

---

# 시맨틱 웹에서 개인화된 선호도를 이용한 의상 코디 시스템 개발

## Development of Apparel Coordination System Using Personalized Preference on Semantic Web

---

은채수\*, 조동주\*\*, 정경용\*\*, 이정현\*\*  
인하대학교 컴퓨터정보공학과\*, 상지대학교 컴퓨터정보공학부\*\*

Chae-Soo Eun(chaesoo79@hanmail.net)\*, Dong-Ju Cho(queen8181@sangji.ac.kr)\*\*,  
Kyung-Yong Jung(kyjung@sangji.ac.kr)\*\*, Jung-Hyun Lee(jhlee@inha.ac.kr)\*\*

---

### 요약

인터넷과 웹이 일상생활의 일부가 되면서 온라인상에는 방대한 양의 정보가 쌓이게 되었다. 이러한 흐름 속에서 정보의 양은 급속도로 늘어나는 현상을 보이며, 개인화를 통해 수많은 데이터들 사이에서 원하는 정보를 자동으로 찾아내는 기술의 중요성이 부각되고 있다. 현재 사용하는 필터링 중에서 콘텐츠를 중심으로 분석하여 사용자에게 추천하는 기법인 내용기반 필터링과 사용자와 유사한 선호도를 가진 사용자 군집의 선호도에 따라 새로운 사용자가 관심을 가질 것으로 생각되는 콘텐츠를 추천해 주는 기법인 협력적 필터링 기법이 있다. 그러나 협력적 필터링 방법으로 추천 받기 위해서는 특정 수 이상의 아이템에 대한 평가가 필요하며, 또한 비슷한 성향을 가지는 일부 사용자 정보에 근거하여 추천함으로써 나머지 사용자 정보를 무시하는 경향이 있다. 따라서 특정 수 이상의 선호정보가 준비되지 않은 사용자들에 대해서도 적절한 추천방법이 필요하다. 본 논문에서는 기존의 필터링들을 조합하고 좀 더 편리하게 정보를 공유하고 학습할 수 있는 시맨틱 웹에서 개인화된 선호도를 이용한 의상코디 시스템을 개발하였다. 이 시스템을 웹에서 제공한 결과 불필요한 검색시간이 줄어들고 사용자의 피드백을 통해 점차 만족도가 향상됨을 알 수 있었다.

■ 중심어 : | 협력적 필터링 | 연관규칙 | 추천 시스템 | 시맨틱 웹 | 의상코디 |

### Abstract

Internet is a part of our common life and tremendous information is cumulated. In these trends, the personalization becomes a very important technology which could find exact information to present users. Previous personalized services use content based filtering which is able to recommend by analyzing the content and collaborative filtering which is able to recommend contents according to preference of users group. But, collaborative filtering needs the evaluation of some amount of data. Also, It cannot reflect all data of users because it recommends items based on data of some users who have similar inclination. Therefore, we need a new recommendation method which can recommend proper items without preference data of users. In this paper, we proposed the apparel coordination system using personalized preference on the semantic web. This paper provides the results which this system can reduce the searching time and advance the customer satisfaction measurement according to user's feedback to system.

■ keyword : | Collaborative Filtering | Association Rule | Recommendation System | Semantic Web | Apparel Coordination |

---

## 1. 서론

최근 시맨틱 웹의 등장으로 자동화 기술을 이용해 좀 더 편리하게 정보를 공유하고 학습할 수 있게 되었다[1]. 현재 추천 시스템은 주로 협력적 필터링을 사용하고 있는데, 이것은 특정 사용자와 유사한 사용자 군집을 구성해 이들의 선호도를 추천에 반영하는 방식이다. 그러나 이 방식은 초기의 콘텐츠에 대해서는 아무런 추천 기능을 수행할 수 없고 어느 정도 사용자들의 프로파일이 모이기 전에는 자신과 유사한 제품만을 사용한 특정 사람을 찾기가 쉽지 않다는데 문제가 있다. 이를 보완하기 위해 내용기반 필터링 등 다양한 방식들을 함께 사용하고 있다[2]. 특히 옷의 경우에는 어느 시대에 어떤 의도로 이를 구매했는가에 따라 삶의 모습과 다양한 문화를 나타내므로 또래의 선호도는 물론 지역, 시대 등의 복합적인 요소를 고려해야 한다. 따라서 본 논문에서는 내용기반 필터링과 협력적 필터링을 적절하게 조합하고 시맨틱 웹에서 해석된 성별, 나이, 직업, 선호도 등의 메타 데이터로부터 콘텐츠의 특성과 사용자들의 유사점을 추출하고, 연관 이웃 마이닝을 사용하여 개인화 시스템에서 중요한 잘못된 정보를 줄이고자 한다.

