



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2008-0035371  
(43) 공개일자 2008년04월23일

(51) Int. Cl.

G06Q 10/0010 (2006.01) G06F 17/00 (2006.01)

(21) 출원번호 10-2006-0101997

(22) 출원일자 2006년10월19일

심사청구일자 없음

(71) 출원인

연세대학교 산학협력단

서울 서대문구 신촌동 134 연세대학교

(72) 발명자

임일

서울 강남구 청담2동 삼성APT 101-207

이준기

경기 용인시 수지구 죽전동 98번지 현대빌라트 205-902

김범수

서울 종로구 창성동 90번지 303호

(74) 대리인

전중학

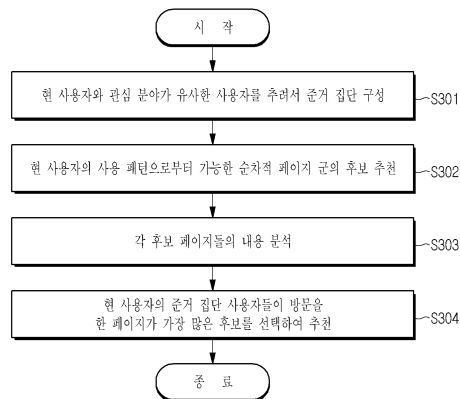
전체 청구항 수 : 총 9 항

(54) 개인화 시스템에서 각 사용자에 대한 맞춤 학습을 위한순서 추천 방법 및 이를 구현할 수 있는 프로그램이 수록된컴퓨터로 읽을 수 있는 기록매체

### (57) 요약

본 발명은 개인화 시스템에서 각 사용자에 대한 맞춤 학습을 위한 순서 추천 방법에 관한 것으로, 특히 상기 사용자와 관심 분야가 유사한 사용자들로 준거 집단을 구성하는 단계; 상기 사용자의 사용 패턴으로부터 가능한 순차적 페이지 군의 후보를 산출하는 단계; 및 상기 후보 페이지들 중에서 상기 구성된 준거 집단 사용자들이 방문을 한 페이지가 가장 많은 순차적 페이지를 선택하여 추천하는 단계를 포함하는 것을 특징으로 한다.

대표도 - 도3



## 특허청구의 범위

### 청구항 1

개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법에 있어서,

상기 사용자와 관심 분야가 유사한 사용자들로 준거 집단을 구성하는 단계;

상기 사용자의 사용 패턴으로부터 가능한 순차적 페이지 군의 후보를 산출하는 단계; 및

상기 후보 페이지들 중에서 상기 구성된 준거 집단 사용자들이 방문을 한 페이지가 가장 많은 순차적 페이지를 선택하여 추천하는 단계를 포함하는 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 2

제1항에 있어서, 상기 순차적 페이지 군의 후보를 산출하는 단계 이후에,

상기 산출된 각 후보 페이지들의 내용을 분석하는 단계를 더 포함하는 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 3

제1항에 있어서, 상기 준거 집단을 구성하는 방법은,

협업 기반 필터링의 방법에 의해 과거 방문한 페이지가 유사한 사용자들로 준거 집단을 구성하는 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 4

제1항에 있어서, 상기 순차적 페이지 군의 후보를 산출하는 방법은,

마코프 모델, 순차 연관 규칙 및 연관 규칙 중에서 선택된 어느 하나 이상의 방법을 이용하여 산출하는 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 5

제1항에 있어서, 상기 각 순차적 페이지 군의 후보에 대해 설정되는 순서의 수는, 기설정된 숫자만큼 산출하는 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 6

제1항에 있어서, 상기 추천되는 순서는,

온라인 학습 사이트에서의 학습 순서인 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 7

제1항에 있어서, 상기 추천되는 순서는,

온라인 쇼핑몰 사이트에서의 물품 구매 순서인 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 8

제1항에 있어서, 상기 추천되는 순서는,

온라인 동호회 사이트에서의 특정 분야에 대한 학습 순서인 것을 특징으로 하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법.

### 청구항 9

제1항 내지 제8항 중에서 선택된 어느 한 항에 따른 방법을 수행할 수 있는 프로그램이 수록된 컴퓨터로 읽을 수 있는 기록매체.

