의 합을 계산하여 사용자가 특정상품에 대해 가지는 선호도를 예측할 수 있다. 이때 두 상품의 유사도에 따라서 각각의 선호도 값이 가중치된다

다음 식 (2) 즉, Weighted Average를 이용하면 쇼핑물2, 5, 6에 대한 추천 대상자 “A”의 예측 선호도를 산출할 수 있다.

$$P_{ik} = \frac{\sum_{j=1}^n r_{ij} \times S_{jk}}{\sum_{j=1}^n r_{ij}} \quad (2)$$

i : 사용자, j : 이웃 사용자, n : 이웃 사용자의 수

P_{ik} : 아이템 k 에 대한 사용자 i 의 예상 평점

r_{ij} : 사용자 i 와 사용자 j 의 유사도

S_{jk} : 아이템 k 에 대한 사용자 j 의 평점

▪ 쇼핑물 2의 경우

[표 3-16] 쇼핑물2의 이웃 아이템 집단 선호 Matrix Table

	쇼핑물2	쇼핑물3	쇼핑물5	쇼핑물6	쇼핑물7	쇼핑물8	쇼핑물10
A	0	4	0	0	2	1	3
유사도	-	0.373	0.939	0.667	0.327	0.685	0.488

$$P_{ik} = \frac{(0.373 \times 4) + (0.939 \times 0) + (0.667 \times 0) + (0.327 \times 2) + (0.685 \times 1) + (0.488 \times 3)}{(0.373 + 0.939 + 0.667 + 0.327 + 0.685 + 0.488)}$$

$$P_{ik} = 1.234$$

▪ 쇼핑물 5의 경우

[표 3-17] 쇼핑물5의 이웃 아이템 집단 선호 Matrix Table

	쇼핑물5	쇼핑물1	쇼핑물2	쇼핑물3	쇼핑물6	쇼핑물8	쇼핑물10
A	0	3	0	4	0	1	3
유사도	-	0.235	0.939	0.210	0.783	0.450	0.413

$$P_{ik} = \frac{(0.235 \times 3) + (0.939 \times 0) + (0.210 \times 4) + (0.783 \times 0) + (0.450 \times 1) + (0.413 \times 3)}{(0.235 + 0.939 + 0.210 + 0.783 + 0.450 + 0.413)}$$

$$P_{ik} = 1.067$$

▪ 쇼핑물 6의 경우

[표 3-18] 쇼핑물6의 이웃 아이템 집단 선호 Matrix Table

	쇼핑물6	쇼핑물2	쇼핑물4	쇼핑물5	쇼핑물9
A	0	0	5	0	3
유사도	-	0.667	0.204	0.783	0.404

$$P_{ik} = \frac{(0.667 \times 0) + (0.204 \times 5) + (0.783 \times 0) + (0.404 \times 3)}{(0.667 + 0.204 + 0.783 + 0.404)}$$

$$P_{ik} = 1.084$$

Weighted Average를 이용하여 추천대상자 "A"의 쇼핑물 2, 5, 6에 대한 선호예측을 수행한 결과, 각각 쇼핑물 2는 1.234, 쇼핑물 5는 1.067, 쇼핑물 6은 1.084로 산출 되었다. 이에 따라 추천대상자 "A"가 선호할 것이라 예상되는 쇼핑물은 순서대로 쇼핑물 2, 쇼핑물 6, 쇼핑물 5라고 볼 수 있다.

본 절에서는 가상의 시나리오를 토대로 본 연구에서 설계 및 구현하고자 하는 추천시스템의 일부 서비스에 대한 일련의 추천과정을 설명하였다. 사용자에게 제공되는 추천의 질을 높이기 위해 상황정보 추출 단계를 통해 사용자의 요구사항 및 위치 정보를 토대로 정보제공의 범위를 결정하였으며, 사용자 기반 협업 필터링을 이용하여 이웃 사용자 집단을 형성하였다. 끝으로 형성된 내용을 토대로 아이템 기반 협업 필터링을 적용하여 사용자에게 선호될 것이라 예측되는 아이템 목록을 도출 하였다. 사용자 기반 협업 필터링 및 아이템 기반 협업 필터링에서 유사성 측정의 척도로 피어슨 상관계수가 이용되며 Weighted Average를 이용하여 특정 아이템에 대한 사용자의 예측 선호도를 산출하고 사용자에게 추천목록을 제공함으로써 추천의 과정이 종료된다.

제 4 장 추천시스템 설계 및 구현

제 1 절 추천시스템 설계

1. 설계 개요

설계는 이해당사자의 요구나 조직의 비즈니스 목표를 만족시키고, 높은 품질의 S/W를 개발하기 위한 방법이며 청사진이다. 또한 S/W 구현에 있어서 요구사항을 올바르게 개발하기 위한 토대가 된다. 일반적으로 S/W에 요구되는 기능과 성능 조건들을 만족하는 내부기능, 구조 및 동적 행위들을 모델링하여 표현, 분석, 검증하는 과정이라 할 수 있다. 일반적인 S/W 설계의 중요성은 다음과 같다.

- 소프트웨어개발, 품질에 영향력이 큼
- 문제의 뚜렷한 공식이 정의되어 있지 않고, 유한법칙이 존재하지 않음
- 명확한 기준이 없고 설계의 품질을 측정할 척도가 없음
- 설계 산출물에 대한 불명확, 누락, 중복, 성능측면에서의 검토만 진행
- S/W 유지보수에 영향

설계는 모든 소프트웨어 공학 단계들과 유지보수 단계들에 대한 기초역할로써 고객 요구사항을 소프트웨어 프로덕트 혹은 시스템으로 정확하게 변환 할 수 있는 유일한 통로이며 품질 평가에 대한 표현을 제공한다. 차후 설계 품질 평가 시 좋은 설계가 되는 기준으로 소프트웨어 요소들 간의 효과적인 제어를 위해 계층적 구조가 제시되어야 하며 모듈적 즉, 논리적으로 특별한 기능과 부 기능을 수행하는 여러 요소로 분리해야 한다. 또한 데이터와 프로시저에 대한 분명하고 분리된 표현을 포함해야 하며 독립적인 기능적 특성을 지닌 모듈화로 유도되어야 한다.

본 절에서는 논리적 설계의 과정으로 요구사항 분석을 통해 시스템이 보유해야 할 기능적/비기능적 요소들을 정의하고 제공 서비스를 도출한다. 또한 유스케이스 명세화 및 다이어그램을 통해 내부적으로 시스템이 어떠한 흐름으로 작동되는지 설명한다.

2. 요구사항 분석

본 연구에서 제시하는 상황정보를 이용한 하이브리드 필터링 추천 시스템의 설계 시 고려되어야 할 요구사항들을 도출하였다. 일반적으로 요구사항은 고객 요구사항/기능/비기능/성능으로 구분되는데 본 과정에서는 제안기법의 평가 및 분석을 위해 시스템을 설계 하므로 고객요구사항은 제외한다.

[표 4-1] 요구사항명세 테이블

순번	요구사항	구분
R01	사용자에게 선호될 것이라 예측되는 POI정보를 추천 및 설명	기능
R02	선호 아이템, 혹은 목적지에 대한 교통수단, 이동경로 등을 안내	기능
R03	제공되는 서비스는 기본적으로 지도화면을 바탕으로 구성.	기능
R04	시스템은 서버-클라이언트 모델로 구성	기능
R05	사용자 별 아이템에 대한 프로파일 구성되며 이는 서버에 저장	기능
R06	사용자의 목적행위를 인식(예: 식당, 병원, 영화, 쇼핑 등)	기능
R07	사용자의 목적지 혹은 현 위치 인식	기능
R08	사용자는 POI정보를 경험 후 평점 정보를 입력	기능
R09	상황 기반, 사용자 기반 협업 필터링, 아이템 기반 협업 필터링을 통해 정보 필터링	기능
R10	필터링 된 정보에 대한 선호도를 예측 및 추천	기능
R11	사용자 친화적 디자인 필요	비기능/디자인
R12	메뉴의 깊이는 3이하여야 함, 최대 3번의 클릭이면 원하는 정보를 얻을 수 있어야 함	비기능/메뉴
R13	모바일 환경에서 동작하며 네트워크 연결망으로 3G/Wifi 망 사용	비기능/환경
R14	사용자 정보요청에 적정시간에 정보를 제공	성능

위 표에서 나타난 요구사항들은 추천시스템이 갖추어야 할 기능/비기능/성능적인 요소들을 나타낸다. 위의 모든 요구사항들은 소프트웨어적으로 설계되

어 시스템에 이식되어야 한다. 요구사항 유형 항목에서 나타나는 기능은 구현되어야 할 필수적인 작업과 동작 등을 정의함으로써 어떤 기능이 구현되어야 하는지를 설명 해당되고, 비기능은 기능 유형이 충족됨으로써 나타날 수 있는 기대 효과나 해결 가능한 문제점들에 해당된다. 마지막으로 성능은 기능이 동작해야 하는 한계를 정의한다. 이는 보통 자료의 양이나 질, 동작의 적시성과 민첩성 등의 척도로 기술된다.

요구사항 분석 단계는 소프트웨어 설계 과정 중 가장 먼저 수행되어야 하는 작업이고 차후 개념/논리 설계, 구현에 기본적인 밑바탕이 된다. 각 요구사항들은 하나 이상의 기능 및 서비스들과 연관관계를 갖으며 요구사항이 만족되기 위한 일련의 시스템 절차들이 기능들로 구현된다. 그러므로 요구사항 분석이 정확하게 이루어지지 않은 경우, 전체적인 시스템 구축 방향이 어긋날 우려가 있다.

3. 서비스 도출

앞서 분석한 요구사항에 따라 각 요구사항을 만족시킬 수 있는 서비스를 도출한다. 기능/비기능적인 측면과 요구사항의 특성에 따라 하나의 요구사항이 실현되기 위해 둘 이상의 서비스가 도출되거나 두 개 이상의 요구사항이 하나의 서비스로 도출 될 수 있다. 도출된 서비스는 개념적인 수준에서 정보 제공 측면에 대한 단어들로 이루어지며 이어질 설계 과정에서 시스템 기본 기능을 선정하기 위한 밑바탕이 된다.

오늘날 모바일 환경에서 제공되는 대다수의 어플리케이션은 사용자의 시스템에 대한 접근성을 위해 가벼우며 핵심적인 소수의 서비스만 제공하고 있다. 또한 본 연구에서의 추천시스템 설계 및 구현은 새롭게 제시된 추천기법에 대한 효용성 및 성능의 평가를 수행하기 위한 목적을 가지고 있으므로 본 단계에서는 이러한 연구 범위에 해당하는 핵심서비스만을 도출한다.

앞의 단락에서 분석된 요구사항 및 본 연구에서 검증하고자 하는 기본적인 시스템의 기능요소 분석에 따라 [표 4-2]와 같은 서비스가 도출되었다. 각 서비스에서 충족될 수 있는 요구사항과 서비스에 대한 간략한 설명을 포함한다.

[표 4-2] 추천시스템 제공 서비스

순번	서비스명	상세설명
S01	협업 필터링 아이템추천	- 추천대상자의 목적행위/현재 위치에 대한 상황인식 후 사용자 협업 기반 필터링을 통해 사용자에게 선호될 것이라 예측되는 아이템을 선별하여 예측 선호도(평점)를 산출하여 제공한다.
S02	하이브리드 필터링 아이템추천	- 추천대상자의 목적행위/현재 위치에 대한 상황인식 후 사용자 협업 기반 필터링 및 아이템 기반 협업 필터링을 결합한 하이브리드 필터링을 통해 사용자에게 선호될 것이라 예측되는 아이템을 선별하여 예측 선호도(평점)를 산출하여 제공한다.
S03	베스트 선호 아이템 추천	- 사용자들의 과거 선호도 기록을 토대로 평균 선호도(평점)를 산출하고 일정 개수의 아이템을 선별하여 추천한다.
S04	내과거 베스트 아이템 추천	- 추천대상자의 과거 선호도 기록을 토대로 가장 높은 선호도(평점)를 표시했던 일정 개수의 아이템을 선별하여 추천한다.
S05	선호도 측정	- 사용자는 추천시스템이 추천해준 아이템 혹은 자신이 경험을 한 아이템에 대해 선호도(평점)를 입력한다.
S06	전체 아이템 보기	- 추천대상자의 목적행위/현재 위치에 대한 상황인식 후 해당 하는 전체 아이템을 보여준다.
S07	목적행위 입력	- 사용자의 목적행위 상황 정보를 입력한다.
S08	목적지(현위치)입력	- 사용자의 목적지 혹은 현위치 상황정보를 입력한다.

