



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년10월16일
 (11) 등록번호 10-1907865
 (24) 등록일자 2018년10월08일

(51) 국제특허분류(Int. Cl.)
 G06Q 50/10 (2012.01) G06F 17/18 (2006.01)
 G06F 17/27 (2006.01) G06F 17/30 (2006.01)
 G06Q 10/10 (2012.01)

(52) CPC특허분류
 G06Q 50/10 (2015.01)
 G06F 17/18 (2013.01)

(21) 출원번호 10-2016-0146002
 (22) 출원일자 2016년11월03일
 심사청구일자 2016년11월03일
 (65) 공개번호 10-2018-0049852
 (43) 공개일자 2018년05월14일
 (56) 선행기술조사문헌
 KR1020090090840 A*
 스마트폰 환경에서 기사 읽기 습관 고려한 뉴스
 추천 시스템 설계 및 구현, 한국정보통신학회논문
 지, Vol. 18, No. 7 : 1628~1633 July. 2014
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
세종대학교산학협력단
 서울특별시 광진구 능동로 209 (군자동, 세종대학
 교)

(72) 발명자
유성준
 대전광역시 유성구 어은로 57, 131동 306호 (어은
 동, 한빛아파트)
구영현
 서울 노원구 상계3동 대림아파트
 (뒷면에 계속)

(74) 대리인
두호특허법인

전체 청구항 수 : 총 11 항

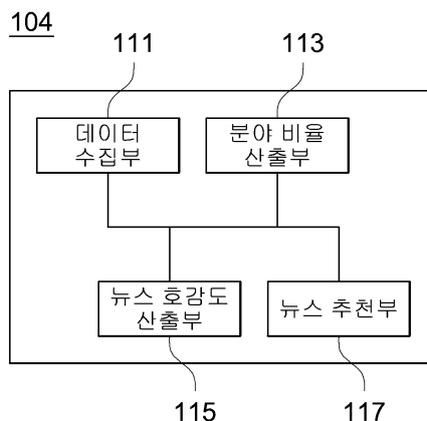
심사관 : 최규돈

(54) 발명의 명칭 **뉴스 추천 서버 및 이를 이용한 뉴스 추천 방법**

(57) 요약

뉴스 추천 서버 및 이를 이용한 뉴스 추천 방법이 개시된다. 예시적인 실시예에 따른 뉴스 추천 서버는, 사용자 단말기로부터 웹 브라우징 히스토리 정보를 수집하는 데이터 수집부, 수집한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야의 비율을 산출하는 분야 비율 산출부, 및 산출한 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하여 사용자 단말기로 뉴스를 추천하는 뉴스 추천부를 포함한다.

대표도 - 도2



(52) CPC특허분류

G06F 17/27 (2013.01)

G06F 17/30861 (2013.01)

G06Q 10/10 (2013.01)

(72) 발명자

박, 철호

서울시 광진구 능동로 21길 33, 101호 (군자동)

윤, 학립

서울특별시 광진구 아차산로33길 75 , 201호 (화양동)

장, 지연

서울특별시 광진구 아차산로 33길 75, 201호 (화양동)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711035312

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 일반 연구자 사업

연구과제명 인메모리 기반 빅데이터의 맞춤형 자동 큐레이션 기술 연구

기 여 율 1/1

주관기관 세종대학교산학협력단

연구기간 2016.01.01 ~ 2016.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

사용자 단말기로부터 웹 브라우징 히스토리 정보 및 로그 데이터를 수집하는 데이터 수집부;

상기 수집한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출하는 분야 비율 산출부;

상기 웹 브라우징 히스토리 정보 및 상기 로그 데이터를 기반으로 상기 사용자 단말기로 제공하는 각 뉴스에 대한 사용자의 호감도를 산출하는 뉴스 호감도 산출부; 및

상기 산출된 뉴스 분야들의 비율 및 상기 산출된 사용자의 호감도에 기초하여 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절한 후 상기 사용자 단말기로 뉴스를 추천하는 뉴스 추천부를 포함하며,

상기 뉴스 호감도 산출부는, 상기 웹 브라우징 히스토리 정보의 웹 페이지들 중 상기 로그 데이터에 포함된 소정 뉴스를 제공한 시점과 동일한 접속 시점을 갖는 웹 페이지를 추출하고, 상기 추출된 웹 페이지의 접속 시간과 상기 추출한 웹 페이지 다음에 접속한 웹 페이지의 접속 시간을 기반으로 상기 뉴스를 읽은 시간을 산출하며, 상기 산출된 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 사용자가 상기 소정 뉴스를 읽는 속도를 산출하고, 상기 산출된 소정 뉴스를 읽는 속도에 기초하여 상기 소정 뉴스에 대한 사용자의 호감도를 산출하는, 뉴스 추천 서버.

청구항 2

청구항 1에 있어서,

상기 분야 비율 산출부는,

상기 웹 브라우징 히스토리 정보를 기반으로 상기 사용자 단말기가 접속한 각 뉴스의 웹 페이지에서 텍스트를 각각 추출하고, 상기 추출한 텍스트에서 단어의 품사를 기반으로 분석 대상 단어를 1차 선별하며, 상기 1차 선별한 분석 대상 단어를 기반으로 상기 접속한 각 뉴스의 분야를 결정하는, 뉴스 추천 서버.

청구항 3

청구항 2에 있어서,

상기 분야 비율 산출부는,

상기 1차 선별한 분석 대상 단어들 각각에 가중치를 부여하고, 상기 1차 선별한 분석 대상 단어들 중 상기 가중치 값이 높은 순서대로 기 설정된 개수의 단어들을 2차 선별하며, 상기 텍스트 별로 상기 2차 선별된 단어들의 포함 정도에 따라 상기 접속한 각 뉴스의 분야를 결정하는, 뉴스 추천 서버.

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

청구항 1에 있어서,

상기 뉴스 호감도 산출부는,

상기 산출한 소정 뉴스를 읽는 속도가 기 설정된 표준 속도에 근접하는 정도에 따라 상기 소정 뉴스에 대한 상기 사용자의 호감도 값을 높게 설정하는, 뉴스 추천 서버.

청구항 8

청구항 1에 있어서,

상기 뉴스 추천부는,

해당 사용자와 뉴스 구독 성향이 유사한 사용자를 검색하고, 상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천하는, 뉴스 추천 서버.

청구항 9

청구항 8에 있어서,

상기 뉴스 추천부는,

상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중 상기 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 순서대로 추출하고, 상기 추출한 뉴스들을 대상으로 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하는, 뉴스 추천 서버.