## 명세서

### 발명의 상세한 설명

#### 발명의 목적

##### 발명이 속하는 기술 및 그 분야의 종래기술

- <11> 본 발명은 개인화 시스템에서의 학습을 위한 추천 방법에 관한 것으로, 보다 상세하게는 각 사용자에게 대한 맞춤 학습을 위하여 학습 순서를 추천하여 제공하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법 및 이를 구현할 수 있는 프로그램이 수록된 컴퓨터로 읽을 수 있는 기록매체에 관한 것이다.
- <12> 최근 인터넷에서 사용자의 선호도를 최대한 반영한 서비스를 제공하기 위해서 개인화(Personalization)의 개념이 등장하고 있다. 예컨대, 이러한 개인화 시스템은 인터넷의 포털 사이트 등에서 기본 화면을 사용자의 구미에 맞게 편집하여 볼 수 있는 기능을 비롯하여 사용자의 스타일에 맞는 정보를 선별하여 볼 수 있게 해준다. 또한, 전자 상거래 업체들도 사용자의 개인적 취향에 따라 자신의 페이지를 구성하고 사용자의 구매기록, 취향에 맞는 제품을 추천받을 수 있는 기능들을 제공한다. 한편, 일반적으로 개인화라고 하면 웹사이트와 연관 지어서 웹사이트 개인화라는 용어를 사용하게 된다.
- <13> 상기 웹 사이트에서의 개인화는 일련의 가치교환 과정이다. 사용자가 자신의 선호, 관심, 구매경험과 같은 정보를 웹 사이트에 제공하면 웹사이트는 사용자가 제공한 자료를 기초로 사용자에게 가장 알맞은 정보를 제공한다. 상기 개인화를 통해서 웹 사이트 운영자는 사용자에게 관한 자료를 얻고 사용자의 지속적인 이용이나 구매를 얻어 낼 수 있게 되며 사용자는 자신에게 가장 알맞은 정보를 편리한 방법으로 얻을 수 있게 된다.
- <14> 상기와 같이 개인화 서비스는, 기존의 대부분의 인터넷 응용 서비스 시스템이 사용자 개개인의 콘텐츠(상품) 이용 특성을 고려하지 않고 모든 사용자에게 동일한 콘텐츠를 제공하는 반면, 데이터마이닝 기술 등을 이용하여 사용자의 콘텐츠 이용 특성을 분석한 후 각각의 사용자에게 적합한 콘텐츠를 제공함으로써 응용 서비스 시스템의 사용자 만족도를 향상시킨다.
- <15> 도 1은 일반적인 개인화 서비스 시스템의 전체적인 구조를 나타낸 블록도이다. 상기 도 1을 참조하면, 웹 응용 서비스 서버(120)와 연동하는 개인화 서비스 서버(140)는 일반적으로 정보 수집/변환 모듈(141), 분석 모듈(143) 및 추천 모듈(142) 등으로 구성될 수 있다.
- <16> 상기 정보 수집/변환 모듈(141)은 웹 응용 서비스 서버(120)로부터 사용자의 정보, 콘텐츠 정보 및 사용자가 콘텐츠를 사용한 내역을 제공받아서 개인화 시스템의 분석 모듈에서 사용하기에 적합한 형태의 데이터로 변환한다. 상기 사용자 정보에는 사용자가 응용 서비스 시스템에 제공한 인구 통계학적 데이터(이름, 나이, 성별, 주소 등)가 포함된다.
- <17> 상기 콘텐츠 정보는 콘텐츠에 대한 메타 정보(예컨대, 장르, 작자, 위치 등)와 사이트의 구조가 포함된다. 그리고 콘텐츠 사용 내역은 사용자가 어떤 콘텐츠(제품)를 사용(구매)했는 지에 대한 정보가 담겨져 있다. 일반적으로 웹서버의 로그 파일이 이에 해당한다. 이와 같은 정보는 웹 응용 서비스 서버(120)마다 다를 수 있으므로 정보 수집 및 변환 모듈(141)에서는 이들을 분석 모듈(143)에서 사용하기에 적합한 형태로 바꾸어 준다.
- <18> 상기 추천 모듈(142)은 웹 응용 서비스 서버(120)의 요구에 따라 현재의 콘텐츠 중에서 분석 모듈(143)에서 생성된 사용자 프로파일과 가장 잘 부합되는 콘텐츠를 추천한다. 상기 분석 모듈(143)이 많은 시간을 필요로 하는 작업이기 때문에 배치(Batch) 작업으로 수행되는데 비하여, 추천 모듈(142)은 실시간으로 처리된다.
- <19> 상기 분석 모듈(143)은 사용자 정보, 콘텐츠 정보 및 콘텐츠 사용 내역을 여러 가지 알고리즘을 사용하여 분석한 후에 사용자 프로파일을 생성한다. 사용자 프로파일은 사용자의 콘텐츠 이용성향에 대한 정보를 가지는 데이터로서 분석 모듈(143)에 의하여 생성 및 갱신되며 추천 모듈(142)에 의해 사용된다. 일반적으로 사용자 프로파일을 만드는 분석 모듈(143)의 종류에 따라서 개인화 시스템은 콘텐츠 기반(Content-based)방식과 협업 필터링(Collaborative Filtering) 방법으로 분류될 수 있다.
- <20> 상기 콘텐츠 기반 방식은 사용자가 구매(또는 사용)한 콘텐츠의 특성을 분석하여 사용자 프로파일을 생성(또는

수정)한 후에 사용자 프로파일과 콘텐츠 특성 정보를 비교하여 가장 유사한 콘텐츠를 추천하는 방식이다. 사용자 프로파일은 사용자별로 해당 사용자가 어떠한 종류의 콘텐츠에 대하여 얼마만큼의 선호도를 가지고 있는 지를 나타내는데 일반적으로(키워드, 가중치) 요소의 집합 형태를 가진다.

