



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0106699
(43) 공개일자 2013년09월30일

(51) 국제특허분류(Int. Cl.)
G06Q 50/10 (2012.01) G06F 17/30 (2006.01)
G06Q 50/30 (2012.01)
(21) 출원번호 10-2012-0028438
(22) 출원일자 2012년03월20일
심사청구일자 2012년03월20일

(71) 출원인
고려대학교 산학협력단
서울 성북구 안암동5가 1
(72) 발명자
정인정
경기도 용인시 수지구 성북동 성동마을LG빌리지1
차아파트 102-1903
송태성
경기도 의정부시 호원동 우성아파트 103동 402호
(뒷면에 계속)
(74) 대리인
특허법인이지

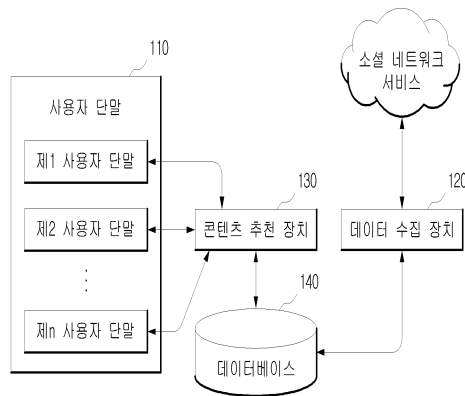
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 콘텐츠 추천 방법 및 그 시스템

(57) 요약

콘텐츠 추천 방법 및 그 시스템이 개시된다. 콘텐츠 추천 방법은 콘텐츠를 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있다.

대표도 - 도1



(72) 발명자

페이 윤평

중국, 랴오닝 프라빈스, 선양 시티, 허핑 디스트릭
트, 푸순 로드, 넘버 66-14, 룸 11-5

손종수

충청남도 연기군 조치원읍 서창리 170-2번지

특허청구의 범위

청구항 1

키워드를 입력받는 단계;

상기 키워드가 포함된 적어도 하나의 콘텐츠 출판 정보를 추출하는 단계;

상기 추출된 콘텐츠 출판 정보를 이용하여 상기 키워드와 콘텐츠간의 유사도를 도출하는 단계;

상기 추출된 콘텐츠 출판 정보에 대응하는 사용자 프로필 정보를 이용하여 상기 콘텐츠를 배포한 사용자의 중요도를 도출하는 단계; 및

상기 도출된 유사도 및 상기 사용자의 중요도를 이용하여 상기 추출된 콘텐츠에 대한 순위를 계산하여, 순위가 높은 순으로 콘텐츠를 추천하는 단계를 포함하는 콘텐츠 추천 방법.

청구항 2

제1 항에 있어서,

상기 유사도를 도출하는 단계는,

상기 콘텐츠 출판 정보를 이용하여 상기 키워드에 대한 빈도수를 도출하고, 전체 콘텐츠 중에서 상기 키워드가 나타난 출현 빈도를 도출하고, 상기 빈도수 및 상기 출현 빈도를 이용하여 상기 키워드와 상기 콘텐츠간의 유사도를 도출하는 것을 특징으로 하는 콘텐츠 추천 방법.

청구항 3

제1 항에 있어서,

상기 사용자의 중요도를 도출하는 단계는,

상기 콘텐츠 출판 정보를 배포한 각 사용자에게 연결된 다른 사용자수를 이용하여 상기 사용자의 중요도를 도출하는 것을 특징으로 하는 콘텐츠 추천 방법.

청구항 4

제1 항에 있어서,

상기 키워드를 입력받는 단계 이전에,

상기 콘텐츠 출판 정보 및 상기 사용자 프로필 정보를 수집하는 단계를 더 포함하는 것을 특징으로 하는 콘텐츠 추천 방법.

청구항 5

사용자 단말로 콘텐츠를 추천하는 콘텐츠 추천 시스템에 있어서,

상기 사용자 단말로부터 키워드를 입력받고, 상기 키워드가 포함된 적어도 하나의 콘텐츠 출판 정보를 추출하는 추출부;

상기 추출된 콘텐츠 출판 정보를 이용하여 상기 키워드와 콘텐츠간의 유사도를 도출하고, 상기 추출된 콘텐츠 출판 정보에 대응하는 사용자 프로필 정보를 이용하여 상기 콘텐츠를 배포한 사용자의 중요도를 도출하는 분석부;

상기 도출된 유사도 및 상기 사용자의 중요도를 이용하여 상기 추출된 콘텐츠에 대한 순위를 계산하여, 순위가 높은 순으로 콘텐츠를 추천하는 추천부를 포함하는 콘텐츠 추천 시스템.

청구항 6

제5 항에 있어서,

상기 분석부는,

상기 콘텐츠 출판 정보를 이용하여 상기 키워드에 대한 빈도수를 도출하고, 전체 콘텐츠 중에서 상기 키워드가 나타난 출현 빈도를 도출하고, 상기 빈도수 및 상기 출현 빈도를 이용하여 상기 키워드와 상기 콘텐츠간의 유사도를 도출하는 것을 특징으로 하는 콘텐츠 추천 시스템.