청구항 10

하나 이상의 프로세서들, 및

상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 방법으로서,

사용자 단말기로부터 웹 브라우징 히스토리 정보 및 로그 데이터를 수집하는 단계;

상기 수집한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출하는 단계;

상기 웹 브라우징 히스토리 정보 및 상기 로그 데이터를 기반으로 상기 사용자 단말기로 제공하는 각 뉴스에 대한 사용자의 호감도를 산출하는 단계; 및

상기 산출된 뉴스 분야들의 비율 및 상기 산출된 사용자의 호감도에 기초하여 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절한 후 상기 사용자 단말기로 뉴스를 추천하는 단계를 포함하며,

상기 사용자의 호감도를 산출하는 단계는, 상기 웹 브라우징 히스토리 정보의 웹 페이지들 중 상기 로그 데이터에 포함된 소정 뉴스를 제공한 시점과 동일한 접속 시점을 갖는 웹 페이지를 추출하는 단계;

상기 추출된 웹 페이지의 접속 시간과 상기 추출한 웹 페이지 다음에 접속한 웹 페이지의 접속 시간을 기반으로 상기 뉴스를 읽은 시간을 산출하는 단계;

상기 산출된 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 사용자가 상기 소정 뉴스를 읽는 속도를 산출하는 단계; 및

상기 산출된 소정 뉴스 읽는 속도에 기초하여 상기 사용자의 호감도를 산출하는 단계를 포함하는, 뉴스 추천 방법.

청구항 11

청구항 10에 있어서,

상기 뉴스 분야의 비율을 산출하는 단계는,

상기 웹 브라우징 히스토리 정보를 기반으로 상기 사용자 단말기가 접속한 각 뉴스의 웹 페이지에서 텍스트를 각각 추출하는 단계;

상기 추출한 텍스트에서 단어의 품사를 기반으로 분석 대상 단어를 1차 선별하는 단계;

상기 1차 선별한 분석 대상 단어들 각각에 가중치를 부여하는 단계;

상기 1차 선별한 분석 대상 단어들 중 상기 가중치 값이 높은 순서대로 기 설정된 개수의 단어들을 2차 선별하는 단계; 및

상기 텍스트 별로 상기 2차 선별된 단어들의 포함 정도에 따라 상기 접속한 각 뉴스의 분야를 결정하는 단계를 포함하는, 뉴스 추천 방법.

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

청구항 10에 있어서,

상기 호감도를 산출하는 단계는,

상기 산출한 소정 뉴스를 읽는 속도가 기 설정된 표준 속도에 근접하는 정도에 따라 상기 소정 뉴스에 대한 상기 사용자의 호감도 값을 높게 설정하는 단계를 포함하는, 뉴스 추천 방법.

청구항 16

청구항 10에 있어서,

상기 뉴스를 추천하는 단계는,

해당 사용자와 뉴스 구독 성향이 유사한 사용자를 검색하는 단계; 및

상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천하는 단계를 포함하는, 뉴스 추천 방법.

청구항 17

청구항 16에 있어서,

상기 뉴스를 추천하는 단계는,

상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중 상기 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 순서대로 추출하는 단계; 및

상기 추출한 뉴스들을 대상으로 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하는 단계를 포함하는, 뉴스 추천 방법.

발명의 설명

기술 분야

[0001] 본 발명의 실시예는 뉴스 추천 기술과 관련된다.

배경 기술

[0003] 스마트 디바이스가 많이 보급됨에 따라 웹 서비스의 중점은 PC(Personal Computer)에서 스마트 디바이스로 옮겨 가고 있다. 이는 뉴스 서비스도 마찬가지이다. 또한, 뉴스 추천 시스템의 발전에 따라 사용자는 선호하는 분야의 뉴스를 자동으로 추천 받고자 하는 욕구가 커지고 있다. 더욱이 선호하는 정도에 따라 추천되는 뉴스의 양이 동적으로 조절되기를 원하고 있으나, 기존의 모바일 뉴스 추천 시스템은 사용자의 선호도에 따라 선호 분야 뉴스의 양을 동적으로 조절하지 못한다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국공개특허공보 제10-2016-0104067호(2016.09.02)

발명의 내용

해결하려는 과제

[0006] 본 발명의 실시예는 사용자의 선호도에 따라 뉴스 분야의 순서 및 각 뉴스 분야의 수량을 조절하여 추천할 수 있는 기법을 제공하기 위한 것이다.

과제의 해결 수단

[0008] 본 발명의 일 실시예에 따른 뉴스 추천 서버는, 사용자 단말기로부터 웹 브라우징 히스토리 정보를 수집하는 데이터 수집부; 상기 수집한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출하는 분야 비율 산출부; 및 상기 산출한 뉴스 분야들의 비율을 기반으로 상기 사용자 단말기로 뉴스를 추천하는 뉴스 추천부를 포함한다.

[0009] 상기 뉴스 추천부는, 상기 산출한 뉴스 분야들의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절할 수 있다.

[0010] 상기 분야 비율 산출부는, 상기 웹 브라우징 히스토리 정보를 기반으로 상기 사용자 단말기가 접속한 각 뉴스의 웹 페이지에서 텍스트를 각각 추출하고, 상기 추출한 텍스트에서 단어의 품사를 기반으로 분석 대상 단어를 1차 선별하며, 상기 1차 선별한 분석 대상 단어를 기반으로 상기 접속한 각 뉴스의 분야를 결정할 수 있다.

[0011] 상기 분야 비율 산출부는, 상기 1차 선별한 분석 대상 단어들 각각에 가중치를 부여하고, 상기 1차 선별한 분석 대상 단어들 중 상기 가중치 값이 높은 기 설정된 개수의 단어들을 2차 선별하며, 상기 텍스트 별로 상기 2차 선별된 단어들의 포함 정도에 따라 상기 접속한 각 뉴스의 분야를 결정할 수 있다.

[0012] 상기 뉴스 추천 서버는, 각 사용자 단말기로 제공하는 각 뉴스에 대한 사용자의 호감도를 산출하는 뉴스 호감도 산출부를 더 포함할 수 있다.

