



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년05월13일
(11) 등록번호 10-1620957
(24) 등록일자 2016년05월09일

(51) 국제특허분류(Int. Cl.)
G06F 17/00 (2006.01) G06F 17/40 (2006.01)
(21) 출원번호 10-2014-0164978
(22) 출원일자 2014년11월25일
심사청구일자 2014년11월25일
(56) 선행기술조사문헌
KR1020070021660 A*
KR1020140118649 A*
KR100929925 B1*
KR1020030048045 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
이상근
서울특별시 강남구 선릉로 221 도곡렉슬아파트 205-1104
이용구
서울특별시 중랑구 사가정로71길 11 용마한신아파트 101동 1602호
(74) 대리인
김동용, 김홍석

전체 청구항 수 : 총 5 항

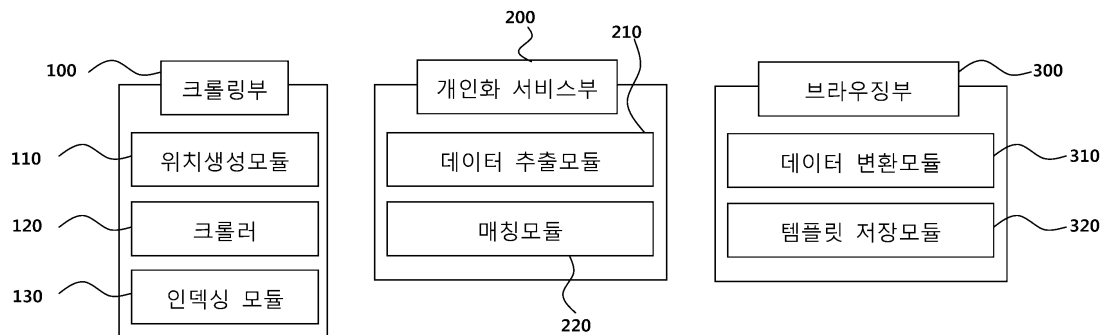
심사관 : 박상현

(54) 발명의 명칭 크롤링 기능을 갖는 웹브라우저 시스템 및 그 방법

(57) 요약

본 발명인 웹브라우저 시스템은 사용자 단말 내에서 크롤링 기능을 갖는 웹브라우저 시스템으로서, 웹사이트에 대한 크롤링을 수행하여 크롤링 웹 데이터를 생성하는 크롤링부, 저장되어 있는 사적(私的)정보인 프라이빗 데이터에서 추출된 문자열 데이터와 크롤링 웹 데이터를 매칭시켜 맞춤형 웹 데이터를 생성하는 개인화 서비스부 및 맞춤형 웹 데이터를 레이아웃 데이터로 변환하고, 기 저장된 템플릿과 결합시키는 브라우징부를 포함한다.

대표도 - 도2



명세서

청구범위

청구항 1

사용자 단말 내에서 크롤링 기능을 갖는 웹브라우저 시스템으로서,
 웹사이트에 대한 크롤링을 수행하여 크롤링 웹 데이터를 생성하는 크롤링부;
 저장되어 있는 사적(私的)정보인 프라이빗 데이터에서 추출된 문자열 데이터와 상기 크롤링 웹 데이터를 매칭시켜 맞춤형 웹 데이터를 생성하는 개인화 서비스부;
 상기 맞춤형 웹 데이터를 레이아웃 데이터로 변환하고, 기 저장된 템플릿과 결합시키는 브라우저부; 및
 상기 크롤링부를 제어하는 크롤링 제어부를 포함하며,
 상기 크롤링 제어부는 와이파이의 연결상태 또는 단말의 배터리 충전량에 따라서 상기 크롤링부를 제어하는 것을 특징으로 하는 웹브라우저 시스템.

