(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0003294
(43) 공개일자 2022년01월10일

(51) 국제특허분류(Int. Cl.)
G06Q 30/06 (2012.01) *G06F 16/951* (2019.01)
G06F 9/455 (2018.01)
(52) CPC특허분류
G06Q 30/0639 (2013.01)
G06F 16/951 (2019.01)
(21) 출원번호 10-2020-0080921
(22) 출원일자 2020년07월01일
심사청구일자 2020년07월01일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
최진영
서울특별시 노원구 동일로248길 16, 506동 1003호 (수락파크빌아파트)
김현조
서울특별시 양천구 목동서로 70, 227동 103호 (목동신시가지아파트2단지)
(74) 대리인
권성현, 유광철, 백두진, 강일신, 김정연

전체 청구항 수 : 총 15 항

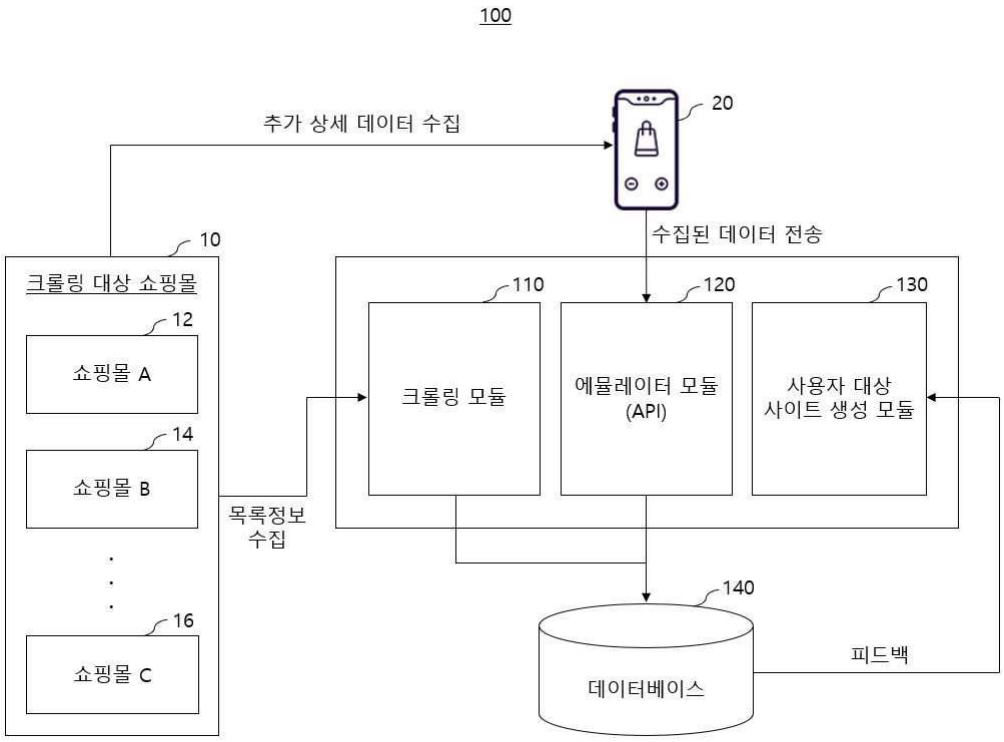
(54) 발명의 명칭 **에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법**

(57) 요 약

통합 플랫폼에 상품 기본 정보 목록을 최신 상태로 유지함과 동시에, 이 과정에서 발생하는 과도하거나 빈번한 크롤링을 감소시킴으로써 크롤링 서버 및 크롤링 대상 시스템의 리소스 부하를 줄일 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법, 기록 매체가 개시된다. 본 발명의 실시예에 따른 에뮬레이터를 활용한 상

(뒷면에 계속)

대 표 도 - 도1



품 데이터 크롤링 장치는 쇼핑몰 사이트에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상품의 상세 정보 중 기본 목록 구성 정보만 크롤링하도록 구성되는 크롤링 모듈; 다양한 쇼핑몰 사이트로부터 다양한 상품에 대해 수집된 기본 목록 구성 정보를 기반으로 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성하고, 사용자 디바이스로부터 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받도록 구성되는 사용자 대상 사이트 생성 모듈; 및 사용자 디바이스를 통해 대상 쇼핑몰 사이트에서 사용자 대상 상품의 상세 정보를 수집하여 데이터베이스에 저장시키도록 구성되는 에뮬레이터 모듈을 포함한다.