- [0013] 상기 뉴스 호감도 산출부는, 해당 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 뉴스에 대한 사용자의 호감도 값을 산출할 수 있다.
- [0014] 상기 데이터 수집부는, 상기 사용자 단말기로부터 상기 뉴스 추천 서버에 접속하는 로그 데이터를 수집하고, 상기 뉴스 호감도 산출부는, 상기 웹 브라우징 히스토리 정보의 웹 페이지들 중 상기 로그 데이터에서 소정 뉴스를 제공한 시점과 동일한 접속 시점을 갖는 웹 페이지를 추출하고, 상기 추출한 웹 페이지의 접속 시간과 상기 추출한 웹 페이지 다음에 접속한 웹 페이지의 접속 시간을 기반으로 상기 뉴스를 읽은 시간을 산출할 수 있다.
- [0015] 상기 뉴스 호감도 산출부는, 해당 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 뉴스를 읽는 속도를 산출하고, 상기 산출한 뉴스 읽는 속도가 기 설정된 표준 속도에 근접하는 정도에 따라 상기 뉴스에 대한 사용자의 호감도 값을 높게 설정할 수 있다.
- [0016] 상기 뉴스 추천부는, 해당 사용자와 뉴스 구독 성향이 유사한 사용자를 검색하고, 상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천할 수 있다.
- [0017] 상기 뉴스 추천부는, 상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중 상기 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 순서대로 추출하고, 상기 추출한 뉴스들을 대상으로 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절할 수 있다.
- [0018] 본 발명의 일 실시예에 따른 뉴스 추천 방법은, 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 방법으로서, 사용자 단말기로부터 웹 브라우징 히스토리 정보를 수집하는 단계; 상기 수집한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출하는 단계; 및 상기 산출한 뉴스 분야들의 비율을 기반으로 상기 사용자 단말기로 뉴스를 추천하는 단계를 포함한다.
- [0019] 상기 뉴스를 추천하는 단계는, 상기 산출한 뉴스 분야들의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하는 단계를 포함할 수 있다.
- [0020] 상기 뉴스 분야들의 비율을 산출하는 단계는, 상기 웹 브라우징 히스토리 정보를 기반으로 상기 사용자 단말기가 접속한 각 뉴스의 웹 페이지에서 텍스트를 각각 추출하는 단계; 상기 추출한 텍스트에서 단어의 품사를 기반으로 분석 대상 단어를 1차 선별하는 단계; 상기 1차 선별한 분석 대상 단어들 각각에 가중치를 부여하는 단계; 상기 1차 선별한 분석 대상 단어들 중 상기 가중치 값이 높은 기 설정된 개수의 단어들을 2차 선별하는 단계; 및 상기 텍스트 별로 상기 2차 선별된 단어들의 포함 정도에 따라 상기 접속한 각 뉴스의 분야를 결정하는 단계를 포함할 수 있다.
- [0021] 상기 뉴스 추천 방법은, 각 사용자 단말기로 제공하는 각 뉴스에 대한 사용자의 호감도를 산출하는 단계를 더 포함할 수 있다.
- [0022] 상기 호감도를 산출하는 단계는, 해당 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 뉴스에 대한 사용자의 호감도 값을 산출할 수 있다.
- [0023] 상기 호감도를 산출하는 단계는, 상기 사용자 단말기로부터 상기 뉴스 추천 서버에 접속하는 로그 데이터를 수집하는 단계; 상기 웹 브라우징 히스토리 정보의 웹 페이지들 중 상기 로그 데이터에서 소정 뉴스를 제공한 시점과 동일한 접속 시점을 갖는 웹 페이지를 추출하는 단계; 및 상기 추출한 웹 페이지의 접속 시간과 상기 추출한 웹 페이지 다음에 접속한 웹 페이지의 접속 시간을 기반으로 상기 뉴스를 읽은 시간을 산출하는 단계를 포함할 수 있다.
- [0024] 상기 호감도를 산출하는 단계는, 해당 사용자가 상기 뉴스를 읽은 시간 및 상기 뉴스의 문장 길이를 기반으로 상기 뉴스를 읽는 속도를 산출하는 단계; 및 상기 산출한 뉴스 읽는 속도가 기 설정된 표준 속도에 근접하는 정도에 따라 상기 뉴스에 대한 사용자의 호감도 값을 높게 설정하는 단계를 포함할 수 있다.
- [0025] 상기 뉴스를 추천하는 단계는, 해당 사용자와 뉴스 구독 성향이 유사한 사용자를 검색하는 단계; 및 상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천하는 단계를 포함할 수 있다.
- [0026] 상기 뉴스를 추천하는 단계는, 상기 검색된 뉴스 구독 성향이 유사한 사용자가 본 뉴스들 중 상기 뉴스 구독 성

향이 유사한 사용자들의 뉴스 호감도가 높은 순서대로 추출하는 단계; 및 상기 추출한 뉴스들을 대상으로 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하는 단계를 포함할 수 있다.

발명의 효과

[0028] 본 발명의 실시예에 의하면, 사용자 별 웹 브라우징 히스토리를 이용하여 사용자가 선호하는 뉴스 분야의 비율을 산출하고, 그에 따라 뉴스를 추천하여 줌으로써, 사용자의 선호도의 변화에 따라 능동적으로 뉴스를 추천할 수 있게 된다.

도면의 간단한 설명

- [0030] 도 1은 본 발명의 일 실시예에 따른 뉴스 추천 시스템을 나타낸 블록도
- 도 2는 본 발명의 일 실시예에 따른 뉴스 추천 서버의 구성을 나타낸 블록도
- 도 3은 본 발명의 일 실시예에 따른 뉴스 추천 서버에서, 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천하는 일 실시예를 나타낸 도면
- 도 4는 본 발명의 일 실시예에 따른 뉴스 추천 서버에서, 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천하는 다른 실시예를 나타낸 도면
- 도 5는 본 발명의 일 실시예에 따른 뉴스 추천 방법을 설명하기 위한 흐름도
- 도 6은 본 발명의 다른 실시예에 따른 뉴스 추천 방법을 설명하기 위한 흐름도
- 도 7은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도

발명을 실시하기 위한 구체적인 내용

[0031] 이하, 도면을 참조하여 본 발명의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.

[0032] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 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. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.

[0033] 이하의 설명에 있어서, 신호 또는 정보의 "전송", "통신", "송신", "수신" 기타 이와 유사한 의미의 용어는 일 구성요소에서 다른 구성요소로 신호 또는 정보가 직접 전달되는 것뿐만이 아니라 다른 구성요소를 거쳐 전달되는 것도 포함한다. 특히 신호 또는 정보를 일 구성요소로 "전송" 또는 "송신"한다는 것은 그 신호 또는 정보의 최종 목적지를 지시하는 것이고 직접적인 목적지를 의미하는 것이 아니다. 이는 신호 또는 정보의 "수신"에 있어서도 동일하다. 또한 본 명세서에 있어서, 2 이상의 데이터 또는 정보가 "관련"된다는 것은 하나의 데이터(또는 정보)를 획득하면, 그에 기초하여 다른 데이터(또는 정보)의 적어도 일부를 획득할 수 있음을 의미한다.

[0034] 또한, 제1, 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로 사용될 수 있다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성 요소는 제2 구성 요소로 명명될 수 있고, 유사하게 제2 구성 요소도 제1 구성 요소로 명명될 수 있다.

[0036] 도 1은 본 발명의 일 실시예에 따른 뉴스 추천 시스템을 나타낸 블록도이다.

[0037] 도 1을 참조하면, 뉴스 추천 시스템(100)은 사용자 단말기(102) 및 뉴스 추천 서버(104)를 포함할 수 있다. 사

용자 단말기(102)는 통신 네트워크(150)를 통해 뉴스 추천 서버(104)와 통신 가능하게 연결된다. 몇몇 실시예들에서, 통신 네트워크(150)는 인터넷, 하나 이상의 로컬 영역 네트워크(local area networks), 광역 네트워크(wire area networks), 셀룰러 네트워크, 모바일 네트워크, 그 밖에 다른 종류의 네트워크들, 또는 이러한 네트워크들의 조합을 포함할 수 있다.