- <21> 이때, 키워드는 사용자가 구매한 콘텐츠로부터 추출한다. 비슷한 종류의 콘텐츠는 같은 키워드를 갖게 되므로, 만약 사용자가 비슷한 종류의 콘텐츠를 자주 구매할 경우 해당 키워드의 가중치가 높아지게 되어 사용자 프로파일은 자주 구매한 콘텐츠의 특성을 반영하게 된다.
- <22> 상기 협업 필터링 방식은 현재 많은 전자 상거래 사이트에서 사용되고 있으며, 협업 필터링을 이용한 개인화 서비스 방식은 사용자와 성향이 유사한 다른 사용자들의 구매 경향을 이용하여 사용자에게 콘텐츠를 추천하는 방식이다. 예컨대, 특정 사용자가 인터넷 서점에서 '골프 배우는 법'에 대한 책을 구매하였다고 가정하자. 그리고 이 사용자가 나중에 다시 이 사이트에 접속하였을 때에 사용자는 '프로 골퍼들로부터의 교훈'이나 '좋은 골프 코스'에 대한 책을 추천받게 된다. 이 사용자가 추천받는 책들은 비슷한 특성(골프에 관심이 있는)을 가진 다른 사용자들이 많이 구매한 것들이다. 이 방식은 사용자들의 일반적인 구매 특성을 잘 반영하기 때문에 추천의 정확도가 높을 뿐만 아니라 구현하기도 쉽다.
- <23> 도 2는 종래의 협업 필터링 방법에 따른 추천 절차를 나타낸 흐름도이다. 상기 도 2를 참조하면, 먼저 현 사용자와 다른 사용자 간의 유사도 지표를 산출(S201)하고, 상기 현 사용자와 유사한 사용자들로 준거 집단을 구성(S202)하게 된다.
- <24> 그런 다음, 상기 준거 집단 사용자들이 높게 평가한 항목 중에서 상기 현 사용자가 아직 보지 않은 항목들을 추출(S203)하고, 상기 추출된 항목들에 대해 현 사용자의 예상 평가 값을 산출(S204)하게 된다. 다음으로, 상기 산출된 예상 평가 값이 높은 순서대로 배열(S205)한 후, 상기 배열된 항목들 중 상위 n 개를 뽑아서 현 사용자에게 추천(S206)하게 된다.
- <25> 위와 같이, 상기 협업 필터링 방식을 이용한 개인화 서비스 시스템은 구현이 쉽고 정확성이 비교적 뛰어나서 대부분의 개인화 서비스 시스템에서 채택하고 있다.
- <26> 그러나 상기 협업 필터링 방식을 포함한 종래의 추천 알고리즘들은 일회성(one-time basis) 추천 방식이다. 즉, 어떤 사용자에게 현재 적합할 것으로 생각되는 물건이나 항목들을 하나의 리스트로 추천해 주는 것으로 끝나게 된다. 한편, 추천 리스트에 있는 항목들은 적합도에 따라 순위가 매겨져 있으나 항목의 순서가 그 이상의 의미를 담고 있지는 않고 있다. 따라서, 어떤 분야에 대한 지식을 쌓고자 하는 사용자가 이를 위해 적합한 학습 순서를 추천받고자 할 경우, 기존의 추천 알고리즘으로서는 추천해 줄 수가 없게 되는 한계점이 있다.

### 발명이 이루고자 하는 기술적 과제

- <27> 본 발명의 목적은 사용자가 특정 분야에 대한 지식을 얻고자 할 때, 요구되는 학습 순서를 추천하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법 및 이를 구현할 수 있는 프로그램이 수록된 컴퓨터로 읽을 수 있는 기록매체를 제공함에 있다.
- <28> 또한, 본 발명의 목적은 협업 필터링 및 내용 기반 필터링 방법을 부분적으로 이용하여 특정 사용자에게 특정 분야에 대한 학습 순서를 추천하는 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법 및 이를 구현할 수 있는 프로그램이 수록된 컴퓨터로 읽을 수 있는 기록매체를 제공함에 있다.

### 발명의 구성 및 작용

- <29> 상기한 목적을 달성하기 위한 본 발명의 방법은, 개인화 시스템에서 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 방법에 있어서, 상기 사용자와 관심 분야가 유사한 사용자들로 준거 집단을 구성하는 단계; 상기 사용자의 사용 패턴으로부터 가능한 순차적 페이지 군의 후보를 산출하는 단계; 및 상기 후보 페이지들 중에서 상기 구성된 준거 집단 사용자들이 방문한 페이지가 가장 많은 순차적 페이지를 선택하여 추천하는 단계를 포함하는 것을 특징으로 한다.
- <30> 상기 순차적 페이지 군의 후보를 산출하는 단계 이후에, 상기 산출된 각 후보 페이지들의 내용을 분석하는 단계를 더 포함하는 것을 특징으로 한다.
- <31> 상기 준거 집단을 구성하는 방법은, 협업 기반 필터링의 방법에 의해 과거 방문한 페이지가 유사한 사용자들로 준거 집단을 구성하는 것을 특징으로 한다.