(52) CPC특허분류
G06F 9/455 (2013.01)
G06Q 30/0641 (2013.01)

명 세 서

청구범위

청구항 1

다양한 쇼핑몰 사이트에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상기 상품의 상세 정보 중 상기 기본 목록 구성 정보만 크롤링하도록 구성되는 크롤링 모듈;

상기 다양한 쇼핑몰 사이트로부터 다양한 상품에 대해 각각 상기 기본 목록 구성 정보를 크롤링하여 수집된 상기 다양한 상품에 관한 상품 데이터를 저장하도록 구성되는 데이터베이스;

상기 데이터베이스에 저장된 상기 상품 데이터로부터 상기 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성하고, 사용자 디바이스로부터 상기 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받도록 구성되는 사용자 대상 사이트 생성 모듈; 및

상기 사용자 대상 사이트에 생성된 사용자 대상 상품에 대해 상기 사용자 디바이스로부터 상기 상세 정보 요청이 입력되면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하여 상기 데이터베이스에 저장시키도록 구성되는 에뮬레이터 모듈;

을 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 2

제1항에 있어서,

상기 에뮬레이터 모듈은,

상기 사용자 디바이스로 하여금 상기 사용자 대상 상품에 해당하는 상품 데이터를 수집한 상기 대상 쇼핑몰 사이트에 상기 상세 정보를 열람하게 하고;

상기 사용자 디바이스에서 상기 대상 쇼핑몰 사이트를 열람한 상기 사용자 대상 상품의 상세 정보를 상기 사용자 디바이스로부터 전달받고; 그리고

상기 사용자 디바이스로부터 전달받은 상기 상세 정보를 상기 사용자 대상 상품에 관한 상품 데이터에 포함시켜 상기 데이터베이스에 저장시키도록 구성되는,

에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 3

제2항에 있어서,

상기 에뮬레이터 모듈은,

상기 사용자 디바이스로부터 상기 사용자 대상 상품에 대해 상기 상세 정보 요청이 입력되면, 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는지 판단하고;

상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는 경우, 상기 데이터베이스에 저장된 상세 정보를 상기 사용자 대상 사이트를 통해 상기 사용자 디바이스로 제공하고;

상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있지 않은 경우, 상기 사용자 디바이스에서 응용 프로그램 어플리케이션을 실행하여 상기 대상 쇼핑몰 사이트에 접근하게 하고; 그리고

상기 사용자 디바이스에서 실행되는 응용 프로그램 어플리케이션에 의해 상기 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 상기 데이터베이스에 저장하도록 구성되는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 4

제1항에 있어서,

상기 상세 정보는, 상기 사용자 대상 상품의 상품 상세 이미지, 크기, 중량, 재질, 색상 및 부가 설명을 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 5

제1항에 있어서,

상기 크롤링 모듈은,

상기 다양한 대상 쇼핑몰 사이트 별로 상기 기본 목록 구성 정보의 배열 패턴을 분석하고, 상기 배열 패턴을 기반으로 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 지정하도록 구성되는 크롤링 대상 분석기; 및

상기 다양한 대상 쇼핑몰 사이트로부터 상기 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 기반으로 상기 기본 목록 구성 정보를 크롤링하도록 구성되는 크롤러;

를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 6

제1항에 있어서,

상기 사용자 대상 사이트 생성 모듈은,

상기 다양한 대상 쇼핑몰 사이트로부터 수집되는 상기 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 상기 사용자 대상 사이트에 표시하도록 구성되는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 7

제1항에 있어서,

상기 에뮬레이터 모듈은,

상기 대상 쇼핑몰 사이트에 포함된 상품들 중, 상기 상세 정보 요청이 입력된 상기 사용자 대상 상품의 선호도 순위를 판단하고;

상기 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단하고;

상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하면, 상기 사용자 디바이스를 통해 상기 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하고; 그리고

상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품의 상세 정보를 수집하지 않도록 구성되는, 에뮬레이터를 활용한 상품 데이터 크롤링 장치.