4. 추천 시스템 논리적 설계

가. 유스케이스 다이어그램 명세

추천시스템의 논리적 설계의 과정으로 앞서 도출된 시스템의 요구사항 및 서비스로부터 유스케이스를 추출하고 정의한다. 유스케이스는 어떤 시스템을 만드느냐를 사용자 입장에서 조망하는 것이라 할 수 있다. 이는 시스템보다 그것을 사용하는 인간의 입장을 우선해서 시스템이 어떠해야 하는가를 알아보는 것이다. 즉 시스템이 제공하는 각 기능을 의미하며, 유스케이스를 정의한다는 것은 시스템의 모든 기능을 찾는 것이고 또한 시스템의 모든 요구사항을

만족시키는 것이 된다. 또한 구체적으로 시스템의 기능을 정의하고, 범위를 결정함으로써 시스템과 외부 환경 변수를 구분하고, 상호관계를 정립하는 것이라 볼 수 있다.

유스케이스 다이어그램을 도출하기 위해 먼저 액터를 추출하고 유스케이스를 정의해야 한다. 액터는 시스템과 상호작용하는 모든 것들을 나타내는 것으로 시스템을 사용하게 될 사람은 물론이고 경우에 따라 외부에서 현재의 시스템과 연계되어 시스템이 요구하는 자원을 제공하거나 제공 받는 외부의 다른 시스템을 의미하기도 한다. 본 추천시스템에서 액터는 시스템을 이용하여 추천서비스를 제공받길 원하는 사람, 즉 시스템 사용자와 사용자의 요청에 대한 처리를 통해 추천정보를 제공하는 센터(서버)로 정의 될 수 있다. 추출된 액터들의 목록 및 세부 내용은 다음과 같다.

[표 4-3] 상황정보를 이용한 하이브리드 추천시스템 액터

구분	유형	내용
시스템 사용자	사용자	- 추천서비스를 제공받길 원하는 일반사람으로 모바일 단말기와 상호작용을 통해 추천시스템에 접근 한다. - 단말기(모바일 기기)를 통해 과거 경험한 아이টে에 선호(평점)정보를 입력할 수 있다. - 위치 및 목적 정보 입력으로 원하는 목적에 대해 자신이 선호할 것이라 예측되는 아이টে를 추천받길 원한다.
센터(서버)	외부시스템	- 모바일 단말기를 통해 사용자 정보를 전송 받고 데이터베이스에 저장하거나, 사용자의 추천요청이 있을 시 데이터베이스의 데이터를 통해 추천정보를 생성한다. - 생성된 정보를 모바일 단말기에 전송한다

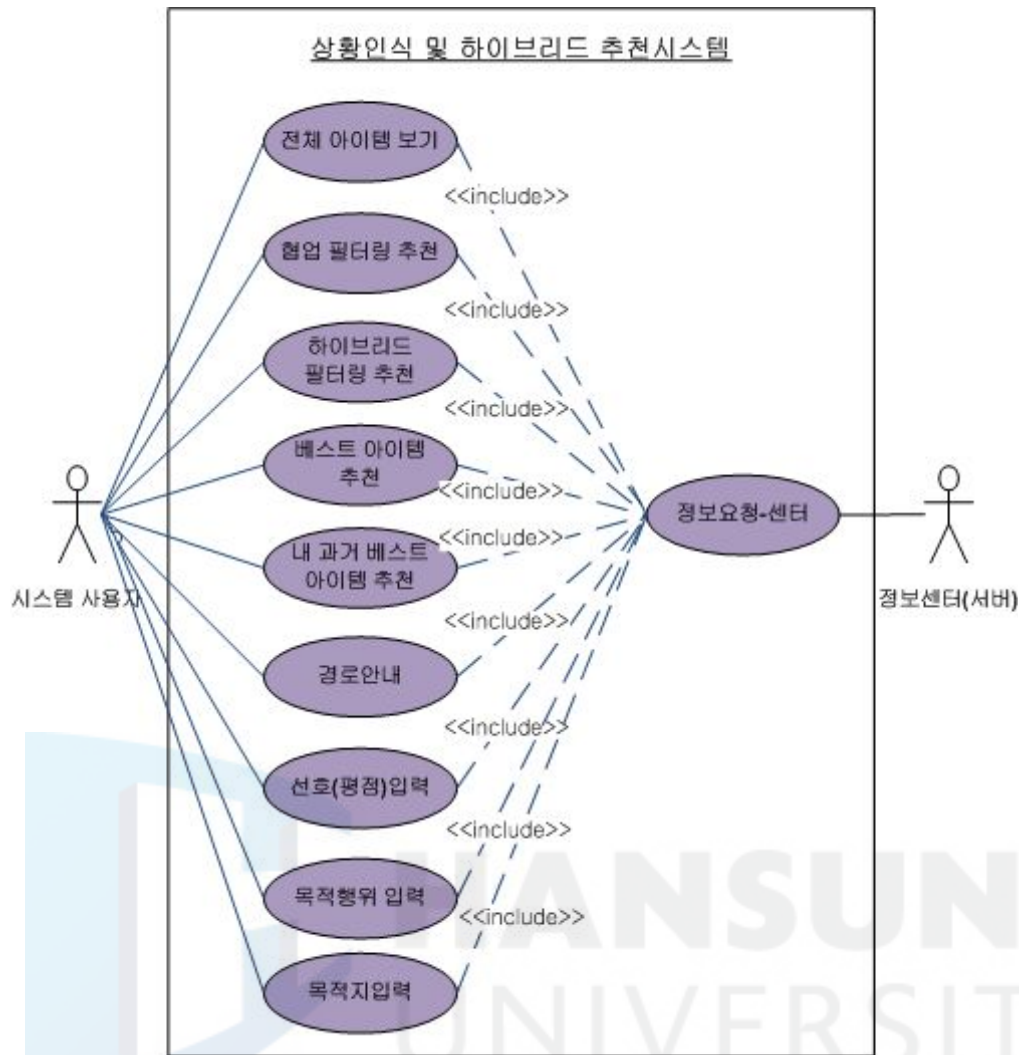
유스케이스 추출은 시스템 기능에 기반하여 이루어진다. 요구사항 분석 과정에서 도출된 시스템 서비스를 참고하여 상황인식 및 하이브리드 추천시스템이 제공해야 할 시스템 기능 즉, 유스케이스를 선정한다. 선정된 유스케이스는 총 8개이며 이에 대한 각 세부내용은 아래 표와 같다.

[표 4-4] 추천시스템 유스케이스

순번	이름	관련액터	세부내용
U01	전체 ITEM 검색	시스템 사용자	- 목적행위/현재지점 혹은 목적지 입력받아 센터(서버)로 전송 - 해당 행위 목적지 대한 모든 가용 아이템 필터링 후 정보 제공
U02	협업필터링 추천	시스템 사용자	- 해당 행위/목적지 대한 아이템에 대해 협업필터링 적용 후 아이템 추천
U03	하이브리드 필터링 추천	시스템 사용자	-해당 행위/목적지 대한 아이템에 대해 하이브리드 필터링 적용 후 아이템 추천
U04	베스트 ITEM 추천	시스템 사용자	-해당 행위/목적지에 대한 아이템 중 평균선호(평점)가 높은 아이템 추천
U05	내 과거 베스트 ITEM 추천	시스템 사용자	-해당 행위/목적지에 대한 아이템 중 추천대상자의 과거정보를 통해 가장 선호(평점)이 높은 아이템 추천
U06	경로안내	시스템 사용자	-사용자가 선택한 아이템에 대한 경로 안내 제공
U07	선호(평점)입력	시스템 사용자	-사용자가 선택한 아이템에 대한 선호(평점) 입력
U08	정보요청-센터	센터(서버)	-사용자 요청에 대한 정보 제공을 위해 필요한 정보들을 요청, 빠른 응답시간이 요구되며 모바일 통신 환경을 위해 무선 네트워크 사용

위 표에서 나타난 관련 액터들은 해당 유스케이스의 기능을 활성화 (Activation) 시킬 수 있는 사용자 또는 외부 시스템을 말한다. 각 유스케이스 수행 과정에서는 관련액터와 추천시스템간의 쌍방향 정보교환이 일어나며 유스케이스 이름은 정보교환의 목적을 나타낸다. 관련액터 중 시스템사용자는 직접적인 추천시스템 조작을 통한 서비스 이용이 가능하다는 것을 의미한다.

지금까지 추출된 액터, 유스케이스들을 토대로 전체적인 시스템 유스케이스 다이어그램을 작성하면 아래 그림과 같다.



<그림 4-1> 상황정보를 이용한 하이브리드 추천시스템 유스케이스 다이어그램

추천시스템 경계 내부로 8개의 유스케이스와 각각에 대한 관련 액터와 연관 관계를 정의하였다. 앞서 언급했듯이 액터와 유스케이스의 연결 의미는 해당 기능의 시작 및 기능 수행 결과의 통보이다. 연결된 모든 액터-유스케이스의 관계들은 해당 액터가 연결된 유스케이스의 기능을 이용할 수 있음을 뜻한다.

유스케이스간의 포함, 확장 관계로 시스템을 구성하는 유스케이스의 수를 최소화하였다. 포함 관계는 시스템 제어 흐름이 항상 넘어감을 의미하고 확장 관계는 선택적인 제어권 이동을 뜻하는데 현재 추천시스템에서 제공되는 서비

스들은 서버에 저장되어 있는 이용자들의 과거선호(평점)정보 데이터를 이용하므로 반드시 정보센터(서버)와의 연계를 통한 정보 제공이 이루어져야 한다. 따라서 본 다이어그램에 기록된 시스템 사용자액터 관련 유스케이스들은 모두 정보요청-센터 유스케이스를 포함하도록 구성하였다.

유스케이스 다이어그램은 시스템의 기능과 그 기능을 사용할 수 있는 이용자와의 관계뿐 아니라 해당 기능을 수행하기 위한 시스템 수행 절차가 내포되어 있다. 예를 들어, 시스템 사용자가 전체아이템 안내를 받기 위해 전체아이템 보기 서비스를 시작하고 서비스에 필요한 목적행위, 목적지 혹은 현위치를 입력하면 이는 정보요청-센터(서버) 유스케이스로 전달된다. 정보센터(서버)로부터 전달된 해당 요청에 대한 결과 정보는 정보요청-센터 유스케이스에서 전체아이템보기 유스케이스로 전달되어 최종적으로 시스템 사용자에게 모바일화면을 통해 표출되게 된다.

나. 유스케이스 명세서 작성

앞서 제시된 추천시스템의 각 유스케이스는 개념적인 레벨에서 시스템의 기능을 제시한 것이기 때문에 각 기능에 대한 구체적인 수행 과정 정의가 필요하다. 각 유스케이스에 대한 시스템적인 실행 과정 정의는 유스케이스 명세서(Use Case Specification)라는 문서를 통해 제시할 수 있다.

유스케이스의 내부 수행과정은 액터와 설계 시스템이 발생시키는 이벤트의 상호작용 관계로 이루어진다. 즉, 액터가 시스템에 입력을 주는 이벤트를 발생시키면 시스템은 이에 대해서 처리를 한 후에 그 결과를 다시 액터에게 표시하는 이벤트를 발생시킨다. 이렇듯 유스케이스 명세서는 액터와 시스템간의 이벤트 흐름으로 작성된다.

본 추천시스템의 설계를 위해 도출된 유스케이스를 통해 제공되는 정보 차별화되어 있지만, 제공정보를 위한 시스템의 일련의 수행과정은 거의 유사하다고 볼 수 있다. 이에 따라 유스케이스 명세서의 작성은 시스템의 핵심 기능이 하리브리드 필터링 추천 유스케이스에 대해서만 수행하였으며 작성된 결과는 다음과 같다.

[표 4-5] 하이브리드 필터링 추천 유스케이스 명세

유스케이스명	하이브리드 필터링 추천 유스케이스
개요	- 추천대상자의 목적행위/목적지 대한 전체 아이템에 대해 하이브리드 필터링 적용 후 아이템 추천
관련 액터	시스템 사용자
우선 순위	상
선행 조건	- 추천시스템(개인단말기)를 통해 추천 대상자의 목적행위/목적지 등 필요한 정보들이 입력된 상태이어야 한다.
이벤트 흐름	기본흐름 1. 추천시스템(개인단말기) 정보요청 - 센터유스케이스를 이용하여 정보센터(서버)에 접속요청을 수행한다.(A1) 2. 정보센터(서버)는 성명에 대한 인증 절차를 통해 접속 요청을 수락한다. 3. 정보센터(서버)는 사용자의 목적행위/목적지에 적합한 아이템들을 필터링 하고, 과거선호(평점)매트릭스 및 기타 사용자 매트릭스를 하이브리드 필터링을 이용해 최종 추천아이템 정보 생성한다.(A2) 4. 생성된 추천 목록은 클라이언트의 화면에 표출된다.
	대안흐름 (A1) 예기치 않은 통신 두절로 정보센터(서버)접근이 불가능할 경우 1. 추천시스템은 접속실패 오류와 재접속 메시지 표출한다. (A2) 사용자 과거 선호(평점)매트릭스가 존재치 않을 경우 1. 정보센터(서버)는 해당 오류 메시지를 추천시스템(개인 단말기)에 전송한다. 2. 추천시스템은 전송된 오류메시지를 표출한다.
후행 조건	추천시스템은 사용자의 요청에 대한 정보를 화면에 표출한다.
기타요구사항	전체적인 접속과정은 3초내로 이루어져야 한다.

