



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0026433
(43) 공개일자 2017년03월08일

(51) 국제특허분류(Int. Cl.)
H04W 8/20 (2009.01) G06F 17/30 (2006.01)
G06Q 10/10 (2012.01) G06Q 30/02 (2012.01)
H04W 4/00 (2009.01)
(52) CPC특허분류
H04W 8/20 (2013.01)
G06F 17/30867 (2013.01)
(21) 출원번호 10-2017-0026102(분할)
(22) 출원일자 2017년02월28일
심사청구일자 2017년02월28일
(62) 원출원 특허 10-2010-0088709
원출원일자 2010년09월10일
심사청구일자 2015년09월04일

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
이정수
경기도 수원시 영통구 도청로 65, 5403동 1203호 (이의동, 자연앤 힐스테이트)
이상근
서울특별시 동대문구 약령시로 25, 101동 203호(제기동, 안암골 벽산아파트)
(뒷면에 계속)
(74) 대리인
권혁록, 이정순

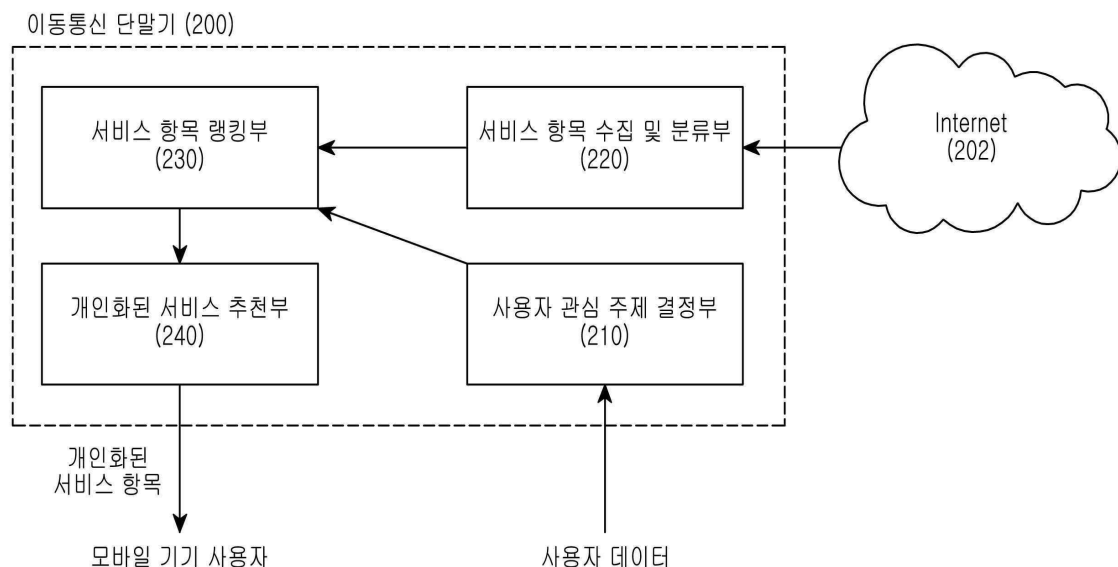
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 이동통신 단말기에서 인터넷 서비스 제공 방법 및 장치

(57) 요약

본 발명은 이동통신 단말기에서 인터넷 서비스 제공 방법 및 장치에 관한 것으로서, 이동통신 단말기에서 인터넷 서비스 제공 방법은, 상기 이동통신 단말기 내에 존재하는 사용자 데이터로부터 사용자 관심 주제를 결정하는 과정과, 네트워크 접속을 통해 서비스 항목을 수집하는 과정과, 상기 수집된 서비스 항목들 각각에 대한 주제를 결정하는 과정과, 상기 사용자 관심 주제와 상기 서비스 항목들 각각의 연관성을 결정하는 과정과, 상기 연관성에 따라 서비스 항목을 추천하는 과정을 포함하여, 사용자의 관심 주제에 대한 별도의 입력 없이도 상기 이동통신 단말기 내의 데이터를 기반으로 사용자 관심 주제를 분석하여 연관된 인터넷 서비스 항목을 추천할 수 있다.

대표도 - 도2



(52) CPC특허분류

G06Q 10/107 (2013.01)

G06Q 30/02 (2013.01)

H04W 4/00 (2013.01)

(72) 발명자

심규선

서울특별시 성북구 안암로9가길 40, 502호 (안암동5가)

이정현

서울특별시 영등포구 도신로4길 14, 102동 1905호 (대림동, 현대아파트)

하중우

서울특별시 성북구 개운사1길 54, 502호 (안암동5가)

명세서

청구범위

청구항 1

이동통신 단말기의 동작 방법에 있어서,

상기 이동통신 단말기 내에 저장된 사용자 데이터에 포함된 텍스트에 기반하여 상기 이동통신 단말기의 사용자의 관심 주제를 결정하는 과정과,

서버로부터 서비스 항목들에 관한 데이터를 수신하는 과정과,

상기 수신된 데이터에 포함된 텍스트에 기반하여 상기 서비스 항목들에 대한 주제들을 각각 결정하는 과정과,

상기 결정된 사용자의 관심 주제와 상기 결정된 서비스 항목들에 대한 주제들에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하는 과정과,

상기 선택된 적어도 하나의 서비스 항목을 나타내는 정보를 표시하는 과정을 포함하는 방법.

청구항 2

제 1항에 있어서,

상기 사용자 데이터는,

단문 메시지, 멀티미디어 메시지, 이메일, 파일, 스케줄, 메모 및 웹 사용 정보 중 적어도 하나를 포함하는 방법.

청구항 3

제 1항에 있어서,

상기 사용자의 관심 주제를 결정하는 과정은,

상기 사용자 데이터로부터 텍스트를 식별하는 과정과,

상기 식별된 텍스트를 기반으로 벡터 스페이스 모델에 따른 단어 벡터를 생성하는 과정과,

상기 생성된 단어 벡터를 주제 분류 트리에 따라 분류하여 상기 사용자의 관심 주제를 결정하는 과정을 포함하는 방법.

청구항 4

제 3항에 있어서,

상기 식별된 텍스트에 대한 정보를 저장하는 과정을 더 포함하며,

상기 식별된 텍스트에 대한 정보는,

상기 식별된 텍스트의 생성 시간, 텍스트 생성 위치, 어플리케이션의 종류, 어플리케이션 특징, 어플리케이션 수정 시간, 어플리케이션 생성 시간 중 적어도 하나를 포함하는 방법.

청구항 5

제 3항에 있어서,

상기 생성된 단어 벡터는, 상기 생성된 단어 벡터에 대응하는 텍스트 내에서 해당 단어의 중요도에 따른 가중치를 반영하는 방법.

청구항 6

제 1항에 있어서,

상기 서비스 항목들에 대한 주제들을 각각 결정하는 과정은,

상기 수신된 데이터에 포함된 텍스트를 식별하는 과정과,

상기 식별된 텍스트를 기반으로 벡터 스페이스 모델에 따른 단어 벡터를 생성하는 과정과,

상기 생성된 단어 벡터를 주제 분류 트리에 따라 분류하여 상기 서비스 항목들에 대한 주제들을 각각 결정하는 과정을 포함하는 방법.

청구항 7

제 6항에 있어서,

상기 식별된 텍스트에 대한 정보를 저장하는 과정을 더 포함하며,

상기 식별된 텍스트에 대한 정보는,

상기 식별된 텍스트의 생성 시간, 텍스트 생성 위치, 어플리케이션의 종류, 어플리케이션 특징, 어플리케이션 수정 시간, 어플리케이션 생성 시간 중 적어도 하나를 포함하는 방법.

청구항 8

제 6항에 있어서,

상기 생성된 단어 벡터는, 상기 생성된 단어 벡터에 대응하는 텍스트 내에서 해당 단어의 중요도에 따른 가중치를 반영하는 방법.

청구항 9

제 1항에 있어서,

상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하는 과정은,

상기 사용자 데이터로부터 생성되는 단어 벡터와 상기 수신된 데이터로부터 생성되는 단어 벡터 간의 구문론적 유사도를 결정하는 과정과,

상기 사용자의 관심 주제와 상기 서비스 항목들에 대한 주제들 사이의 의미론적 유사도를 결정하는 과정과,

상기 구문론적 유사도와 상기 의미론적 유사도를 기반으로 전체 유사도를 결정하는 과정과,

상기 전체 유사도에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하는 과정을 포함하는 방법.

청구항 10

제 9항에 있어서,

상기 전체 유사도에 기반하여 상기 서비스 항목들 중에서 상기 적어도 하나의 서비스 항목을 선택하는 과정은,

상기 결정된 전체 유사도에 기반하여 상기 서비스 항목들에 대한 주제들의 순위를 결정하는 과정과,
상기 결정된 순위에 기반하여 상기 서비스 항목들 중 상기 적어도 하나의 서비스 항목을 선택하는 과정을 포함하는 방법.

청구항 11

제 9항에 있어서,
상기 의미론적 유사도를 결정하는 과정은,
주제 분류 트리에서 각 카테고리에 대한 학습 데이터들의 평균 단어 벡터를 결정하는 과정과,
상기 평균 단어 벡터를 이용하여 상기 각 카테고리의 계층적 관계를 반영한 병합 벡터를 결정하는 과정과,
각 카테고리의 병합 벡터 간의 코사인 유사도를 이용하여 카테고리들 간의 링크 가중치를 결정하는 과정과,
상기 링크 가중치를 바탕으로 각 카테고리들 간의 이동 확률을 결정하는 과정과,
상기 각 카테고리들 간의 이동 확률을 이용하여 의미론적 유사도를 결정하는 과정을 포함하는 방법.