청구항 8

크롤링 모듈에 의해, 다양한 쇼핑몰 사이트에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상기 상품의 상세 정보 중 상기 기본 목록 구성 정보 만을 크롤링하는 단계;

상기 다양한 쇼핑몰 사이트로부터 다양한 상품에 대해 각각 상기 기본 목록 구성 정보를 크롤링하여 수집된 상기 다양한 상품에 관한 상품 데이터를 데이터베이스에 저장하는 단계;

사용자 대상 사이트 생성 모듈에 의해, 상기 데이터베이스에 저장된 상기 상품 데이터로부터 상기 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성하는 단계;

상기 사용자 대상 사이트 생성 모듈에 의해, 사용자 디바이스로부터 상기 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받는 단계; 및

에뮬레이터 모듈에 의해, 상기 사용자 대상 상품에 대해 상기 사용자 디바이스로부터 상기 상세 정보 요청이 입력되면 상기 사용자 디바이스를 통해 상기 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하는 단계;

를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 9

제8항에 있어서,

상기 사용자 대상 상품의 상세 정보를 수집하는 단계는,

상기 사용자 디바이스로 하여금 상기 사용자 대상 상품에 해당하는 상품 데이터를 수집한 상기 대상 쇼핑몰 사이트에 상기 상세 정보를 열람하게 하는 단계;

상기 사용자 디바이스에서 상기 대상 쇼핑몰 사이트를 열람한 상기 사용자 대상 상품의 상세 정보를 상기 사용자 디바이스로부터 전달받는 단계; 및

상기 사용자 디바이스로부터 전달받은 상기 상세 정보를 상기 사용자 대상 상품에 관한 상품 데이터에 포함시켜 상기 데이터베이스에 저장시키는 단계;

를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 10

제9항에 있어서,

상기 사용자 대상 상품의 상세 정보를 수집하는 단계는,

상기 사용자 디바이스로부터 상기 사용자 대상 상품에 대해 상기 상세 정보 요청이 입력되면, 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는지 판단하는 단계;

상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는 경우, 상기 데이터베이스에 저장된 상세 정보를 상기 사용자 대상 사이트를 통해 상기 사용자 디바이스로 제공하는 단계;

상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있지 않은 경우, 상기 사용자 디바이스에서 응용 프로그램 어플리케이션을 실행하여 상기 대상 쇼핑몰 사이트에 접근하게 하는 단계; 및

상기 사용자 디바이스에서 실행되는 응용 프로그램 어플리케이션에 의해 상기 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 상기 데이터베이스에 저장하는 단계;를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 11

제8항에 있어서,

상기 상세 정보는, 상기 사용자 대상 상품의 상품 상세 이미지, 크기, 중량, 재질, 색상 및 부가 설명을 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 12

제8항에 있어서,

상기 크롤링하는 단계는,

크롤링 대상 분석기에 의해, 상기 다양한 대상 쇼핑몰 사이트 별로 상기 기본 목록 구성 정보의 배열 패턴을 분석하고, 상기 배열 패턴을 기반으로 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 지정하는 단계; 및

크롤러에 의해, 상기 다양한 대상 쇼핑몰 사이트로부터 상기 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 기반으로 상기 기본 목록 구성 정보를 크롤링하는 단계;

를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 13

제8항에 있어서,

상기 사용자 대상 사이트를 생성하는 단계는,

상기 다양한 대상 쇼핑몰 사이트로부터 수집되는 상기 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 상기 사용자 대상 사이트에 표시하는 단계를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 14

제8항에 있어서,

상기 사용자 대상 상품의 상세 정보를 수집하는 단계는,

상기 대상 쇼핑몰 사이트에 포함된 상품들 중, 상기 상세 정보 요청이 입력된 상기 사용자 대상 상품의 선호도 순위를 판단하는 단계;

상기 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단하는 단계;

상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하면, 상기 사용자 디바이스를 통해 상기 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하는 단계; 및

상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품의 상세 정보를 수집하지 않는 단계;

를 포함하는, 에뮬레이터를 활용한 상품 데이터 크롤링 방법.