본 절에서는 추천시스템 개념 설계의 과정으로 시스템의 요구사항 분석, 서비스 도출을 수행하였으며 실제적인 논리설계의 과정으로 유스케이스 및 액터 도출, 유스케이스 다이어그램, 유스케이스 명세화를 수행하였다. 이를 통해 본 연구에서 개발하는 추천시스템의 전체적인 개념을 정립했다고 볼 수 있다.

다음 절에서는 추천시스템의 논리적 설계를 위한 과정인 소프트웨어 아키텍처 명세에 대해 설명한다.

제 2 절 소프트웨어 아키텍처

1. 소프트웨어 아키텍처 개요

애플리케이션이나 시스템의 규모가 커지고 복잡해지면서 아키텍처의 중요성이 점점 더 커지고 있다. 이에 따라 소프트웨어 아키텍처는 매우 중요한 연구 분야로 자리를 잡은 상태이다. 일반적으로 “아키텍처란 그 시스템의 구조나 구조들로 각 요소들과 외부의 보이는 특성들 그리고 그들 간의 관계를 절충한다.” 라고 정의된다.

기본적으로 소프트웨어 아키텍처는 시스템 구축 및 개발을 위한 하나의 도구로 볼 수 있으며 시스템의 개발 초기 단계에 수행된다. 소프트웨어 아키텍처를 명세함으로써 시스템의 향후 유지보수 및 확장성을 크게 제고할 수 있는 장점이 있다. 본 절에서는 소프트웨어 아키텍처의 명세 기법인 C&C View, Module View, Allocation View 등을 이용하여 본 연구에서 제시하는 추천시스템의 소프트웨어 아키텍처를 표현한다.

소프트웨어 아키텍처는 시스템 구조를 설명하며 소프트웨어의 구성 요소들을 정의한다. 소프트웨어 아키텍처의 중요성은 다음과 같다.

- 소프트웨어 아키텍처는 시스템의 공통적인 추상화를 제공함으로써 관련 이해당사자들 간의 상호이해와 협상, 일치, 의사소통의 기본을 제공한다.
- 시스템 분석, 설계 시의 초기 결정사항 및 청사진을 제공한다.
- 시스템의 요약된 형태로서 향후 상세한 시스템 구성의 기본을 제공한다.

아키텍처는 시스템이 요구하는 품질속성의 달성을 위해 분석을 시도해 볼 수 있는 첫 번째 문서일 뿐만 아니라, 프로젝트의 청사진이 된다. 또한 아키텍처는 의사소통의 도구이자 최초의 설계 결정사항을 표현한 것이며 새로운 시스템에 재사용할 수 있도록 추상화된 결과물이다. 아키텍처 구조는 각 요소가 보여주는 의미에 따라 크게 3가지 그룹으로 나뉘지며 대략적인 개념을 요약하면 다음과 같다.

- 모듈 구조(module structure)

모듈구조는 구현과 관련된 모듈을 아키텍처 요소로 정의한다. 모듈은 시스템과 관련 있는 코드 기반 방식을 묘사한다. 모듈은 기능 영역으로 구분된다. 모듈 구조는 시스템 실행 시에 소프트웨어가 동작하는 방식을 보여주기에 적합하지 않다. “각 모듈에 부여된 주요 기능은 무엇인가?”, “어떤 소프트웨어 요소가 사용되도록 허가된 모듈인가?”, “실제로 사용되는 소프트웨어는 무엇인가?” “어떤 모듈이 일반화나 특수화(상속 관계 같은)관계인가?”라는 질문에 대한 답을 구할 수 있다. 논리적 구조의 요소는 “핵심 추상화”로 이 구조는 객체지향 분야에서 객체와 객체 클래스로 분명히 표현된다.

- 컴포넌트와 커넥터 구조(component-and-connector structure)

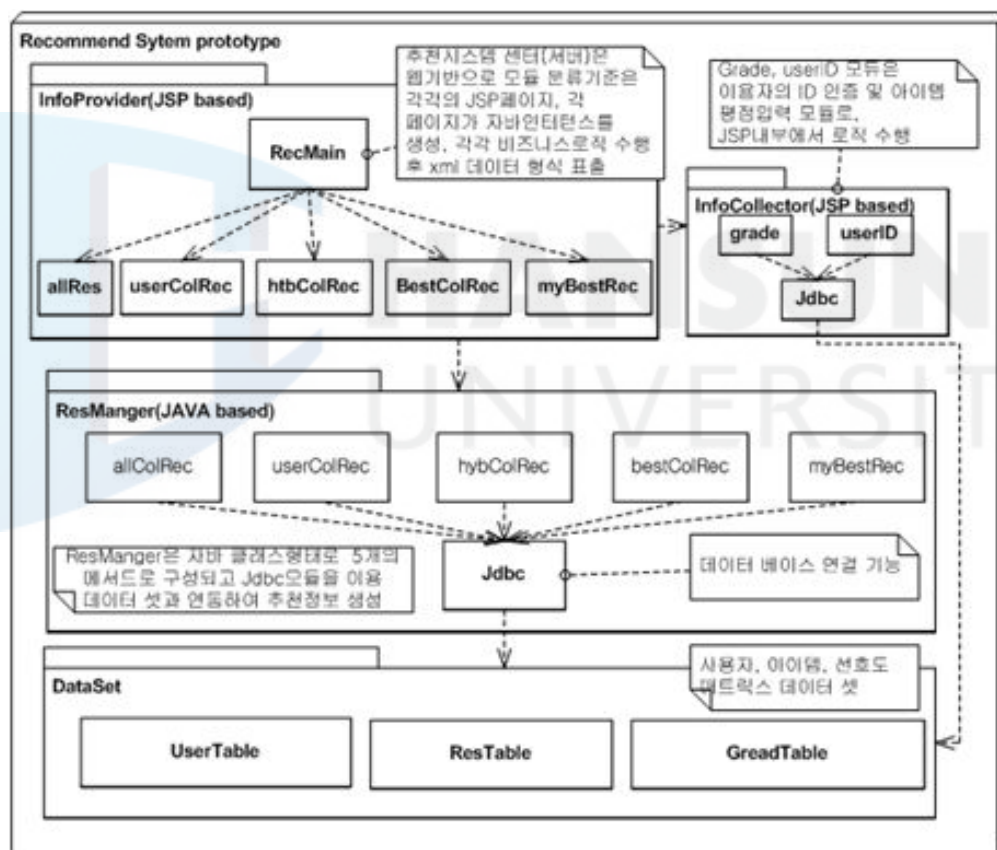
컴포넌트와 커넥터 구조의 아키텍처 요소는 실행 시에 발생하는 컴포넌트(연산을 하는 가장 중요한 단위)와 커넥터(컴포넌트 간의 대화 수단)이다. 컴포넌트와 커넥터 구조를 통해 “가장 중요한 실행 컴포넌트는 무엇이고 이것들은 어떻게 연동되는가?”, “가장 중요한 공유 데이터 저장소는 무엇인가?”, “어떤 부분이 복제됐는가?”, “시스템을 통해 데이터는 어떻게 변화하는가?”, “어떤 부분이 시스템에 병렬로 실행되는가?”, “어떻게 시스템의 구조가 실행 시에 변경되는가?”에 대한 답을 얻을 수 있다. 프로세스 관점에서 동시성과 기능분산에 초점을 둔다.

- 할당 구조(allocation structure)

할당 구조는 소프트웨어가 개발되고 실행되는 외부환경과 소프트웨어 요소 간에 어떤 관계가 있는지를 보여준다. 할당 구조를 통해 “어떤 프로세스가 각 소프트웨어 요소가 실행되는가?”, “시스템이 개발되고 테스트되고 시스템이 운용될 때 각 요소는 어느 파일에 저장되는가?”, “개발팀에 부여되는 소프트웨어 요소는 무엇인가”라는 질문에 답을 구할 수 있다. 소프트웨어 모듈과 라이브러리, 서브시스템, 개발 단위에 대한 구조를 설명하는 관점으로 소프트웨어와 개발 조직의 관계를 표현한다.

2. 추천시스템의 모듈 뷰

모듈뷰는 크게 분할(decomposition), 사용(use), 계층(layered), 클래스(class) 혹은 일반화(generalization)구조로 구분된다. 본 과정에서는 이중 사용(use) 구조를 이용하여 추천시스템의 아키텍처를 표현하였다. 사용구조는 모듈이나 (아주 잘게 나누는 경우) 프로시저, 모듈의 인터페이스에 있는 자원 등의 단위로 구성된다. 이런 단위들은 사용 관계로 연결된다. 사용 구조에서는 A의 정확성(correctness)을 위해 B의 의 정확한 버전이 존재해야 하는 경우, A가 B를 사용한다고 정의한다. 사용구조는 새로운 기능을 시스템에 추가하거나 유용한 기능의 부분집합을 추출하는데 사용된다. 시스템의 부분집합을 쉽게 개발할 수 있다는 것은 좀 더 손쉽게 개발을 점진적으로 진행할 수 있다는 것을 의미한다.



<그림 4-2> 추천시스템 서버 모듈뷰 아키텍처

위 그림은 상황정보를 이용한 하이브리드 추천시스템의 정보센터(서버)의 모듈류를 나타낸다. 서버는 클라이언트(모바일)와 HTTP 방식으로 통신으로 구성된다. 클라이언트 시스템을 통해 서버의 각 JSP 모듈에 정보를 요청하면 JSP 페이지는 JAVA 클래스의 메서드를 호출하고 각 메서드는 요청에 대한 각각의 비즈니스 로직을 수행한다.

정보의 추출을 위해 각 메서드는 사용자, 아이템, 선호도 매트릭스 데이터 셋에 접근하여 요청하는 정보에 대한 필터링을 실행 한 후 그 결과를 JSP에 XML 데이터 형식으로 표출한다. 클라이언트는 해당 요청에 대한 JSP의 XML 데이터 전송을 요청 후 데이터 전송이 이루어지면 클라이언트의 정보 표출 비즈니스 로직을 통해 결과를 화면에 표출한다.

클라이언트 시스템에 대한 모듈 구조는 사용자 요구정보에 대한 요청, 요청 결과에 대한 표출의 단순한 로직을 수행하므로 본 모듈 뷰에서 자세히 다루지 않았으며, 이는 차후 정의될 컴포넌트와 커넥터 뷰, 할당 뷰에서 다뤄진다.

3. 추천시스템의 컴포넌트와 커넥터 뷰

컴포넌트와 커넥터 구조는 모듈 구조와 상호 독립적이다. 이는 시스템 실행 관점을 표현하는 동적인 모습을 표현한다. 크게 프로세스(process) 혹은 통신 프로세스(communicating process), 동시성(concurrency), 공유 데이터(shared data) 혹은 저장소(repository), 클라이언트/서버(client-server)로 구분된다. 본 과정에서는 이중 클라이언트/서버(client-server) 구조를 이용하여 BIT prototype의 아키텍처를 표현하였다. 클라이언트/서버 개념의 시스템이라면 이 구조를 사용하는 것이 좋다. 컴포넌트는 클라이언트와 서버이고 커넥터는 시스템의 기능을 수행하기 위해 클라이언트/서버가 사용하는 통신규약과 메시지다. 기능을 분화하기가 용이(변경용이성 지원)하고 물리적으로 분산처리 기능을 표현하기에 유리하며 부하 분산에 유용하게 사용된다.

c&c view의 주요 설계요소들은 컴포넌트(Component), 커넥터(Connector), 인터페이스(Interface), 포트(Port) 이며, 요소별 세부 내용은 다음과 같다.