청구항 12

이동통신 단말기에 있어서,
상기 이동통신 단말기 내에 저장된 사용자 데이터에 포함된 텍스트에 기반하여 상기 이동통신 단말기의 사용자의 관심 주제를 결정하는 사용자 관심 주제 추출부와,
서버로부터 서비스 항목들에 관한 데이터를 수신하고, 상기 수신된 데이터에 포함된 텍스트에 기반하여 서비스 항목들에 대한 주제들을 각각 결정하는 서비스 항목 수집 및 분류부와,
상기 결정된 사용자의 관심 주제와 상기 결정된 서비스 항목들에 대한 주제들에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하고, 상기 선택된 적어도 하나의 서비스 항목을 나타내는 정보를 표시하는 서비스 추천부를 포함하는 이동통신 단말기.

청구항 13

제 12항에 있어서,
상기 사용자 데이터는,
단문 메시지, 멀티미디어 메시지, 이메일, 파일, 스케줄, 메모 및 웹 사용 정보 중 적어도 하나를 포함하는 이동통신 단말기.

청구항 14

제 12항에 있어서,
상기 사용자 관심 주제 결정부는,
상기 사용자 데이터로부터 텍스트를 식별하고, 상기 식별된 텍스트를 기반으로 벡터 스페이스 모델에 따른 단어 벡터를 생성한 후, 상기 생성된 단어 벡터를 주제 분류 트리에 따라 분류하여 상기 사용자의 관심 주제를 결정하는 이동통신 단말기.

청구항 15

제 14항에 있어서,
 상기 사용자 관심 주제 결정부는,
 상기 식별된 텍스트에 대한 정보를 저장하며,
 상기 식별된 텍스트에 대한 정보는,
 상기 식별된 텍스트의 생성 시간, 텍스트 생성 위치, 어플리케이션의 종류, 어플리케이션 특징, 어플리케이션 수정 시간, 어플리케이션 생성 시간 중 적어도 하나를 포함하는 이동통신 단말기.

청구항 16

제 14항에 있어서,
 상기 사용자 관심 주제 결정부는,
 상기 생성된 단어 벡터에 대응하는 텍스트 내에서 해당 단어의 중요도에 따른 가중치를 반영하는 이동통신 단말기.

청구항 17

제 12항에 있어서,
 상기 서비스 항목 수집 및 분류부는,
 상기 수신된 데이터에 포함된 텍스트를 식별하고, 상기 식별된 텍스트를 기반으로 벡터 스페이스 모델에 따른 단어 벡터를 생성한 후, 상기 생성된 단어 벡터를 주제 분류 트리에 따라 분류하여 상기 서비스 항목들에 대한 주제들을 각각 결정하는 이동통신 단말기.

청구항 18

제 17항에 있어서,
 상기 서비스 항목 수집 및 분류부는,
 상기 식별된 텍스트에 대한 정보를 저장하며,
 상기 식별된 텍스트에 대한 정보는,
 상기 식별된 텍스트의 생성 시간, 텍스트 생성 위치, 어플리케이션의 종류, 어플리케이션 특징, 어플리케이션 수정 시간, 어플리케이션 생성 시간 중 적어도 하나를 포함하는 이동통신 단말기.

청구항 19

제 17항에 있어서,
 상기 서비스 항목 수집 및 분류부는,
 상기 생성된 단어 벡터에 대응하는 텍스트 내에서 해당 단어의 중요도에 따른 가중치를 반영하는 이동통신 단말기.

청구항 20

제 12항에 있어서,

상기 이동통신 단말기는,

상기 사용자 데이터로부터 생성되는 단어 벡터와 상기 수신된 데이터로부터 생성되는 단어 벡터 간의 구문론적 유사도를 결정하고, 상기 사용자의 관심 주제와 상기 서비스 항목들에 대한 주제들 사이의 의미론적 유사도를 결정하고, 상기 구문론적 유사도와 상기 의미론적 유사도를 기반으로 전체 유사도를 결정하는 서비스 항목 랭킹 부를 더 포함하고,

상기 서비스 추천부는,

상기 전체 유사도에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하는 이동통신 단말기.

청구항 21

제 20항에 있어서,

상기 서비스 추천부는,

상기 결정된 전체 유사도에 기반하여 상기 서비스 항목들에 대한 주제들의 순위를 결정하고, 상기 결정된 순위 에 기반하여 상기 서비스 항목들 중 상기 적어도 하나의 서비스 항목을 선택하는 이동통신 단말기.

청구항 22

제 20항에 있어서,

상기 의미론적 유사도는,

주제 분류 트리에서 각 카테고리에 대한 학습 데이터들의 평균 단어 벡터를 결정하고, 상기 평균 단어 벡터를 이용하여 상기 각 카테고리의 계층적 관계를 반영한 병합 벡터를 결정한 후, 각 카테고리의 병합 벡터 간의 코사인 유사도를 이용하여 카테고리들 간의 링크 가중치를 결정하고, 상기 링크 가중치를 바탕으로 각 카테고리들 간의 이동 확률을 결정하여 의미론적 유사도를 결정하는 이동통신 단말기.

발명의 설명

기술 분야

[0001] 본 발명은 이동통신 단말기에서 인터넷 서비스를 제공하는 방법 및 장치에 관한 것으로서, 특히 이동통신 단말기의 사용자 데이터를 분석하여 서비스를 추천하는 방법 및 장치에 관한 것이다.

배경 기술

[0003] 이동통신 기술이 급격하게 발달함에 따라 이동통신 단말기를 통한 다양한 서비스가 제공되고 있다. 특히, 최근에는 이동통신 단말기에서 사용자의 관심 분야를 고려하여 해당 분야에 대한 정보를 제공하는 서비스 즉, 온 디맨드 방식의 개인화된 정보 서비스가 제공되고 있다.

[0004] 도 1은 종래 기술에 따른 온 디맨드 방식의 개인화된 정보 서비스 제공 방식을 도시하고 있다. 상기 도 1에 도시된 바와 같이, 종래의 개인화된 정보 서비스는 이동통신 단말기(100)에서 사용자 정보 저장 서버(102)에 사용자 관심 분야를 등록하면(110단계), 상기 사용자 정보 서버(102)가 상기 등록된 사용자 관심 분야를 개인화된 정보 제공 서버(104)로 제공하여 상기 개인화된 정보 제공 서버(104)에서 상기 사용자 관심 분야에 해당하는 서비스 정보들을 검색한 후, 검색된 서비스 정보들을 상기 이동통신 단말기(100)로 제공하는(112단계) 방식으로 이루어진다.

[0005] 상기와 같이, 종래의 개인화된 정보 서비스는 이동통신 단말기에서 사용자의 관심 분야를 서버에 미리 등록해야만 한다. 즉, 종래의 개인화된 정보 서비스는 상기 이동통신 단말기에서 해당 서버에 사용자의 관심 분야를 미리 등록하지 않은 경우, 원하는 정보를 제공받을 수 없는 단점이 있다. 또한, 상기 종래의 개인화된 정보 서비

스는 사용자가 자신의 관심 분야가 변경될 때마다 상기 서버에 변경된 관심 분야를 직접 등록해야 하므로, 상기 사용자 입장에서 번거롭고 불편한 문제점이 있다.

발명의 내용

해결하려는 과제

- [0007] 본 발명은 상술한 바와 같은 문제점을 해결하기 위해 도출된 것으로서, 본 발명의 목적은 이동통신 단말기에서 사용자 데이터를 분석하여 인터넷 서비스를 추천하는 방법 및 장치를 제공함에 있다.
- [0008] 본 발명의 또 다른 목적은 이동통신 단말기에서 사용자 데이터로부터 사용자 관심 주제를 결정하고, 인터넷 서비스 항목을 수집하여 서비스 항목의 주제를 결정하여 상기 사용자 관심 주제에 대응하는 서비스 항목을 추천하는 방법 및 장치를 제공함에 있다.
- [0009] 본 발명의 또 다른 목적은 이동통신 단말기에서 사용자 데이터와 서비스 항목 각각으로부터 단어 벡터를 추출하여 각각의 단어 벡터 간의 구문론적 유사도를 결정하는 방법 및 장치를 제공함에 있다.
- [0010] 본 발명의 또 다른 목적은 이동통신 단말기에서 사용자 데이터와 서비스 항목 각각으로부터 주제를 추출하여 각각의 주제 간의 의미론적 유사도를 결정하는 방법 및 장치를 제공함에 있다.
- [0011] 본 발명의 또 다른 목적은 이동통신 단말기에서 사용자 관심 주제와 서비스 항목 주제 간의 유사도를 계산하고, 상기 유사도에 따라 서비스를 추천하는 방법 및 장치를 제공함에 있다.
- [0012] 본 발명의 또 다른 목적은 이동통신 단말기에서 주제 분류 트리의 각 카테고리에 대해 카테고리들 간의 계층적 구조의 특징을 반영하는 단어 벡터를 계산하는 방법 및 장치를 제공함에 있다.
- [0013] 본 발명의 또 다른 목적은 이동통신 단말기에서 주제 분류 트리의 각 카테고리에 대해 다른 카테고리들에 대한 연관성을 결정하고, 상기 연관성을 바탕으로 사용자 관심 주제에 대응하는 서비스를 추천하는 방법 및 장치를 제공함에 있다.