청구항 2

제1항에 있어서,
 상기 크롤링부는 사용자가 직접 정해주거나 사용자가 자주 방문한 웹사이트 정보를 통해 크롤링을 수행할 웹사이트의 위치를 생성하는 위치생성모듈; 및
 생성된 웹사이트의 제목과 내용을 크롤링하는 크롤러를 포함하는 것을 특징으로 하는 웹브라우저 시스템,

청구항 3

제1항에 있어서,
 상기 개인화 서비스부는 상기 프라이빗 데이터에서 문자열 데이터를 추출하는 데이터 추출모듈; 및
 상기 문자열 데이터와 상기 크롤링 웹 데이터를 매칭시켜 맞춤형 웹데이터를 생성하는 매칭모듈을 포함하는 것을 특징으로 하는 웹브라우저 시스템.

청구항 4

제1항에 있어서,
 상기 브라우저부는 상기 맞춤형 웹 데이터를 상기 레이아웃 데이터로 변환하는 데이터 변환모듈; 및
 상기 레이아웃 데이터와 결합하는 레이아웃 템플릿을 저장하는 템플릿 저장모듈을 포함하는 것을 특징으로 하는 웹브라우저 시스템.

청구항 5

제4항에 있어서,
 상기 레이아웃 데이터는 화면에 출력할 메타정보를 포함하는 것을 특징으로 하는 웹브라우저 시스템.

청구항 6

삭제

청구항 7

삭제

발명의 설명

기술 분야

[0001] 본 발명은 크롤링 기능을 갖는 웹브라우저 시스템 및 그 방법에 관한 발명으로서 보다 상세하게는 단말이 크롤링을 통해 웹사이트의 정보를 저장하고, 저장된 웹사이트 정보에서 브라우저를 수행하는 웹브라우저 시스템 및 그 방법에 관한 발명이다.

배경 기술

[0002] 최근, 웹(Web)의 빠른 성장과 더불어 엄청난 양의 정보들이 웹을 통해 제공되고 있다. 웹이 처음 등장하였을 때에는 웹이 갖고 있는 정보의 양이 상대적으로 매우 적었기 때문에 이러한 초기의 웹은 정적 페이지(static page) 형태로 표현되었다. 이러한 정적 페이지의 URL(Uniform Resource Locator)은 대개 지속적으로 존재함으로써 하이퍼링크(hyperlink)를 따라 방문이 가능했다. 이러한 특징을 갖는 웹을 서피스웹(Surface Web)이라 부른다.

[0003] 이러한 서피스웹을 검색하기 위해 많은 검색 엔진들이 개발되었다. 이러한 검색 엔진들은 크롤러(crawler)라 불리는 프로그램을 이용하여 하이퍼링크를 따라 웹 페이지를 방문하고 방문한 웹 페이지의 인덱스를 생성한다. 이때, 생성된 인덱스는 사용자의 질의를 만족하는 페이지를 찾는 데 사용된다.

[0004] 그러나 현재의 웹은 초기의 웹보다 훨씬 많은 정보들을 갖고 있다. 이러한 정보들을 효율적으로 관리하고 검색하기 위해 현재 웹은 자신의 후단(back-end) 데이터베이스(database)에 정보들을 저장한다. 일반적으로 웹 후단의 데이터베이스로부터 정보를 가져오기 위해 사용자들은 웹 사이트의 검색 기능을 이용한다. 여기서, 웹사이트들은 이러한 검색 기능을 HTML(Hypertext Markup Language)의 <FORM> 태그를 사용하여 제공한다. HTML의 <FORM> 태그는 사용자의 입력을 웹 서버에 전달하는 역할을 한다. 즉, 사용자가 검색 폼(search form)에 찾고자 하는 질의를 입력하면 웹 사이트는 이러한 질의와 연관된 정보들을 자신의 후단 데이터베이스에서 검색한다. 웹 사이트는 검색된 결과들을 포함한 페이지를 동적으로 생성하고, 이 페이지를 사용자에게 보여준다. 따라서, “깊은(deep)” 혹은 “숨겨진(hidden)” 데이터베이스로부터 정보가 검색되어 나오기 때문에 이러한 웹을 “딥웹(Deep Web)” 혹은 “히든웹(Hidden Web)”이라 부른다.