청구항 7

제5 항에 있어서,

상기 분석부는,

상기 콘텐츠 출판 정보를 배포한 각 사용자에 연결된 다른 사용자수를 이용하여 상기 사용자의 중요도를 도출하는 것을 특징으로 하는 콘텐츠 추천 시스템.

청구항 8

제5 항에 있어서,

상기 콘텐츠 출판 정보 및 상기 사용자 프로필 정보를 수집하는 정보 수집단계를 장치를 더 포함하는 것을 특징으로 하는 콘텐츠 추천 시스템.

청구항 9

제1 항 내지 제4항 중 어느 하나의 항에 따른 콘텐츠 추천 방법을 수행하기 위한 프로그램을 기록한 기록매체.

명세서

기술분야

[0001] 본 발명은 콘텐츠 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있는 방법 및 시스템에 관한 것이다.

배경기술

[0002] 인터넷이 발달하고 이에 수반되는 각종 정보기술 및 통신 네트워크가 발달이 눈부시게 진행됨에 따라서, 수반되는 각종 매체와 콘텐츠의 양이 정보의 홍수(Deluge of information)로 표현될 만큼 폭발적으로 증가하고 있다.

[0003] 이에 따라, 사용자들이 접할 수 있는 아이템의 수와 종류가 대단히 많아지고 있으며, 사용자들이 원하고 필요로 하는 아이템을 얻기가 점점 어려워지고 있다. 따라서 사용자가 원하고 필요로 하는 아이템을 알맞게 추천하는 콘텐츠 추천 시스템(CRS, Contents Recommendation System)이 최근 중요한 이슈로 대두되었으며, 이에 대한 콘텐츠 추천 방법(CRM, Contents Recommendation Method)이 활발히 연구, 소개되어왔다. 종래의 콘텐츠 추천 방법에는 내용 기반 추천(Content-based recommendation) 방법과 협업적 추천(Collaborative recommendation) 방법이 있다. 그리고, 위 두 가지 방법을 혼합한 하이브리드 추천(Hybrid recommendation) 방법이 있다. 내용기반 추천 방법은 사용자의 선호도와 콘텐츠의 텍스트, 태그 등의 정보를 활용하여 사용자에게 적절한 콘텐츠를

추천하는 방법이다. 내용기반 추천 방법을 활용하면 사용자의 선호도에 기초한 콘텐츠를 선별하고 추천할 수 있지만 사용자들 간의 상호작용이 반영되지 못한다는 단점이 있다. 그리고 협업적 추천 방법은 어떤 콘텐츠에 대해 여러 사용자가 협업적으로 평가하고 그 평가된 값에 의하여 콘텐츠를 추천하는 방법이다. 협업적 추천 방법은 콘텐츠를 추천하는데 있어 여러 사용자의 주관적인 평가값을 반영하기 때문에 사용자의 경험과 지식의 축적하는데 효율적이다. 그러나 주제에 대한 콘텐츠의 적합성 등을 평가하기 위한 객관적 근거가 없는 한계를 가지고 있다. 마지막으로 하이브리드 추천 방법은 인공지능, 데이터마이닝 등에서 활용되는 알고리즘들을 콘텐츠 추천에 활용함으로써 언급한 두 추천 방법의 한계점을 극복하고자 한 방법들이다. 그러나 대부분의 하이브리드 콘텐츠 추천 방법들은 특정 도메인에 국한된 콘텐츠를 추천하는 것에만 활용되며 콘텐츠를 생성한 사용자의 신뢰성 및 위상이 반영되고 있지 않다.

발명의 내용

해결하려는 과제

- [0004] 본 발명은 콘텐츠 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있는 방법 및 시스템을 제공하기 위한 것이다.