과제의 해결 수단

- [0015] 본 발명의 다양한 실시 예에 따른 이동통신 단말기의 동작 방법은, 이동통신 단말기 내에 저장된 사용자 데이터에 포함된 텍스트에 기반하여 상기 이동통신 단말기의 사용자의 관심 주제를 결정하는 과정과, 서버로부터 서비스 항목들에 관한 데이터를 수신하는 과정과, 상기 수신된 데이터에 포함된 텍스트에 기반하여 상기 서비스 항목들에 대한 주제들을 각각 결정하는 과정과, 상기 결정된 사용자의 관심 주제와 상기 결정된 서비스 항목들에 대한 주제들에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하는 과정과, 상기 선택된 적어도 하나의 서비스 항목을 나타내는 정보를 표시하는 과정을 포함할 수 있다.
- [0016] 본 발명의 다양한 실시 예에 따른 이동통신 단말기는, 상기 이동통신 단말기 내에 저장된 사용자 데이터에 포함된 텍스트에 기반하여 상기 이동통신 단말기의 사용자의 관심 주제를 결정하는 사용자 관심 주제 추출부와, 서버로부터 서비스 항목들에 관한 데이터를 수신하고, 상기 수신된 데이터에 포함된 텍스트에 기반하여 서비스 항목들에 대한 주제들을 각각 결정하는 서비스 항목 수집 및 분류부와, 상기 결정된 사용자의 관심 주제와 상기 결정된 서비스 항목들에 대한 주제들에 기반하여 상기 서비스 항목들 중에서 적어도 하나의 서비스 항목을 선택하고, 상기 선택된 적어도 하나의 서비스 항목을 나타내는 정보를 표시하는 서비스 추천부를 포함할 수 있다.

발명의 효과

- [0018] 본 발명은 이동통신 단말기에서 사용자 데이터로부터 사용자 관심 주제를 결정하고 인터넷 서비스 항목을 수집하여 서비스 항목의 주제를 결정한 후, 사용자 관심 주제와 서비스 항목 주제 간의 유사도를 계산하여 상기 유사도에 따라 서비스를 추천함으로써, 상기 사용자의 관심 주제에 대한 별도의 입력 없이도 상기 이동통신 단말기 내의 데이터를 기반으로 사용자 관심 주제를 분석하여 연관된 인터넷 서비스 항목을 추천할 수 있는 효과가 있다. 또한, 상기 이동통신 단말기 내의 데이터를 기반으로 사용자 관심 주제를 분석하기 때문에, 실시간으로 변

경되는 사용자의 관심 주제에 적절하게 대응하여 적합한 인터넷 서비스 항목을 추천할 수 있는 효과가 있다. 또한, 상기 이동통신 단말기에서 사용자 개인의 정보를 네트워크를 통해 외부로 전송하거나 서버에 저장하지 않고 상기 이동통신 단말기 내의 데이터를 기반으로 사용자 관심 주제를 분석하여 적합한 인터넷 서비스 항목을 추천함으로써, 사용자 개인의 데이터를 보호할 수 있는 효과가 있다.

도면의 간단한 설명

- [0020] 도 1은 종래 기술에 따른 온 디맨드 방식의 개인화된 정보 서비스 제공 방식을 도시하는 도면,
 도 2는 본 발명에 따라 개인화된 정보 서비스를 제공하기 위한 이동통신 단말기의 블록 구성을 도시하는 도면,
 도 3은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 관심주제 결정부의 상세한 블록 구성을 도시하는 도면,
 도 4는 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목 수집 및 분류부의 상세한 블록 구성을 도시하는 도면,
 도 5는 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목 랭킹부의 상세한 블록 구성을 도시하는 도면,
 도 6은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 데이터를 도시하는 도면,
 도 7은 본 발명의 실시 예에 따른 이동통신 단말기에서 생성한 XML 파일을 도시하는 도면,
 도 8은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 데이터를 기반으로 판단한 사용자 관심 주제를 도시하는 도면,
 도 9는 본 발명의 실시 예에 따른 이동통신 단말기에서 수집하는 인터넷 서비스 항목의 예를 도시하는 도면,
 도 10은 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목의 주제들 간의 유사도를 도시하는 도면,
 도 11은 본 발명의 실시 예에서 두 번째 순회 패턴을 따를 때 상기 tw_{ij} 값을 구하는 방법을 도시하는 도면,
 도 12는 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목을 추천하는 화면 구성을 도시하는 도면, 및
 도 13은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 관심 주제에 따라 서비스 항목을 추천하는 절차를 도시하는 도면.

발명을 실시하기 위한 구체적인 내용

- [0021] 이하 본 발명의 바람직한 실시 예를 첨부된 도면을 참조하여 설명한다. 그리고, 본 발명을 설명함에 있어서, 관련된 공지기능 혹은 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단된 경우 그 상세한 설명은 생략할 것이다. 또한, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0023] 이하 본 발명에서는 본 발명의 다른 목적은 이동통신 단말기에서 사용자 데이터를 분석하여 서비스를 추천하는 방법 및 장치에 관해 설명할 것이다.
- [0024] 도 2는 본 발명에 따라 개인화된 정보 서비스를 제공하기 위한 이동통신 단말기의 블록 구성을 도시하고 있다.
- [0025] 상기 도 2를 참조하면, 이동통신 단말기(200)는 사용자 관심 주제 결정부(210), 서비스 항목 수집 및 분류부(220), 서비스 항목 랭킹부(230) 및 개인화된 서비스 추천부(240)를 포함하여 구성된다.
- [0026] 상기 사용자 관심 주제 결정부(210)는 상기 이동통신 단말기 내에 존재하는 사용자 데이터를 분석하여 사용자 관심 주제를 결정한다. 여기서, 상기 사용자 데이터는 상기 이동통신 단말기 내에 존재하는 단문 메시지, 멀티미디어 메시지, 이메일, 파일, 스케줄, 메모 및 웹 사용 정보 등과 같은 데이터들을 의미한다. 즉, 상기 사용자 관심 주제 결정부(210)는 상기 이동통신 단말기 내에 존재하는 사용자 데이터로부터 텍스트를 추출하고, 텍스트

를 분석하여 단어 벡터를 생성한 후, 상기 단어 벡터를 상기 이동통신 단말기에 임베디드된 주제 분류 트리에 따라 분류하여 사용자 관심 주제를 결정한다. 여기서, 상기 주제 분류 트리는 사용자의 관심 분야를 나타내기에 적합한 개념들을 주제별로 분류하여 트리 구조로 나타낸 것으로서, 종래에 잘 알려진 오픈 디렉토리 프로젝트(Open Directory Project)일 수 있다. 물론, 일반적인 오픈 디렉토리 프로젝트는 웹 디렉토리인 만큼 광범위하기 때문에, 상기 오픈 디렉토리 프로젝트에서 사용자 관심 분야를 나타내기에 적합한 일부 카테고리들만을 추출하여 사용할 수 있을 것이다. 상기 사용자 관심 주제 결정부(210)는 하기에서 도 3을 참조하여 상세히 설명하기로 한다.

[0027] 상기 서비스 항목 수집 및 분류부(220)는 인터넷(202)에 접속하여 인터넷 서비스 항목을 수집하고, 수집한 서비스 항목들의 텍스트를 분석하여 단어 벡터를 생성한 후, 상기 단어 벡터를 상기 이동통신 단말기에 임베디드된 주제 분류 트리에 따라 분류하여 상기 수집된 서비스 항목들 각각의 주제를 결정한다. 상기 서비스 항목 수집 및 분류부(220)는 하기에서 도 4를 참조하여 상세히 설명하기로 한다.

[0028] 상기 서비스 항목 랭킹부(230)는 상기 사용자 데이터에 대한 단어 벡터와 사용자 관심 주제 및 상기 서비스 항목에 대한 단어 벡터와 서비스 항목 주제들을 이용하여 구문론적 유사도와 의미론적 유사도를 결정하고, 상기 구문론적 유사도와 의미론적 유사도를 이용하여 상기 사용자 관심 주제와 서비스 항목 주제에 대한 전체 유사도를 결정하여 상기 사용자 관심 주제에 대한 서비스 항목 각각의 연관성 순위를 계산한다. 상기 서비스 항목 랭킹부(230)는 하기에서 도 5를 참조하여 상세히 설명하기로 한다.

[0029] 상기 개인화된 서비스 추천부(240)는 서비스 항목 추천 이벤트가 발생될 시, 화면을 통해 추천 서비스 항목의 종류를 선택하는 창을 디스플레이하고, 서비스 항목의 종류가 선택되면, 상기 서비스 항목 랭킹부(230)에서 계산된 연관성 순위를 고려하여 상기 선택된 종류에 해당하는 서비스 항목들을 결정하고, 결정된 서비스 항목들을 포함하는 목록을 화면에 디스플레이하기 위한 기능을 제어한다. 또한, 상기 개인화된 서비스 추천부(240)는 추천한 서비스 항목들 중 어느 하나의 서비스 항목이 선택되는 경우, 해당 서비스 항목의 상세 내용을 화면에 디스플레이하기 위한 기능을 제어 및 처리한다. 예를 들어, 상기 개인화된 서비스 추천부(240)는 도 12에 도시된 바와 같이, 뉴스 혹은 모바일 앱 중에서 어떤 종류의 서비스 항목을 추천받을 것인지 선택하는 창을 디스플레이하고, 사용자로부터 모바일 앱이 선택되면 사용자 관심 주제와 연관성이 큰 모바일 앱 서비스들을 포함하는 목록을 디스플레이한 후, 사용자에게 의해 선택된 모바일 앱 서비스에 대한 상세 정보를 화면에 디스플레이하기 위한 기능을 제어 및 처리한다.

[0031] 그러면, 하기에서 도 3 내지 도 5를 참조하여 상기 이동통신 단말기의 상세한 구성에 대하여 살펴보기로 한다.

[0032] 도 3은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 관심주제 결정부의 상세한 블록 구성을 도시하고 있다.