- [0038] 사용자 단말기(102)는 웹 브라우징(Web Browsing)이 가능한 통신 장치로서, 예를 들어 모바일 단말기, 스마트 디바이스, 데스크 탑, 랩탑, 웨어러블 기기 등을 포함할 수 있다. 사용자 단말기(102)에는 웹 브라우저가 탑재된다.
- [0039] 사용자 단말기(102)는 웹 브라우징 히스토리 정보를 뉴스 추천 서버(104)로 전송할 수 있다. 웹 브라우징 히스토리 정보는 예를 들어, 사용자 단말기(102)가 접속한 웹 페이지의 ID 인덱스, 해당 웹 페이지의 타이틀(제목), 해당 웹 페이지의 주소(즉, 웹 페이지의 URL), 해당 웹 페이지의 즐겨 찾기 설정 여부, 해당 웹 페이지를 방문한 시점 등이 포함될 수 있다. 사용자 단말기(102)는 소정 웹 페이지에 접속할 때마다 웹 브라우징 히스토리 정보를 뉴스 추천 서버(104)로 전송할 수 있다.
- [0040] 사용자 단말기(102)는 뉴스를 제공하는 웹 서버(미도시)(뉴스 추천 서버(104)와는 다른 서버일 수 있음)에 접속하여 뉴스를 제공받을 수 있다. 또한, 사용자 단말기(102)는 뉴스 추천 서버(104)에 접속하고 뉴스 추천 서버(104)로부터 사용자가 선호하는 분야의 뉴스를 추천 받을 수 있다.
- [0041] 뉴스 추천 서버(104)는 각 사용자 단말기(102)로부터 웹 브라우징 히스토리 정보를 수신할 수 있다. 뉴스 추천 서버(104)는 각 사용자 단말기(102)로부터 수신한 웹 브라우징 히스토리 정보를 기반으로 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출할 수 있다. 즉, 뉴스 추천 서버(104)는 각 사용자 단말기(102)로부터 수신한 일정 기간 동안의 웹 브라우징 히스토리 정보를 분석하여 기 설정된 뉴스 분야들 중 해당 사용자가 선호하는 뉴스 분야들의 비율을 산출할 수 있다.
- [0042] 뉴스 추천 서버(104)는 사용자 단말기(102)가 접속하는 경우, 해당 사용자가 선호하는 분야의 뉴스를 추천(즉, 사용자 단말기(102)로 제공)할 수 있다. 이때, 뉴스 추천 서버(104)는 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절할 수 있다.
- [0043] 뉴스 추천 서버(104)는 뉴스 추천 서버(104)가 제공하는 각 뉴스 별로 사용자들의 호감도를 산출할 수 있다. 뉴스 추천 서버(104)는 사용자가 해당 뉴스를 읽는 속도를 기반으로 해당 뉴스의 호감도를 산출할 수 있다. 뉴스 추천 서버(104)는 뉴스 분야 별로 호감도가 높은 순으로 뉴스들을 분류하여 추천 뉴스 셋(Set)을 생성할 수 있다. 이때, 추천 뉴스 셋(Set)은 뉴스의 호감도가 기 설정된 값 이상인 뉴스들을 대상으로 할 수 있다.
- [0044] 뉴스 추천 서버(104)는 해당 사용자와 뉴스 구독 성향이 유사한 사용자들이 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천할 수 있다. 뉴스 추천 서버(104)는 해당 사용자와 뉴스 구독 성향이 유사한 사용자들이 본 뉴스들 중에서 해당 사용자와 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 뉴스들을 추출한 후, 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스 분야의 순서를 배열하고, 분야 별 추천하는 뉴스의 수량을 조절할 수 있다.
- [0045] 본 발명의 실시예에 의하면, 사용자 별 웹 브라우징 히스토리를 이용하여 사용자가 선호하는 뉴스 분야의 비율을 산출하고, 그에 따라 뉴스를 추천하여 줌으로써, 사용자의 선호도의 변화에 따라 능동적으로 뉴스를 추천할 수 있게 된다.
- [0047] 도 2는 본 발명의 일 실시예에 따른 뉴스 추천 서버의 구성을 나타낸 블록도이다.
- [0048] 도 2를 참조하면, 뉴스 추천 서버(104)는 데이터 수집부(111), 분야 비율 산출부(113), 뉴스 호감도 산출부(115), 및 뉴스 추천부(117)를 포함할 수 있다.
- [0049] 데이터 수집부(111)는 각 사용자 단말기(102)들로부터 웹 브라우징 히스토리 정보를 수집할 수 있다. 데이터 수집부(111)는 수집한 웹 브라우징 히스토리 정보를 데이터베이스(미도시)에 저장할 수 있다. 여기서, 데이터베이스(미도시)는 뉴스 추천 서버(104)에 구비되는 데이터 저장 매체일 수도 있고, 뉴스 추천 서버(104)의 외부에 구비되는 데이터 저장 매체일 수도 있다.
- [0050] 또한, 데이터 수집부(111)는 각 사용자 단말기(102) 별로 뉴스 추천 서버(104) 접속 시 로그 데이터를 수집할 수 있다. 예를 들어, 사용자 단말기(102)가 뉴스 추천 서버(104)에 접속하여 뉴스를 제공 받는 경우, 데이터 수집부(111)는 그에 따른 로그 데이터(해당 뉴스의 ID, 해당 뉴스의 타이틀, 해당 뉴스의 URL, 해당 뉴스를 제공

한 시점 등)를 수집할 수 있다. 데이터 수집부(111)는 수집한 로그 데이터를 데이터베이스(미도시)에 저장할 수 있다.

[0051] 여기서, 웹 브라우징 히스토리 정보는 해당 사용자의 웹 브라우징에 따른 전체 웹 페이지를 본 기록이고, 로그 데이터는 해당 사용자가 뉴스 추천 서버(104)로부터 제공 받은 웹 페이지를 본 기록일 수 있다.

[0052] 분야 비율 산출부(113)는 각 사용자 단말기(102)들로부터 수집한 웹 브라우징 히스토리 정보를 기반으로 사용자 별 선호하는 뉴스 분야들의 비율을 산출할 수 있다. 예시적인 실시예에서, 뉴스 분야는 스포츠(sports), 엔터테인먼트(entertainment), 과학(science), 글로벌(global), 정치(politics), 경제(economy), 라이프(life), 기타 등의 8개의 분야로 분류될 수 있다. 여기서, 기타는 해당 웹 페이지가 이미지로만 되어 있거나 로그인에 필요한 이유로 분류가 불가능한 경우에 해당될 수 있다.

[0053] 구체적으로, 분야 비율 산출부(113)는 일정 기간 동안의 사용자 단말기(102)의 웹 브라우징 히스토리에서 해당 사용자 단말기(102)가 접속한 각 웹 페이지를 방문하여 해당 웹 페이지의 본문(즉, 텍스트)을 추출할 수 있다. 만약, 사용자 단말기(102)가 접속한 웹 페이지가 로그인이 필요한 서비스이거나 이미지로만 이루어져 본문 추출이 어려운 경우, 분야 비율 산출부(113)는 해당 웹 페이지의 타이틀(즉, 제목)을 텍스트로 추출할 수 있다.