[0005] 현재 웹 크롤링을 위해 많은 검색 엔진 기업들은 엄청난 양의 분산된 컴퓨팅 자원을 사용하고 있다. 예를 들어, 구글(Google, <http://www.google.com>)은 수십만 대의 컴퓨터들을 전 세계에 직접 설치하여 웹 크롤링에 이용하고 있다.

[0006] 도1은 일반적인 웹사이트 검색을 위한 시스템도이다.

[0007] 도1에 도시된 바와 같이 컴퓨터, 노트북, 테블릿, 스마트 폰 등과 같은 사용자 단말(20)은 인터넷 망을 통해서 서버(10)와 연결되고, 사용자 단말(20)을 통해서 검색어가 입력되면 서버(10)에서 웹 크롤링을 통해 데이터베이스에 저장되어 있는 웹 사이트내에 검색어와 관련된 콘텐츠를 사용자 단말(20)로 제공한다. 최근에는 사용자가 직접 검색어 등을 입력하지 않더라도 사용자가 다른 사용자와 주고 받는 문자 메시지나 사진, 위치 정보 등에서 추출한 정보, 특히 텍스트 정보 등을 이용하여 사용자가 원하는 웹 페이지의 자동검색을 실행한다.

[0008] 다만 이 경우 사용자가 주고 받은 문자 메시지나 자신의 단말기에 저장되어 있는 사진이나 위치정보 등은 보호가 필요한 개인정보에 해당함에도 이러한 개인정보가 웹 크롤링을 위해서 서버(10)로 전송되거나 공개되어야 한다. 이 과정에서 자칫 사용자 단말(20)에 저장되어 있는 프라이빗 데이터(문자 메시지, 사진, 위치정보 등)가 공중에 오픈되어 사생활 침해의 위험이 발생한다.

발명의 내용

해결하려는 과제

[0009] 본 발명의 목적은 웹브라우저 사용자의 보안을 획득하고 개인정보를 보호하기 위한 크롤링 기능을 갖는 웹브라우저 시스템 및 그 방법을 제공하는 데 있다.

과제의 해결 수단

[0010] 본 발명의 실시예에 따른 웹브라우저 시스템은 사용자 단말 내에서 크롤링 기능을 갖는 웹브라우저 시스템으로서, 웹사이트에 대한 크롤링을 수행하여 크롤링 웹 데이터를 생성하는 크롤링부, 저장되어 있는 사적(私的)개인 프라이빗 데이터에서 추출된 문자열 데이터와 상기 크롤링 웹 데이터를 매칭시켜 맞춤형 웹 데이터를 생성하는 개인화 서비스부 및 상기 맞춤형 웹 데이터를 레이아웃 데이터로 변환하고, 기 저장된 템플릿과 결합시키는

브라우징부를 포함한다.

[0011] 본 발명의 실시예에 따른 웹브라우징 방법은 사용자 단말에 의해서 구현되는 웹브라우징 방법으로서, 크롤링을 수행할 웹사이트 위치를 생성하는 단계, 상기 웹사이트 위치에서 크롤링을 수행하여 크롤링 웹데이터를 생성하는 단계, 프라이빗 데이터에서 문자열 데이터를 추출하는 단계, 상기 문자열 데이터와 상기 크롤링 웹데이터를 매칭시켜 맞춤형 웹데이터를 생성하는 단계, 상기 맞춤형 웹데이터를 메타데이터를 포함하는 레이아웃 데이터로 변환하는 단계 및 상기 레이아웃 데이터와 레이아웃 템플릿을 결합하여 웹페이지를 출력하는 단계를 포함한다.

발명의 효과

[0012] 본 발명의 실시예에 따른 웹브라우징 시스템 및 그 방법에 의하면 브라우징을 통해 노출위험이 있는 개인정보를 보호할 수 있다. 또한 크롤링이 단말내에서 이루어짐에도 불구하고, 배터리의 소모를 최소화하고 무선데이터의 사용으로 인한 과금을 방지할 수 있다.

도면의 간단한 설명

[0013] 도1은 일반적인 웹사이트 검색을 위한 시스템도이다.

도2는 본 발명의 실시예에 따른 웹브라우징 시스템의 블록도이다.