[0033] 상기 도 3을 참조하면, 상기 사용자 관심 주제 결정부(210)는 사용자 데이터 텍스트 추출부(310), 사용자 데이터 텍스트 분석부(320) 및 사용자 데이터 단어 벡터 분류부(330)를 포함하여 구성된다.

[0034] 상기 사용자 데이터 텍스트 추출부(310)는 상기 이동통신 단말기 내에 존재하는 사용자 데이터로부터 사용자 관심 주제를 나타내는 텍스트를 추출한다. 예를 들어, 상기 사용자 데이터 텍스트 추출부(310)는 도 6에 도시된 바와 같이 이동통신 단말기 내에 저장된 단문 메시지, 멀티미디어 메시지, 이메일, 파일, 스케줄, 메모 및 웹 사용 정보 등으로부터 텍스트를 추출할 수 있다. 이때, 상기 사용자 데이터 텍스트 추출부(310)는 상기 추출된 텍스트에 대한 생성 시간, 생성 위치, 해당 어플리케이션의 메타 정보를 함께 추출하여 XML 형태로 저장한다. 여기서, 상기 어플리케이션의 메타 정보는 어플리케이션의 종류, 특징, 수정 시간, 생성 시간 및 문맥 정보 중 적어도 하나를 포함한다. 여기서, 상기 문맥 정보는 사용자 데이터 텍스트 분석부(320)에서 단어 벡터 생성 시에 가중치를 적용하기 위해 활용할 수 있다. 즉, 상기 사용자 데이터 텍스트 추출부(310)는 상기 도 6에 도시된 바와 같은 사용자 데이터로부터 텍스트를 추출하고, 이때 도 7에 도시된 바와 같은 XML 파일을 생성할 수 있다.

[0035] 상기 사용자 데이터 텍스트 분석부(320)는 상기 사용자 데이터 텍스트 추출부(310)에서 추출된 텍스트 데이터를 분석하여 벡터 스페이스 모델(Vector Space Model)에 따른 단어 벡터(Term Vector)를 생성한다. 여기서, 상기 단어 벡터는 텍스트 데이터에 존재하는 개별적인 단어들로 이루어지며, 추출된 텍스트 내에서 각 단어들의 중요도에 따른 가중치를 반영하여 생성될 수 있다. 이때, 상기 각 단어들의 중요도에 따른 가중치는 추출된 텍스트 내에서의 빈도 수, 해당 텍스트의 생성 시간과 생성 위치를 고려하여 계산할 수 있다. 예를 들어, 상기 사용자 데이터 텍스트 분석부(320)는 상기 사용자 데이터 텍스트 추출부(310)로부터 제공된 텍스트들에서 특정 단어의

출현 빈도수가 높을 경우, 해당 단어를 사용자의 관심 주제를 표현하는 중요 단어로 판단하여 상기 해당 단어에 대한 가중치를 높게 설정할 수 있다. 또한, 상기 사용자 데이터 텍스트 분석부(320)는 종래에 알려진 TFIDF(Term Frequency Inverse Document Frequency) 가중치 할당 방식을 이용하여 각 단어들의 가중치를 설정할 수도 있고, 상기 사용자 데이터 텍스트 추출부(310)에서 생성된 XML 파일에 기록된 문맥 정보를 활용하여 가중치를 설정할 수 있다. 예를 들어, 상기 데이터 텍스트 분석부(320)는 상기 XML 파일에 기록된 문맥 정보 중에서 생성 시간을 바탕으로 최근에 생성된 단어들에 보다 높은 가중치를 설정할 수 있으며, 이를 통해 사용자의 최근 관심 주제를 반영할 수 있는 효과를 얻을 수 있다.

[0036] 사용자 데이터 단어 벡터 분류부(330)는 상기 사용자 데이터 텍스트 분석부(320)에서 텍스트 내 단어들 각각에 대한 단어 벡터가 생성되면, 상기 이동통신 단말기에 임베디드된 주제 분류 트리(340)를 기반으로 상기 생성된 단어 벡터들을 분류하여 사용자 관심 주제를 결정한다. 여기서, 상기 주제 분류 트리(340)는 사용자의 관심 분야를 나타내기에 적합한 개념들을 주제별로 분류하여 트리 구조로 나타낸 것으로서, 종래에 사용되는 오픈 디렉토리 프로젝트(Open Directory Project)일 수 있다. 물론, 종래에 사용되는 일반적인 오픈 디렉토리 프로젝트는 웹 디렉토리인 만큼 광범위하기 때문에, 상기 오픈 디렉토리 프로젝트에서 사용자 관심 분야를 나타내기에 적합한 일부 카테고리들만을 추출하여 사용할 수 있을 것이다. 이때, 상기 주제 분류 트리(340)의 각 카테고리는 각 카테고리에 해당하는 웹 페이지들의 리스트를 포함할 수 있으며, 상기 웹 페이지들의 리스트는 해당 카테고리의 특성을 나타내는 단어들을 포함할 수 있을 것이다.

[0037] 상기 사용자 데이터 단어 벡터 분류부(330)는 상기 주제 분류 트리(340)의 각 카테고리에 포함된 웹 페이지들의 리스트를 참조하여 단어 벡터 분류를 위한 기계학습을 수행할 수 있으며, 이때 기계학습 알고리즘으로는 종래에 잘 알려진 Rocchio, K-NN, NB, SVM 등을 이용할 수 있다. 예를 들어, 상기 사용자 데이터 단어 벡터 분류부(330)는 A 카테고리에 a,b,c 웹 페이지가 포함된 경우, a 웹 페이지에 해당하는 단어 벡터를 상기 A 카테고리로 분류하도록 학습될 것이다. 학습이 완료된 이후 사용자 데이터 단어 벡터가 입력되면, 상기 사용자 데이터 단어 벡터 분류부(330)는 상기 주제 분류 트리(340)를 참조하여 상기 사용자 데이터 단어 벡터가 해당하는 카테고리를 결정하여, 결정된 카테고리의 주제를 상기 사용자 데이터 단어 벡터에 대응되는 사용자 관심 주제로 결정할 수 있다. 예를 들어, 상기 사용자 데이터 단어 벡터 분류부(330)는 도 6에 도시된 바와 같은 사용자 데이터로부터 추출된 단어 벡터들을 분류하여 도 8에 도시된 바와 같이, 다섯 개의 카테고리로 분류하여 사용자의 관심 주제로 결정할 수 있을 것이다.

[0039] 도 4는 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목 수집 및 분류부의 상세한 블록 구성을 도시하고 있다.

[0040] 상기 도 4를 참조하면, 상기 서비스 항목 수집 및 분류부(220)는 모바일 인터넷 서비스 항목 수집부(410), 서비스 항목 텍스트 분석부(420) 및 서비스 항목 단어 벡터 분류부(430)를 포함하여 구성된다.

[0041] 상기 모바일 인터넷 서비스 항목 수집부(410)는 모바일 인터넷에 접속하여 사용자에게 추천 가능한 서비스 항목(예: 뉴스, 모바일 앱(Mobile APP))들을 수집한다. 예를 들어, 상기 모바일 인터넷 서비스 항목 수집부(410)는 도 9에 도시된 바와 같이 상기 이동통신 단말기의 운영환경에 적합한 모바일 앱 사이트에서 최신 모바일 앱 정보를 수집하고, 정보 수집이 가능한 뉴스 포털 사이트에서 최신 뉴스들을 수집한다. 이때, 상기 모바일 인터넷 서비스 항목 수집부(410)는 상기 사용자 관심 주제 결정부(210)에서 결정된 사용자 관심 주제를 이용하여 관련된 서비스 항목들을 수집할 수 있다.

[0042] 상기 서비스 항목 텍스트 분석부(420)는 상기 모바일 인터넷 서비스 항목 수집부(410)에서 수집된 서비스 항목들로부터 텍스트를 추출하고, 추출된 텍스트를 분석하여 벡터 스페이스 모델에 따른 단어 벡터를 생성한다. 여기서, 상기 단어 벡터는 상기 수집된 서비스 항목들의 텍스트에 존재하는 개별적인 단어들로 이루어지며, 추출된 텍스트 내에서 각 단어들의 중요도에 따른 가중치를 반영한다. 이때, 상기 각 단어들의 중요도에 따른 가중치는 추출된 텍스트 내에서의 빈도 수, 해당 텍스트의 생성 시간과 생성 위치를 고려하여 계산할 수 있다. 즉, 상기 서비스 항목 텍스트 분석부(420)는 상기 사용자 데이터 텍스트 분석부(320)와 동일한 방식으로 상기 각 단어들의 중요도에 따른 가중치를 설정할 수 있을 것이다.

[0043] 상기 서비스 항목 단어 벡터 분류부(430)는 상기 이동통신 단말기에 임베디드된 주제 분류 트리(440)를 기반으로 상기 서비스 항목 텍스트 분석부(420)에서 생성된 단어 벡터들을 분류하여 상기 수집된 서비스 항목들 각각의 주제를 결정한다. 여기서, 상기 서비스 항목 단어 벡터 분류부(430)는 상기 사용자 데이터 단어 벡터 분류부

(330)와 동일한 방식으로 단어 벡터를 분류하여 해당 주제를 결정한다. 또한, 상기 서비스 항목 단어 벡터 분류부(430)에서 참조하는 주제 분류 트리(440)와 상기 사용자 데이터 단어 벡터 분류부(330)에서 참조되는 주제 분류 트리(340)는 동일한 주제 분류 트리이다.

[0045] 도 5는 본 발명의 실시 예에 따른 이동통신 단말기에서 서비스 항목 랭킹부의 상세한 블록 구성을 도시하고 있다.