과제의 해결 수단

- [0005] 본 발명의 일 측면에 따르면, 콘텐츠 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있는 방법 및 그 방법을 수행하기 위한 프로그램을 기록한 기록매체가 제공된다.
- [0006] 본 발명의 일 실시예에 따르면, 키워드를 입력받는 단계; 상기 키워드가 포함된 적어도 하나의 콘텐츠 출판 정보를 추출하는 단계; 상기 추출된 콘텐츠 출판 정보를 이용하여 상기 키워드와 콘텐츠간의 유사도를 도출하는 단계; 상기 추출된 콘텐츠 출판 정보에 대응하는 사용자 프로필 정보를 이용하여 상기 콘텐츠를 배포한 사용자의 중요도를 도출하는 단계; 및 상기 도출된 유사도 및 상기 사용자의 중요도를 이용하여 상기 추출된 콘텐츠에 대한 순위를 계산하여, 순위가 높은 순으로 콘텐츠를 추천하는 단계를 포함하는 콘텐츠 추천 방법이 제공될 수 있다.
- [0007] 상기 유사도를 도출하는 단계는, 상기 콘텐츠 출판 정보를 이용하여 상기 키워드에 대한 빈도수를 도출하고, 전체 콘텐츠 중에서 상기 키워드가 나타난 출현 빈도를 도출하고, 상기 빈도수 및 상기 출현 빈도를 이용하여 상기 키워드와 상기 콘텐츠간의 유사도를 도출할 수 있다.
- [0008] 상기 사용자의 중요도를 도출하는 단계는, 상기 콘텐츠 출판 정보를 배포한 각 사용자에게 연결된 다른 사용자수를 이용하여 상기 사용자의 중요도를 도출할 수 있다.
- [0009] 상기 키워드를 입력받는 단계 이전에, 상기 콘텐츠 출판 정보 및 상기 사용자 프로필 정보를 수집하는 단계를 더 포함할 수 있다.
- [0010] 본 발명의 다른 측면에 따르면, 콘텐츠 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있는 시스템이 제공된다.
- [0011] 본 발명의 일 실시예에 따르면, 사용자 단말로 콘텐츠를 추천하는 콘텐츠 추천 시스템에 있어서, 상기 사용자 단말로부터 키워드를 입력받고, 상기 키워드가 포함된 적어도 하나의 콘텐츠 출판 정보를 추출하는 추출부; 상기 추출된 콘텐츠 출판 정보를 이용하여 상기 키워드와 콘텐츠간의 유사도를 도출하고, 상기 추출된 콘텐츠 출판 정보에 대응하는 사용자 프로필 정보를 이용하여 상기 콘텐츠를 배포한 사용자의 중요도를 도출하는 분석부; 상기 도출된 유사도 및 상기 사용자의 중요도를 이용하여 상기 추출된 콘텐츠에 대한 순위를 계산하여, 순위가 높은 순으로 콘텐츠를 추천하는 추천부를 포함하는 콘텐츠 추천 시스템이 제공될 수 있다.
- [0012] 상기 분석부는, 상기 콘텐츠 출판 정보를 이용하여 상기 키워드에 대한 빈도수를 도출하고, 전체 콘텐츠 중에서 상기 키워드가 나타난 출현 빈도를 도출하고, 상기 빈도수 및 상기 출현 빈도를 이용하여 상기 키워드와 상기

콘텐츠간의 유사도를 도출할 수 있다.

[0013] 상기 분석부는, 상기 콘텐츠 출판 정보를 배포한 각 사용자에게 연결된 다른 사용자수를 이용하여 상기 사용자의 중요도를 도출할 수 있다.

[0014] 상기 콘텐츠 출판 정보 및 상기 사용자 프로필 정보를 수집하는 정보 수집단계를 장치를 더 포함할 수 있다.

발명의 효과

[0015] 본 발명의 일 실시예에 따른 콘텐츠 추천 방법 및 시스템을 제공함으로써, 본 발명은 콘텐츠 추천을 요청한 사용자가 입력한 키워드와 콘텐츠간의 유사도와 콘텐츠를 배포한 사용자의 소셜 네트워크상에서의 위상을 반영하여 콘텐츠를 추천할 수 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 콘텐츠 추천 시스템의 구성을 나타낸 블록도.

도 2는 본 발명의 일 실시예에 따른 콘텐츠 출판 정보를 도시한 도면.

도 3은 본 발명의 일 실시예에 따른 사용자 프로필 정보를 도시한 도면.

도 4는 본 발명의 일 실시예에 따른 콘텐츠 추천 장치의 내부 구성을 나타낸 블록도.

도 5는 본 발명의 일 실시예에 따른 분석부의 내부 구성을 나타낸 블록도.

도 6은 본 발명의 일 실시예에 따른 콘텐츠 추천 방법을 나타낸 순서도.

발명을 실시하기 위한 구체적인 내용

[0017] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0018] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0019] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0020] 이하, 본 발명의 실시예를 첨부한 도면들을 참조하여 상세히 설명하기로 한다.

[0021] 도 1은 본 발명의 일 실시예에 따른 콘텐츠 추천 시스템의 구성을 나타낸 블록도이고, 도 2는 본 발명의 일 실시예에 따른 콘텐츠 출판 정보를 도시한 도면이고, 도 3은 본 발명의 일 실시예에 따른 사용자 프로필 정보를 도시한 도면이며, 도 4는 본 발명의 일 실시예에 따른 콘텐츠 추천 장치의 내부 구성을 나타낸 블록도이고, 도 5는 본 발명의 일 실시예에 따른 분석부의 내부 구성을 나타낸 블록도이다.

[0022] 도 1을 참조하면, 콘텐츠 추천 시스템은 복수의 사용자 단말(110), 데이터 수집 장치(120), 콘텐츠 추천 장치(130) 및 데이터베이스(140)를 포함하여 구성된다.