## 2. 관련연구

### 2.1 시맨틱 웹

사용자가 원하는 내용과 무관한 많은 양의 내용으로 인해 사용자들은 웹의 검색을 효율적으로 사용할 수가 없었다. 그래서 표현은 같지만 뜻이 다른 데이터를 본래 의미로 해석하고 인식하며, 표준화된 정보를 제공해 줄 수 있는 환경이 시맨틱 웹이다[3]. 시맨틱 웹은 XML과 RDF 기술에 크게 의지하고 있다. 하지만 시맨틱 웹의 표준 언어로 사용 하는 데는 한계가 있다. XML로 정의한 태그는 사람이 이해하기는 어렵지 않으나 기계가 자동적으로 그 의미를 이해하기는 어렵다. 특히 태그 사이의 의미 연관성을 추론하기란 매우 어렵다. 임의로 만든 태그의 사용법은 알 수 있어도 어떤 의미를 담고 있는 태그인지 파악하기는 어려운 것이 XML의 단점인 것이다. RDF는 이런 문제를 해결하기 위해 제시된 기술이다. RDF는

특정 자원에 대한 숨은 자료를 설명하는 XML 기반의 프레임워크이다. RDF는 자원, 속성, 속성값을 묶어서 하나의 단위로 취급하는 기술로 정보를 구성하는 자원에 대한 좀 더 세밀한 설명과 관계 파악을 쉽게 해준다. 즉 RDF라는 기술을 이용하면 문서에서 사용되는 요소의 의미와 문서 사이의 관련성 표시가 쉬워지고, 기계끼리 자동화 처리가 한결 쉬워진다.

### 2.2 개인화

개인화는 사용자의 프로파일에 근거 하여 내용을 선정하고, 사용자에게 새롭거나 다른 사용자가 미처 생각하지 못했던 것을 채택하는 것이다. 즉, 사용자의 특성을 대상으로 하여 이와 관련한 다양한 통계의 분석과 비교를 통해 개별화 할 수 있는 지식 및 규칙을 찾아내는 과정이다. 이러한 개인화를 통해서 웹 사이트 운영자는 사용자에 관한 자료를 얻고 사용자의 지속적인 이용이나 구매를 얻어낼 수 있게 되며 사용자는 자신에게 가장 알맞은 정보를 편리한 방법으로 얻을 수 있게 된다. 대부분의 개인화 기술은 어느 정도 사용자의 개인 정보를 필요로 한다. 사용자 등록 과정을 통해 사용자의 인적사항이나 선호도에 관한 질문을 하는 경우도 있고, 사용자의 사이트 내에서의 행동을 관찰하는 방법도 있지만 개인화 과정 자체가 사용자의 선호도나 습관, 구매행태와 같은 정보를 수집하여 사용자에게 알맞은 정보를 제공하는 것이 제안하는 방법의 목표이다.

### 2.3 협력적 필터링 기술

협력적 필터링은 사용자의 선호도에 대한 데이터를 기반으로 새로운 사용자가 관심을 가질 것으로 생각되는 아이템(상품, 광고, 웹 페이지 등)을 추천해 주는 기법이다. 협력적 필터링은 아이템에 대한 다른 사용자들의 선호를 기반으로 하기 때문에 협력적이라는 용어를 사용하게 된다. 협력적 필터링 시스템은 시스템이 사용자의 암묵적인 데이터를 사용하는지 명시적인 데이터를 사용하는지에 따라 구분을 한다. 명시적인 데이터는 사용자에게 특정 아이템에 대한 선호를 0에서 5까지 정도의 이산형 척도로 입력받는 경우를 말한다. 암묵적 데이터를 사용하는 경우는 사용자의 선호를 대변하는 사용자의 웹

사이트 클릭 패턴이나 구매 패턴 등을 웹 로그나 구매 이력 데이터에서 발견하여 특정 아이템에 대한 선호를 예측하는 것을 말한다.