[0046] 상기 도 5를 참조하면, 상기 서비스 항목 랭킹부(230)는 구문론적 매칭부(510)와 의미론적 매칭부(520) 및 통합 랭킹부(530)를 포함하여 구성된다.

[0047] 상기 구문론적 매칭부(510)는 상기 사용자 데이터 텍스트 분석부(320)에서 생성된 단어 벡터와 상기 서비스 항목 텍스트 분석부(420)에서 생성된 단어 벡터들간의 구문론적 유사도를 계산한다. 상기 구문론적 매칭부(510)는 하기 수학식 1과 같이 벡터 스페이스 모델의 코사인 유사도를 계산함으로써, 사용자 데이터에 대한 단어 벡터와 서비스 항목에 대한 단어 벡터 간의 구문론적 유사도를 계산한다.

[0048] 하기 수학식 1은 코사인 유사도 계산식을 나타낸다.

수학식 1

$$SyntaticScore(\vec{u}, \vec{s}_i) = \cos(\vec{u}, \vec{s}_i) = \frac{\vec{u} \cdot \vec{s}_i}{\|\vec{u}\| \|\vec{s}_i\|}$$

[0049]

[0050] 여기서, $\vec{u}$ 는 사용자 데이터에 대한 단어 벡터를 나타내고, $\vec{s}_i$ 는 서비스 항목에 대한 단어 벡터를 나타낸다.

[0051] 상기 의미론적 매칭부(520)는 상기 사용자 데이터 단어 벡터 분류부(330)에서 추출된 사용자 관심 주제와 상기 서비스 항목 단어 벡터 분류부(430)에서 파악된 서비스 항목의 주제들 간의 의미론적 유사도를 계산한다. 상기 의미론적 매칭부(520)는 주제 분류 트리에서 각 카테고리들 간의 의미적 유사도를 나타내는 유사도 그래프(540)에 wPPR(weighted Personalized PageRank) 알고리즘을 적용하여 상기 사용자 관심 주제와 서비스 항목 주제들 간의 의미적 유사도를 계산한다. 여기서, 상기 유사도 그래프(540)는 상기 주제 분류 트리에서 각 카테고리들의 계층적 트리 구조를 카테고리들 사이의 의미적인 유사도에 따라 연결된 그래프 구조로 변환한 것으로서, 상기 유사도 그래프(540)의 각 노드는 상기 주제 분류 트리의 각 카테고리를 의미하며, 각 노드 간의 연결고리는 해당 카테고리 간에 의미론적 유사도가 존재함을 의미한다. 또한, 상기 wPPR 알고리즘은 종래에 잘 알려진 개인화된 페이지랭크(Personalized PageRank) 알고리즘에 가중치를 적용한 것으로서 하기에서 자세히 설명하기로 한다.

[0052] 상기 유사도 그래프(540)의 각 노드들 간의 링크를 생성하는 방법은 하기와 같이 3 단계로 이루어진다. 상기 유사도 그래프(540)는 상기 의미론적 매칭부(520)에서 생성될 수도 있으며, 상기 이동통신 단말기의 다른 기능 블록에서 생성될 수도 있을 것이다.

[0053] 1 단계는 상기 주제 분류 트리의 각 카테고리에 대한 센트로이드 벡터(Centroid Vector)를 하기 수학식 2와 같이 계산하는 단계이다. 상기 센트로이드 벡터는 각 카테고리의 학습 데이터들을 대표하는 벡터로서, 학습 데이터들의 평균 단어 벡터이다. 여기서 상기 학습 데이터는 상기 사용자 데이터 단어 벡터 분류부(330) 혹은 서비스 항목 단어 벡터 분류부(430)의 기계학습 시에 사용된 오픈 디렉토리 프로젝트의 웹 페이지들 일 수 있다.

[0054] 하기 수학식 2는 각 카테고리에 대한 센트로이드 벡터 계산식을 나타낸다.

수학식 2

$$\vec{\mu}(c) = \frac{1}{|D_c|} \sum_{d \in D_c} \vec{v}(d)$$

[0055]

[0056] 여기서, 상기 c 는 카테고리를 의미하며, 상기 $\vec{\mu}(c)$ 는 카테고리 c 의 센트로이드 벡터를 의미하고, D_c 는 카테고리 c 의 학습 데이터 집합을 의미하며, $\vec{v}(d)$ 는 학습 데이터 d 에 대한 단어 벡터를 의미한다.

[0057] 2 단계는 상기 주제 분류 트리의 각 카테고리에 대한 병합 센트로이드 벡터(Merge Centroid Vector)를 하기 수학식 3과 같이 계산하는 단계이다. 상기 병합 센트로이드 벡터는 상기 각 카테고리들의 센트로이드 벡터에 주제 분류 트리의 계층적 구조의 특징을 반영하여 나타낸다. 즉, 상기 센트로이드 벡터는 해당 카테고리에 대한 특징만을 포함하고 주제 분류 트리 내에서 카테고리들 간의 계층적 관계를 반영하지는 않으나, 상기 병합 센트로이드 벡터는 부모 카테고리가 자식 카테고리의 특징을 포함할 수 있도록 부모 카테고리의 센트로이드 벡터에 후손 카테고리의 센트로이드 벡터들의 특징을 포함시켜 나타낸다.

[0058] 하기 수학식 3은 병합 센트로이드 벡터 계산식을 나타낸다.

수학식 3

$$\vec{\mu}'(c) = \frac{1}{1+|child(c)|} \left(\frac{\vec{\mu}(c)}{\|\vec{\mu}(c)\|} + \sum_{c_k \in child(c)} \frac{\vec{\mu}'(c_k)}{\|\vec{\mu}'(c_k)\|} \right)$$

[0059]

[0060] 여기서, 상기 $\vec{\mu}'(c)$ 는 카테고리 c 에 대한 병합 센트로이드 벡터를 의미하고, $child(c)$ 는 카테고리 c 의 자식 카테고리들의 집합을 의미하며, $\vec{\mu}(c)$ 는 카테고리 c 에 대한 센트로이드 벡터를 의미한다.

[0061] 3단계는 모든 카테고리들 간의 의미적 유사도를 계산하는 단계이다. 상기 카테고리들 간의 의미적 유사도는 상기 카테고리들에 대한 병합 센트로이드 벡터(Merge Centroid Vector)들 간의 코사인 유사도(Cosine Similarity)를 의미한다. 여기서 상기 의미론적 매칭부(520)는 상기 카테고리들 간의 의미적 유사도를 임계값과 비교하여 카테고리들 간의 의미적 유사도가 임계값 이상인 경우, 해당 카테고리들 간의 링크를 생성하여 유사도 그래프를 생성한다. 이때 상기 의미론적 매칭부(520)는 도 10에 도시된 바와 같이 상기 의미적 유사도를 생성된 링크에 대한 가중치로 설정한다. 여기서, 상기 링크에 대한 가중치는 병합 센트로이드 벡터들 간의 코사인 유사도로 계산되었으므로, 상기 주제 분류 트리의 계층적 구조의 특징을 나타낼 수 있다.

[0062] 상기 의미론적 매칭부(520)는 본 발명에서 제안하는 wPPR 알고리즘에 따라 상기와 같이 생성된 유사도 그래프(540)에서 각 카테고리에 대한 다른 카테고리의 의미론적 유사도를 랭킹한다.

[0063] 이를 위해, 상기 의미론적 매칭부(520)는 먼저 상기 유사도 그래프(540)의 링크 가중치에 기초하여 연관성 행렬(Relevance Matrix) R 을 산출한다. 상기 연관성 행렬 R 의 (i, j) 성분인 r_{ij} 는 j 번째 카테고리에 대한 i 번째 카테고리의 의미론적 유사도를 의미한다. 즉, 상기 의미론적 매칭부(520)는 종래에 잘 알려진 마코브 랜덤 워크 모델(Markov Random Walk Model) 중 하나인 개인화된 페이지랭크(Personalized PageRank)를 활용하여 유사도 그래프를 순회하는 랜덤 서퍼(Random Surfer)가 각각의 카테고리를 방문할 확률을 계산한다. 좀더 자세히 말해, 상기 의미론적 매칭부(520)는 랜덤 서퍼가 j 번째 카테고리에서 i 번째 카테고리를 방문할 확률을 계산하고, 계산된 값을 두 카테고리 간의 의미론적 유사도로 결정하며, 상기 의미론적 유사도의 크기에 따라 상기 j 번째 카테

고리에 대한 다른 카테고리들의 의미론적 유사도를 랭킹할 수 있다.

[0064] 여기서, 상기 랜덤 서퍼가 유사도 그래프를 순회하는 패턴은 두 가지의 경우로 정의할 수 있으며, 상기 랜덤 서퍼는 매 순간마다 (1-d)의 확률로 첫 번째 순회 패턴에 따라 유사도 그래프를 순회하며, d의 확률로 두 번째 순회 패턴에 따라 상기 유사도 그래프를 순회한다. 여기서, 상기 d는 댐핑 팩터(damping factor)로써 0에서 1 사이의 실수 값을 가지며 실험에 따라 경험적으로 최적의 값을 찾을 수 있다. 상기 랜덤 서퍼가 첫 번째 순회 패턴을 따를 경우, 상기 랜덤 서퍼가 방문하는 카테고리는 유사도 그래프 내 신뢰하는 카테고리, 즉 본 발명에서는 현재 방문한 j번째 카테고리가 되고, 상기 랜덤 서퍼가 두 번째 순회 패턴을 따를 경우, 상기 랜덤 서퍼는 현재 방문 중인 카테고리나 링크로 연결된 카테고리를 링크 가중치와 비례하는 확률로 방문하게 된다. 상기 랜덤 서퍼가 두 번째 순회 패턴을 따를 때 다음 카테고리로 이동할 확률은 하기 수학식 4와 같이 계산한다.