도3은 본 발명의 또 다른 실시예에 따른 웹브라우징 시스템의 블록도이다.

도4는 본 발명의 실시예에 따른 웹브라우징 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명은 단지 본 발명의 개념에 따른 실시 예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시 예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시 예들에 한정되지 않는다.

[0015] 본 발명의 개념에 따른 실시 예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시 예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시 예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.

[0016] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0017] 본 명세서에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0018] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0019] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시 예들을 상세히 설명한다.

[0020] 도2는 본 발명의 실시예에 따른 웹브라우징 시스템의 블록도이다.

[0021] 도2에 도시된 바와 같이 본 발명의 실시예에 따른 웹브라우징 시스템은 사용자 단말 내에서 크롤링 기능을 갖는

웹브라우저 시스템으로서, 웹사이트에 대한 크롤링을 수행하여 크롤링 웹 데이터를 생성하는 크롤링부(100), 저장되어 있는 사적(私的)정보인 프라이빗 데이터에서 추출된 문자열 데이터와 크롤링 웹 데이터를 매칭시켜 맞춤형 웹 데이터를 생성하는 개인화 서비스부(200) 및 맞춤형 웹 데이터를 레이아웃 데이터로 변환하고, 기 저장된 템플릿과 결합시키는 브라우저부(300)를 포함한다.

- [0022] 크롤링부(100)는 웹사이트에 대한 크롤링을 수행한다. 크롤링은 무수히 많은 컴퓨터에 분산 저장되어 있는 문서를 수집하여 검색대상의 색인으로 포함시키는 것을 말한다. 특히 웹크롤링은 검색 엔진과 같은 여러 사이트에서 데이터의 최신 상태를 유지하기 위해 수행한다.
- [0023] 도2와 같이 크롤링부는 위치생성모듈(110)과 크롤러(120)를 포함한다.
- [0024] 위치생성모듈(110)은 사용자가 직접 정하거나 사용자가 자주 방문한 웹사이트 정보를 통해 크롤링을 수행할 웹사이트의 위치를 생성한다. 위치생성모듈(110)에서 생성한 웹사이트는 사용자가 자주 방문한 웹사이트 또는 관리자가 직접 정해준 웹사이트 일 수 있으며, 위치생성모듈(110)을 통해 획득된 웹사이트에 대한 크롤링을 수행하게 된다.
- [0025] 크롤러(120)는 앞서 위치생성모듈에 의해서 생성된 웹사이트의 제목과 내용을 크롤링한다. 크롤러(120)는 대체로 생성된 웹사이트의 모든 페이지의 복사본을 저장한다. 크롤러(120)는 주기적으로 웹사이트의 페이지를 복사하여 최신 데이터를 우선적으로 저장한다. 크롤러(120)는 크롤링을 통해서 웹페이지에 대한 크롤링 웹 데이터를 생성한다.
- [0026] 또한 크롤러(120)는 링크 체크나 HTML코드 검증과 같은 웹사이트의 자동유지관리작업을 위해 사용될 수 있으며, 자동 이메일 수집과 같은 웹페이지의 특정 형태의 정보를 수집할 수도 있다. 여기서, 검색엔진은 월드와이드웹과 FTP서버에서 정보를 찾을 수 있도록 만들어진 소프트웨어를 의미한다. 또한 정보들은 웹페이지, 이미지, 영상, 다른 파일형식의 정보 등을 모두 포함한다.
- [0027] 더불어 크롤링부(100)는 이렇게 생성된 페이지에 대한 빠른 검색을 위해서 각 페이지를 인덱싱하는 인덱싱모듈(130)을 더 포함할 수 있다
- [0028] 이와 같이 크롤링부(100)에서 생성된 크롤링 웹 데이터는 본 발명에 따른 웹브라우저 시스템에 포함되어 있는 메모리(미도시)에 저장될 수 있다.
- [0029] 개인화 서비스부(200)는 데이터 추출모듈(210)과 매칭모듈(220)을 포함한다.
- [0030] 데이터 추출모듈(210)은 프라이빗 데이터, 즉 개인의 단말에 저장되어 있는 문자 메시지, 사진, 영상 등과 같은 각종 데이터에서 문자열 데이터를 추출한다. 이러한 문자열 데이터는 사용자가 검색을 위한 별도의 검색어 입력 없이도 자동으로 관련 웹사이트에 연결하여 사용자에게 정보를 제공하기 위한 자동으로 입력되는 검색어가 된다.
- [0031] 매칭모듈(220)은 앞서 데이터 추출모듈(210)에서 추출된 문자열 데이터와 크롤링부(100)에서 생성되었던 크롤링 웹 데이터를 매칭시켜 맞춤형 웹데이터를 생성한다.
- [0032] 맞춤형 웹데이터는 크롤링 웹 데이터 내에서 검색된 데이터로서 추출된 문자열 데이터와 매칭되는 웹 데이터이다. 특히 추출된 문자열을 포함하고 있는지 여부와 노출되는 빈도수 등 정확도를 활용하여 나열될 수 있으며, 최근에 크롤링된 데이터를 위주로 나열될 수도 있다.
- [0033] 브라우저부(300)는 데이터 변환모듈(310)과 템플릿 저장모듈(320)을 포함한다.
- [0034] 데이터 변환모듈(310)은 맞춤형 웹데이터를 레이아웃 데이터로 변환한다. 또한 템플릿 저장모듈(320)은 레이아웃 데이터와 결합하는 레이아웃 템플릿을 저장한다.
- [0035] 레이아웃 데이터는 화면에 출력할 메타 정보를 보유한 데이터로서, 레이아웃 데이터는 지정된 레이아웃 템플릿에 의해 화면에 출력될 형태가 결정된다. 레이아웃 템플릿은 적용되는 디바이스마다 다르게 적용될 수 있으며, 위 템플릿은 수정이 가능하다.
- [0036] 본 발명의 실시예에 따른 웹브라우저 시스템에 의하면 크롤링부(100)에서 생성된 크롤링 웹 데이터를 개인화 서비스부(200)에서 추출된 문자열 데이터와 매칭시켜 맞춤형 웹데이터를 생성하고, 브라우저부(300)에서는 맞춤형 웹데이터를 단말의 시스템에 적용하여 레이아웃 데이터로 변환하여 개인 단말의 화면에 출력하게 된다.
- [0037] 본 발명의 실시예에 따른 웹브라우저 시스템을 실제 적용한 예를 살펴보면, 스마트폰 등 개인단말을 활용하여