[0054] 분야 비율 산출부(113)는 형태소 분석기(Part of Speech Tagger)를 이용하여 상기 추출한 웹 페이지의 텍스트에 포함된 단어들의 품사를 확인할 수 있다. 분야 비율 산출부(113)는 웹 페이지의 텍스트에 포함된 단어 중 품사가 동사, 형용사, 부사인 단어 및 단어의 문자열 길이가 1인 단어를 제외한 단어(이하, “분석 대상 단어”라 지칭될 수 있음)들을 대상으로 단어 별 빈도수를 측정할 수 있다. 즉, 품사가 동사, 형용사, 부사인 단어는 명사에 비해 상대적으로 정보의 양이 적고, 문자열 길이가 1인 단어는 뜻이 없거나 무의미한 경우가 대다수이므로 이러한 단어들은 제외시킬 수 있다.

[0055] 분야 비율 산출부(113)는 분석 대상 단어들에 가중치를 적용하여 가중치 값이 높은 순서대로 기 설정된 개수(예를 들어, 1,000개)의 단어(이하, 분야 산출 기초 단어로 지칭될 수 있음)를 추출하여 단어 벡터(Term Vector) (또는 벡터 공간(Vector Space))를 구성할 수 있다. 예시적인 실시예에서, 분야 비율 산출부(113)는 각 분석 대상 단어들에 TF-IDF(Term Frequency - Inverse Document Frequency) 가중치를 적용할 수 있다. 여기서, TF-IDF 가중치는 단어의 빈도수(Term Frequency)와 역 문헌 빈도수(Inverse Document Frequency)의 곱으로 나타낼 수 있다. 단어의 빈도수(Term Frequency)는 어떤 텍스트 내에서 특정 단어가 나타나는 횟수를 텍스트 내에 있는 모든 단어 수로 나눈 값일 수 있다. 역 문헌 빈도수(Inverse Document Frequency)는 전체 텍스트의 개수(N)를 특정 단어가 포함된 텍스트의 개수(n)로 나눈 값에 로그(log)를 취한 값(즉, $\log(N/n)$)일 수 있다.

[0056] 분야 비율 산출부(113)는 가중치 값이 높은 기 설정된 개수의 단어(분야 산출 기초 단어)들로 구성된 단어 벡터를 이용하여 각 텍스트 및 뉴스 분야에 대한 매트릭스(Matrix)를 구성한 후, 각 텍스트의 뉴스 분야를 분류할 수 있다. 즉, 분야 비율 산출부(113)는 표 1에 나타낸 바와 같이 단어 벡터를 이용하여 각 텍스트 및 뉴스 분야에 대한 매트릭스(Matrix)를 구성할 수 있다.

표 1

[0057]

	N_money	V_interest rate	N_family	...	N_football	뉴스 분야
텍스트1	1	0	0	...	0	Sports
텍스트2	1	1	0	...	0	Economy
텍스트 3	0	0	1	...	1	Life
...
텍스트 j	1	0	0	...	1	Sports

[0059] 분야 비율 산출부(113)는 상기 매트릭스에서 각 뉴스의 텍스트 별로 분야 산출 기초 단어들의 어느 정도 포함되어 있는지 여부를 확인(예를 들어, 해당 단어가 포함되어 있으면 1, 포함되어 있지 않으면 0을 부여)하고, 이를 머신 러닝(Machine Learning) 기법에 적용하여 각 뉴스의 분야가 어느 분야인지를 결정할 수 있다. 즉, 이미 분야들이 분류된 뉴스들에서 어떤 분야 산출 기초 단어들이 어느 정도 포함되어 있을 때 해당 분야로 분류되었는지를 머신 러닝(Machine Learning) 기법으로 분석하면, 상기 매트릭스를 이용하여 각 뉴스의 분야가 어느 분야인지를 결정할 수 있게 된다. 이때, 사용되는 알고리즘으로는 Naive Bayes, K-Nearest Neighbor, Decision Tree, Artificial Neural Network, Support Vector Machine 등이 있을 수 있다.