[0065] 하기 수학식 4는 랜덤 서퍼의 이동 확률 계산식이다.

수학식 4

$$tw_{ij} = \frac{sim(c_i, c_j)}{\sum_{c_k \in N(c_j)} sim(c_k, c_j)}$$

[0066] 여기서, 상기 tw_{ij} 는 상기 랜덤 서퍼가 j번째 카테고리에서 i번째 카테고리로 이동할 확률을 나타내며, $sim(c_i, c_j)$ 는 상기 카테고리 간의 유사도 즉, 카테고리 간의 링크 가중치를 나타내고, 상기 $N(c_j)$ 는 c_j 와 연결된 카테고리의 집합을 나타낸다.

[0068] 도 11은 본 발명의 실시 예에서 두 번째 순회 패턴을 따를 때 상기 tw_{ij} 값을 구하는 방법을 도시하고 있다. 상기 도 11을 참조하면, 분류 트리 내 카테고리들 간의 이동 확률(tw_{ij} 값)은 도 10에 도시된 바와 같은 각 카테고리들 간의 링크 가중치에 기초하여 결정된다. 예를 들어, 현재 랜덤 서퍼가 sports(1001) 카테고리를 방문하고 있으며, d의 확률로 두 번째 순회 패턴에 따라 유사도 그래프를 순회한다고 가정하면, 상기 Sports(1001) 카테고리와 링크로 연결된 Soccer(1003), Baseball(1005) 및 Shopping(1007) 카테고리 각각으로의 이동 확률은 상기 Sports(1003) 카테고리와 연결된 전체 링크의 가중치 합($1.7 = 0.7(1013) + 0.7(1015) + 0.3(1017)$)과 해당 링크의 가중치의 비율에 따라 결정된다. 즉, 도 11에 도시된 바와 같이, 상기 Sports(1001) 카테고리에서 Soccer(1003) 카테고리로 랜덤 서퍼가 이동할 확률은 $0.7/1.7(1113)$ 이 되고, 상기 Sports(1001) 카테고리에서 Baseball(1005) 카테고리로 랜덤 서퍼가 이동할 확률은 $0.7/1.7(1115)$ 이 되며, 상기 Sports(1001) 카테고리에서 Shopping(1007) 카테고리로 랜덤 서퍼가 이동할 확률은 $0.3/1.7(1117)$ 이 된다.

[0069] 상기 랜덤 서퍼의 두 가지 순회 패턴에 대한 정의를 바탕으로 하여 상기 r_{ij} 는 하기 수학식 5와 같이 계산할 수 있다.

수학식 5

$$r_{ij} = d[\sum_{c_k \in I(c_i)} tw_{ik} \cdot r_{kj}] + (1-d)t_{ij}$$

[0070] 여기서, $I(c_i)$ 는 c_i 에 대한 링크를 가지는 카테고리들의 집합을 나타내며, t_{ij} 는 첫 번째 순회 패턴을 계산하기 위한 것으로, 현재 카테고리가 c_j 로 설정된 경우 c_i 의 신뢰도(trusted weight)를 나타낸다. 따라서, i가 j일 경우에는 1 나머지 경우에는 0으로 설정된다.

[0072] 상기 수학식 5를 행렬 표기법으로 재정의하면, 하기 수학식 6과 같이 나타낼 수 있다.

수학식 6

$$R_t = d[W \cdot R_{t-1}] + (1-d)T$$

[0073]

[0074]

여기서, 상기 R은 wPPR 알고리즘에 따라 계산하는 연관성 행렬을 나타내며, W는 변화 행렬(Transition Matrix)을 나타내는 것으로, (i, j)번째 성분은 상기 수학식 4의 tw_{ij} 와 동일하다. 상기 T는 신뢰 행렬로서 (i, j)번째 성분은 상기 수학식 4의 t_{ij} 와 동일하므로, 상기 T는 단위 행렬이 된다.

[0075]

즉, 상기 의미론적 매칭부(520)는 상기 수학식 6과 같이 연관성 행렬을 바탕으로 임의의 카테고리 간의 의미론적 유사도를 수치화할 수 있다. 이때, 상기 카테고리는 사용자 관심 주제 혹은 서비스 항목의 주제에 대응되는 카테고리일 수 있다. 따라서, 상기 의미론적 매칭부(520)는 상기 사용자 관심 주제가 카테고리(c_j)로 결정된 경우, 상기 서비스 항목의 주제에 대응하는 카테고리(c_i)와의 의미론적 유사도를 연관성 행렬의 (i, j) 성분 값으로 결정할 수 있다.

[0076]

상기 통합 랭킹부(530)는 상기 구문론적 매칭부(510)와 의미론적 매칭부(520)에서 각각 계산된 구문론적 유사도와 의미론적 유사도를 선형 조합하여 하기 수학식 7과 같이 전체 유사도를 계산한다.

[0077]

하기 수학식 7은 전체 유사도 계산식을 나타낸다.

수학식 7

$$TotalScore(u, s_i) = (1 - \lambda) \times SyntacticScore(\vec{u}, \vec{s}_i) + \lambda \times SemanticScore(uc, sc_i)$$

[0078]

[0079]

여기서, 상기 uc는 사용자 데이터로부터 추출된 사용자 관심 주제에 대응되는 카테고리이고, 상기 sc_i 는 서비스 항목의 주제에 대응되는 카테고리를 나타내며, λ 는 선형 조합에서 의미론적 유사도에 대한 가중치를 나타내는 것으로 0~1.0의 값으로 실험을 통해 결정될 수 있다.

[0080]

상기 통합 랭킹부(530)는 상기 수학식 7을 통해 계산된 전체 유사도를 바탕으로 사용자에게 추천할 서비스 항목들의 연관성 순위를 결정한다.

[0082]

도 13은 본 발명의 실시 예에 따른 이동통신 단말기에서 사용자 관심 주제에 따라 서비스 항목을 추천하는 절차를 도시하고 있다.

[0083]

상기 도 13을 참조하면, 상기 이동통신 단말기는 1301단계에서 미리 설정된 사용자 데이터 수집 주기인지 여부를 검사한다. 상기 사용자 데이터 수집 주기인 경우, 상기 이동통신 단말기는 1303단계로 진행하여 상기 이동통신 단말기 내에 존재하는 사용자 데이터로부터 텍스트를 추출한다. 여기서, 상기 이동통신 단말기는 단문 메시지, 멀티미디어 메시지, 이메일, 파일, 스케줄, 메모 및 웹 사용 정보 등과 같은 상기 사용자 데이터로부터 텍스트를 추출할 수 있다. 이후, 상기 이동통신 단말기는 1305단계에서 상기 추출된 사용자 데이터의 텍스트를 기반으로 단어 벡터를 생성하고, 1307단계로 진행하여 상기 이동통신 단말기에 임베디드된 주제 분류 트리에 따라 상기 단어 벡터를 분류하여 사용자 관심 주제를 결정한 후, 하기 1317단계로 진행한다.

[0084]

반면, 상기 사용자 데이터 수집 주기가 아닌 경우 상기 이동통신 단말기는 1309단계로 진행하여 기 설정된 서비스 항목 수집 주기인지 여부를 검사한다. 상기 이동통신 단말기는 상기 서비스 항목 수집 주기가 아닐 경우, 상기 1301단계로 되돌아가 이하 단계를 재수행하는 반면 상기 서비스 항목 수집 주기일 경우, 1311단계로 진행하여 모바일 인터넷에 접속하여 인터넷 서비스 항목을 수집하고, 수집된 서비스 항목으로부터 텍스트를 추출한다. 이후, 상기 이동통신 단말기는 1313단계로 진행하여 상기 추출된 서비스 항목의 텍스트를 기반으로 단어 벡터를 생성하고, 1315단계로 진행하여 상기 주제 분류 트리에 따라 상기 단어 벡터를 분류하여 서비스 항목에 대한 주제를 결정한 후, 하기 1317단계로 진행한다.

[0085]

상기 이동통신 단말기는 1317단계에서 사용자에게 의해 서비스 항목 추천 이벤트가 발생되는지 여부를 검사한다.

상기 이동통신 단말기는 상기 서비스 항목 추천 이벤트가 발생되지 않을 시 상기 1301단계로 되돌아가 이하 단계를 재수행하는 반면, 상기 서비스 항목 추천 이벤트가 발생될 시 1319단계로 진행하여 상기 사용자 데이터의 단어 벡터와 상기 서비스 항목의 단어 벡터 간의 구문론적 유사도를 계산한다. 이때, 상기 이동통신 단말기는 상기 수학적 식 1을 이용하여 사용자 데이터의 단어 벡터와 상기 서비스 항목의 단어 벡터 간의 구문론적 유사도를 계산할 수 있다.

[0086] 이후, 상기 이동통신 단말기는 1321단계에서 상기 사용자 관심 주제와 상기 서비스 항목에 대한 주제 간의 의미론적 유사도를 계산한다. 여기서, 상기 이동통신 단말기는 주제 분류 트리에서 각 카테고리들을 의미적인 유사도에 따라 연결하고, 각 카테고리들 간의 연결 링크에 대한 가중치를 계산한 후, 종래의 잘 알려진 일반화된 페이지랭크 알고리즘에 기초하여 랜덤 서퍼의 이동 확률을 계산하여 각 카테고리들 간의 의미론적 유사도를 결정함으로써, 상기 사용자 관심 주제와 서비스 항목에 대한 주제 간의 의미론적 유사도를 계산할 수 있다. 즉, 상기 이동통신 단말기는 상기 수학적 식 5를 이용하여 상기 사용자 관심 주제와 상기 서비스 항목에 대한 주제 간의 의미론적 유사도를 계산할 수 있다.