컴포넌트 : 하나의 모듈 또는 시스템을 의미하며 시스템 구축 후 전체 시스

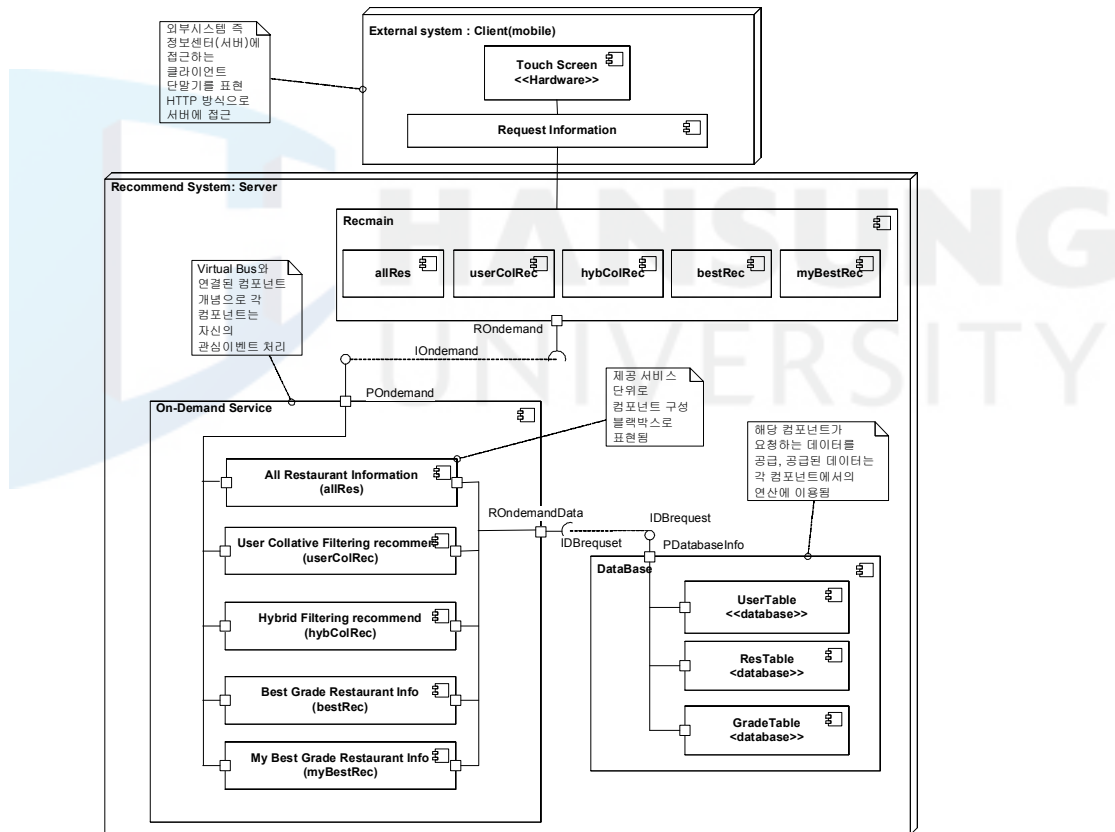
템 배치에 이용될 수 있는 구조이다. 컴포넌트 내부는 블랙박스 형태로 간주하고 입출력 및 주변 환경과의 접점을 표현해야 한다.

커넥터 : 주요 설계 요소끼리 관계를 맺고 있는 경우에 이용될 수 있는 일종의 연결자 형태이다. 이러한 커넥터는 컴포넌트와 컴포넌트, 연결 접점과 컴포넌트 등 다양한 의미에서 사용될 수 있다.

인터페이스 : 하나의 컴포넌트와 외부 환경과의 상호작용을 표현할 때 사용되며 정보 및 기능을 제공하는 의미를 담은 제공 인터페이스, 정보 제공을 기대하는 뜻을 나타내는 요구 인터페이스로 구분된다.

포트 : 하나의 컴포넌트와 외부 및 내부 환경과의 상호작용일 일어나는 접점을 v표시한다. 한 컴포넌트는 여러 개의 포트를 가질 수 있으며 포트는 인터페이스와 연결된다.

추천시스템의 c&c view 도출 결과는 아래와 같다.



<그림 4-3> 추천시스템 서버 컴포넌트 & 커넥터 뷰 아키텍처

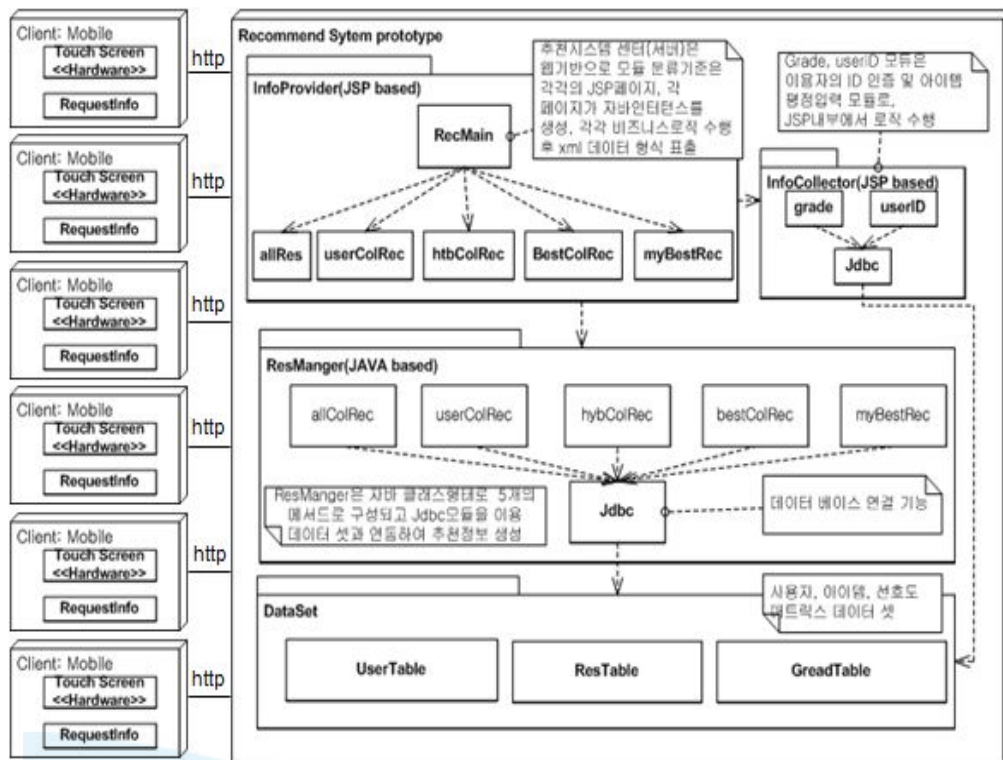
추천시스템은 시스템 사용자와 정보센터(서버)의 접점이 되는 모바일 단말기, 사용자의 요청에 대한 비즈니스 로직을 처리하여 정보를 제공하는 정보센터(서버), 그리고 사용자 및 아이템, 선호(평점)매트릭스 등의 데이터베이스로 구성된다. 시스템 사용자는 단말기의 애플리케이션을 통해 서버에 접근하여 원하는 범주(목적행위, 목적지 혹은 현위치)에 대한 추천아이템 정보를 선별할 수 있다.

C&C view에서 중점은 정보흐름에 중점을 두어 각 컴포넌트들 간의 관계를 표현한다는 것이다. 클라이언트 단말기에서 비 주기적인 사용자의 요청이 있을 경우, 클라이언트 시스템은 HTTP 방식으로 서버시스템과의 통신을 요청하고 서버는 이 정보요청에 대해 블랙박스 형태로 표현된 각 세부 컴포넌트를 이용하여 데이터베이스에 접근 및 정보를 추출 한 후 적절한 비즈니스 로직을 수행하여 추천정보를 생성 한 뒤 이를 클라이언트에게 전송한다.

4. 추천시스템의 할당 뷰

할당 구조는 소프트웨어가 개발되고 실행되는 외부환경과 소프트웨어 요소간에 어떤 관계가 있는 지를 보여준다. 할당구조에는 배치(deployment), 구현(implementation), 업무 할당(work assignment) 구조가 있으며 본 과정에서는 구현구조를 사용하여 BIT prototype의 할당 뷰를 표현한다. 구현 구조는 소프트웨어 요소(일반적으로 모듈)가 시스템의 개발환경, 통합환경, 형상 통제환경에서 파일 구조에 어떻게 대응되는지 보여준다. 이것은 개발 활동 관리와 프로세스 생성 관리에 매우 중요하다. 다음 그림은 상황인식 및 하이브리드 추천시스템의 할당 뷰를 나타낸다.

할당 뷰는 일반적으로 모듈 뷰의 구조를 표현된 구성요소를 하드웨어 혹은 개발 단위 등의 노드에 할당하여 표현한다. 추천시스템의 할당 뷰 도출 결과는 아래와 같다.



<그림 4-4> 추천시스템 서버 할당부 아키텍처

소프트웨어 아키텍처는 시스템 개발의 초기 설계 결정사항을 표현한다. 이러한 초기 설계 결정사항을 이후에 고친다는 것은 매우 힘든 일이다. 특히 개발 단계에서 수정하는 것은 거의 불가능하다. 아키텍처 명세를 통해 결정될 수 있는 사항 및 이점에는 1) 아키텍처는 개발의 제약 및 구조 결정, 2) 시스템의 품질속성 장려 하거나 억제, 3)시스템의 품질을 예측 가능, 4)점진적 프로토타입 개발용이 등이 있다. 소프트웨어 아키텍처의 구성은 차후 소프트웨어의 개발 및 구현 측면에서 다양한 방법으로 참조되며 실질적인 시스템의 청사진이 될 뿐 아니라 시스템의 유지보수, 변경용이성에 큰 영향을 미칠 수 있다. 다음 절에서는 현 단계까지 수행된 시스템의 설계 및 소프트웨어 아키텍처를 기반으로 실제 시스템의 구현을 수행한 결과에 대해 설명한다.

제 3 절 상황정보를 이용한 하이브리드 필터링 추천 시스템

1. 개발 환경

가. iPhone SDK - Objective-C

상황정보를 이용한 하이브리드 필터링 추천시스템의 클라이언트 소프트웨어는 iPhone 3G 모바일 환경에서 Objective-C 기반으로 하여 개발되었다.

아이폰(iPhone)은 2007년 1월 9일에 애플이 미국 샌프란시스코에서 열린 맥월드 2007에서 발표한 터치 스크린 기반의 아이팟, 휴대전화, 모바일 인터넷이라는 세 가지 주요 기능을 가진 스마트폰이다.

아이폰의 주요 장점으로 애플에서 운영하는 앱스토어를 통해 애플리케이션은 세계 최대 수준인 20만개에 달해 다양한 기능들을 사용할 수 있다. 유료로 제공되는 애플리케이션의 수익 구조는 제작자가 70%, 애플측이 30%를 가져가게 되기 때문에 애플리케이션 제작 경쟁도 치열하다.

오브젝티브-C (Objective-C, 종종 ObjC로 표기) 언어는 C 프로그래밍 언어에 스톱토크 스타일의 메시지 구문을 추가한 객체 지향 언어로서 이를 이용하여 아이폰의 애플리케이션을 개발할 수 있다. 오브젝티브-C는 1980년대 초에 브래드 콕스(Brad Cox)에 의해 만들어졌다. 1988년에 넥스트사의 워크스테이션 NeXT의 사용자 인터페이스인 NextStep에 사용됨으로써 널리 알려지기 시작하였으며, 1996년 애플이 넥스트 사를 인수한 후 맥 오에스 텐의 코코아 애플리케이션 프레임워크의 표준 프로그래밍 언어로서 쓰이고 있다.

애플에서는 기업 및 일반 유저가 아이폰용 애플리케이션하여 실행 할 수 있도록 iPhone SDK를 제공하고 있기 때문에 원하는 누구든 iPhone SDK를 이용하여 다양한 애플리케이션을 개발할 수 있다.

나. JAVA, JSP

추천시스템의 정보센터(서버)는 JAVA와 JSP를 기반으로 하여 개발되었다.

JAVA는 썬마이크로시스템즈의 제임스 고슬링 및 연구원들에 의해 개발된

객체 지향적 프로그래밍 언어이며, 썬마이크로시스템즈에서 무료로 제공하고 있다. 1991년 그린 프로젝트라는 이름으로 시작해 1995년에 발표했다.

Java의 구문은 기존의 대표적인 프로그래밍 언어인 C, C++과 매우 유사하다. 다시 말하면 JAVA는 전혀 새로운 프로그래밍 언어가 아니며, 기존의 프로그래밍 언어에 익숙한 이용자는 좀 더 쉽게 JAVA에 접근할 수 있다.

그런데 일반적으로 JAVA하면 웹 브라우저 상에 움직이는 애플릿을 생각하는 경우가 많으나 애플릿은 JAVA로 개발할 수 있는 것들 중 일부에 지나지 않는다. (참고로 applet은 web server에 존재하고 있다가 사용자 pc의 web browser로 download되어 그 안에서 수행되는 Java 프로그램을 말한다).

JSP는 Java Server Page의 약자이며 자바를 기반으로 한 웹 사이트 개발 기술 언어이다, 서블릿(JAVA 클래스 구조) 기술을 바탕으로 만들어져 있으며, HTML, XML, WML 등에 의해 표현되는 정적인 자료와 JSP 요소에 의해 생성되는 동적인 내용을 함께 사용함으로써 동적인 웹 페이지를 쉽게 개발할 수 있는 썬마이크로시스템즈의 기술이다. JSP를 이용해 웹 환경 상에서 100% 순수한 자바만으로 서버 사이드 모듈을 개발할 수 있으며 뿐만 아니라 DBMS와 같은 백 엔드와 연동하여 그들의 데이터를 가공해 웹상의 최종적 사용자에게 디스플레이 할 수 있다.

다. 오라클 10 G

추천시스템의 정보센터(서버) 내부에 데이터를 저장하고 관리하기 위한 데이터베이스는 오라클 10g를 기반으로 개발되었다.