- [0060] 분야 비율 산출부(113)는 일정 기간 동안 해당 사용자 단말기(102)가 접속한 각 웹 페이지의 텍스트들을 스포츠(sports), 엔터테인먼트(entertainment), 과학(science), 글로벌(global), 정치(politics), 경제(economy), 라이프(life), 및 기타 중 어느 하나의 분야로 분류할 수 있다. 분야 비율 산출부(113)는 일정 기간 동안 해당 사용자 단말기(102)가 접속한 각 웹 페이지의 텍스트들에서 기타 분야를 제외한 나머지 분야 빈도의 합(이하, 전체 분야 빈도의 합이라 지칭될 수 있음)을 계산할 수 있다. 분야 비율 산출부(113)는 각 분야(스포츠, 엔터테인먼트, 과학, 글로벌, 정치, 경제, 라이프)의 빈도를 전체 분야 빈도의 합으로 나누어서 각 분야의 비율을 산출할 수 있다. 분야 비율 산출부(113)는 주기적으로 각 사용자 단말기(102)로부터 수집한 웹 브라우징 히스토리 정보를 분석하여 각 사용자들의 선호하는 뉴스 분야의 비율을 갱신할 수 있다.
- [0061] 뉴스 호감도 산출부(115)는 뉴스 추천 서버(104)에서 사용자 단말기(102)로 뉴스를 제공하는 경우, 각 뉴스에 대한 사용자의 호감도를 산출할 수 있다. 뉴스 호감도 산출부(115)는 사용자가 해당 뉴스를 읽는 속도를 기반으로 해당 뉴스에 대한 사용자의 호감도를 산출할 수 있다.
- [0062] 구체적으로, 뉴스 호감도 산출부(115)는 데이터 수집부(111)에서 수집한 웹 브라우징 히스토리 정보와 로그 데이터들을 비교하여 웹 브라우징 히스토리 정보들의 웹 페이지 주소(URL)들 중 로그 데이터에서 소정 뉴스를 제공한 시점과 동일한 접속 시점을 갖는 웹 페이지를 추출할 수 있다. 여기서, 추출한 웹 페이지의 주소는 뉴스 추천 서버(104)에서 사용자 단말기(102)로 제공한 뉴스의 웹 페이지 주소(즉, 로그 데이터에서 해당 뉴스의 URL)와 동일하다.
- [0063] 이때, 웹 브라우징 히스토리 정보의 웹 페이지 URL과 로그 데이터의 웹 페이지 URL을 접속 시간에 근거하여 매칭하는 이유는, 사용자 단말기(102)로부터 수집되는 웹 브라우징 히스토리 정보에는 뉴스 추천 서버(104)에 접속한 기록뿐만 아니라 해당 사용자의 웹 브라우징에 따른 전체 웹 페이지 접속 기록을 포함되기 때문이다.
- [0064] 뉴스 호감도 산출부(115)는 웹 브라우징 히스토리 정보에서 상기 추출한 웹 페이지 다음으로 접속한 웹 페이지의 접속 시간을 확인하여 해당 사용자가 상기 추출한 웹 페이지(즉, 뉴스 추천 서버(104)에서 사용자 단말기(102)로 제공한 뉴스의 웹 페이지)를 읽은 시간을 산출할 수 있다.
- [0065] 구체적으로, 뉴스 호감도 산출부(115)는 해당 사용자의 웹 브라우징 히스토리 정보에서 해당 뉴스의 웹 페이지 접속 시간 $T(i)$ 및 해당 뉴스의 웹 페이지 다음에 접속한 웹 페이지의 접속 시간 $T(i+1)$ 의 차이(즉, $T(i+1) - T(i)$)를 해당 뉴스를 읽은 시간으로 산출할 수 있다. 이때, 뉴스 호감도 산출부(115)는 해당 뉴스를 읽은 시간이 기 설정된 시간(예를 들어, 10분)을 초과하는 경우, 해당 뉴스를 노이즈로 분류할 수 있다. 즉, 사용자들이 사용자 단말기(102)(예를 들어, 스마트 디바이스)를 통해 뉴스를 보는 경우, 웹 페이지 또는 웹 브라우저를 종료하고 끝내는 경우보다는 스마트 디바이스에서 홈 버튼을 통해 홈 화면(메인 화면)으로 돌아가는 경우가 많다. 이때, 해당 웹 페이지는 종료된 것이 아니라 백 그라운드(Background)에서 계속 대기하는 중이므로, 실제적으로는 사용자가 뉴스를 읽고 있지 않음에도 불구하고 해당 뉴스를 읽는 시간이 길게 나올 수 있기 때문에, 뉴스를 읽은 시간이 기 설정된 시간(예를 들어, 10분)을 초과하는 경우, 해당 뉴스를 읽은 건을 노이즈로 분류할 수 있다.
- [0066] 뉴스 호감도 산출부(115)는 상기 산출한 뉴스 읽는 시간과 해당 뉴스의 문장 길이를 기반으로 해당 뉴스를 읽는 속도를 산출할 수 있다. 이때, 뉴스 호감도 산출부(115)는 해당 뉴스의 문장 길이가 기 설정된 문장 길이 미만(예를 들어, 50자 미만)인 경우, 해당 뉴스를 읽은 건을 노이즈로 분류할 수 있다.
- [0067] 뉴스 호감도 산출부(115)는 사용자의 해당 뉴스를 읽는 속도와 기 설정된 표준 속도(예를 들어, 20 ~ 30자/sec)를 비교하여 사용자의 해당 뉴스에 대한 호감도를 산출할 수 있다. 구체적으로, 뉴스 호감도 산출부(115)는 사용자의 해당 뉴스를 읽는 속도가 기 설정된 표준 속도에 근접할수록 해당 뉴스에 대한 호감도 값을 높게 할 수 있다. 예시적인 실시예에서, 해당 뉴스에 대한 호감도 값의 범위는 0 ~ 5로 설정할 수 있다. 뉴스 호감도 산출부(115)는 사용자의 해당 뉴스를 읽는 속도가 기 설정된 표준 속도에 근접하는 정도를 코사인 유사도(Cosine Similarity) 등을 이용하여 구할 수 있다.
- [0068] 뉴스 호감도 산출부(115)는 사용자 단말기(102)로 제공하는 각 뉴스에 대한 사용자의 호감도를 데이터베이스(미도시)에 저장할 수 있다. 즉, 뉴스 호감도 산출부(115)는 표 2에 나타낸 바와 같이, 사용자 별 각 뉴스에 대한 호감도를 데이터베이스화하여 저장할 수 있다.

표 2

	사용자 ID	뉴스 ID	호감도
1	1	521	4.7
2	1	620	3.9
3	1	258	3.4
4	2	521	3
...
n	n	521	5

[0069]

[0071]

[0072]

[0073]

[0074]

[0076]

[0077]

[0078]

[0079]

[0081]

[0082]

뉴스 추천부(117)는 뉴스 추천 서버(104)에 로그인하여 접속하는 사용자 단말기(102)로 해당 사용자가 선호하는 분야의 뉴스를 추천(즉, 사용자 단말기(102)로 제공)할 수 있다. 뉴스 추천부(117)는 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 배열 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절할 수 있다.

예를 들어, 해당 사용자가 선호하는 뉴스 분야의 비율이 스포츠:글로벌:정치:경제:라이프:과학:엔터테인먼트 = 1:4:4:4:2:7:2인 경우, 뉴스 추천부(117)는 도 3에 도시된 바와 같이, 뉴스를 과학, 글로벌, 정치, 경제, 라이프, 엔터테인먼트, 스포츠의 순서로 사용자에게 추천해 줄 수 있다. 그리고, 뉴스 추천부(117)는 추천해주는 뉴스의 양을 해당 사용자가 선호하는 뉴스 분야의 비율(즉, 1:4:4:4:2:7:2)에 따라 스포츠 1건, 글로벌 4건, 정치 4건, 경제 4건, 라이프 2건, 과학 7건, 엔터테인먼트 2건으로 추천하는 뉴스 수량을 결정할 수 있다.

또한, 해당 사용자의 선호하는 뉴스 분야의 비율이 스포츠:글로벌:정치:경제:라이프:과학:엔터테인먼트 = 6:3:3:3:1:5:1로 달라진 경우, 뉴스 추천부(117)는 도 4에 도시된 바와 같이, 뉴스를 스포츠, 과학, 글로벌, 정치, 경제, 라이프, 엔터테인먼트의 순서로 사용자에게 추천해 줄 수 있다. 그리고, 뉴스 추천부(117)는 추천해주는 뉴스의 양을 해당 사용자가 선호하는 뉴스 분야의 비율(즉, 6:3:3:3:1:5:1)에 따라 스포츠 6건, 글로벌 3건, 정치 3건, 경제 3건, 라이프 1건, 과학 5건, 엔터테인먼트 1건으로 추천하는 뉴스 수량을 결정할 수 있다.

뉴스 추천부(117)는 사용자 단말기(102)의 로그 데이터를 분석하여 해당 사용자가 본 뉴스와 유사한 뉴스를 본 다른 사용자들을 검색한 후, 검색된 다른 사용자들이 본 뉴스 중에서 뉴스를 추천할 수도 있다. 구체적으로, 뉴스 추천부(117)는 검색된 다른 사용자들(즉, 해당 사용자와 뉴스 구독 성향이 유사한 사용자)이 본 뉴스들 중에서 해당 사용자가 보지 않은 뉴스들을 추출한 후, 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 추천할 수 있다. 이때, 뉴스 추천부(117)는 해당 사용자가 선호하는 뉴스 분야의 뉴스들 중 해당 사용자와 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 순서대로 추천할 수 있다.

도 5는 본 발명의 일 실시예에 따른 뉴스 추천 방법을 설명하기 위한 흐름도이다. 도 5에 도시된 방법은 예를 들어, 전술한 뉴스 추천 서버(104)에 의해 수행될 수 있다. 도시된 흐름도에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.