- <32> 상기 순차적 페이지 군의 후보를 산출하는 방법은, 마코프 모델, 순차 연관 규칙 및 연관 규칙 중에서 선택된 어느 하나 이상의 방법을 이용하여 산출하는 것을 특징으로 한다.
- <33> 상기 각 순차적 페이지 군의 후보에 대해 설정되는 순서의 수는, 기설정된 숫자만큼 산출하는 것을 특징으로 한다.
- <34> 상기 추천되는 순서는, 온라인 학습 사이트에서의 학습 순서 또는 온라인 쇼핑몰 사이트에서의 물품 구매 순서 또는 온라인 동호회 사이트에서의 특정 분야에 대한 학습 순서인 것을 특징으로 한다.
- <35> 한편, 상기 개인화 시스템에서 각 사용자에게 맞춤 학습을 위한 순서 추천 방법에 대한 정보는 서버 컴퓨터로 읽을 수 있는 기록 매체에 저장될 수 있다. 이러한 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있도록 프로그램 및 데이터가 저장되는 모든 종류의 기록매체를 포함한다. 그 예로는, 롬(Read Only Memory), 램(Random Access Memory), CD(Compact Disk), DVD(Digital Video Disk)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들면, 인터넷을 통한 전송)의 형태로 구현되는 것도 포함된다. 또한, 이러한 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.
- <36> 본 발명은, 개인화 시스템에서 각 사용자에게 맞춤화되고 특정 분야에 대한 지식을 얻기 위한 학습 순서를 추천하는 방법을 제안한다. 이때, 상기 사용자는 현 상태에서 필요한 항목을 추천받고 끝나는 것이 아니라, 상기 해당 분야에 대해서 후후 학습할 경우 적합한 학습 순서에 대한 추천을 받게 된다.
- <37> 이하 본 발명의 바람직한 실시 예에 따른 상세한 설명을 첨부된 도면들을 참조하여 설명한다. 하기에는 본 발명을 설명함에 있어 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다.
- <38> 한편, 본 발명에 따른 순서 추천 방법은 종래의 협업 필터링(Collaborative filtering; CF) 방법과 내용 기반 필터링(Content-based filtering; CB) 방법에서 일부 개념을 가져오고 순서의 개념을 추가함으로써 새로운 알고리즘을 생성하게 된다.
- <39> 따라서, 본 발명에 따른 순서 추천 방법을 설명하기에 앞서 본 발명에 부분적으로 개념을 이용하는 상기 CF 방법과 CB 방법을 개략적으로 설명하기로 한다.
- <40> <협업 필터링 (Collaborative filtering: CF)>
- <41> 상기 협업 필터링은 정보를 사용자의 평가나 행동을 기반으로 필터링을 하는 기술이다. 상기 CF는 어떤 항목에 대한 사용자의 평가 점수와 같은 연속적인 값(continuous values)에 적용될 수도 있고, 사용자의 마우스 클릭과 같은 단속적(discrete) 혹은 이분적(binary) 값에 적용될 수도 있다.
- <42> - 연속적인 값에 적용되는 CF
- <43> 연속적인 값을 다루는 CF 엔진은 하기와 같은 알고리즘을 이용한다.
- <44> 1 단계 : 현 사용자와 다른 사용자 간의 유사도 지표(similarity index)를 계산한다.
- <45> 2 단계 : 현 사용자와 가장 유사한 사용자들을 찾아낸다. 이들이 현 사용자의 준거 집단(reference group)이 된다.
- <46> 3 단계 : 상기 준거 집단의 사용자들이 높게 평가한 항목 중에서 현 사용자가 아직 보지 않은 항목들을 골라서 현 사용자가 어떻게 평가할지를 나타내는 예상 평가 값(estimated evaluation)을 계산한다.
- <47> 4 단계 : 이들 항목을 예상 평가 값이 높은 순서대로 배열한다. 이중 상위 n개를 뽑아서 현 사용자에게 추천한다.
- <48> - 이분 값에 적용되는 CF
- <49> 상기 CF는 사용자의 일련의 마우스 클릭을 나타내는 클릭스트림(clickstream)에 적용되어 사용자의 다음 페이지를 추천하는데 사용될 수 있다. 상기 클릭스트림에 적용되는 CF는 클릭 스트림 CF(clickstream-based CF; CCF)라고 불리운다. 대부분의 CCF 알고리즘들은 일단 오프라인에서 배치 모드로 예측 모형을 수립한 후 온라인 추천을 위해 사용된다.
- <50> 가장 널리 쓰이는 3가지 예측 모형은 마코프 모델(Markov model), 연관규칙(association rule), 군집 분석



(clustering)이다. 상기 각 알고리즘은 장단점이 있다. 대개의 경우 정확도와 커버리지(전체 사용자 중에서 추천을 할 수 있는 사용자의 비율)라는 두 개의 측면에서 상충 관계에 있다.