SMS를 사용하는 경우 SMS 대화내용은 개인 단말에 저장된다. 스마트폰 브라우저를 통해 방문한 웹페이지들에 대한 기록은 개인 단말에 저장되어 있다. 이러한 SMS대화내용에는 개인 관심사를 담고 있을 가능성이 높다. SMS대화 내용은 공개된 API를 통해 어플리케이션에서 접근이 가능하다. 예를 들어 사용자가 "영화"에 대한 대화를 다른 친구와 SMS를 통해 주고 받게 되면, 개인화 서비스부(200)는 "영화"라는 문자열 데이터를 추출한다.

[0038] 스마트폰 브라우저를 통해 방문하여 남아 있는 웹페이지 기록을 통해서 크롤링부(100)는 해당 웹페이지에 대한 크롤링을 주기적으로 수행한다. 사용자가 자주 방문하는 웹페이지는 해당 경로를 통해 사용자의 방문이 없더라도 주기적으로 크롤링을 수행하여 크롤링 웹 데이터를 생성하고, 이를 저장한다.

[0039] 본 발명에 따른 시스템에 의하면 앞서 "영화"라는 문자열 데이터와 생성된 크롤링 웹 데이터를 서로 매칭시켜 맞춤형 웹 데이터를 생성하고, 레이아웃 데이터로 변환하고 템플릿과 결합시켜 단말의 화면에 출력한다. 만일 사용자가 www.naver.com에 주로 방문하였을 경우, 이 페이지는 실시간으로 내용이 업데이트 되고, 크롤러를 통해 받아온 데이터 중에서 "영화"라는 텍스트가 포함된 링크주소, 혹은 페이지 내용을 사용자가 볼 수 있도록 브라우징이 가능하다.