[0087] 이후, 상기 이동통신 단말기는 1323단계에서 상기 계산된 구문론적 유사도와 의미론적 유사도를 바탕으로 사용자 관심 주제와 서비스 항목 주제 간의 전체 유사도를 나타내는 연관성 순위를 결정하고, 1325단계로 진행하여 상기 결정된 연관성 순위에 따라 서비스 항목을 추천한다. 여기서, 상기 이동통신 단말기는 상기 수학적 식 7을 이용하여 사용자 관심 주제와 서비스 항목 주제 간의 전체 유사도를 계산할 수 있으며, 사용자 관심 주제에 대한 전체 유사도 값이 큰 서비스 항목 주제일수록 높은 연관성 순위를 갖도록 결정할 수 있다.

[0088] 이후, 상기 이동통신 단말기는 본 발명에 따른 알고리즘을 종료한다.

[0089] 여기서는, 사용자 데이터 및 서비스 항목을 소정 주기마다 수집하는 것에 대해 설명하였으나, 사용자에 의해 서비스 항목 추천 이벤트가 발생하는 경우에만 상기 사용자 데이터 및 서비스 항목을 수집할 수도 있을 것이다.

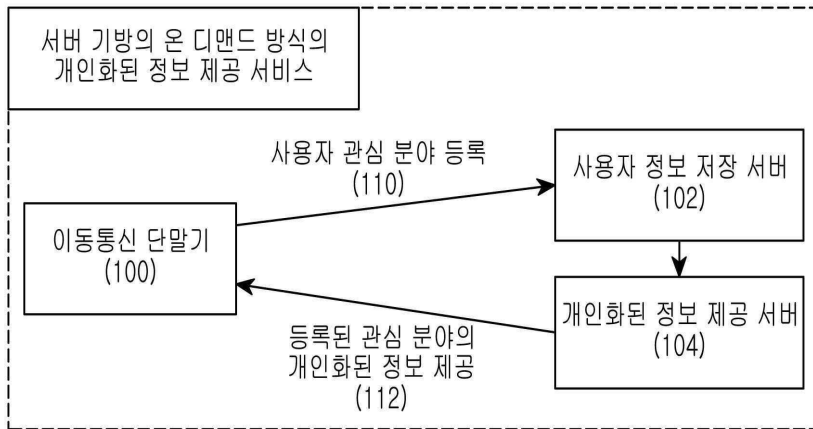
[0091] 한편 본 발명의 상세한 설명에서는 구체적인 실시 예에 관해 설명하였으나, 본 발명의 범위에서 벗어나지 않는 한도 내에서 여러 가지 변형이 가능하다. 그러므로 본 발명의 범위는 설명된 실시 예에 국한되어 정해져서는 아니 되며 후술하는 특허청구의 범위뿐만 아니라 이 특허청구의 범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

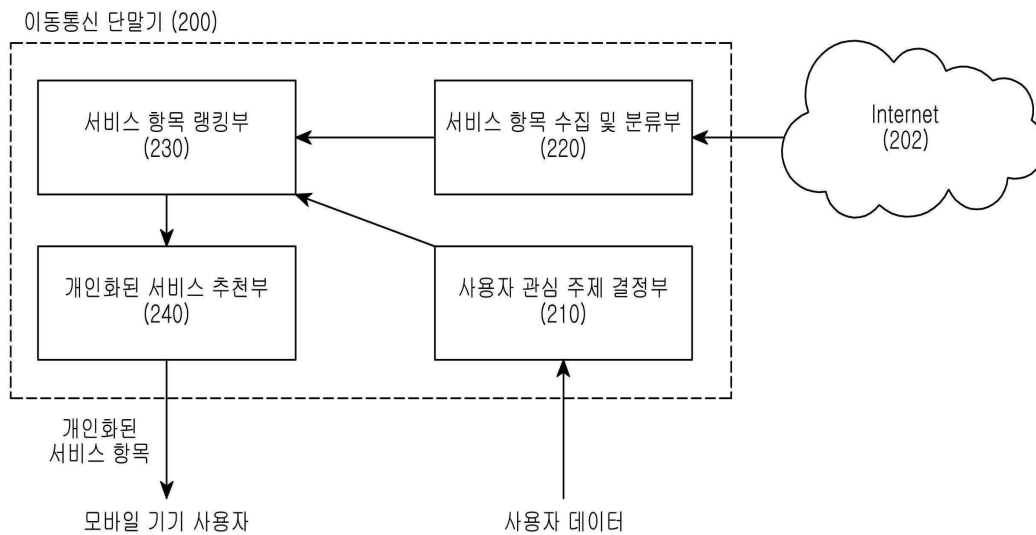
- [0093]
- 210: 사용자 관심 주제 결정부
 - 220: 서비스 항목 수집 및 분류부
 - 230: 서비스 항목 랭킹부
 - 240: 개인화된 서비스 추천부