청구항 15

제8항 내지 제14항 중 어느 한 항에 기재된 에뮬레이터를 활용한 상품 데이터 크롤링 방법을 실행하기 위한 프로그램이 기록된 컴퓨터로 판독 가능한 기록 매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 에뮬레이터(emulator)를 활용한 상품 데이터 크롤링 장치 및 방법에 관한 것으로, 보다 상세하게는 통합 플랫폼에 상품 기본 정보 목록을 최신 상태로 유지함과 동시에, 이 과정에서 발생하는 과도하거나 빈번한 크롤링을 감소시킴으로써 크롤링 서버 및 크롤링 대상 시스템의 리소스 부하를 줄일 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 최근 인터넷을 통해 쇼핑몰 등의 서비스를 제공하는 서비스 공급자의 경우 사용자 유입 채널을 늘리고 유저 풀을 공유할 수 있으며, 수요자의 경우에는 한 곳에서 여러 플랫폼의 데이터를 모아서 볼 수 있는 메타 서비스가 각광받고 있다. 이와 같은 메타 서비스는 일반적으로 크롤링(crawling)을 통해 웹 사이트로부터 정보를 수집한다. 현재의 웹 사이트 크롤링 방식은 응용 프로그램 인터페이스(API; Application Program Interface)를 사용하여 호출된 값을 받아오거나 웹 사이트의 각 페이지로 직접 접근하여 해당 페이지 내 데이터를 모으는 방식으로 구현되고 있다.

[0003] 이와 같은 종래의 웹 사이트 직접 접근 방식은 일정 시간 내 빈번한 접근 시도시 웹 사이트로의 접근이 쉽게 차단되는 문제가 있다. 웹 사이트는 기본적으로 불특정 상대가 자신의 웹 페이지를 스크랩하지 못하게 막도록 설정될 수 있으며, 이로 인해 웹 사이트로부터 접근이 차단되거나 데이터 크롤링이 제한될 수 있는 한계가 있다.

[0004] 또한, 종래에는 사용자가 비교를 원하는 여러 사이트에 각각 접근하여 데이터를 수집하고, 각 사이트에서 수집한 데이터를 비교, 분석하는 과정을 거쳐야 하므로, 조회 과정이 번거롭고 불편한 문제가 있었다. 또한, 트렌드에 민감한 커머스의 경우 그 특성상 웹 사이트에 항상 최신 목록을 유지해야 할 필요성이 있으나, 이를 유지하는 과정에서 발생하는 과도하거나 빈번한 크롤링은 크롤링 주체 시스템 뿐 아니라 크롤링 대상인 시스템에도 리소스 부하를 주게 되는 문제가 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 통합 플랫폼에 상품 기본 정보 목록을 최신 상태로 유지함과 동시에, 이 과정에서 발생하는 과도하거나 빈번한 크롤링을 감소시킴으로써 크롤링 서버 및 크롤링 대상 시스템의 리소스 부하를 줄일 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법, 기록 매체를 제공하기 위한 것이다.

[0006] 또한, 본 발명은 상품의 기본 목록 구성만을 크롤링한 후, 상품의 상세 정보를 요청한 사용자 디바이스를 통해 웹 사이트의 상품 상세 정보를 간접적으로 전달받아 상품 상세 데이터를 수집함으로써 서버의 크롤링 부담을 줄이고 웹 사이트에 의한 크롤링 제한을 방지할 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법, 기록 매체를 제공하기 위한 것이다.

[0007] 또한, 본 발명은 크롤링 대상 쇼핑몰의 상품 선호도(우선 순위) 기반으로 선택적으로 상품 상세 정보를 수집하여, 일부 상품에 대한 상세 정보만을 수집하여 전체 상품에 대한 상세 정보를 수집한 것에 가까운 효과를 발휘할 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법, 기록 매체를 제공하기 위한 것이다.