### 2.4 ARHP 알고리즘에 의한 연관 사용자

ARHP 알고리즘은 연관 규칙과 하이퍼 그래프 분할을 이용하여 트랜잭션 기반의 데이터베이스에서 연관된 아이템들을 군집하는 방법이다. 하이퍼 그래프는 일반 그래프와 달리 각 하이퍼 간선이 두 개 이상의 정점에 연결될 수 있다는 점에서 그래프의 확장으로 볼 수 있다. 하이퍼 그래프 분할 알고리즘은 항목들 간의 거리가 아닌 가중치를 이용하기 때문에 아이템들 간의 거리 계산이 어려운 다차원 데이터 집합에 대한 군집에 유용하다. 이때 하이퍼 그래프 분할을 위한 가중치로 연관 규칙의 신뢰도를 사용한다. 본 논문에서 사용한 ARHP 알고리즘은 군집하기 위한 연관 사용자 집합들의 모든 연관 규칙과 신뢰도를 구한 후, 연관 규칙에 포함되는 사용자를 정점으로, 연관 관계를 하이퍼 간선으로 매핑 한다. 하이퍼 그래프 분할에서의 클러스터는 유사한 사용자들을 분류하거나 예측할 때 사용되고 관심없는 연관 규칙을 제거함으로써 연관 규칙의 차원 수를 감소시키는데 사용된다.

관리자는 아이템 데이터베이스의 새로운 정보의 생성, 불필요한 정보의 삭제, 기존 정보의 편집을 담당한다. 사용자 웹 인터페이스는 사용자의 요청을 의상 추천 서버를 통해 응답한다. 사용자 계정 데이터베이스는 인증에 사용되고 사용자 프로파일 데이터베이스는 개인화된 서비스를 제공하기 위해 각 사용자의 정보들을 저장한다. 의상 추천 결과에 따른 사용자의 피드백을 통해 프로파일은 계속 업데이트 되어 의상 추천시 반영된다. 의상 추천 서버는 필터링 된 의상 데이터베이스에서 사용자 관련 정보들을 기반으로 연관 이웃 마이닝을 적용하여 추천 결과를 보여준다.

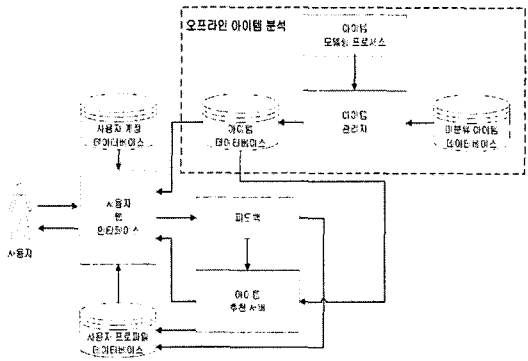


그림 1. 의상코드 시스템 구조도

### 3. 개인화된 선호도를 이용한 의상 코드 시스템

이 장에서는 사용자 정보와 선호도를 고려하여 의상을 추천하고 코딩하기 위해 상황을 정의하고 내용기반 필터링 및 협력적 필터링 기법과 연관 이웃 마이닝을 사용한다. 시스템의 기본적인 구성은 [그림 1]과 같다.

아이템 관련 정보들을 안정적으로 이용하기 위해서 오프라인 분석이 이루어진다. 만약 웹 서비스에서 인증된 사용자가 어떠한 아이템에 대해 추천을 해달라고 요청하면 아이템 추천 서버에서 분석된 아이템 데이터베이스를 기반으로 선호도를 고려하여 아이템을 추천한다. 이러한 정보들은 사용자 프로파일로 저장되고 추천 결과의 피드백을 통해 업데이트가 된다. 아이템 모델링 프로세스는 정의된 상황에 따라 미분류 아이템 데이터베이스로부터 필터링을 통해 아이템 데이터베이스를 추출한다. 아이템

### 3.1 의상 콘텐츠 정보 생성

본 시스템에서는 콘텐츠의 내용으로 의상 콘텐츠 정보 데이터를 구축하고 사용자 체형에 맞는 프로파일을 사용하여 의상을 추천한다. 내용 정보 데이터를 구축하기 위해 웹문서에서 자동으로 구축하는 방법과 사용자가 정보를 직접 입력하는 두 가지 방법을 사용한다. 자동구축에서는 웹 로봇 에이전트에 의해 웹문서를 추출하여 데이터베이스를 구축한다. [표 1]은 내용 정보 데이터베이스의 예이다.