[0023] 사용자 단말(110)은 통신망을 통해 콘텐츠 추천 장치(130)에 특정 키워드를 입력하고, 그에 상응하는 콘텐츠 추

천 서비스를 제공받을 수 있는 장치이다.

- [0024] 예를 들어, 사용자 단말(110)의 유형으로는 컴퓨터, 노트북, 이동통신 단말기, PDA, 태블릿 PC 등일 수 있다.
- [0025] 콘텐츠 추천 시스템은 데이터를 수집하는 구성과 수집된 데이터를 분석하여 콘텐츠를 추천하는 구성으로 구성될 수 있다. 이하, 본 명세서에서는 데이터를 수집하는 구성을 데이터 수집 장치로 통칭하여 설명하고, 수집된 데이터를 분석하여 특정 키워드에 상응하는 콘텐츠를 추천하는 구성을 콘텐츠 추천 장치로 통칭하여 설명하기로 한다.
- [0026] 데이터 수집 장치(120)는 소셜 네트워크 서비스에서 콘텐츠 출판 정보와 사용자 프로필 정보를 수집하여 데이터베이스(140)에 저장하기 위한 수단이다. 예를 들어, 데이터 수집 장치(120)는 페이스북, 트위터, 라이브저널 등과 같은 소셜 네트워크 서비스를 통해 콘텐츠 출판 정보와 사용자에게 대한 사용자 프로필 정보를 각각 수집하여 이를 데이터베이스(140)에 저장할 수 있다.
- [0027] 본 명세서에서, 콘텐츠 출판 정보는 소셜 네트워크 서비스에서 제공되는 RSS(Really Simple Syndication or Rich Site Summary)일 수 있다. 콘텐츠 출판 정보는 도 2에 도시된 바와 같이, 콘텐츠의 제목, 작성 날짜, URI, 글 내용, 카테고리 중 적어도 하나를 포함할 수 있다.
- [0028] 또한, 사용자 프로필 정보는 사용자에게 대한 개인 정보와 함께 사용자 선호도 정보(즉, 사용자가 선호하는 특정 분야 등)를 포함할 수 있다. 예를 들어, 사용자 프로필 정보는 FOAF(Friends of a friend)일 수 있으며, 도 3에 FOAF가 도시되어 있다.
- [0029] 도 3에서, 사용자와 사용자간의 관계는 <foaf: knows>로 나타내고, 사용자의 선호도 정보, 즉 관심 분야와 같은 관심사는 <foaf: interest>로 나타낼 수 있다. 사용자 프로필 정보에 포함된 <foaf: interest>는 사용자의 사회적 연결 정도 중심성 분석을 수행시, 사용자 검색에 활용될 수 있다.
- [0030] 도 3에 도시된 사용자 프로필 정보에 포함된 사용자와 사용자간의 관계를 나타내는 정보(이하에서는 관계 정보라 칭하기로 함)(foaf: know)와 사용자의 관심 분야(이하에서는 관심 분야 정보라 칭하기로 함)(foaf: interest)는 각각 데이터베이스(140)에 별도의 테이블에 저장된다.
- [0031] 또한, 도 3에 도시된 나머지 사용자에게 대한 개인 정보를 나타내는 항목들(예를 들어, 닉네임 정보(foaf: nick), 이름(foaf: name) 등등)은 데이터베이스(140)에 별도의 동일 테이블을 통해 저장될 수 있다.
- [0032] 페이스북과 같이, 사용자 프로필 정보를 제공하지 않는 소셜 네트워크 서비스에 대해서는 데이터 수집 장치(120)는 OPEN API를 이용하여 사용자 프로필을 획득한 후, 이를 미리 정해진 형식에 맞춰 변환하여 데이터베이스(140)에 저장할 수 있다.
- [0033] 다시, 도 1을 참조하여, 콘텐츠 추천 장치(130)는 사용자 단말(110)을 통해 입력된 키워드에 가장 관련된 콘텐츠를 추천하기 위한 수단이다.
- [0034] 예를 들어, 콘텐츠 추천 장치(130)는 도 4에 도시된 바와 같이, 추출부(410), 분석부(420) 및 추천부(430)를 포함하여 구성된다.
- [0035] 추출부(410)는 사용자 단말(110)을 통해 입력된 키워드에 상응하는 적어도 하나 이상의 콘텐츠 및 해당 콘텐츠를 배포한 사용자에게 대한 사용자 프로필 정보를 데이터베이스(140)에서 추출하기 위한 수단이다. 즉, 추출부(410)는 키워드를 포함하는 콘텐츠 출판 정보 및 해당 콘텐츠를 배포한 사용자에게 대한 사용자 프로필 정보를 데이터베이스(140)에서 추출한다.
- [0036] 분석부(420)는 추출부(410)를 통해 추출된 콘텐츠 출판 정보 및 사용자 프로필 정보를 이용하여, 키워드와 콘텐츠간의 유사도와 소셜 네트워크상에서의 해당 콘텐츠를 배포한 사용자의 중요도를 분석하기 위한 수단이다.
- [0037] 도 5에 도시된 바와 같이, 분석부(420)는 제1 분석 수단(510) 및 제2 분석 수단(520)을 포함할 수 있다.
- [0038] 제1 분석 수단(510)은 사용자 단말(110)을 통해 입력된 키워드와 콘텐츠간의 유사도를 분석하기 위한 수단이다. 본 명세서에서는 이해와 설명의 편의를 도모하기 위해 콘텐츠 출판 정보를 구문에 따라 콘텐츠로 칭하여 설명하기로 한다.
- [0039] 이를 위해, 제1 분석 수단(510)은 키워드가 특정 콘텐츠에서 차지하는 비중을 도출한다. 예를 들어, 제1 분석 수단(510)은 하기 수학적 식 1을 이용하여 키워드가 특정 콘텐츠에서 차지하는 비중을 도출할 수 있다.