도 5를 참조하면, 뉴스 추천 서버(104)는 각 사용자 단말기(102)들로부터 웹 브라우징 히스토리 정보를 수집한다(S 101).

다음으로, 뉴스 추천 서버(104)는 각 사용자 단말기(102)들로부터 수집한 웹 브라우징 히스토리 정보를 기반으로 사용자 별 선호하는 뉴스 분야들의 비율을 산출한다(S 103).

다음으로, 뉴스 추천 서버(104)는 각 사용자가 선호하는 뉴스 분야의 비율에 따라 추천하는 뉴스 분야의 순서 및 분야 별 추천하는 뉴스의 수량 중 적어도 하나를 조절하여 제공한다(S 105).

도 6은 본 발명의 다른 실시예에 따른 뉴스 추천 방법을 설명하기 위한 흐름도이다. 도 6에 도시된 방법은 예를 들어, 전술한 뉴스 추천 서버(104)에 의해 수행될 수 있다. 도시된 흐름도에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.

도 6을 참조하면, 뉴스 추천 서버(104)는 각 사용자 단말기(102)들로부터 웹 브라우징 히스토리 정보를 수집하

고, 각 사용자 단말기(102) 별 로그 데이터를 수집한다(S 201).

- [0083] 다음으로, 뉴스 추천 서버(104)는 각 사용자 단말기(102)들로부터 수집한 웹 브라우징 히스토리 정보를 기반으로 사용자 별 선호하는 뉴스 분야의 비율을 산출한다(S 203).
- [0084] 다음으로, 뉴스 추천 서버(104)는 각 사용자 단말기(102)로 제공하는 각 뉴스에 대한 사용자의 호감도(즉, 뉴스 호감도)를 산출한다(S 205).
- [0085] 다음으로, 뉴스 추천 서버(104)는 소정 사용자 단말기(102)가 접속하는 경우(S 207), 해당 사용자와 뉴스 구독 성향이 유사한 사용자들을 검색한다(S 209). 예를 들어, 뉴스 추천 서버(104)는 Nearest Neighborhood 알고리즘을 이용하여 해당 사용자와 뉴스 구독 성향이 유사한 사용자들을 검색할 수 있다.
- [0086] 다음으로, 뉴스 추천 서버(104)는 해당 사용자와 뉴스 구독 성향이 유사한 사용자들이 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스를 추천한다(S 211). 예를 들어, 뉴스 추천 서버(104)는 해당 사용자와 뉴스 구독 성향이 유사한 사용자들이 본 뉴스들 중에서 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스 분야의 순서를 배열하고, 분야 별 추천하는 뉴스의 수량을 조절할 수 있다. 또한, 뉴스 추천 서버(104)는 해당 사용자와 뉴스 구독 성향이 유사한 사용자들이 본 뉴스들 중에서 해당 사용자와 뉴스 구독 성향이 유사한 사용자들의 뉴스 호감도가 높은 뉴스들을 추출한 후, 해당 사용자가 선호하는 뉴스 분야의 비율에 따라 뉴스 분야의 순서를 배열하고, 분야 별 추천하는 뉴스의 수량을 조절할 수 있다. 이때, 각 분야 별로 상기 뉴스 호감도가 가장 높은 뉴스의 순서대로 추출하여 해당 수량에 맞게 뉴스를 추천할 수 있다.
- [0088] 도 7은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되는 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.
- [0089] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 사용자 단말기(예를 들어, 사용자 단말기(102))일 수 있다. 또한, 컴퓨팅 장치(12)는 뉴스를 추천하는 서버(예를 들어, 뉴스 추천 서버(104))일 수 있다.
- [0090] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0091] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0092] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0093] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.

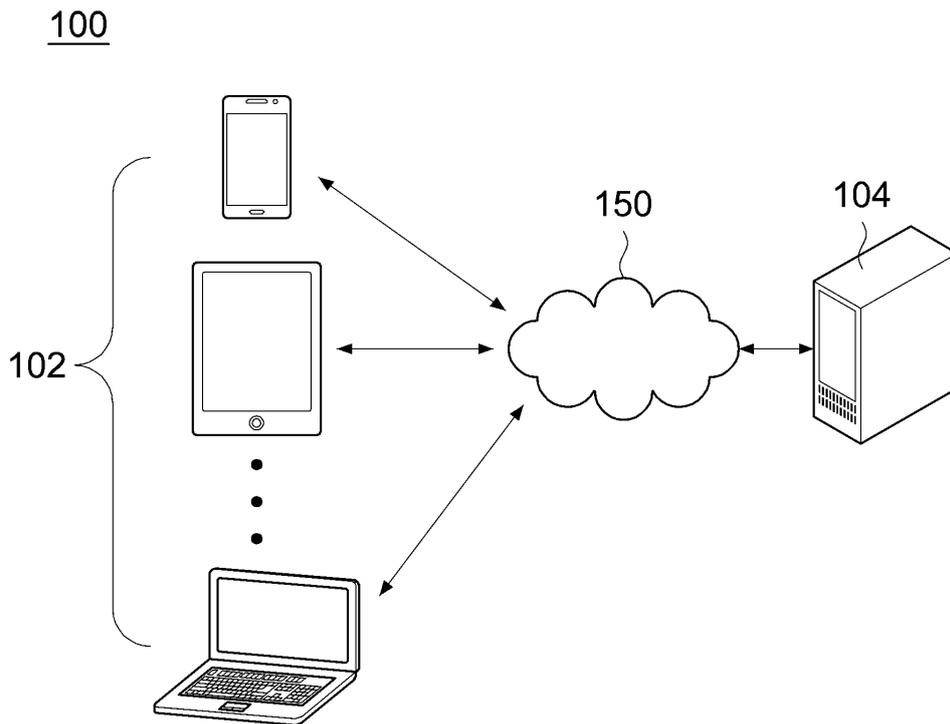
[0095] 이상에서 본 발명의 대표적인 실시예들을 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허 청구범위뿐만 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

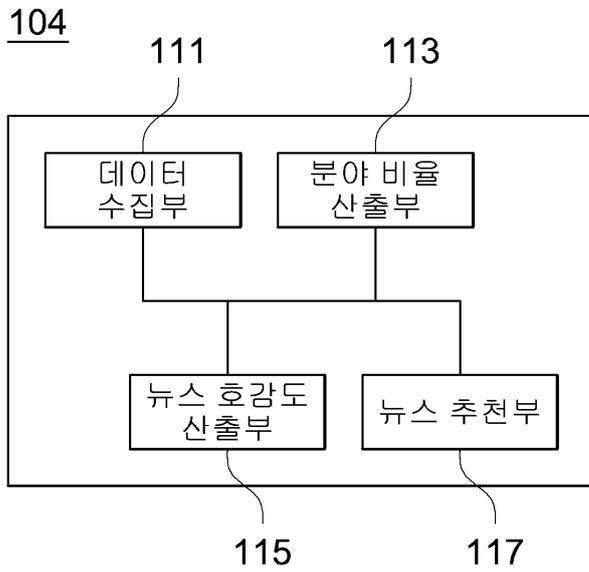
- [0097] 100 : 뉴스 추천 시스템
- 102 : 사용자 단말기
- 104 : 뉴스 추천 서버
- 111 : 데이터 수집부
- 113 : 분야 비율 산출부
- 115 : 뉴스 호감도 산출부
- 117 : 뉴스 추천부