표 1. 내용정보 데이터베이스의 예

의상 정보	내용
의상류	정장, 난방, 티셔츠, 후드티, 나팔바지, 자켓...
상의/하의	상의, 하의
색깔	빨강, 노랑, 파랑, 초록, 보라, 검정.....
계절	봄, 여름, 가을, 겨울

### 3.2 사용자간의 관련성 정보 생성

본 시스템에서는 사용자의 성별, 연령, 직업, 계절, 지역 등의 정보로 의상 콘텐츠 데이터를 구축하고 사용자 체형과 맞는 프로파일을 사용하여 의상을 추천한다. 협력 정보 데이터베이스를 구축하기 위해서는 웹문서에서 자동으로 구축하는 방법을 사용한다. [표 2]는 협력 정보 데이터베이스이다. 본 논문에서 제안한 개인화된 선호도를 이용한 의상 코디 시스템에서는 상황의 정확한 표현, 다양한 관계 기술 등 명시적 규정을 위하여 시맨틱 웹에서 사용되는 온톨로지로 상황을 정의하였다.

표 2. 협력 정보 데이터베이스

성별	연령	직업	계절	지역	선호하는 옷	색깔
남	10~20	학생	가을	서울	상의:가디건 하의:청바지	파란색 청 색
여	10~20	학생	여름	경기도	상의:나시 하의:치마	노란색 흰 색
남	20~30	회사원	봄	전라도	상의:난방 하의:정장치마	흰 색 검정색
여	20~30	선생님	가을	경상도	상의:브라우스 하의:정장치마	노란색 갈 색
남	30~40	자영업	겨울	강원도	상의:점퍼 하의:플렌바지	흰 색 갈 색
여	30~40	주부	여름	제주도	상의:반팔티 하의:반바지	빨간색 흰색

### 3.2 연관 사용자 군집

Apriori 알고리즘을 이용하여 사용자 트랜잭션으로부터 연관 사용자의 군집하는 절차를 보인다. 의상 데이터에서 사용자에 의해 선호도를 평가한 아이템들을 [표 3]의 사용자 트랜잭션으로 재구성한다.

표 3. 사용자 트랜잭션

트랜잭션 번호 (평가된 콘텐츠)	추출된 사용자
1 - 청바지	U <sub>1</sub> , U <sub>2</sub> , U <sub>3</sub> , U <sub>4</sub>
2 - 정장상의	U <sub>2</sub> , U <sub>3</sub> , U <sub>1</sub> , U <sub>5</sub> , U <sub>6</sub> , U <sub>7</sub> , U <sub>12</sub> , U <sub>8</sub>
3 - 정장하의	U <sub>9</sub> , U <sub>3</sub> , U <sub>2</sub> , U <sub>10</sub> , U <sub>5</sub> , U <sub>11</sub>
4 - 민소매티	U <sub>13</sub> , U <sub>3</sub> , U <sub>14</sub> , U <sub>15</sub> , U <sub>16</sub> , U <sub>17</sub>
5 - 점퍼	U <sub>18</sub> , U <sub>13</sub> , U <sub>3</sub>
6 - 반팔티	U <sub>13</sub> , U <sub>3</sub> , U <sub>19</sub> , U <sub>20</sub> , U <sub>15</sub>
7 - 후드티	U <sub>21</sub> , U <sub>22</sub>
8 - 면바지	U <sub>23</sub> , U <sub>24</sub> , U <sub>25</sub>

트랜잭션 번호는 사용자가 평가한 아이템을 의미하며 추출된 사용자는 후보 사용자의 집합과 고빈도 사용자 집합을 구성한다.

[표 3]의 사용자 트랜잭션으로부터 Apriori 알고리즘으로 연관 규칙을 [표 4]의 방법으로 마이닝한다[4][5]. Apriori 알고리즘은 첫 단계에서 후보 사용자 집합(C1)을 구성하며 이들의 지지도를 확인하기 위해 데이터베이스를 검색하고, 고빈도 사용자 집합(L1)을 구성할 수 있다. 이와 같은 방법으로 두번째 단계에서는 C2, L2 을 구성하며, 세번째 단계에서는 C3, L3를 구성한다. [표 5]에서 제시한 바와 같이 L3의 연관 사용자 집합, {U1, U2, U3}, {U2, U3, U5}, {U2, U5, U15}, {U3, U13, U15}으로 추출된다.