오라클은(Oracle)은 Database/ERP와 이에 대한 솔루션을 제공하는 회사의 이름이기도 하며 Oracle DBMS를 지칭하는 말로도 쓰인다. Oracle은 관계형 데이터베이스 모델이 기초하여 만들어졌으며 Unix, Linux, Window 등의 대부분의 OS를 지원한다.

오라클은 현재 사용하는 컴퓨터 시스템이나 네트워크를 최대한 활용할 수 있도록 데이터베이스 서버와 클라이언트 응용 프로그램에 처리를 분산시킨다. 즉, 데이터베이스 관리 시스템을 실행하는 컴퓨터는 데이터베이스 서버 역할

을 담당하고 데이터베이스 응용 프로그램을 실행하는 워크스테이션은 데이터 해석과 표시를 담당한다.

오라클은 테라바이트 크기의 데이터를 저장할 수 있는 대규모 데이터베이스를 지원한다. 고가의 하드웨어 장치를 효율적으로 활용할 수 있도록 영역 사용을 완벽하게 제어한다. 또한 데이터 경합을 최소화하고 데이터 동시성을 보장함으로써 여러 사용자가 동일한 데이터에서 다양한 데이터베이스 응용 프로그램을 실행하도록 지원한다. 뿐만 아니라 전체 시스템의 성능을 높게 유지하면서 고성능 트랜잭션 처리를 제공하므로 신속한 처리속도를 제공한다.

2. 추천 시스템 구성 및 실행과정

다음은 iPhone SDK, Objective-C를 이용하여 개발된 상황정보를 이용한 하이브리드 필터링 추천시스템을 iPhone SDK에서 제공하는 iPhone Simulator로 구동한 모습이다.



<그림 4-5> 상황정보를 이용한 하이브리드 필터링 추천시스템

iPhone Simulator 상의 Recommend 아이콘을 클릭하면 상황정보를 이용한 하이브리드 필터링 추천시스템 애플리케이션을 실행 할 수 있다.

본 시스템의 개발은 새롭게 제시되는 추천기법에 대한 효용성 및 정확성을 평가하기 위함이므로 별도의 사용자 보안관련 모듈을 탑재하지 않았다. 애플리케이션을 실행하면 미리 구현된 클라이언트의 비즈니스 로직을 통해 iPhone 기기에 할당된 고유번호를 Http통신 방식으로 정보센터(서버)로 전송함으로써 사용자 인증단계를 수행한다. 이 과정을 통해 해당 사용자가 새로운 사용자일 경우 정보센터(서버)의 비즈니스 로직을 통해 새로운 전체사용자 테이블에 새로운 사용자를 추가하며 해당사용자의 선호도(평점)매트릭스 테이블을 생성한다.

사용자 인증을 마치면 해당 추천대상자의 상황정보를 추출하기 위해 사용자의 목적행위 및 현재위치 혹은 목적지의 선택과정을 수행한다. 이러한 과정을 다음 그림에 표현하였다.



<그림 4-6> 추천시스템 목적행위 및 목적지/현위치 선택화면

- 목적행위 선택 : 추천대상자는 자신이 하고자 하는 목적행위를 선택한다. 음식점, 병원, 쇼핑의 카테고리로 분류되어 있으며, 이는 시스템의 확

장을 통해 차후 추가될 수 있다. 음식점을 선택할 경우, 서버는 아이템 테이블에서 음식점 아이템외의 데이터들을 필터링하여 음식점에 해당하는 데이터를 제공한다.

- 목적지 및 현재 위치 선택 : 추천대상자는 목적행위를 실행하기 위해 이동하고자 할 때 혹은 현재 위치에서 추천정보를 얻고자 할 때 해당 지점을 선택한다. 이는 현재 텍스트로 표현되어 있으나, 차후 시스템이 확장될 시 목적지 검색의 모듈이 추가될 수 있다. 해당 지점을 선택할 경우 서버는 해당 지점에서 일정거리 이상에 존재하는 아이템 데이터들을 필터링하고 일정거리 내부에 존재하는 데이터를 제공한다.

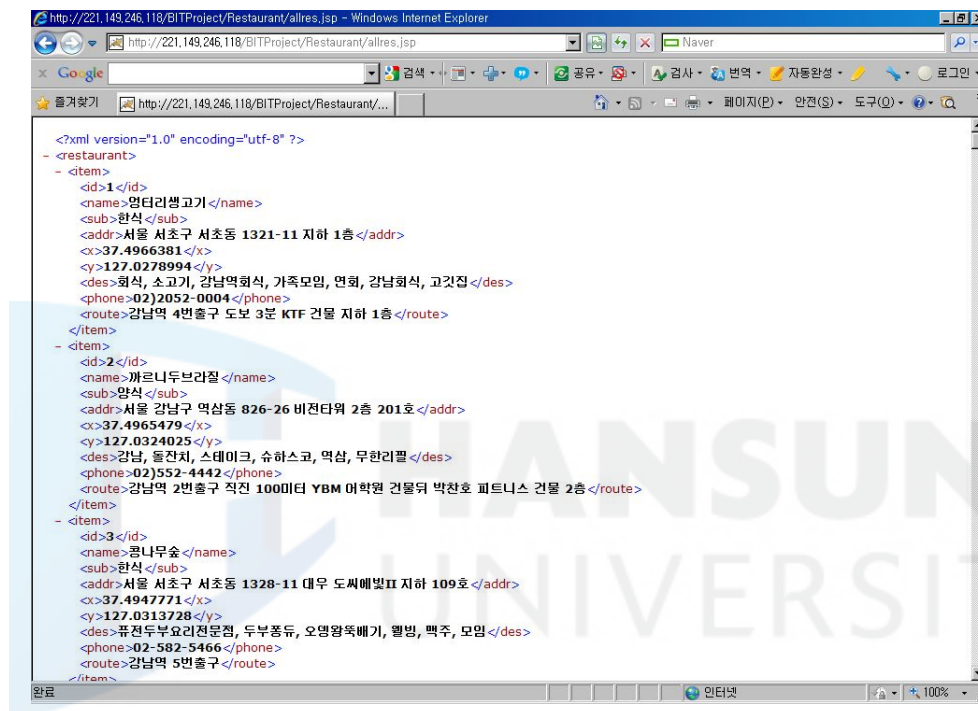
목적행위, 목적지(현위치) 선택 과정이 종료되면, 시스템은 해당 목적행위 및 목적지에 대한 아이템 목록을 제공한다. 아이템 목록은 두 개의 인터페이스로 사용자에게 제공되며 이는 리스트 방식과 지도표출 방식이다. 추천대상자는 “리스트 보기”, “지도보기” 버튼을 통해 원하는 인터페이스로 아이템정보를 제공 받을 수 있다.



<그림 4-7> 추천시스템 아이템 표출 화면(왼쪽: 리스팅, 오른쪽: 지도)

이 때 해당 지점/목적지(현위치)에 대한 모든 아이템을 표시하며 이는 클라이언트와 정보센터(서버)와의 통신으로 정보를 요청하고, 정보센터를 요청에 대한 처리를 위해 특정 비즈니스 로직을 수행한 후 결과를 xml 데이터 형식으로 JSP에 표출한다. 클라이언트는 해당 xml 데이터를 클라이언트의 자료구조에 저장 한 후 이를 화면에 표출한다.

다음은 정보센터(서버)에서 클라이언트의 요청으로 추천대상자의 목적행위 및 목적지(현 위치)에 대한 모든 아이템을 xml 데이터 표출한 결과를 나타낸다.



<그림 4-8> 추천시스템 정보센터(서버)의 xml 데이터 아이템 표출 화면

추천대상자는 아이템 리스트보기 및 지도보기 화면에서 제공되는 “전체보기”, “베스트 아이템 추천”, “내 과거 베스트 아이템 추천”, “사용자협업필터링 추천”, “하이브리드 필터링 추천” 다섯 개의 버튼 선택을 통해 원하는 형태의 추천을 받을 수 있다. 각각에 대한 간략한 설명은 다음과 같다.

- 전체보기(All) : 추천대상자의 목적행위 및 목적지(현 위치)에 대한 모든 아이템 정보를 제공한다. 디폴트로 설정되어 있다.

- 임의추천(Rand) : 추천대상자의 목적행위 및 목적지(현 위치)에 대한 임의 추천리스트를 생성 하여 아이템 정보를 제공한다.
- 베스트 아이템 추천(Best) : 추천대상자의 목적행위 및 목적지(현 위치)에 대한 아이템 중 과거 선호도(평점) 평균이 가장 높은 순으로 일정 개수의 아이템 정보를 제공한다.
- 사용자 협업 필터링 추천(User) : 추천대상자의 목적행위 및 목적지(현 위치)에 서의 모든 사용자의 아이템에 대한 과거 선호(평점) 정보를 토대로 추천대상자와 유사도가 높은 사용자(이웃 사용자 집단)들을 추출해내고 추천대상자가 경험하지 않은 아이템에 대해 예측선호(평점)를 산출하고, 높은 순으로 일정 개수의 아이템 정보를 제공한다.(3장 사용자협업 필터링 단계 참조)
- 하이브리드 필터링 추천(HybC) : 추천대상자의 목적행위 및 목적지(현 위치)에의 사용자 협업 필터링의 적용으로 추천대상자와 유사도가 높은 사용자(이웃 사용자 집단)들을 추출해내고 추천대상자가 경험하지 않은 아이템에 대해 이웃 사용자 집단이 경험했던 아이템들과의 유사도 비교를 통해 유사 아이템 군을 추출함으로써 추천대상자가 경험하지 않은 아이템에 대해 예측선호(평점)를 산출하고 높은 순으로 일정 개수의 아이템 정보를 제공한다.
- 내 과거 베스트(Mybe) : 추천대상자의 목적행위 및 목적지(현 위치)에 대한 아이템 중 추천대상자가 과거 입력한 선호도(평점)가 가장 높은 순으로 일정 개수의 아이템 정보를 제공한다.

상기 버튼 입력을 통해 사용자는 원하는 추천방식으로 필터링된 정보들을 제공 받을 수 있다. 사용자는 클라이언트 시스템을 통해 각 추천서비스를 통해 제공된 아이템을 경험한 후 아이템에 대한 선호도(평점), 혹은 추천서비스를 제공 받지 않는했으나 기존의 경험을 통해 얻어진 아이템 선호도(평점)를 입력 할 수 있다. 선호도(평점)입력을 통해 얻어진 데이터는 클라이언트-서버간의 통신을 통해 정보센터(서버)로 전달되며 서버의 해당사용자의 선호도(평점)매트릭스 테이블에 입력된다.

사용자들의 선호도(평점)입력 과정을 통해 지속적으로 확장되는 사용자의 선호도(평점)매트릭스 테이블은 차후 추천서비스의 실행 시 각각의 추천방식 (베스트 아이템, 내 과거 베스트 아이템, 사용자 협업 필터링 추천, 하이브리드 필터링 추천)에 반영된다.

다음은 아이템에 대한 세부정보 확인 및 평점입력 화면을 나타낸다.



<그림 4-9> 아이템 세부 정보 및 평점 입력 화면

상황정보를 이용한 하이브리드 필터링 추천시스템의 동작 및 실행 여부를 확인하기 위해 시뮬레이션 테스트로 가상의 데이터를 생성하여 각 추천서비스들을 수행하였다. 시뮬레이션 테스트 환경은 다음과 같다.

[표 4-6] 시뮬레이션 테스트 환경

추천 대상자	실험내용	실험 데이터
C	추천대상자 C의 목적행위(음식점방문) 및 목적지(강남역)에 대해 아이템 추천 서비스 실행	C포함 10개의 가상 사용자선호도(평점) 매트릭스 데이터(랜덤) 및 5개의 음식점 데이터

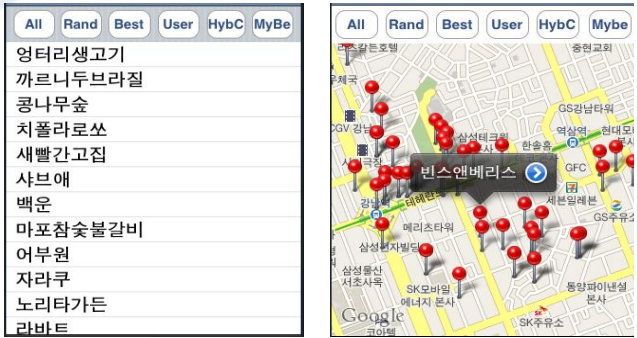
가상의 추천대상자 C와 기타 9명의 가상 사용자를 생성한 뒤 5개의 음식점에 대해 랜덤한 선호도를 입력하였다. 이 때 C는 5개의 음식점 중 “까르니두 브라질”, “새빨간 고집”에 대한 경험이 없다고 가정하여 선호도 표시를 하지 않은 상태이다. 가상사용자 10명의 5개 음식점에 대한 선호도(평점)매트릭스는 다음과 같다.