[0040] 도3은 본 발명의 또 다른 실시예에 따른 웹브라우징 시스템의 블록도이다.

[0041] 도3에 도시된 바와 같이 본 발명의 실시예에 따른 웹브라우징 시스템은 크롤링부(100)를 제어하는 크롤링 제어부(400)를 더 포함하고 있고, 크롤링 제어부(400)는 와이파이(Wi-Fi)(500)의 연결상태 또는 단말의 배터리(600) 충전량을 실시간으로 모니터링하여 크롤링부(100)를 제어한다.

[0042] 본 발명의 실시예에 따른 웹브라우징 시스템은 크롤링을 통해서 수신하는 데이터에 대한 과금을 방지하고, 배터리 소모를 최소화하기 위하여 크롤링 제어부(400)를 포함한다.

[0043] 크롤링을 사용자의 단말을 통해 직접 실행할 경우 인터넷 데이터 전송과 배터리 자원이 많이 소모된다. 크롤링 제어부(400)는 Wifi의 연결상태를 확인하여 크롤링 시간을 결정한다. 예를 들어 크롤링 시간을 1시간 단위로 설정한 경우 크롤링 시간이 되었으나 Wifi에 연결상태가 끊긴 경우에는 크롤링부로 하여금 크롤링을 수행하지 못하도록 한다. 또한 크롤링 제어부(400)는 배터리 충전상태를 확인하여 크롤링 시간을 결정한다. 예를 들어 크롤링 시간을 1시간 단위로 결정하였을 경우 크롤링 시간이 되었지만 배터리 충전중인 상태가 아닐 경우 또는 기 지정된 %미만의 충전량인 경우에는 크롤링부(100)로 하여금 크롤링을 수행하지 못하도록 한다.

[0044] 사용자는 크롤링 시간 간격, 연결 제한상태를 옵션으로 결정함으로써 사용자가 리소스를 직접 관리할 수 있도록 한다. 물론 사용자가 직접 해당 웹사이트를 크롤링하고 싶은 경우에는 크롤링 제어부(400)를 수동으로 동작시켜 크롤링이 가능하다.

[0045] 이상 본 발명의 실시예에 따른 웹브라우징 시스템에 대해서 살펴보았다.

[0046] 이하 본 발명의 또 다른 양태에 따른 웹브라우징 방법에 대해서 살펴본다.

[0047] 도4는 본 발명의 실시예에 따른 웹브라우징 방법의 순서도이다.

[0048] 도4에 도시된 바와 같이 본 발명의 실시예에 따른 웹브라우징 방법은 사용자 단말에 의해서 구현되는 웹브라우징 방법으로서, 크롤링을 수행할 웹사이트 위치를 생성하는 단계(S100), 웹사이트 위치에서 크롤링을 수행하여 크롤링 웹데이터를 생성하는 단계(S200), 프라이빗 데이터에서 문자열 데이터를 추출하는 단계(S300), 문자열 데이터와 크롤링 웹데이터를 매칭시켜 맞춤형 웹데이터를 생성하는 단계(S400), 맞춤형 웹데이터를 메타데이터를 포함하는 레이아웃 데이터로 변환하는 단계(S500) 및 레이아웃 데이터와 레이아웃 템플릿을 결합하여 웹페이지를 출력하는 단계(S600)를 포함한다.

[0049] 크롤링을 수행할 웹사이트 위치를 생성하는 단계(S100)는 사용자가 방문한 기록이 있거나 사용자의 설정에 의해서 정해진 웹사이트의 위치를 생성할 수 있다. 크롤링을 수행할 웹사이트 위치가 생성된 경우, 해당 웹사이트 위치에서 크롤링을 수행하여 크롤링 웹데이터를 생성하는 단계(S200)를 거친다. 이 단계는 크롤링되는 해당 웹사이트의 모든 페이지에 대해서 복사 및 저장하는 단계이다.