표 4. 연관 사용자 유추과정

후보 사용자 집합(C1)	U <sub>1(2)</sub> , U <sub>2(3)</sub> , U <sub>3(6)</sub> , U <sub>4(1)</sub> , U <sub>5(2)</sub> , U <sub>6(1)</sub> , U <sub>7(1)</sub> , U <sub>8(1)</sub> , U <sub>9(1)</sub> , U <sub>10(1)</sub> , U <sub>11(1)</sub> , U <sub>12(1)</sub> , U <sub>13(1)</sub> , U <sub>14(1)</sub> , U <sub>15(2)</sub> , U <sub>16(1)</sub> , U <sub>17(1)</sub> , U <sub>18(1)</sub> , U <sub>19(1)</sub> , U <sub>20(1)</sub> , U <sub>21(1)</sub> , U <sub>22(1)</sub> , U <sub>23(1)</sub> , U <sub>24(1)</sub> , U <sub>25(1)</sub>
고빈도 사용자 집합(L1)	U <sub>1(2)</sub> , U <sub>2(3)</sub> , U <sub>3(6)</sub> , U <sub>5(2)</sub> , U <sub>13(3)</sub> , U <sub>15(2)</sub>
후보 사용자 집합(C2)	(U <sub>1</sub> , U <sub>2</sub> ) <sub>(2)</sub> , (U <sub>1</sub> , U <sub>3</sub> ) <sub>(2)</sub> , (U <sub>1</sub> , U <sub>5</sub> ) <sub>(1)</sub> , (U <sub>1</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>1</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>2</sub> , U <sub>3</sub> ) <sub>(3)</sub> , (U <sub>2</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>2</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>3</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>3</sub> , U <sub>13</sub> ) <sub>(3)</sub> , (U <sub>3</sub> , U <sub>15</sub> ) <sub>(2)</sub> , (U <sub>5</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>5</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>13</sub> , U <sub>15</sub> ) <sub>(2)</sub>
고빈도 사용자 집합(L2)	(U <sub>1</sub> , U <sub>2</sub> ) <sub>(2)</sub> , (U <sub>1</sub> , U <sub>3</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>3</sub> ) <sub>(3)</sub> , (U <sub>2</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>3</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>3</sub> , U <sub>13</sub> ) <sub>(3)</sub> , (U <sub>3</sub> , U <sub>15</sub> ) <sub>(2)</sub> , (U <sub>13</sub> , U <sub>15</sub> ) <sub>(2)</sub>
후보 사용자 집합(C3)	(U <sub>1</sub> , U <sub>2</sub> , U <sub>3</sub> ) <sub>(2)</sub> , (U <sub>1</sub> , U <sub>2</sub> , U <sub>5</sub> ) <sub>(0)</sub> , (U <sub>1</sub> , U <sub>2</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>1</sub> , U <sub>2</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>1</sub> , U <sub>3</sub> , U <sub>5</sub> ) <sub>(1)</sub> , (U <sub>1</sub> , U <sub>3</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>1</sub> , U <sub>3</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>2</sub> , U <sub>3</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>3</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>2</sub> , U <sub>3</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>2</sub> , U <sub>5</sub> , U <sub>15</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>5</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>3</sub> , U <sub>5</sub> , U <sub>13</sub> ) <sub>(0)</sub> , (U <sub>3</sub> , U <sub>5</sub> , U <sub>15</sub> ) <sub>(0)</sub> , (U <sub>3</sub> , U <sub>13</sub> , U <sub>15</sub> ) <sub>(2)</sub> , (U <sub>13</sub> , U <sub>15</sub> , U <sub>1</sub> ) <sub>(0)</sub> , (U <sub>13</sub> , U <sub>15</sub> , U <sub>2</sub> ) <sub>(1)</sub> , (U <sub>13</sub> , U <sub>15</sub> , U <sub>3</sub> ) <sub>(0)</sub> , (U <sub>13</sub> , U <sub>15</sub> , U <sub>5</sub> ) <sub>(0)</sub>
고빈도 사용자 집합(L3)	(U <sub>1</sub> , U <sub>2</sub> , U <sub>3</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>3</sub> , U <sub>5</sub> ) <sub>(2)</sub> , (U <sub>2</sub> , U <sub>5</sub> , U <sub>15</sub> ) <sub>(2)</sub> , (U <sub>3</sub> , U <sub>13</sub> , U <sub>15</sub> ) <sub>(2)</sub>