과제의 해결 수단

[0008] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치는, 다양한 쇼핑몰 사이트에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상기 상품의 상세 정보 중 상기 기본 목록 구성 정보만 크롤링하도록 구성되는 크롤링 모듈; 상기 다양한 쇼핑몰 사이트로부터 다양한 상품에 대해 각각 상기 기본 목록 구성 정보를 크롤링하여 수집된 상기 다양한 상품에 관한 상품 데이터를 저장하도록 구성되는 데이터베이스; 상기 데이터베이스에 저장된 상기 상품 데이터로부터 상기 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성하고, 사용자 디바이스로부터 상기 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받도록 구성되는 사용자 대상 사이트 생성 모듈; 및 상기 사용자 대상 사이트에 생성된 사용자 대상 상품에 대해 상기 사용자 디바이스로부터 상기 상세 정보 요청이 입력되면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하여 상기 데이터베이스에 저장시키도록 구성되는 에뮬레이터 모듈;을 포함한다.

[0009] 상기 에뮬레이터 모듈은, 상기 사용자 디바이스로 하여금 상기 사용자 대상 상품에 해당하는 상품 데이터를 수집한 상기 대상 쇼핑몰 사이트에 상기 상세 정보를 열람하게 하고; 상기 사용자 디바이스에서 상기 대상 쇼핑몰 사이트를 열람한 상기 사용자 대상 상품의 상세 정보를 상기 사용자 디바이스로부터 전달받고; 그리고 상기 사용자 디바이스로부터 전달받은 상기 상세 정보를 상기 사용자 대상 상품에 관한 상품 데이터에 포함시켜 상기 데이터베이스에 저장시키도록 구성될 수 있다.

[0010] 상기 에뮬레이터 모듈은, 상기 사용자 디바이스로부터 상기 사용자 대상 상품에 대해 상기 상세 정보 요청이 입력되면, 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는지 판단하고; 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는 경우, 상기 데이터베이스에 저장된 상세 정보를 상기 사용자 대상 사이트를 통해 상기 사용자 디바이스로 제공하고; 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있지 않은 경우, 상기 사용자 디바이스에서 응용 프로그램 어플리케이션을 실행하여 상기 대상 쇼핑몰 사이트에 접근하게 하고; 그리고 상기 사용자 디바이스에서 실행되는 응용 프로그램 어플리케이션에 의해 상기 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 상기 데이터베이스에 저장하도록 구성될 수 있다. 상기 상세 정보는, 상기 사용자 대상 상품의 상품 상세 이미지, 크기, 중량, 재질, 색상 및 부가 설명을 포함할 수 있다.

[0011] 상기 크롤링 모듈은, 상기 다양한 대상 쇼핑몰 사이트 별로 상기 기본 목록 구성 정보의 배열 패턴을 분석하고, 상기 배열 패턴을 기반으로 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 지정하도록 구성되는 크롤링 대상 분석기; 및 상기 다양한 대상 쇼핑몰 사이트로부터 상기 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 기반으로 상기 기본 목록 구성 정보를 크롤링하도록 구성되는 크롤러;를 포함할 수 있다.

[0012] 상기 사용자 대상 사이트 생성 모듈은, 상기 다양한 대상 쇼핑몰 사이트로부터 수집되는 상기 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 상기 사용자 대상 사이트에 표시하도록 구성될 수 있다.

[0013] 상기 에뮬레이터 모듈은, 상기 대상 쇼핑몰 사이트에 포함된 상품들 중, 상기 상세 정보 요청이 입력된 상기 사용자 대상 상품의 선호도 순위를 판단하고; 상기 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단하고; 상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하면, 상기 사용자 디바이스를 통해 상기 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하고; 그리고 상기

사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품의 상세 정보를 수집하지 않도록 구성될 수 있다.