수학식 1

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

[0040]

[0041] 여기서, i 는 키워드를 나타내고, j 는 전체 콘텐츠 중 특정 콘텐츠를 나타내며, $n_{i,j}$ 는 키워드가 특정 콘텐츠에 나타난 횟수를 나타내고, k 는 콘텐츠에 포함된 각 단어를 나타내며, $\sum_k n_{k,j}$ 는 특정 콘텐츠가 포함하고 있는 k 개의 전체 단어들이 나타내는 횟수의 합을 의미한다.

[0042] 이어, 제1 분석 수단(510)은 전체 콘텐츠 집합에서 특정 키워드가 출현하는 빈도를 도출한다. 즉, 어떤 단어에 대한 출현 빈도가 높다면 많은 콘텐츠에서 해당 단어를 사용하고 있음을 의미하며, 이는 관사 등과 같이 흔하게 사용되는 단어를 나타낸다.

[0043] 이로 인해, 제1 분석 수단(510)은 도출된 출현 빈도의 역수를 도출한다. 제1 분석 수단(510)은 하기 수학식 2를 이용하여 출현 빈도에 대한 역수를 도출할 수 있다.

수학식 2

$$idf_i = \log \frac{|D|}{|\{d : t_i \in d\}|}$$

[0044]

[0045] 여기서, i 는 키워드를 나타내고, $|D|$ 는 전체 콘텐츠의 개수를 나타내고, $|\{d : t_i \in d\}|$ 는 키워드 i 를 포함한 콘텐츠의 개수를 나타낸다.

[0046] 제1 분석 수단(510)은 도출된 키워드 i 가 특정 콘텐츠에 차지하는 비중과 출현 빈도에 대한 역수를 곱하여 최종 키워드와 특정 콘텐츠와의 유사도를 각각 도출할 수 있다.

[0047] 예를 들어, 데이터 수집 장치(120)를 통해 수집된 콘텐츠 출판 정보가 1000개라고 가정하고, 사용자 단말(110)을 통해 입력된 키워드가 “computer”라고 가정하자. 이때, 수집된 1000개의 콘텐츠 출판 정보들 중에서 키워드 “computer”를 포함하는 콘텐츠가 5개라고 가정하자.

표 1

[0048]

콘텐츠	사용자	전체단어의 개수	키워드가 나타난 횟수
C1	H1	100	5
C2	H2	80	3
C3	H3	200	8
C4	H4	50	3
C5	H5	60	1

[0049] 전체 콘텐츠 출판 정보 중에서 “computer”를 포함하는 콘텐츠의 전체 단어의 개수와 해당 콘텐츠에서 키워드가 나타난 횟수가 표 1과 같다고 하자. 표 1을 참조하여 키워드가 나타난 빈도를 수학식 1을 이용하여 계산하면, 하기와 같다.

[0050] $tf_{computer,C1} = 5/100 = 0.05$

[0051] 이어, 키워드가 나타난 출현 빈도에 대한 역수를 도출하면, 하기와 같다.

- [0052] $\text{Idf}_{\text{computer}} = \log(1000/5) = 2.3$
- [0053] 따라서, 제1 분석 수단(510)은 키워드에 대한 빈도 정보와 출현빈도를 이용하여 키워드와 콘텐츠에 대한 유사도를 도출할 수 있다.
- [0054] 하기 표 2에 키워드와 콘텐츠에 대한 유사도를 도출한 예가 예시되어 있다.