표 5. 연관 사용자 군집

연관 사용자 군집	각 군집에 군집된 사용자
1	{U <sub>13</sub> , U <sub>14</sub> , U <sub>15</sub> , U <sub>16</sub> , U <sub>17</sub> , U <sub>18</sub> , U <sub>19</sub> }
2	{U <sub>2</sub> , U <sub>3</sub> , U <sub>4</sub> , U <sub>5</sub> , U <sub>9</sub> }
3	{U <sub>1</sub> , U <sub>6</sub> , U <sub>7</sub> , U <sub>8</sub> , U <sub>12</sub> , U <sub>21</sub> }
4	{U <sub>20</sub> , U <sub>22</sub> , U <sub>23</sub> }
5	{U <sub>10</sub> , U <sub>11</sub> , U <sub>24</sub> , U <sub>25</sub> }

연관 사용자 집합에서 모든 연관 규칙과 신뢰도를 구한 후, ARHP 알고리즘을 이용하여 연관 사용자 군집을



의상 코디에 관한 예로서 사용자가 자신에게 어울리는 옷을 구입하기를 원하며 분위기에 맞는 옷을 추천 받고 싶은 경우를 생각해 볼 수 있다. 이때 사용자는 특별히 원하는 종류의 의상이 있을 경우 추천된 옷 중에서 특정 의상을 선택하고 그에 적합한 상의나 하의를 추천받을 수 있다. 새로운 사용자에 대한 선호도 예측은 기존의 협력적 필터링을 이용한 피어슨 상관 계수(PCC)방식과 제안한 연관 관계 군집에 의한 협력적 필터링 방식(ARCF) 두 가지를 모두 실험하였다.

#### 4.2 분석 및 성능평가

성능평가는 두 가지로 나누어서 하였다. 예측 값과 실제 값의 차이를 표시하여 정확성 측면에서 성능을 평가하기 위해 MAE 방식과 재현율과 정확도를 함께 사용해서 시스템의 전체적인 성능을 평가할 수 있는 F-measure 방식을 사용하였다. MAE는 예측의 정확성을 판단하는데 가장 많이 쓰이는 방법이고, F-measure는 값이 클수록 분류가 우수함을 의미한다[7]. 본 연구에서는 MAE를 기반으로 제안하는 방식(ARCF)과 협력적 필터링에 의한 피어슨 상관 계수에 의한 방식(PCC)을 실험하여 MAE에 의해 예측의 성능을 평가한 것이다.

표 6. MAE에 의한 성능평가

	P_count	ARCF	PCC
MAE	10	1.169	1.391
	20	0.938	1.122
	100	0.912	0.94
	150	1.074	1.152

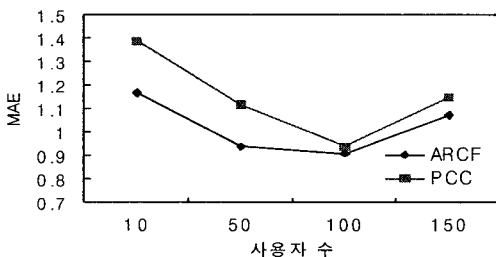


그림 4. ARCF/PCC의 MAE에 의한 성능평가

[표 6]의 MAE에 의한 평가에 의한 결과 값으로 사용

자 수를 증가시키기에 따른 MAE에 의한 성능 평가 결과는 [그림 4]와 같다. 한 아이템에 대해 평가한 사용자가 적을 경우는 대체로 상관관계가 높게 나타나는데, 이것은 그 아이템을 평가한 사용자가 평가한 다른 아이템에 대해 모두 높은 관계를 가지기 때문이다. 이러한 경우 이 아이템이 속한 군집의 크기는 매우 큰 반면, 이 아이템이 속한 군집의 수는 매우 적대[8]. 따라서 군집을 평가한 사용자들이 많아지게 되고, 예측 값의 정확도는 높아진다. 반면, 많은 사용자들이 평가한 아이템에 대해서는 많은 군집에 속하게 되어 평가의 정확도가 낮아진다. 그러므로 아이템들이 군집 속에 들어가는 수를 제한하고, 그 수보다 많아질 경우 기존의 피어슨 상관 계수를 이용하여 예측하고 그 수보다 적을 경우 제안하는 방법을 사용하면 정확도 면에서 좋은 성능이 나타난다. [표 7]은 정확도와 재현율을 분석한 결과를 나타낸다.