<51> 예를 들어, 상기 마코프 모델은 다른 모델에 비해 정확도는 매우 높으나 커버리지는 낮고, 연관분석 방법은 커버리지는 높으나 정확도가 낮다. 이 두 가지를 동시에 높이기 위해 몇몇 연구에서는 두 개 이상의 알고리즘이 결합될 수 있다. 이러한 잡종모형(hybrid model)은 웹 사이트의 구조를 알아야 한다는 제한점이 있다.

<52> 상기 모형에서는 세가지 모형(즉, 마코프 모델, 순차연관규칙(sequential association rule), 그리고 연관규칙(association rule)) 중에서 주어진 상황에서 가장 정확한 방법을 선택하여 적용하기 때문에 정확도와 커버리지를 동시에 높일 수 있다는 장점이 있다. 여기서 군집 분석 방법은 정확도가 떨어지기 때문에 제외시키기로 한다.

<53> <내용기반 필터링 (Content-based filtering ; CB)>

<54> 상술한 CF는 어떤 형태의 항목(문서 혹은 사진)도 추천할 수 있고 사용자와 평가 정보가 충분히 많은 경우에는 잘 작동한다. 그러나 사용자 숫자가 적은 경우나 항목에 대한 평가가 적은 경우에는 추천이 불가능하거나 추천이 가능하더라도 매우 부정확하다는 단점이 있다. 따라서 본 발명에서 상술한 CF의 단점을 보완하기 위해 CB의 일부 개념을 추가로 이용하게 된다.

<55> 상기 CB는 사용자가 높게 평가하는 문서나 페이지와 비슷한 페이지를 추천하는 시스템이다. 즉, 상기 CB는 사용자의 선호나 평가가 아니라 문서 자체의 내용을 이용해서 추천을 한다. 온라인 환경에서는 문서의 내용이 이미 코드화가 되어 입력이 되어 있으므로 더 사용하기 용이하다는 장점이 있다. 추천에 사용되는 내용 항목은 제목, 저자, 본문 등 어떤 형태의 텍스트 정보라도 가능하다.

<56> 현재 사용되는 CB 알고리즘들은 대부분 문헌정보 검색 (information retrieval ; IR)과 기계학습 (machine learning ; ML) 기법을 결합하고 있다. 많은 CB 기법들은 추천을 “감독된 학습을 통한 분류(classification problem with supervised learning)” 라고 본다.

<57> 이 말의 의미는 다음과 같다. 컴퓨터는 우선 사용자가 좋아하는 혹은 유용하다고 생각하는 문서와 그렇지 않은 문서라는 두 가지 카테고리로 구분하는 방법을 학습한다. 그 다음 어떤 새로운 문서가 있는 경우 이를 두 카테고리에 속하는지 분류해서 전자에 해당하는 문서를 추천한다는 의미이다. 또 다른 CB 기법들은 추천을 회귀분석(regression) 문제로 간주하기도 한다. 컴퓨터는 훈련데이터(training dataset)로부터 학습을 통해서 새로운 문서를 분류하는 방법을 배우게 된다.

<58> 한편, 상기 CB에서 문서 간의 유사도를 계산하는 대표적인 기법은 "tf.idf(term frequency - inverse document frequency)" 방법이다. 즉, 어떤 두 문서 간의 유사도는 두 문서에 포함되어 있는 단어(term)의 유사도에서 추론할 수 있다. 따라서, 상기 CB 기법에서는 우선문서에 포함되어 있는 조사나 관사와 같이 필요 없는 단어를 없애고 명사형의 중요 키워드를 추출한다. 이때, 추출된 키워드의 빈도가 그 문서의 특징을 규정한다고 볼 수 있다.

<59> 키워드 중 그 문서에서 가장 많이 나타나는 것이 중요한 키워드라고 볼 수 있다. 이것이 단어 빈도(term frequency)이다. 그런데, 그 문서에서 아무리 많이 출현하는 키워드라 하더라도 다른 문서에서도 많이 나타나는 것은 그 문서의 특징을 규정하는 키워드라 보기 힘들다. 이것이, 역문서 빈도(inverse document frequency)이다. 상기 "tf.idf" 지표는 이 두 가지 모두를 고려해서 어떤 문서에서 추출된 키워드들의 중요도를 나타내는 수치이다. 어떤 키워드가 한 문서에서 많이 출현하면서 다른 문서에서는 적게 출현할수록 그 문서에서 더 큰 중요도를 갖는다. 이렇게 추출된 키워드와 중요도를 비교해서 문서 간의 유사도를 계산할 수 있다.