표 2

콘텐츠	사용자	빈도정보	출현빈도에 대한 역수	유사도
C1	H1	0.0500	2.3	0.115
C2	H2	0.0375	2.3	0.086
C3	H3	0.0400	2.3	0.092
C4	H4	0.0600	2.3	0.138
C5	H5	0.0170	2.3	0.039

- [0056] 제2 분석 수단(520)은 추출된 콘텐츠 출판 정보에 대응하는 사용자 프로필 정보를 이용하여 사용자에게 대한 중요도를 분석하기 위한 수단이다.
- [0057] 예를 들어, 제2 분석 수단(520)은 사용자에게 대한 중요도를 분석하기 위해, 사용자 프로필 정보를 이용하여 사용자에게 대한 연결 정보 중심성을 도출할 수 있다.
- [0058] 본 명세서에서, 연결 정보 중심성이란 소셜네트워크에서 특정사용자가 다른사용자와 얼마나 많이 연결되어 있는지를 나타낸다. 즉, 연결 정보 중심성은 소셜 네트워크상에서 특정 사용자의 친구가 어느 정도인지를 나타낼 수 있다.
- [0059] 이로 인해, 결과적으로, 연결 정도 중심성은 소셜 네트워크상에서의 특정 사용자에게 대한 중요도를 알 수 있다. 즉, 연결 정도 중심성을 이용하여 특정 콘텐츠를 배포한 사용자가 소셜 네트워크상에서 차지하는 중요도가 어느 정도인지, 파급력이 어느 정도인지 가능할 수 있다.
- [0060] 연결 정도 중심성은 하기 수식식 3을 이용하여 도출될 수 있다.

수식식 3

$$C_D(v) = \frac{\text{deg}(v)}{n - 1}$$

- [0061]
- [0062] 여기서, $\text{deg}(v)$ 는 사용자 v 와 연결된 다른 사용자의 수(즉, 친구수)를 나타내고, n 은 전체 사용자의 수를 나타낸다.
- [0063] 즉, 연결 정도 중심성은 사용자 v 와 연결된 사용자의 수를 전체 사용자 수에서 자신을 제외한 전체 사용자의 수($n-1$)로 나눈 값으로 도출될 수 있다.
- [0064] 예를 들어, 제2 분석 수단(520)은 수집된 사용자 프로필 정보 중에서 사용자와 사용자간의 관계를 나타내는 <foaf: knows>를 이용하여 연결 정도 중심성을 계산할 수 있다.
- [0065] 연결 정도 중심성은 소셜네트워크 상에서 사용자와 연결된 다른 사용자가 많을수록(즉, 친구가 많을수록) “1”에 가까운 값으로 도출되며, 다른 사용자와의 연결이 적을수록 “0”에 가까운 값이 도출될 수 있다.
- [0066] 이어, 제2 분석 수단(520)은 키워드에 관련된 콘텐츠를 배포한 사용자들 중에서 해당 사용자가 차지하는 비율을 계산한다.
- [0067] 제2 분석 수단(520)은 하기 수식식4를 이용하여 사용자가 차지하는 비율을 도출할 수 있다.

수학식 4

$$P_{i,m} = \frac{C_D(v_{im})}{\sum_{r=1}^n C_D(v_{ir})} \times 100$$

[0068]

[0069] 여기서, $CD(v_{i,m})$ 는 키워드 i 와 관련된 콘텐츠를 배포한 사용자 m 의 연결 정도 중심성 값을 나타내고, $\sum_{r=1}^n C_D(v_{ir})$ 는 키워드 i 에 관련된 사용자가 총 n 명 있을 때 n 명의 $CD(v)$ 의 합을 나타낸다. 따라서, $P_{i,m}$ 은 키워드 i 에 관련된 전체 사용자들의 연결 정도 중심성 합에서 특정 사용자 m 의 연결 정도 중심성이 차지하는 비율이다. $P_{i,m}$ 의 계산 결과를 통해 사용자가 입력한 키워드에 대한 콘텐츠 배포자의 순위를 결정한다.

[0070] 예를 들어, 표 1에 예시된 각 사용자에게 연결된 다른 사용자의 수가 하기 표 3과 같다고 가정하면, 각 사용자별 중요도는 표 3의 세번째 칼럼과 같이 도출될 수 있다. 예를 들어, 전체 사용자수가 500이라고 가정하자.

표 3

사용자	연결된 다른 사용자수	연결 정도 중요성
H1	30	0.06
H2	60	0.12
H3	25	0.05
H4	10	0.02
H5	80	0.16

[0072] 이를 이용하여, 각 사용자가 차지하는 연결 정도 중심성 비율을 도출하면, 도 4와 같다.

표 4

사용자	연결 정도 중요성	비율
H1	0.06	14.63
H2	0.12	29.26
H3	0.05	12.19
H4	0.02	4.87
H5	0.16	39.02

[0074] 다시, 도 4를 참조하여, 추천부(430)는 분석부(420)를 통해 도출된 유사도 및 중요도를 이용하여 콘텐츠에 대한 순위를 도출할 수 있다.

[0075] 예를 들어, 추천부(430)는 하기 수학식 5를 이용하여 콘텐츠에 대한 순위를 도출할 수 있다.