[표 4-7] 10명의 가상사용자에 대한 5개 음식점의 선호(평점)매트릭스

	까르니두브 라질	एंटर리생고 기	치폴라로쏘	콩나무숲	새빨간 고집
A	3	4	5	0	0
B	5	1	0	4	2
C	0	2	1	4	0
D	3	4	1	0	2
E	1	5	0	3	0
F	2	5	3	3	5
G	2	0	2	5	3
H	0	3	2	5	4
I	5	4	2	2	3
J	2	5	3	2	0

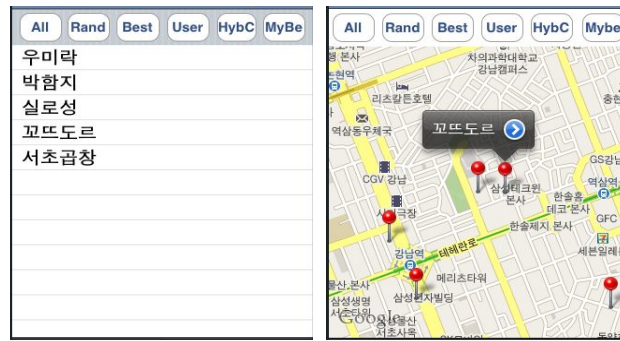
가상의 데이터를 바탕으로 상황인식 및 하이브리드 추천시스템을 이용하여 추천대상자 C에게 추천정보를 생성한 결과는 다음과 같다.

▪ 전체보기



<그림 4-10> 가상 사용자 C의 전체 아이템 리스트

- 임의추천



<그림 4-11> 가상 사용자 C의 임의 추천 리스트

- 베스트 아이템 추천



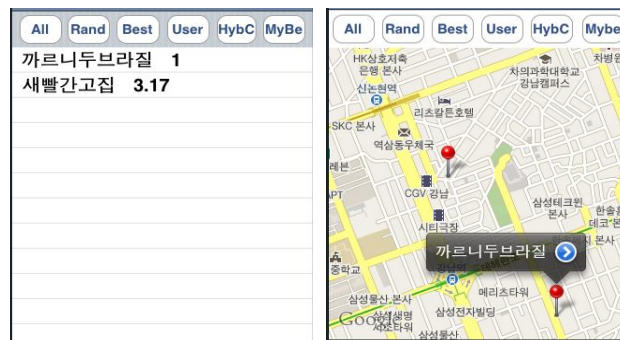
<그림 4-12> 가상 사용자 C의 베스트 아이템 추천 리스트

- 사용자 협업 필터링 추천



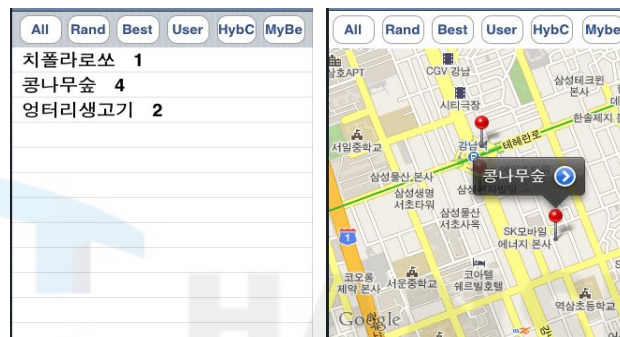
<그림 4-13> 가상 사용자 C의 사용자 협업 필터링 추천 리스트

- 하이브리드 필터링 추천



<그림 4-14> 가상 사용자 C의 하이브리드 필터링 추천 리스트

- 내 과거 베스트



<그림 4-15> 가상 사용자 C의 내 과거 베스트 추천 리스트

제 5 장 상황정보를 이용한 하이브리드 추천시스템 평가

제 1 절 실험 배경 및 절차

1. 실험 배경

기존의 웹상에서 다양한 추천 시스템에서 주로 사용된 기법인 협업 필터링을 이용하여 모바일 환경에서의 적용하기 위해 추천시스템을 설계하고 구현하였다. 이 때 단순 협업필터링의 적용으로 추천서비스를 제공하는 것이 아닌, 추천의 질을 높이기 위한 방안으로, 모바일에서 활용 가능한 사용자의 상황정보(목적, 위치)에 대한 추출을 통해 적절한 추천 아이템을 선별하고 협업 필터링의 대표적인 두 기법인 사용자 기반 협업 필터링과 아이템 기반 협업 필터링을 혼합한 하이브리드 필터링을 적용하여 추천서비스를 제공하는 시스템을 설계 및 구축하였으며, 이를 상황정보를 이용한 하이브리드 필터링 추천시스템이라 명하였다.

본 장에서는 구축된 추천시스템의 효용성 및 정확성을 판단하기 위한 실험을 실시하고자 한다.

본 연구에서는 고객의 아이템에 대한 선호(평점) 정보를 활용하므로 실험을 통해 자료를 수집하고 추천시스템이 제공하는 추천 목록에 대한 만족도를 조사하여 상황정보를 이용한 하이브리드 필터링 추천시스템을 평가한다.

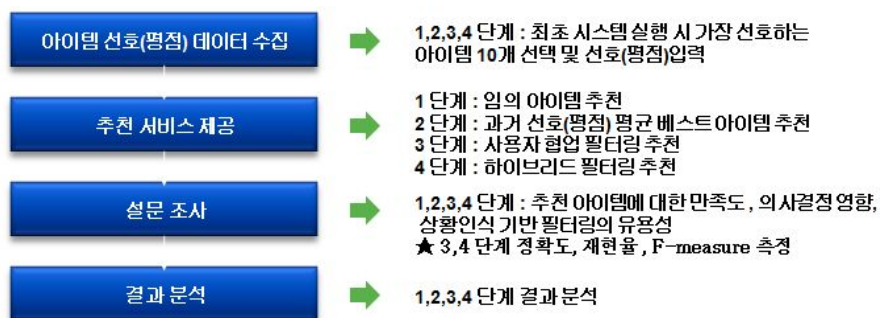
또한 추천 시스템은 상황정보를 이용한 제공정보의 범위 결정 후 단순 협업 필터링 방식의 추천서비스와 하이브리드 필터링 방식의 추천 서비스를 동시에 제공하도록 구성된다. 두 가지 방식의 추천서비스를 별도의 정확성 및 재현율의 평가를 통해 결과를 도출하여 이를 두 추천방식의 차이 및 상황정보를 이용한 하이브리드 필터링 추천시스템 검증 자료로 사용한다.

또한 추천의 질을 높이기 위한 수단으로 제시된 상황정보를 이용한 제공정보의 범주 및 범위 결정의 적용이 정보 추천의 질에 대한 영향을 미치는가에 대한 평가를 설문을 통해 파악하도록 한다.

2. 실험 절차

평가를 위한 실험은 2010년 12월 27일부터 2010년 12월 31일까지 총 60명의 을 대상으로 이루어 졌다. 실험은 총 4단계로 이루어지며 각 단계마다 15명의 실험대상에게 대해 상이한 방식의 아이템 추천을 실행한다. 1단계 실험은 임의로 아이템을 추천한다. 2단계 실험은 과거 사용자 별 아이템 선호(평점) 기록을 토대로 산출된 베스트 아이템을 추천한다. 3단계 실험은 사용자 협업 필터링을 기반으로 이웃 사용자 집단을 형성 한 후 사용자가 경험하지 않은 아이템에 대해 예측 선호를 산출한 후 아이템을 추천한다. 4단계 실험은 하이브리드 필터링을 기반으로 이웃 사용자 집단 및 이웃 아이템 집단을 형성 한 후 사용자가 경험하지 않은 아이템에 대해 예측 선호를 산출한 후 아이템을 추천한다. 이는 기본적으로 사용자의 목적, 위치 상황정보를 이용해 정보의 범주 결정이 된 상태로 가정한다. 각 방식의 추천절차가 이루어지면 아이템 추천의 결과가 만족스러운지, 그리고 사용자의 의사결정에 도움이 되었는지 여부에 대한 설문을 실행한다. 또한 사용자 협업 필터링 및 하이브리드 필터링 전에 수행되는 상황정보의 활용이 정보의 필터링이 정보추천의 질에 미치는 영향에 대한 설문을 실행한다.

실험의 각 단계 마다 사용자가 추천 시스템을 처음 사용할 시 제공되는 목록 중 가장 선호하는 10개의 아이템을 선택하여 평점을 입력하도록 한다. 이는 실험의 총 과정에서 아이템에 대한 개인 사용자 및 전체 사용자 집단의 성향을 분석함으로써 베스트 아이템 추천, 사용자 협업 필터링 추천, 하이브리드 필터링 추천을 실행하기 위한 기초 데이터로 사용된다.



<그림 5-1> 아이템 세부 정보 및 평점 입력 화면

제 2 절 실험 데이터 및 실험군 구성

1. 실험 데이터 구성

추천 아이템 데이터는 강남역 인근에서 현재 운영 중인 100개의 음식점으로 구성하였다. 본 실험에서는 사용자의 목적행위 정보를 “음식점 정보요청”으로 가정하였으며 목적지(현위치) 정보를 강남역 인근으로 가정하였다. 이는 연구의 내용적 범위 측면에서 본 실험의 주요 목적이 본 연구에서 제시하는 추천 기법의 질 및 효용성을 평가하기 위함임을 감안할 때 적절한 수준의 방식이라 판단되었다.

이에 따라 강남역 인근에서 실제 운영 중에 있는 100여개의 음식점 데이터를 조사 및 분석하여 실험 데이터를 구성하였으며 다음은 실험 데이터의 세부 구성을 나타낸다.

[표 5-1] 실험 데이터 구성(지면 한계 상 10여개의 데이터만을 표시하였음)

구분	명칭	종류	주소	위도	경도	특징	전화번호	오시는길
1	영터리생고기	한식	...	37.4966381	127.0278994	...	02)2052-0004	...
2	까르니두브라질	양식	...	37.4965479	127.0324025	...	02)552-4442	...
3	콩나무숲	한식	...	37.4947771	127.0313728	...	02-582-5466	..
4	치폴라로소	양식	...	37.498297	127.0280021	...	02)599-1256	..
5	새빨간고집	카페 주점	...	37.501828	127.0274142	...	02-730-6263	...
6	샤브애	한식	...	37.5009825	127.0278494	...	02)594-4955	...
7	백운	한식	...	37.4995208	127.0320219	...	02)3453-4985	...
8	마포참숯불갈비	한식	...	37.5003623	127.0305686	...	02-553-2075	...
9	어부원	한식		37.5035903	127.0263169		02-3486-9995	
....	...	...	...	...	...	...	...	...
100	츠루하시 후게츠	일식	...	37.5022537	127.0301834	...	02)564-5922	...

2. 실험군 구성

본 연구에서 제시하는 상황정보를 이용한 하이브리드 필터링 추천시스템의 효용성 및 안정성을 평가하기 위해 실시한 실험에 참여한 인구통계학적 빈도 분포와 실험군의 기본적인 특성은 다음과 같다.

[표 5-2] 실험군 구성

분류	구분	빈도	비율
성별	남	41	68%
	여	19	32%
연령	20~30 대	32	53%
	30~40 대	28	47%
스마트폰 사용 여부	사용	40	67%
	미사용	20	33%
스마트폰 POI추천 애플 사용 경험 유무	경험 유	30	50%
	경험 무	30	50%
온라인 추천시스템(쇼핑몰, 서점 등) 사용 경험 유무	경험 유	27	45%
	경험 무	33	55%

성별 분포로 살펴보면 남성이 68%, 여성이 32%로 남성의 비중이 높다. 연령대의 경우 20~30 대의 비중이 높으며 스마트 폰 사용 여부는 67%로 최근 스마트폰의 확장이 가속화되고 있음을 알 수 있다. 스마트폰 POI(Point of Interest) 관련 애플리케이션 경험 유무는 50%로 나타났으며 과거 온라인 쇼핑몰 혹은 온라인 서점 등의 사이트에서 추천시스템을 이용한 경험은 45%로 나타났다.

제 3 절 실험 및 실험 결과

1. 실험

가. 1단계 실험

추천 시스템 애플리케이션 실행 이후 모든 아이템 보기에서 제공되는 아이템 중 추천 대상자가 가장 선호하는 아이템 10개를 선택하고 선호(평점)를 입력하게 하였다. 이는 추천 대상자 및 모든 사용자의 성향 및 상관관계를 파악하기 위한 기초데이터로 활용된다.



<그림 5-2> 임의 아이템 추천화면

10개의 선호 아이템 선정 및 선호(평점)입력을 마친 후 추천시스템을 재실행 하여 임의의 추천목록을 생성한 후 추천대상자에게 제공하였다. 사용자가 아이템 추천 결과에 만족하는지, 사용자의 의사결정에 영향을 미쳤는지 설문 을 조사를 실시하였다.

나. 2단계 실험

1단계 실험에서 수집된 기초 데이터를 바탕으로 하여 2단계 실험을 실시하였다. 1단계 실험과 마찬가지로 최초 애플리케이션 실행 시 가장 선호하는 아이템 10개를 선정하고 선호(평점)정보를 입력하도록 하였다. 두 번째 추천시스템 실행 시 실험과정에서 수집된 모든 기초 데이터를 바탕으로 가장 사용자들의 선호(평점) 평균이 가장 높은 순으로 5개의 아이템을 선별하여 추천하였다.