<60> <잡종 시스템(Hybrid system)>

<61> 상술한 CF와 CB를 결합한 방법이 잡종 시스템(hybrid systems)이다. 상기 잡종 시스템은 하나의 기술을 사용한 경우보다 더 정확한 추천 결과를 보여줄 수 있다. 즉, 상술한 CF와 CB를 결합하면 각 기술을 단독으로 사용한 경우보다 정확도와 커버리지(전체 사용자 중 추천을 제공할 수 있는 사용자의 비율)가 높아질 수 있다.

<62> 한편, 후술하는 본 발명에 따른 순서 추천 시스템은 상술한 CF를 기반으로 하지만 CB의 기법으로 보강이 되므로, 하나의 기법만 단독으로 사용된 경우보다 더 정확한 추천이 가능해지게 된다. 이하, 본 발명에 따른 순서 추천 알고리즘을 상세히 설명하기로 한다.

<63> <순서 추천 시스템>

- <64> 본 발명에서는 일반적인 CF와 CB 개념을 일부 혼합하고, 단순히 추천 리스트만을 제공하는 것이 아니라 순서를 추천하는 방법을 제안한다.
- <65> 도 3은 본 발명에 따른 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 절차를 나타낸 흐름도이다. 상기 도 3을 참조하면, 먼저 현 사용자와 관심 분야가 유사한 사용자를 추려서 준거 집단을 구성(S301)하게 된다. 그런 다음, 상기 현 사용자의 사용 패턴으로부터 가능한 순차적 페이지 군의 후보를 추천(S302)한다.
- <66> 이때, 상기 각 후보 페이지들의 내용을 분석(S303)한 후, 현 사용자의 준거 집단 사용자들이 방문을 한 페이지가 가장 많은 후보를 선택하여 추천(S304)함으로써 순서 추천 절차가 완료된다.
- <67> 이하, 상술한 순서 추천 알고리즘을 보다 상세히 설명하기로 한다.
- <68> - 1 단계 : 현 사용자와 비슷한 사용자를 추려서 현 사용자의 준거 집단을 구한다. 이때, 비슷한 사용자를 고르는 기준은 CF에서 사용되는 사용자의 평가에 의한 것이 아니고, CB의 기법을 이용한 과거 방문한 페이지의 유사성에 의한다. 즉, 준거 집단을 구하는 것은 CF에서 일부 이용되는 개념이나, 본 발명에서는 상기 준거 집단을 구할 때 CB에서 일부 이용하는 과거 방문한 페이지의 유사성을 이용하여 구하게 된다. 따라서, 본 발명에서 사용되는 과거 방문한 페이지의 유사성은 관심 분야의 유사성을 반영하는 것이라고 볼 수 있다.
- <69> - 2 단계 : 현 사용자의 사용 패턴을 보고 상술한 마코프 모델, 순차 연관 규칙, 연관 규칙을 결합한 CCF 방법을 사용하여 추천되는 후보 순서를 기설정된 숫자만큼(예컨대, 5 내지 10개 정도) 구한다. 예컨대, 현 사용자가 A-->B-->C의 순서로 페이지를 방문하였다면 그 다음으로 가능한 순차적 페이지 군의 후보(예컨대, D-->F-->K, D-->G-->L, 등등)를 구하게 된다.
- <70> - 3 단계 : 상기 2 단계에서 구한 후보 페이지들의 내용을 분석해서 현 사용자의 준거 집단 사용자들이 방문한 페이지가 가장 많은 후보를 선택해서 추천한다.
- <71> 상술한 본 발명에 따른 순서 추천 알고리즘의 기본 아이디어는 다음과 같다. 단순히 CCF로 순서를 추천하는 경우의 가장 큰 문제는 사용자의 취향과 필요가 너무 다양하다는 것이다. 따라서 CCF로 구한 후보 중에서 현 사용자와 취향과 필요가 비슷한 사용자의 것만 골라낼 필요가 있다. 이를 위해 현 사용자와 관심사항이 비슷한 준거 집단 사용자를 과거 방문 페이지의 내용을 바탕으로 골라내서 이를 후보 순서 중 최적의 순서를 선택하는데 사용한다.
- <72> 한편, 기존의 추천 알고리즘은 일회성(one-time basis) 추천이다. 다시 말해서, 어떤 사용자에게 현재 적합할 것으로 생각되는 물건이나 항목들을 하나의 리스트로 추천해 주고 끝이 난다. 따라서, 추천 리스트에 있는 항목들은 적합도에 따라 순위가 매겨져 있지만 항목의 순서가 그 이상의 의미가 있지는 않다. 만일 어떤 분야에 대한 지식을 쌓고자 하는 사용자가 이를 위해 적합한 학습 순서를 추천받고자 해도 기존의 추천 알고리즘은 이를 추천해 줄 수가 없다.
- <73> 그러나 상술한 바와 같이 본 발명에 따른 순서 추천 알고리즘은 각 사용자에게 맞춤화되고 어떤 분야에 대한 지식을 얻기 위해 필요한 학습 순서를 추천해 준다는 점에서 차이가 있다. 따라서, 사용자는 현 상태에서 필요한 항목을 추천받고 끝나는 것이 아니고 어떤 분야에 대해서 앞으로 학습을 할 경우 적합한 학습 순서에 대한 추천을 받게 된다.
- <74> <적용 분야>
- <75> 순서 추천 시스템의 적용분야는 여러 가지가 가능하지만 그중 가장 적합할 것으로 생각되는 분야로는 다음과 같은 것이 있다.
- <76> - 온라인 학습 사이트 : 온라인 학습을 제공하는 사이트에서 사용자의 과거 학습행동과 선호도를 고려하여 각 사용자에게 적합한 학습 순서를 추천해 준다.
- <77> - 온라인 쇼핑몰 : 어떤 분야에서는 꼭 필요한 제품이 있고 소비자의 필요에 따라 선택적으로 구입할 필요가 있는 제품이 있다. 만일 제품의 가짓수가 많고 가능한 조합이 많은 경우 소비자는 적합한 구매 순서 결정에 혼란을 느끼게 된다. 예를 들면, 화장품 같은 경우 영양크림과 같이 기본적으로 필요한 제품도 있지만 소비자의 선호나 필요에 따라 구입할 수도 있고 안 할 수도 있는, 그리고 소비자의 필요에 따라 어떤 사람은 꼭 필요한 제품이 다른 사람에게는 별 필요가 없는 제품이 되기도 한다. 이와 같은 경우 온라인 쇼핑몰에서 소비자가 어떤 제품분야에 대한 선호와 필요를 감안해서 그 소비자에게 적합한 구매 순서를 추천해 줄 수 있다.
- <78> - 동호회 사이트 : 온라인에서 비슷한 취미를 가진 사람들이 모여서 활동하는 동호회의 경우 동일한 취미 안에