수학식 5

$$SR_{i,j,m} = tfidf_{i,j} \times P_{i,m}$$

[0076]

[0077] 예를 들어, 키워드를 포함하는 콘텐츠에 대한 유사도가 표 2와 같고, 해당 콘텐츠를 배포한 사용자에게 대한 중요도가 표 4와 같다고 가정하자.

[0078] 추천부(430)는 예를 들어, 해당 콘텐츠에 대한 유사도와 해당 콘텐츠를 배포한 사용자에게 대한 중요도를 곱한 결과값을 이용하여 결과값이 높은 순으로 콘텐츠의 순위를 결정하고, 순위가 높은 순으로 콘텐츠를 정렬한 후 콘

텐츠를 추천할 수 있다.

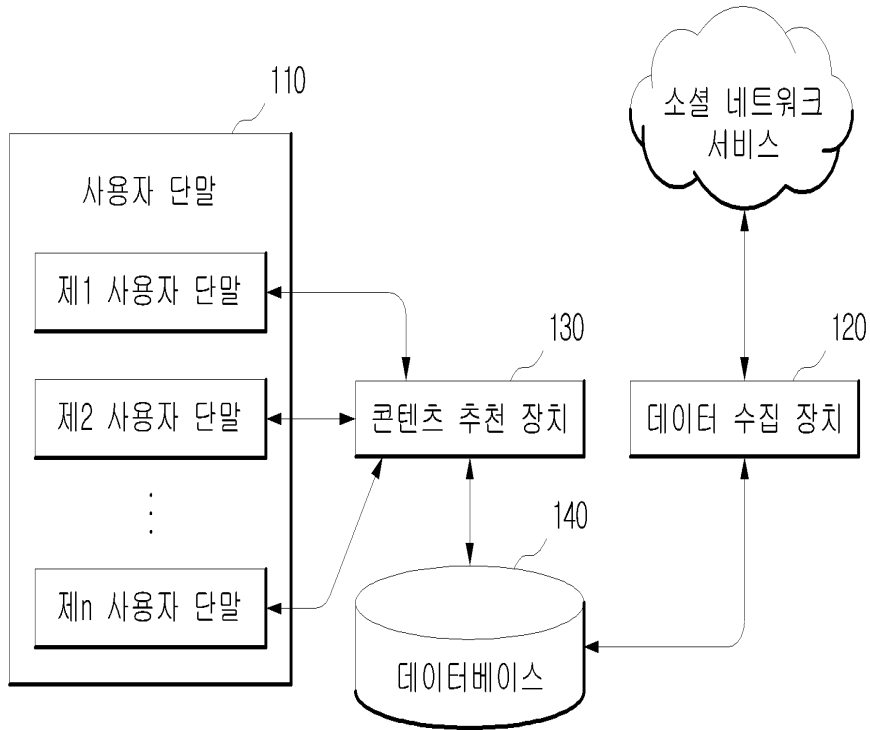
- [0079] 도 6은 본 발명의 일 실시예에 따른 콘텐츠 추천 방법을 나타낸 순서도이다.
- [0080] 단계 610에서 데이터 수집 장치(120)는 소셜 네트워크에서 적어도 하나의 콘텐츠에 대한 콘텐츠 출판 정보와 해당 콘텐츠를 배포한 사용자에 대한 사용자 프로필 정보를 수집하여 데이터베이스(140)에 저장한다.
- [0081] 전술한 바와 같이, 콘텐츠 출판 정보는 RSS 일 수 있으며, 사용자 프로필 정보는 FOAF 형식일 수 있다.
- [0082] 또한, 페이스북과 같이, RSS 또는 FOAF를 지원하지 않는 소셜 네트워크 서비스에 대해서, 데이터 수집 장치(120)는 OPEN API를 이용하여 콘텐츠에 대한 정보, 사용자에 대한 정보를 획득한 후 이를 RSS 또는 FOAF 형식으로 변환하여 데이터베이스(140)에 저장할 수도 있다.
- [0083] 이어, 단계 615에서 콘텐츠 추천 장치(130)는 사용자 단말(110)로부터 키워드를 입력받는다.
- [0084] 단계 620에서 콘텐츠 추천 장치(130)는 키워드를 포함하는 콘텐츠에 대한 콘텐츠 출판 정보 및 해당 콘텐츠를 배포한 사용자 프로필 정보를 데이터베이스(140)에서 추출한다.
- [0085] 단계 625에서 콘텐츠 추천 장치(130)는 추출된 콘텐츠 출판 정보를 이용하여 키워드와 콘텐츠와의 유사도를 분석한다.
- [0086] 이는 이미 전술한 바와 동일하므로, 중복되는 설명은 생략하기로 한다.
- [0087] 단계 630에서 콘텐츠 추천 장치(130)는 추출된 콘텐츠에 대한 사용자 프로필 정보를 이용하여 소셜 네트워크에서의 사용자의 중요도를 분석한다.
- [0088] 예를 들어, 콘텐츠 추천 장치(130)는 소셜 네트워크에서의 각 사용자에게 연결된 다른 사용자의 수(즉, 친구수)를 이용하여 사용자의 중요도를 분석할 수 있다.
- [0089] 이는 이미 전술한 바와 동일하므로, 중복되는 설명은 생략하기로 한다.
- [0090] 단계 635에서 콘텐츠 추천 장치(130)는 키워드와 콘텐츠간의 유사도 및 사용자의 중요도를 이용하여 콘텐츠에 대한 순위를 결정하고, 순위가 높은 콘텐츠 순으로 사용자 단말(110)로 추천할 수 있다.
- [0091] 한편, 본 발명의 실시예에 따른 콘텐츠를 추천하는 방법은 다양한 전자적으로 정보를 처리하는 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 저장 매체에 기록될 수 있다. 저장 매체는 프로그램 명령, 데이터 파일, 데이터 구조등을 단독으로 또는 조합하여 포함할 수 있다.
- [0092] 저장 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 소프트웨어 분야 당업자에게 공지되어 사용 가능한 것일 수도 있다. 저장 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 또한 상술한 매체는 프로그램 명령, 데이터 구조 등을 지정하는 신호를 전송하는 반송파를 포함하는 광 또는 금속선, 도파관 등의 전송 매체일 수도 있다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 전자적으로 정보를 처리하는 장치, 예를 들어, 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0093] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0094] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야에서 통상의 지식을 가진 자라면 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