추천서비스 제공 후 사용자가 아이템 추천 결과에 만족하는지, 사용자의 의사결정에 영향을 미쳤는지 설문을 조사를 실시하였다.



<그림 5-3> 베스트 아이템 추천화면

다. 3단계 실험

1, 2단계 실험에서 수집된 기초 데이터를 바탕으로 하여 3단계 실험을 실시하였다. 전 단계 실험과 마찬가지로 최초 애플리케이션 실행 시 가장 선호하는 아이템 10개를 선정하고 선호(평점)정보를 입력하도록 하였다. 두 번째 추천시스템 실행 시 실험과정에서 수집된 모든 기초 데이터를 바탕으로 사용자간의 상관관계를 분석하고 추천 대상 아이템의 예측 선호를 산출하는 사용자 협업 필터링을 기반으로 추천 아이템을 선정하여 추천하였다. 이 때 해당 추

천 아이템과 아이템에 대한 추천대상자의 예측 선호(평점)를 함께 표시하여 사용자의 의사결정을 돕도록 하였다. 추천서비스 제공 후 사용자가 아이템 추천 결과에 만족하는지, 사용자의 의사결정에 영향을 미쳤는지 설문을 조사를 실시하였다.



<그림 5-4> 사용자 협업 필터링 추천화면

라. 4단계 실험

1, 2, 3단계 실험에서 수집된 기초 데이터를 바탕으로 하여 4단계 실험을 실시하였다. 전 단계 실험과 마찬가지로 최초 애플리케이션 실행 시 가장 선호하는 아이템 10개를 선정하고 선호(평점)정보를 입력하도록 하였다. 두 번째 추천시스템 실행 시 실험과정에서 수집된 모든 기초 데이터를 바탕으로 사용자 간 및 아이템간의 상관관계를 분석하고 추천 대상 아이템의 예측 선호를 산출하는 하이브리드 필터링을 기반으로 추천 아이템을 선정하여 추천하였다. 이 때 해당 추천 아이템과 아이템에 대한 추천대상자의 예측 선호(평점)를 함께 표시하여 사용자의 의사결정을 돕도록 하였다. 추천서비스 제공 후 사용자가 아이템 추천 결과에 만족하는지, 사용자의 의사결정에 영향을 미쳤는지 설문을 조사를 실시하였다.



<그림 5-5> 하이브리드 필터링 추천화면



2. 실험 결과

가. 1단계 실험

실험의 전 과정에서 사용자 및 아이템 상관관계 분석에 유용하게 이용될 기초데이터 생성을 위해 추천대상자에게 1차 추천시스템 애플리케이션 실행 시 모든 아이템 보기 서비스를 통해 가장 선호하는 10개의 아이템을 선정하고 선호(평점) 정보를 입력하도록 하였다. 다음 단계로 사용자에게 추천서비스를 제공하기 위해서 2차 추천시스템 애플리케이션 실행을 권유하고 임의의 추천목록을 생성하여 제공하였다. 추천된 아이템에 대해 만족하는지, 사용자의 의사결정에 영향을 미쳤는지에 대한 결과는 다음과 같다.

[표 5-3] 1단계 실험 결과

분류	매우만족 (%)	만족 (%)	보통 (%)	불만족 (%)	매우불만족 (%)
추천 아이템 리스트 만족도	0	27	27	33	13
사용자 의사 결정 영향	0	20	40	27	13

나. 2단계 실험

1단계 실험과 동일하게 기초데이터 생성을 위해 추천대상자에게 1차 추천시스템 애플리케이션 실행 시 모든 아이템 보기 서비스를 통해 가장 선호하는 10개의 아이템을 선정하고 선호(평점) 정보를 입력하도록 하였다.

사용자에게 추천서비스를 제공하기 위해서 2차 추천시스템 애플리케이션 실행을 권유하고 기 수집된 모든 사용자 및 아이템 선호(평점)정보를 이용하여 평균 선호(평점)가 가장 높은 5개의 추천목록을 생성하여 제공하였다. 추천된 아이템에 대해 만족하는지, 사용자의 의사결정에 영향을 미쳤는지에 대한 결과는 다음과 같다.

[표 5-4] 2단계 실험 결과

분류	매우만족 (%)	만족 (%)	보통 (%)	불만족 (%)	매우불만족 (%)
추천 아이템 리스트 만족도	6	27	40	27	0
사용자 의사 결정 영향	14	20	33	20	13

2단계 실험에서는 1단계 실험과 비교하여 추천 아이템 리스트에 대한 만족도가 높은 것으로 나타났다. 이는 1단계 및 2단계 실험에서 가장 많이 선택되었으며 평점이 높은 아이템 위주로 추천이 되었기 때문이다.

다. 3단계 실험

1, 2단계 실험과 동일하게 기초데이터 생성을 위해 추천대상자에게 1차 추천시스템 애플리케이션 실행 시 모든 아이템 보기 서비스를 통해 가장 선호하는 10개의 아이템을 선정하고 선호(평점) 정보를 입력하도록 하였다.

사용자에게 추천서비스를 제공하기 위해서 2차 추천시스템 애플리케이션 실행을 권유하고 기 수집된 모든 사용자 및 아이템 선호(평점)정보를 이용하여 사용자 협업 필터링을 적용하여 추천대상 아이템을 선정하고 추천대상 아이템에 대한 사용자의 예측선호(평점) 값을 산출 한 뒤 추천대상 아이템 리스트와 함께 제공하였다. 이 때 사용자가 특정 아이템에 대해 선호(평점)정보를 기 입력했을 시 이 아이템은 추천의 대상에서 자동적으로 제외하도록 하였다. 즉 사용자가 경험하지 않았던 아이템들을 대상으로 예측선호(평점) 값을 산출하고 추천을 실행 하는 것이다. 추천된 아이템에 대해 만족하는지, 사용자의 의사결정에 영향을 미쳤는지에 대한 결과는 [표 5-5]와 같다.

추천 아이템에 대한 만족도 및 사용자 의사결정에 영향은 2단계의 실험인 베스트 아이템 추천 실험보다 높은 것으로 평가 되었다.

[표 5-5] 3단계 실험 결과

분류	매우만족 (%)	만족 (%)	보통 (%)	불만족 (%)	매우불만족 (%)
추천 아이템 리스트 만족도	27	33	27	13	0
사용자 의사 결정 영향	20	47	20	13	0

라. 4단계 실험

1, 2, 3단계 실험과 동일하게 기초데이터 생성을 위해 추천대상자에게 1차 추천시스템 애플리케이션 실행 시 모든 아이템 보기 서비스를 통해 가장 선호하는 10개의 아이템을 선정하고 선호(평점) 정보를 입력하도록 하였다.

사용자에게 추천서비스를 제공하기 위해서 2차 추천시스템 애플리케이션 실행을 권유하고 기 수집된 모든 사용자 및 아이템 선호(평점)정보를 이용하여 하이브리드협업 필터링을 적용하여 추천대상 아이템을 선정하고 추천대상 아이템에 대한 사용자의 예측선호(평점) 값을 산출 한 뒤 추천대상 아이템 리스트와 함께 제공하였다. 3단계와 마찬가지로 사용자가 선호(평점)를 기 입력한 아이템은 추천대상에서 자동적으로 제외된다. 추천된 아이템에 대해 만족하는지, 사용자의 의사결정에 영향을 미쳤는지에 대한 결과는 다음과 같다.

[표 5-6] 4단계 실험 결과

분류	매우만족 (%)	만족 (%)	보통 (%)	불만족 (%)	매우불만족 (%)
추천 아이템 리스트 만족도	20	47	27	6	0
사용자 의사 결정 영향	20	53	20	7	0

추천 아이템 리스트에 대한 만족도는 매우만족 및 만족을 합하여 약 67%의 사용자가 긍정적인 반응을 보였으며 이는 3단계 실험과 비교 할 때 조금 높아진 수치이다. 사용자 개인의 상황인식과 하이브리드 필터링을 이용하여 구성된 추천리스트를 제공하였기 때문이라고 볼 수 있다.

1, 2, 3, 4 단계 실험에서 중에서 하이브리드 필터링을 이용하여 제공한 추천리스트에 대한 피실험자의 반응이 가장 좋았다.

본 연구에서는 기존의 추천시스템에 관한 연구와의 차별화를 위해 상황인식 필터링을 기반으로 사용자의 상황정보를 이용하여 사용자의 상황에 적합한 정보를 추천함으로써 추천의 질을 높이기 위한 방안을 제시했다. 각 실험 단계의 설문과정에서 사용자 협업 필터링 혹은 하이브리드 필터링 전 단계에 수행되는 상황인식을 통한 정보의 필터링이 정보추천의 질에 미치는 영향에 대한 설문을 실행하였으며 이에 대한 결과는 다음과 같다.

[표 5-7] 상황정보 활용 유용성 평가 결과

분류	매우만족 (%)	만족 (%)	보통 (%)	불만족 (%)	매우불만족 (%)
상황인식 필터링 유용성	22	51	25	2	0



제 4 절 하이브리드 필터링 검증

1. 정확도와 재현율

정확도(Precision)와 재현율(Recall)은 검색엔진의 질을 판단하는 척도로서 활용되어 왔으며 Billsus와 Basu의 연구에서 협업 필터링 기법을 적용한 추천 시스템을 평가하는 도구로 사용되어졌다. 정확도와 재현율에 대한 검색시스템에서의 정의는 다음과 같다.

[표 5-8] 검색 시스템에서 정확도와 재현율의 정의

분류		정의
검색시스템	정확도	전체 웹의 검색키워드 관련 문서집합이 A, 검색 결과 문서 집합이 B(관련문서, 비관련문서 혼합) 일 때 정확도는 $ A \cap B / B $
	재현율	전체 웹의 검색키워드 관련 문서집합이 A, 검색 결과 문서 집합이 B(관련문서, 비관련문서 혼합) 일 때 정확도는 $ A \cap B / A $

예를 들어 정확도의 경우, 검색결과가 한 개이며 그 문서가 관련된 문서라면 정확도는 1(=100%)이 된다. 만약 검색 시스템이 특정 키워드에 대해 전체 웹상의 모든 문서를 검색해냈다면 재현율은 100%가 되나, 정확도는 상당히 낮을 것이다. 이에 따라 정확도가 높아지면, 재현율은 낮아지고, 재현율이 높아지면 정확도는 낮아지는 반비례 관계가 성립한다. 이를 요약 및 정리하면 정확도는 시스템이 오답을 찾지 않을 확률이고 재현율은 시스템이 얼마나 많은 정답을 찾아내는가에 대한 확률이라 할 수 있다.

이처럼 주로 정보검색 시스템의 성능을 평가하는 방법으로 많이 사용되었던 정확도와 재현율은 Billsus와 Basu에 의해 추천시스템의 성능을 평가하는 방법으로 확장되었다. 추천시스템에 적용된 정확도와 재현율의 주로 추천 리스트를 평가하기 위한 방법으로 사용되며, 정확도는 추천 리스트 중 몇 개의 항목을 실제로 좋아했는지를 나타내는 평가 방법이며, 재현율은 고객이 좋아하

는 항목 중 얼마나 많은 항목이 추천되었는지 나타내는 평가 방법이다. 이를 정리하면 다음과 같다.

[표 5-9] 추천 시스템에서 정확도와 재현율의 정의

분류		정의
추천시스템	정확도	추천리스트 중 몇 개의 항목을 사용자가 실제로 만족하였는지 나타내는 평가 방법
	재현율	고객이 좋아하는 항목 중에서 얼마나 많은 항목이 추천이 되었는지 나타내는 평가 방법

정확도와 재현율에 동등한 가중치를 부여하여 하나의 평가방법으로 사용하는 F-measure는 다음과 같이 구할 수 있다.

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

2. 검증 결과

추천 시스템의 성능을 평가하기 위해 본 연구에서 개발된 추천시스템에서 제공하는 사용자 협업 필터링 추천 및 하이브리드 필터링 추천 서비스를 통해 사용자에게 제공되는 추천 리스트의 정확도와 재현율 그리고 F-measure에 대한 분석을 수행하였다. 이를 위해 본 실험에서 행해진 1, 2, 3, 4 단계의 실험을 통해 수집된 데이터를 이용하였으며 그 결과는 다음과 같다.

[표 5-10] 상황정보를 이용한 하이브리드 추천 시스템 정확도 및 재현율

기법	정확도	재현율	F-measure
사용자 협업 필터링	53.8	42.2	47.2
하이브리드 필터링	63.3	49	55.2

정확도, 재현율, F-measure에 대한 분석 결과에 따르면 본 연구에서 개발

된 상황정보를 이용한 하이브리드 필터링 추천시스템에서 제공되는 서비스 중 사용자 협업 필터링으로 제공되는 추천 목록보다 하이브리드 필터링을 이용하여 제공되는 추천 목록에 대한 사용자의 만족도가 높은 것으로 평가 되었다.