서도 특정 분야(예를 들어 사진 동호회 안에서도 인물사진, 풍경 등과 같은 특정 분야)별로 각 회원의 경험이나 지식의 정도가 차이가 많이 난다. 이 경우 각 사용자의 지식 정도와 선호도를 고려해서 어떤 특정 분야에 대한 지식을 높이하고자 할 때 공부해야 하는 자료의 적합한 순서를 추천해 준다.

- <79> 한편, 본 발명의 실시 예에서는 구체적인 실시 예에 관해 설명하였으나, 본 발명의 범위에서 벗어나지 않는 한도 내에서 여러 가지 변형이 가능함은 물론이다. 그러므로 본 발명의 범위는 설명된 실시 예에 국한되어 정해져서는 안되며 후술하는 특허 청구의 범위뿐만 아니라 이 특허 청구의 범위와 균등한 것들에 의해 정해져야 한다.

### 발명의 효과

- <80> 본 발명에 따르면, 각 개인에 적합한 학습 순서에 대한 추천을 제공함으로써 특정 분야에 대한 학습을 할 경우 효과적인 학습을 할 수가 있게 되는 장점이 있다.
- <81> 또한, 본 발명이 온라인 쇼핑몰 등에 적용이 될 경우, 각 사용자의 구매 패턴 및 성향을 분석하여 다음 구매할 상품들의 구매 순서를 추천해 줄 수가 있게 되는 장점이 있다.