- [0095]
- 110: 사용자 단말
 - 120: 데이터 수집 장치
 - 130: 콘텐츠 추천 장치
 - 140: 데이터베이스

도면

도면1



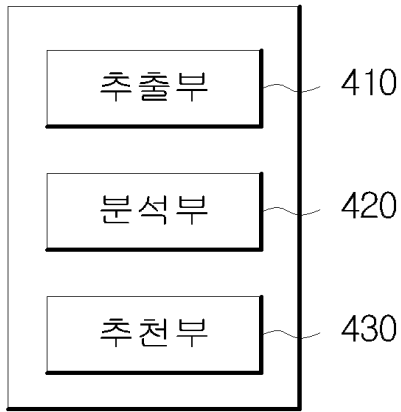
도면2

```
- <item>
  <guid isPermaLink="true">http://adafrog.livejournal.com/337514.html</guid>
  <pubDate>Tue, 14 Sep 2010 17:22:49 GMT</pubDate>
  <title>A03 Invite</title>
  <link>http://adafrog.livejournal.com/337514.html</link>
  <description>Actually still have it. Leave a comment with your email, and I'll give
    screened.</description>
  <comments>http://adafrog.livejournal.com/337514.html</comments>
  <category>ao3</category>
  <category>flist</category>
  <lj:security>public</lj:security>
  <lj:reply-count>0</lj:reply-count>
</item>
```

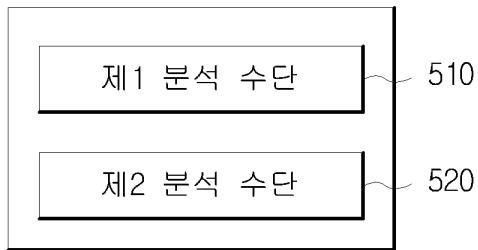
도면3

```
- <foaf:Person>
  <foaf:nick>clannuisnigh</foaf:nick>
  <foaf:name>clannuisnigh</foaf:name>
  <foaf:openid rdf:resource="http://clannuisnigh.livejournal.com/" />
  <ya:country dc:title="UK" rdf:resource="http://www.livejournal.com/directory.bml?o
  <ya:city dc:title="London" rdf:resource="http://www.livejournal.com/directory.bml?
    opt_sort=ut&s_loc=1&loc_cn=UK&loc_st=&loc_ci=London" />
  <foaf:dateOfBirth>1983-02-19</foaf:dateOfBirth>
+ <foaf:page>
  <foaf:icqChatID />
  <foaf:aimChatID />
  <foaf:jabberID />
  <foaf:msnChatID />
  <foaf:yahooChatID />
+ <ya:blogActivity>
  <foaf:weblog rdf:resource="http://clannuisnigh.livejournal.com/" lj:dateCreated="200
    03-26T12:17:12" />
- <foaf:knows>
  - <foaf:Person>
    <foaf:nick>elfflame</foaf:nick>
    <foaf:member_name>elfflame</foaf:member_name>
    <foaf:tagLine>Little Fall of Rain</foaf:tagLine>
    <foaf:image>http://l-userpic.livejournal.com/33239411/3938525</foaf:image>
    <rdfs:seeAlso rdf:resource="http://elfflame.livejournal.com/data/foaf" />
    <foaf:weblog rdf:resource="http://elfflame.livejournal.com/" />
  </foaf:Person>
</foaf:knows>
+ <foaf:knows>
+ <foaf:knows>
```

도면4



도면5



도면6