본 5장에서는 상황정보를 이용한 하이브리드 필터링 추천시스템 평가의 일환으로 실험 시나리오를 설정하여 총 4단계의 절차로 수행하였다. 정보필터링이 수행되기 전 목적 및 위치에 대한 상황정보를 활용하여 정보의 범주를 결정되었다고 가정하고, 임의추천, 베스트아이템 추천, 협업필터링추천, 하이브리드 필터링 추천을 실행하였다. 실험결과 본 추천시스템에서 제공하는 추천 기법 중 하이브리드 필터링을 이용한 추천목록에 대한 사용자의 만족도가 가장 높은 것으로 나타났다. 또한 추천시스템의 일반적인 평가척도인 정확도(Precision), 재현율(Recall), F-measure에 대한 분석 결과 역시 하이브리드 필터링 기법을 이용하여 추천리스트를 제공하는 것이 좋은 결과를 나타내는 것으로 평가되었다. 본 연구에서는 100여개의 제한된 음식점 정보와 적은 수의 실험군을 대상으로 실험하였지만 더 많은 아이템 정보를 수집한다면 다양한 사용자 목적 및 목적지 혹은 현 위치 등의 상황에 따른 사용자의 선호도에 적합한 아이템을 추천할 수 있을 것이다. 또한 여러 사용자들을 대상으로 실험할수록 사용자간의 유사도를 산출하기가 수월해짐으로 더 정확한 추천 서비스가 가능할 것으로 예상된다.

제 6 장 결 론

본 연구에서는 대표적인 추천기법인 협업필터링의 두 방식 즉, 사용자 기반 협업 필터링 및 아이템 기반 협업 필터링을 결합한 하이브리드 필터링 이용하여 모바일 환경에서 추천시스템을 설계하고 구축했다. 이 때 추천시스템의 예측력을 향상시키고 및 추천의 질을 높이기 위한 방안으로 모바일에서 활용 가능한 사용자의 목적행위 및 위치의 상황정보를 이용하여 제공되어야 할 정보의 범주를 결정한 후 하이브리드 필터링을 적용하도록 했다. 부가적으로 본 연구에서 설계된 추천시스템을 토대로 제공되는 몇 가지 추천방식에 대한 평가를 수행했다. 사용자의 목적행위 및 위치가 일정하게 가정된 상황에서 일반적인 협업 필터링보다 하이브리드 필터링의 추천결과가 더 좋은 만족도를 보이는 것으로 나타났다.

본 연구 결과의 몇 가지 한계점으로 먼저 다양한 상황정보의 활용이 미비하다는 점을 들 수 있다. 상황인식은 현재 상황의 상태를 실체화, 특성화 하여 이용하는 것을 의미하며 크게 사용자 상황(User context), 물리적 상황(Physical context), 컴퓨터 상황(Computing context), 사용자 특성 및 상호이력 등으로 구성된다. 세부적으로 네트워크 연결 상태, 통신 대역폭, 그리고 프린터·디스플레이·워크스테이션과 같은 컴퓨팅 상황, 사용자의 프로파일·위치·주변의 사람들을 비롯한 사용자 상황, 조명·소음 레벨·교통 상태·온도 등 물리적 상황 시간·주·달·계절 등 시간적 상황 등이 있다. 본 연구에서 하이브리드 필터링의 매개체로 이용되는 아이템에 대한 사용자의 선호 정보 역시 상황 정보의 한 형태로 볼 수 있다.

본 연구에서 설계 및 구현된 추천시스템은 목적, 위치, 선호 세 가지를 이용하나 좀 더 다양한 상황 정보를 활용하지 못한다는 점이 한계라 할 수 있다. 또한 시스템이 사용자에게 대한 상황을 동적으로 인식하는 단계가 아닌 사용자의 입력을 토대로 상황정보가 추출된다는 데에 한계가 있다.

모바일 디바이스의 특성 상 기기를 사용하는 개인의 나이, 성별, 지역, 직업 등의 사용자 상황 뿐 아니라 시간·주·달·계절 등 시간적 상황 등에 대한 정보를 활용할 수 있는 가능성이 존재 한다. 차후 이러한 다각적인 상황정보를 활

용함으로써 사용자에게 대한 고차원적인 분석을 통해 추천 서비스를 제공할 필요가 있다.

두 번째 한계점은 본 연구에서 사용자 및 아이템간의 유사도를 측정하기 위해 주로 사용되는 피어슨 상관계수의 고질적인 한계점인 희박성의 문제를 해결하지 못했다는 점이다. 희박성의 문제는 시스템의 데이터베이스에 아이템의 수가 기하급수적으로 늘어날 시 고객의 선호도가 입력되지 않은 콘텐츠의 수가 상대적으로 증가함으로써 새로운 아이템에 대한 추천을 원천적으로 봉쇄한다는 것이다. 연구 진행의 과정 중 이를 해결하기 위한 새로운 방법론의 필요성이 인식되었으나, 이를 개발하지 못한 것이 본 연구의 한계점이라 할 수 있다. 희박성의 문제를 해결하기 위한 방법 중 하나의 예로 가상 머신에 의한 아이템 평점 입력을 들 수 있다. 이는 새롭게 아이템 데이터베이스에 입력된 관계로 사용자의 선호도가 입력되지 않은 아이템에 대해 특정 가상 머신이 일정 평점을 입력하는 형태로서 단순한 통계적 데이터를 바탕으로 하기보다는 사용자와 새로운 아이템간의 특성에 대한 상관관계 분석을 통해 수행되어야 한다. 향후 사용자의 특성 및 아이템에 대한 특성 정보를 활용함으로써 희박성의 문제에 대한 한계점을 해결해야 할 필요가 있다.

HANSUNG
UNIVERSITY

【참고문헌】

1. 국내문헌

국토연구원(2010) “이용자맞춤형 대중교통서비스 기술개발 최종보고서”

김병만, 이경, 김시관, 임은기, 김주연(2004), “추천시스템을 위한 내용기반 필터링과 협력필터링의 새로운 결합 기법,” 『정보과학회논문지 : 소프트웨어 및 응용』, 제 31 권, 제 3 호, pp. 332-342

김재경, 안도현, 저운호(2005), “개인별 상품 추천 시스템, WebCF-PT; 웹마케팅과 상품계층도를 이용한 협업필터링”, 『경영정보학연구』, 제 15권, 제1호, pp.63-79

김종우, 배세진, 이홍주(2004), “협업 필터링 기반 개인화 추천에서의 평가 자료의 희소정도의 영향”, 『경영정보학연구』, 제14권, 제2호, pp. 131-149

김진, 윤정섭, 조근식(2000), “다중 정보 여과 방법을 이용한 동적 정보 우선순위 결정”, 한국지능정보시스템 학회, pp. 323-332.

박지선, 김택현, 류영석, 양성봉(2002), “추천시스템을 위한 2-way 협동적 필터링 방법을 이용한 예측 알고리즘”, 『정보과학회지 논문지』, 29권, 9,10호, pp.669-675

박철제, 전종근(2002), “웹상에서 정보검색 효율성 제고를 위한 검색결과 제시 방법의 개선 방안”, 『대한경영학회지』, 제34호, pp.225-237

손창환(2006), “web 상에서 개인화된 상품추천을 위한 Hybrid 추천시스템에 관한 연구” 영남대 대학원

솔트룩스(2010), <http://semanticwiki.saltlux.com/index.php/추천기술>

송미랑, 김교정(1999), “사용자 그룹을 이용한 효과적인 정보 여과 및 학습방법에 관한 연구”, 한국정보과학회 『가을 학술발표논문집』, Vol. 26, No. 2., pp. 63-65

- 안신현(2007), “콘텐츠 추천 시스템에 관한 이슈”, KAIST 문화기술대학원
- 이도경(2010) “상황인식”, 제너시스템즈
- 정보통신용어사전(2010), “상황인식 컴퓨팅”, 한국정보통신기술협회
- Atlee, T. (2008). Reflections on the evolution of choice and collective intelligence, Retrieved August 26, 2008
- Breese, J.S., Heckerman, D. and Kadie, C(1998), "Empirical analysis of predictive collaborative filtering“, In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence(UAI-98), pp. 43-52.
- Burke, R(2002), "Hybrid Recommender Systems : Survey and Experiment, User Modeling and User-Adapted Interaction, Vol. 12, Iss. 4, pp.331-370
- CIA. (2008). The World Factbook. (accessed 3 September 2008)
- Cristiana Bolchini and Carlo A. Curino and Elisa Quintarelli and Fabio A. Schreiber and Letizia Tanca (2007). "A data-oriented survey of context models" (PDF). SIGMOD Rec. (ACM) 36 (4): 19--26.
- Dey, Anind K. (2001). "Understanding and Using Context". Personal Ubiquitous Computing 5 (1): 4 - 7
- Flew, Terry (2008). New Media: an introduction. Melbourne: Oxford University Press.
- George Pór, Blog of Collective Intelligence
- Goldberg, D., Nichols, D., oKI, B.M. and Terry,D.(1992), "Using collaborative filtering to weave an information tapestry," Communications of the ACM, Vol. 35, No. 12, pp. 61-70.
- Howard Bloom(1995), The Lucifer Principle: A Scientific Expedition Into the Forces of History
- Leiner, Barry, Cerf, Vinton, Clark, David, Kahn, Robert, Kleinrock, Leonard, Lynch, Daniel, Postel, Jon, Roberts, Larry and Wolff, Stephen(2003)

A Brief History of the Internet. Version 3.32 (accessed 3 September 2008)

Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.(2000)", "Analysis of Recommendation Algorithms for E-Commerce," Proceedings of the 2nd ACM conference on Electronic commerce, pp.158-167

_____ (2001), "Item-based Collaborative Filtering Recommender Algorithms," In Proceeding of 10th International WWW Conference, pp.285-295.

Schmidt, Albrecht (2003). "Ubiquitous Computing - Computing in Context". PhD dissertation, Lancaster University.

Shardanand, U., and Maes, p.(1995) Social Information Filtering: Algorithms for Automating "World of Mouth". Proceeding of the SIGCHI conference on Human factors in computing systems, pp 210-217

Tapscott, D., & Williams, A. D. (2008). Wikinomics: How Mass Collaboration Changes Everything, USA: Penguin Group

Weiser, Mark(1991). "The Computer for the 21st Century". <http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html>. Retrieved 2008-05-07.

_____ (1996). "Ubiquitous computing". <http://www.ubiq.com/hypertext/weiser/UbiHome.html>. Retrieved 2007-11-03.)

Weiser; Gold;Brown(1999). "Ubiquitous computing". <http://www.research.ibm.com/journal/sj/384/weiser.html>. Retrieved 2008-05-07.

wikipedia(2010) <http://www.wikipedia.org>, "Collaborative Filtering"

ABSTRACT

Development of Hybrid Filtering Recommendation System using Context-Information in Mobile Environments

Ko Jung-Min

Major in Information System Engineering

Dept. of Information System Engineering

Graduate School, Hansung University

According to today's rapid growth and development of telecommunication information technology, interest has been amplified regarding ubiquitous network computing. In addition, interest has been also being increasing regarding appropriate service based on user's propensity, interest such as Internet, e-commerce, telecommunications or regarding customized services in various fields associated with public life and activities.

the rapid development of related technologies and has been a big spotlight. Smart phone, with features such as a PC with advanced features is a mobile phone. The rapid development of mobile technology, anywhere, anytime to receive the desired information and services can mean that the arrival of ubiquitous network environment. According to environment and infrastructure development, a variety of mobile-based application software to provide various kinds of information and services has been released in recent mobile environment. However, most of them are provider-driven information systems and aim to provide large amounts of information simply to an unspecified number of users. Therefore, customized or personalized provision of information and service explained earlier for

individual users has been hardly come true. Personalized or customized information and services are needed to be provided throughout interaction with users by escaping these limitations. In other words, more advanced types of user-centered information system are necessary in step with ubiquitous computing and information technology with rapid development.

According to background and need, this study wants to design and implement recommendations system for personalization and customization in mobile environments. Information filtering called key element of recommendation system uses Collaborative Filtering known as the most effective techniques in the latest Web and e-commerce recommendation system. To acquire more accurate recommendation results, recommendation system shall be composed using the Hybrid Filtering with advantage's combination of each kind of information filtering technique as well as simple collaborative filtering. Effective information recommendation according to user's situation by using user's context-information of purpose and location that are available in mobile devices before running the filtering of the information to improve the quality of recommendations.

Through this study, efficient and convenient obtaining information and providing services would be possible based on communications equipment including future mobile terminals by implementing information recommendation system using user situation information and information filtering techniques. In addition, the results of this study may be expected to show the development direction of study regarding personalized and customized systems interlinked with ubiquitous network computing technologies ultimately.

Keywords Mobile, Hybrid filtering, Recommendation system, Context