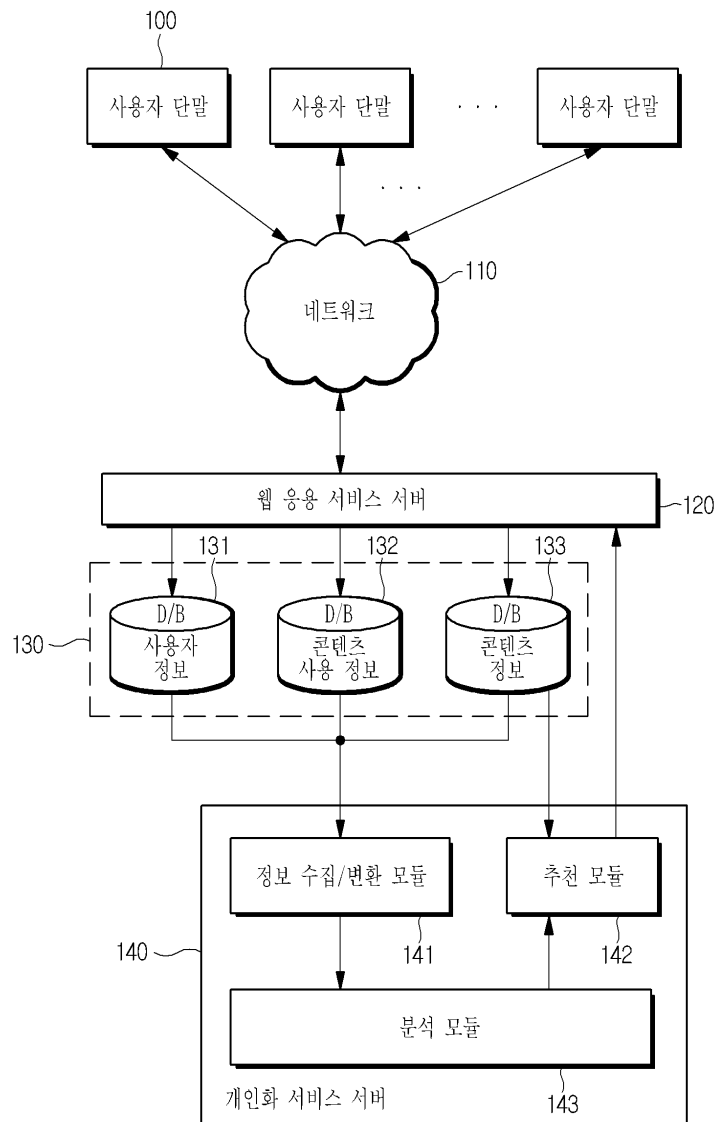
### 도면의 간단한 설명

- <1> 도 1은 일반적인 개인화 서비스 시스템의 전체적인 구조를 나타낸 블록도.
- <2> 도 2는 종래의 협업 필터링 방법에 따른 추천 절차를 나타낸 흐름도.
- <3> 도 3은 본 발명에 따른 각 사용자에게 대한 맞춤 학습을 위한 순서 추천 절차를 나타낸 흐름도.
- <4> <도면의 주요 부분에 대한 부호의 설명>
- |                         |                        |
|-------------------------|------------------------|
| <5> 100 : 사용자 단말        | 110 : 네트워크             |
| <6> 120 : 웹 응용 서비스 서버   | 130 : 데이터베이스           |
| <7> 131 : 사용자 정보 데이터베이스 | 132 : 콘텐츠 사용 정보 데이터베이스 |
| <8> 133 : 콘텐츠 정보 데이터베이스 | 140 : 개인화 서비스 서버       |
| <9> 141 : 정보 수집/변환 모듈   | 142 : 추천 모듈            |
| <10> 143 : 분석 모듈        |                        |

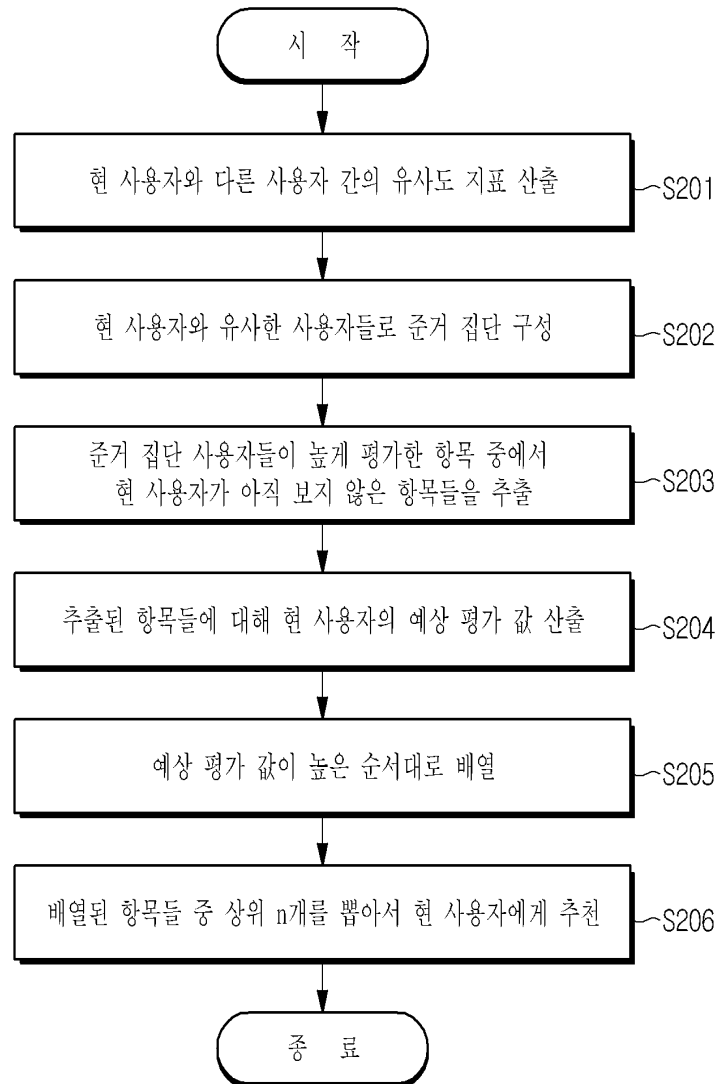


도면

도면1



도면2



도면3