[0050] 프라이빗 데이터에서 문자열데이터를 추출하는 단계(S300)는 개인단말에 저장되어 있는 SMS문자, 사진, 동영상 등 다양한 형태의 개인의 사생활 정보인 프라이빗 데이터에서 사용자가 선호하거나 관심을 갖는 테마 등을 프라이빗 데이터의 텍스트를 분석하여 문자열 데이터를 추출한다.

[0051] 문자열 데이터와 크롤링 웹데이터를 매칭시켜 맞춤형 웹 데이터를 생성하는 단계(S400)는 앞서 크롤링을 통해서 저장된 크롤링 웹데이터에서 사용자의 프라이빗 데이터에서 추출된 문자열 데이터와 일치하거나 유사한 내용을

포함하고 있는 웹데이터를 매칭시켜 맞춤형 웹 데이터를 생성한다.

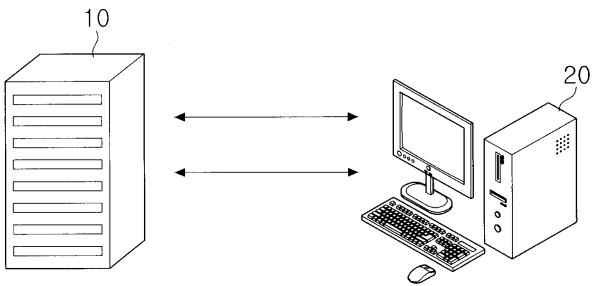
[0052] 맞춤형 웹데이터를 레이아웃 데이터로 변환하는 단계(S500)는 맞춤형 웹데이터를 메타정보 등을 포함하는 레이아웃 데이터로 변환하고, 이 레이아웃 데이터를 기 저장된 레이아웃 템플릿과 결합시켜 화면에 출력될 형태를 결정하고 웹페이지를 출력한다(S600).

부호의 설명

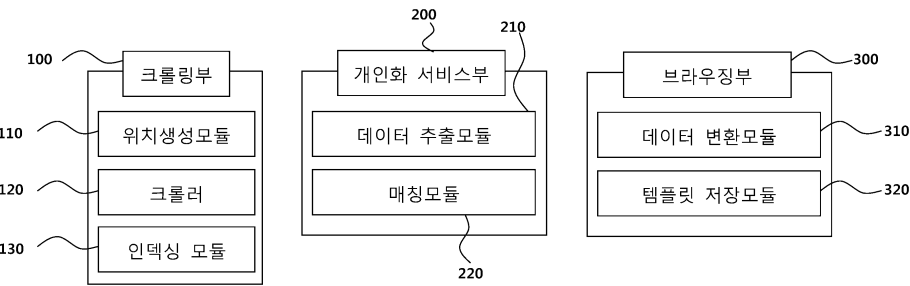
- [0053]
- | | |
|--------------|--------------|
| 100 크롤링부 | 110 위치생성모듈 |
| 120 크롤러 | 130 인덱싱 모듈 |
| 200 개인화 서비스부 | 210 데이터 추출모듈 |
| 220 매칭모듈 | 300 브라우징부 |
| 310 데이터 변환모듈 | 320 템플릿 저장모듈 |
| 400 크롤러 제어부 | 500 Wifi |
| 600 배터리 | |

도면

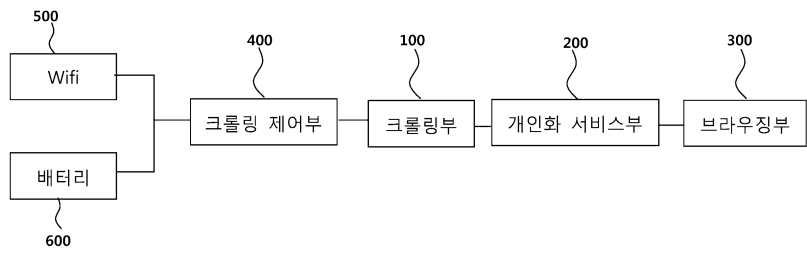
도면1



도면2



도면3



도면4