표 7. 사용자 군집의 정확도와 재현율 분석

군집	ARCF			PCC		
	정확도	재현율	F-measure	정확도	재현율	F-measure
1	90.91	90.91	90.91	90.91	86.96	83.33
2	90.00	84.62	84.62	89.57	81.48	84.62
3	88.89	88.89	88.89	88.89	84.21	80.00
4	91.20	87.50	87.50	77.78	82.35	87.50
5	90.91	90.91	90.91	90.91	86.96	83.33
6	85.71	85.71	85.71	85.71	80.00	75.00
7	85.71	85.71	85.71	92.86	85.71	85.71
8	92.86	89.66	86.67	76.50	89.66	86.67
9	84.00	80.00	85.71	75.00	74.30	75.00
10	75.00	75.00	75.00	86.00	75.00	75.00
평균	89.49	86.48	86.87	83.83	82.34	82.78

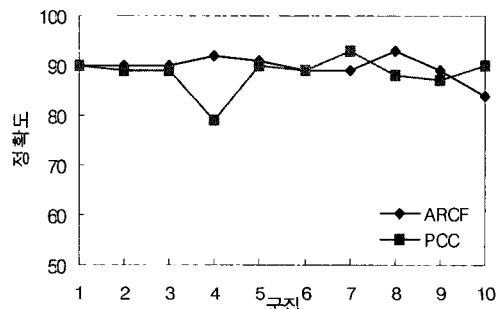


그림 5. 사용자 군집에 따른 정확도

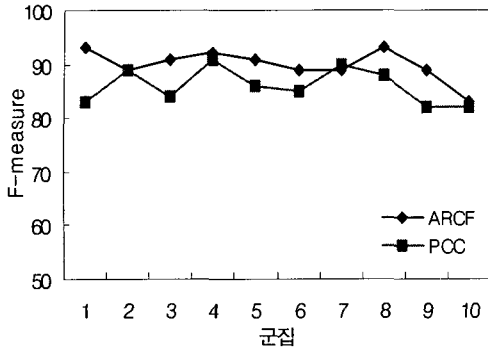


그림 6. 사용자 군집에 따른 F-measure

[그림 5]와 [그림 6]은 [표 7]의 결과를 바탕으로 사용자 군집에 따른 정확도와 F-measure의 성능 평가이다. AR CF는 기존 PCC보다 5.07% 높은 정확도와 4.21% 향상된 F-measure의 결과를 나타낸다. [그림 7]은 시맨틱 웹에서 개인화된 선호도를 이용한 의상 코드 시스템에서 추천된 아이템을 보여주고 있다.

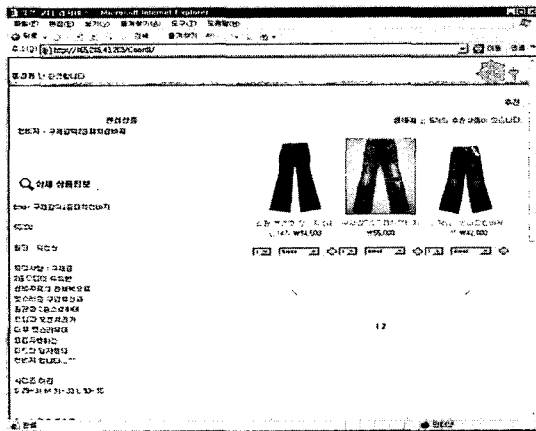


그림 7. 의상 추천 결과

기존의 추천 시스템은 사용자가 많은 정보를 제공해야 하며, 선호도에 대한 정보를 축적하기까지 많은 시간이 걸리고 다른 사용자들이 먼저 선호도를 평가한 아이템에 대해서만 추천이 가능하다는 제약이 있다. 반면 제한한 방법은 의미 있는 아이템들간의 연관 규칙에 의해 추천을 제공하는 것이다. 즉 기존의 대규모 데이터베이스의 거래 데이터 중 서로 일정 기준점 이상의 지지도와 신뢰

도를 만족하는 모든 연관 규칙을 찾아내는 것으로 특정 사용자의 선호에 대한 정보가 없더라도 단지 아이템간의 연관성에 의해 추천이 가능하다. 따라서 사용자가 선호하는 아이템을 휴리스틱한 방법에 의해 추출함으로써 데이터베이스 접속횟수와 계산횟수를 줄이는 방법이다.