도면

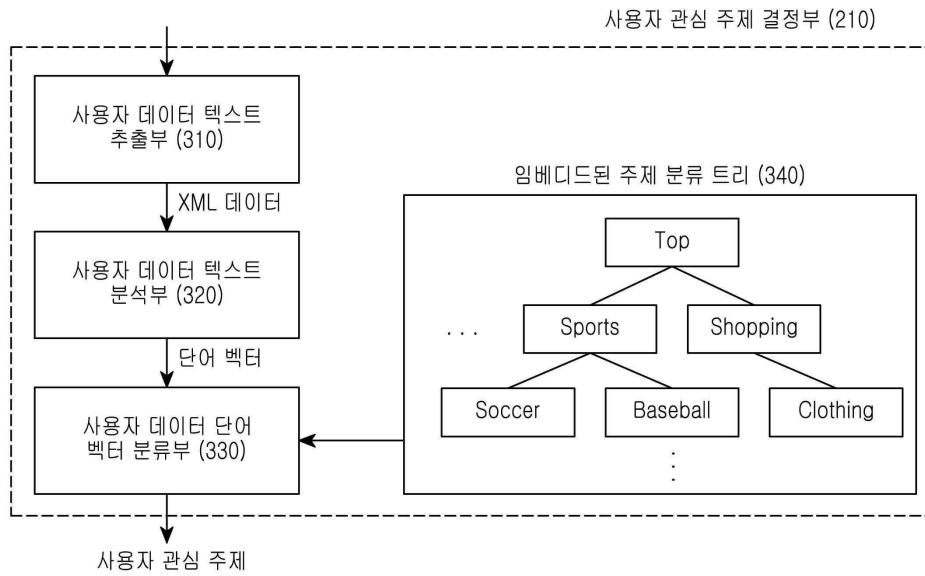
도면1



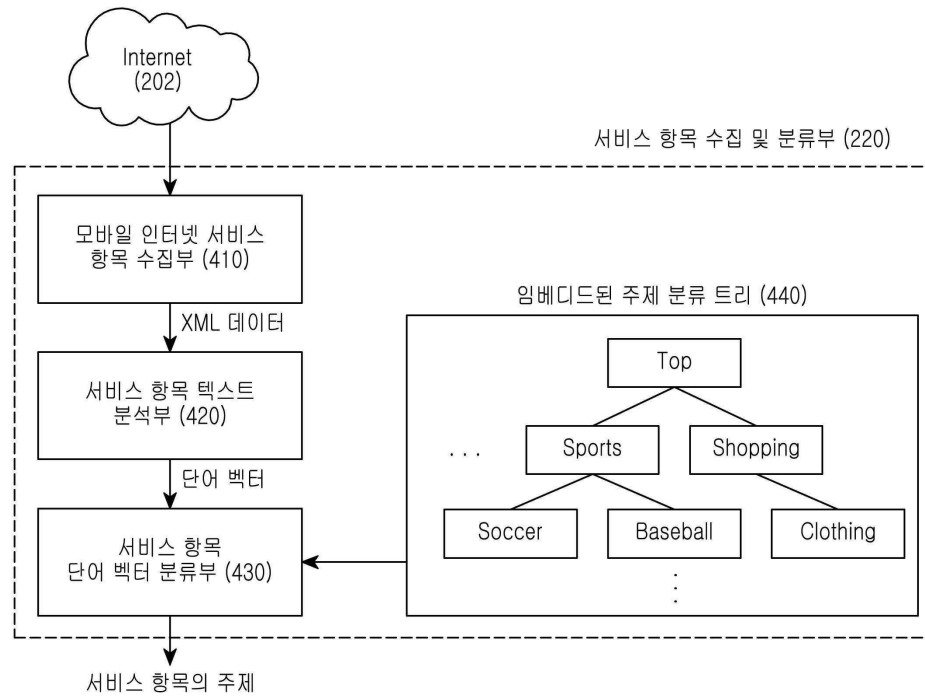
도면2



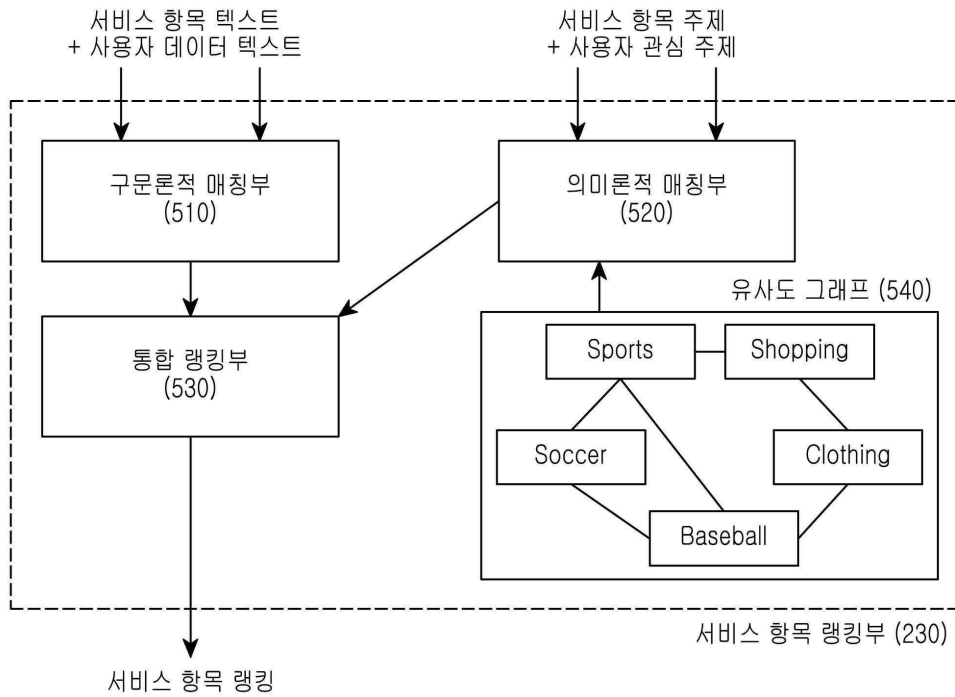
도면3



도면4



도면5



도면6



도면7

```
<xml version="1.0" encoding="utf-16">
<ExtractData>
  (ReceiveSMSList)
    <SMS date="20091027 09:49:36" text="Ok. I will give you some information. I"
    <SMS date="20091027 09:49:06" text="Welcom. If you have any question, please"
    <SMS date="20091026 03:48:21" text="Ok. I will give you some information ab"
    <SMS date="20091026 03:48:06" text="Yes, car vendor said that they have su"
    <SMS date="20091026 03:47:44" text="Yes, picture and other information. Do"
    <SMS date="20091026 03:47:30" text="I think so, you can choose e2l, e2b, e3"
    <SMS date="20091026 03:47:16" text="Yes, like 7-series. And the youngsters li"
    <SMS date="20091026 03:47:03" text="Ah- my friend. Lee has BMW. It is very"
  </ReceiveSMSList>
  (SendSMSList)
    <SMS date="20091027 09:49:48" text="Yes, ok. bye- name = 'Kevin' />
    <SMS date="20091027 09:49:23" text="Oh- sorry, please give my review about"
    <SMS date="20091026 03:46:38" text="Thank you. I will wait your reply-." name="Kevin"
    <SMS date="20091026 03:46:23" text="Oh- I am interested that. Is there 740i"
    <SMS date="20091026 03:46:10" text="Sure, especially 540i and 740i. And do"
    <SMS date="20091026 03:45:56" text="Ok. Do you have any information of that"
    <SMS date="20091026 03:45:42" text="Oh, but mini is very small for me. how"
    <SMS date="20091026 03:45:28" text="Really? however, their are many model"
    <SMS date="20091026 03:45:11" text="I have a plan to buy a new car. Would y"
  </SendSMSList>
  (FileInfoList)
    (FileInfo date="20091027 11:17:24" text="2009-bmw-7-series-2.jpg" app="I"
    (FileInfo date="20091027 11:12:04" text="BMW 740i.jpg" app="Image" exten"
    (FileInfo date="20091027 11:24:50" text="BMW MOTORRAD DAYS 2009-Gb6" app="p"
    (FileInfo date="20091027 11:25:36" text="BMW List" app="Document" exten"
    (FileInfo date="20091027 11:23:08" text="BMW_Finance_final_" app="Docu"
    (FileInfo date="20091027 11:12:34" text="540i picture.jpg" app="Image" e"
  </FileInfoList>
  (EmailList)
    (EmailInfo date="20091025 20:01:54" subject="2011 BMW 5 Series F10 sold in"
    (EmailInfo date="20091025 20:01:41" subject="New set of 2011 BMW 5 Series"
    (EmailInfo date="20091025 20:01:28" subject="BMW believes in 'Green'" ser"
    (EmailInfo date="20091025 20:01:13" subject="Video with the rumored BMW M3"
    (EmailInfo date="20091025 20:00:56" subject="Video: WLMV Driver vs. Non WLM"
    (EmailInfo date="20091025 20:00:28" subject="New interior and exterior spy"
  </EmailList>
  (WebUsageInfoList)
    (WebUsageInfo date="20091027 08:00:00" text="BMW 5 Series Sedan Introduce"
    (WebUsageInfo date="20091027 08:00:00" text="BMW M models Overview" url="

```

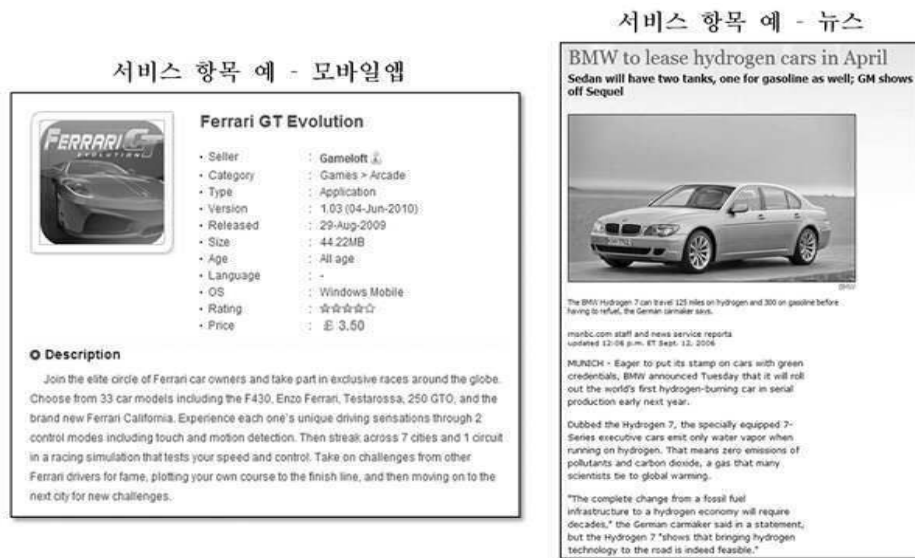
도면8

	Similarity
es and Models/BMW	0.71219
es and Models/BMW/3 Series	0.62295
es and Models/BMW/Clubs	0.60430
es and Models/BMW/5 Series	0.56298
es and Models/BMW/7 Series	0.49664

확인

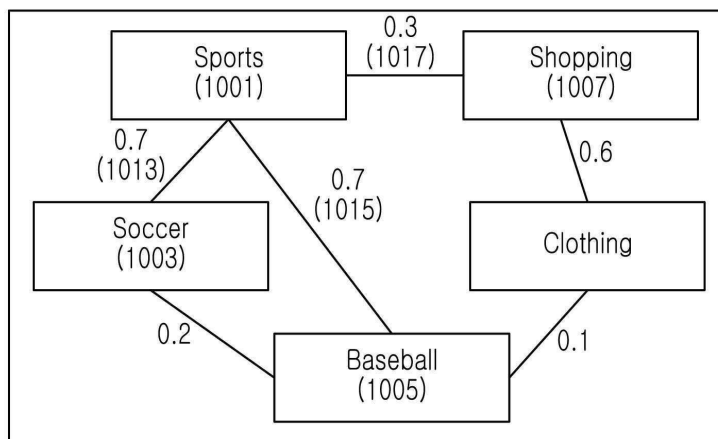
가

도면9

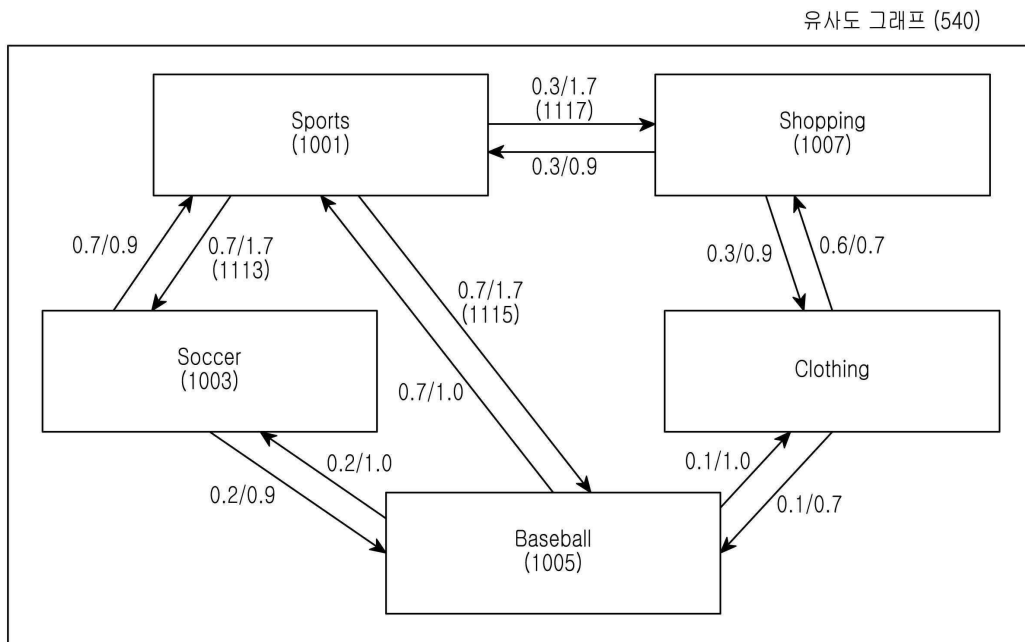


도면10

유사도 그래프 (540)



도면11



도면12



도면13