도면

도면1



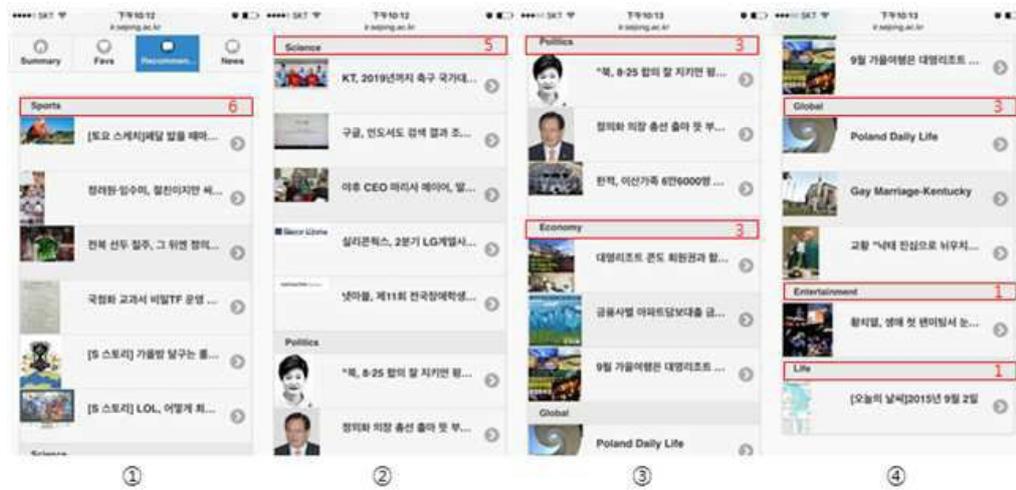
도면2



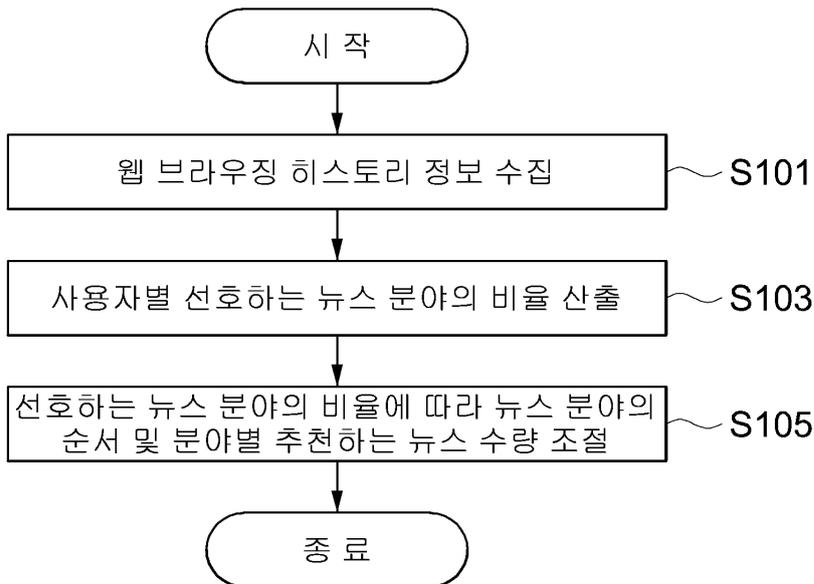
도면3



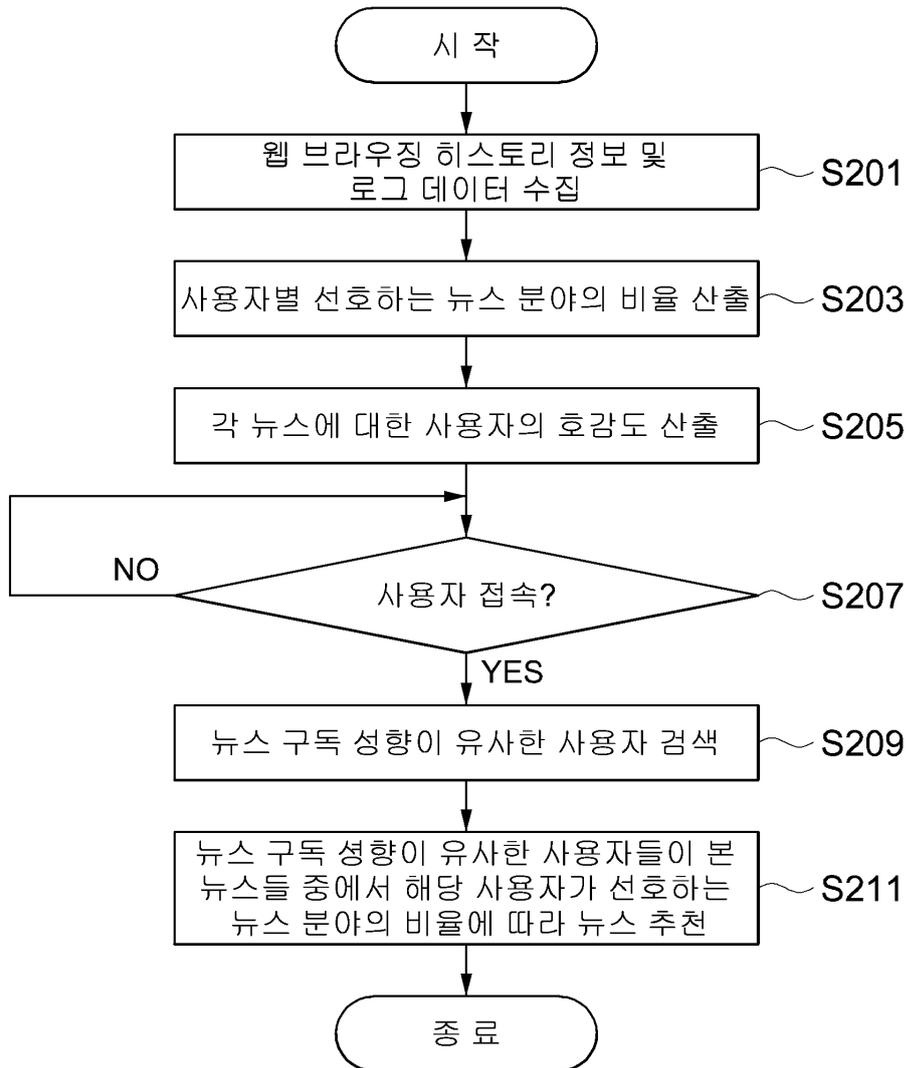
도면4



도면5



도면6



도면7