### 5. 결론

본 논문에서는 추천의 효율성과 정확성을 증진시키기 위해 시맨틱 웹에서 개인화된 선호도를 이용한 의상코드 시스템을 개발하였다. 의상을 검색시 데이터베이스에 저장된 낱말 관련 정보가 같은지에 대해 분석을 위해 내용 기반 필터링 기술을 사용하였다. 초기 추천의상을 선택 후 적절한 코드를 해주기 위해 사용자의 선호도와 의상 관련 정보들의 정확도를 높여줄 수 있는 협력적 필터링 및 연관 사용자 마이닝을 이용하여 이전보다 향상된 추천을 할 수 있었다. 기존의 필터링만을 통한 추천 모델이 지니고 있는 문제점을 해결하기 위한 방안으로, 시맨틱 웹을 이용하여 더 나은 만족도를 통해 추천 방법의 타당성을 입증하였다. 향후 연구로는 정보의 생성, 배포, 수집, 활용을 관리하는 시스템을 설계하여 예외적인 상황들을 고려하여 더욱 만족할만한 추천에 대한 연구가 필요하다.

### 참고 문헌

- [1] A. M. Tjoa, A. Andjomshoa, F. Shayeganfar, and R. Wagner, "Semantic Web Challenges and New Requirements," Database and Expert Systems Applications, 2005.
- [2] J. Salter and N. Antonopoulos, "CinemaScreen Recommender Agent: Combining Collaborative and Content-Based Filtering," IEEE Intelligent Systems and Their Applications, 2006.
- [3] L. V. Lin and L. Y. Shu, "Research of English Text Classification Methods Based on Semantic

Meaning," Information and Communications Technology : ITI 3rd International Conference, 2005.

- [4] S. J. Ko and J. H. Lee, "User Preference Mining through Collaborative Filtering and Content Based Filtering in Recommender System," LNCS 2455, Proc of International Conference on E-Commerce and Web Technologies, pp.244-253, 2002.
- [5] 정경용, 김진현, 정현만, 이정현, "텍스타일 기반의 협력적 필터링 기술과 디자인 요소에 따른 감성 분석을 이용한 패션 디자인 추천 에이전트 시스템", 정보과학회지 : 컴퓨팅의 실제, 제10권, 제2호, pp. 174-188, 2004.
- [6] 은채수, 송창우, 이승근, 이정현, "시맨틱 웹에서 다중 혼합필터링을 이용한 개인화된 의상 코디 시스템", 정보과학회 가을 학술발표논문집, 제33권, 제2호(B), pp.178-182, 2006.
- [7] 정경용, 김진현, 정현만, 이정현, "개인화 추천 시스템에서 연관 관계 군집에 의한 아이템 기반의 협력적 필터링 기술", 정보과학회논문지 : 소프트웨어 및 응용, 제31권, 제4호, pp.467-477, 2004.
- [8] 정영미, 정보검색론, 구미무역(주) 출판부, 1993.

저 자 소개

은 채 수(Chae-Soo Eun)

준회원



- 2005년 2월 : 선문대학교 산업공학과(공학사)
- 2007년 2월 : 인하대학교 컴퓨터정보공학과(공학석사)

<관심분야> : 지능시스템, 시맨틱웹, 데이터마이닝

조 동 주(Dong-Ju Cho)

준회원



- 2004년 2월 : 상지대학교 행정학과 (행정학사)
- 2007년 3월 ~ 현재 : 상지대학교 컴퓨터정보공학과 석사과정

<관심분야> : 지능시스템, 시맨틱웹, 인공지능

정 경 용(Kyung-Yong Jung)

정회원



- 2000년 2월 : 인하대학교 전자계산공학과 (공학사)
- 2002년 2월 : 인하대학교 컴퓨터정보공학과(공학석사)
- 2005년 8월 : 인하대학교 컴퓨터정보공학과(공학박사)

• 2006년 3월 ~ 현재 : 상지대학교 컴퓨터정보공학부 교수

<관심분야> : 데이터마이닝, 지능시스템, 인공지능

이 정 현(Jung-Hyun Lee)

정회원



- 1977년 : 인하대학교 전자공학과(공학사)
- 1980년 : 인하대학교 대학원 전자공학과(공학석사)
- 1988년 : 인하대학교 대학원 전자공학과(공학박사)

• 1979년 ~ 1981년 : 한국전자기술연구소 시스템 연구원

• 1984년 ~ 1989년 : 경기대학교 전자계산학과 교수

• 1989년 ~ 현재 : 인하대학교 컴퓨터공학부 교수

<관심분야> : 자연어처리, HCI, 정보검색, 컴퓨터구조