[0014] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법은, 크롤링 모듈에 의해, 다양한 쇼핑몰 사이트에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상기 상품의 상세 정보 중 상기 기본 목록 구성 정보 만을 크롤링하는 단계; 상기 다양한 쇼핑몰 사이트로부터 다양한 상품에 대해 각각 상기 기본 목록 구성 정보를 크롤링하여 수집된 상기 다양한 상품에 관한 상품 데이터를 데이터베이스에 저장하는 단계; 사용자 대상 사이트 생성 모듈에 의해, 상기 데이터베이스에 저장된 상기 상품 데이터로부터 상기 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성하는 단계; 상기 사용자 대상 사이트 생성 모듈에 의해, 사용자 디바이스로부터 상기 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받는 단계; 및 에뮬레이터 모듈에 의해, 상기 사용자 대상 상품에 대해 상기 사용자 디바이스로부터 상기 상세 정보 요청이 입력되면 상기 사용자 디바이스를 통해 상기 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하는 단계;를 포함한다.

[0015] 상기 사용자 대상 상품의 상세 정보를 수집하는 단계는, 상기 사용자 디바이스로 하여금 상기 사용자 대상 상품에 해당하는 상품 데이터를 수집한 상기 대상 쇼핑몰 사이트에 상기 상세 정보를 열람하게 하는 단계; 상기 사용자 디바이스에서 상기 대상 쇼핑몰 사이트를 열람한 상기 사용자 대상 상품의 상세 정보를 상기 사용자 디바이스로부터 전달받는 단계; 및 상기 사용자 디바이스로부터 전달받은 상기 상세 정보를 상기 사용자 대상 상품에 관한 상품 데이터에 포함시켜 상기 데이터베이스에 저장시키는 단계;를 포함할 수 있다.

[0016] 상기 사용자 대상 상품의 상세 정보를 수집하는 단계는, 상기 사용자 디바이스로부터 상기 사용자 대상 상품에 대해 상기 상세 정보 요청이 입력되면, 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는지 판단하는 단계; 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있는 경우, 상기 데이터베이스에 저장된 상세 정보를 상기 사용자 대상 사이트를 통해 상기 사용자 디바이스로 제공하는 단계; 상기 사용자 대상 상품의 상세 정보가 상기 데이터베이스에 저장되어 있지 않은 경우, 상기 사용자 디바이스에서 응용 프로그램 어플리케이션을 실행하여 상기 대상 쇼핑몰 사이트에 접근하게 하는 단계; 및 상기 사용자 디바이스에서 실행되는 응용 프로그램 어플리케이션에 의해 상기 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 상기 데이터베이스에 저장하는 단계;를 포함할 수 있다.

[0017] 상기 크롤링하는 단계는, 크롤링 대상 분석기에 의해, 상기 다양한 대상 쇼핑몰 사이트 별로 상기 기본 목록 구성 정보의 배열 패턴을 분석하고, 상기 배열 패턴을 기반으로 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 지정하는 단계; 및 크롤러에 의해, 상기 다양한 대상 쇼핑몰 사이트로부터 상기 크롤링 대상 블록 및 상기 크롤링 대상 블록의 속성 및 필드를 기반으로 상기 기본 목록 구성 정보를 크롤링하는 단계;를 포함할 수 있다.

[0018] 상기 사용자 대상 사이트를 생성하는 단계는, 상기 다양한 대상 쇼핑몰 사이트로부터 수집되는 상기 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 상기 사용자 대상 사이트에 표시하는 단계를 포함할 수 있다.

[0019] 상기 사용자 대상 상품의 상세 정보를 수집하는 단계는, 상기 대상 쇼핑몰 사이트에 포함된 상품들 중, 상기 상세 정보 요청이 입력된 상기 사용자 대상 상품의 선호도 순위를 판단하는 단계; 상기 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단하는 단계; 상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하면, 상기 사용자 디바이스를 통해 상기 대상 쇼핑몰 사이트에서 상기 사용자 대상 상품의 상세 정보를 수집하는 단계; 및 상기 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 상기 사용자 디바이스를 통해 상기 사용자 대상 상품의 상세 정보를 수집하지 않는 단계;를 포함할 수 있다.

[0020] 본 발명의 실시예에 따르면, 상기 에뮬레이터를 활용한 상품 데이터 크롤링 방법을 실행하기 위한 프로그램이 기록된 컴퓨터로 판독 가능한 기록 매체가 제공된다.

발명의 효과

[0021] 본 발명의 실시예에 의하면, 통합 플랫폼에 상품 기본 정보 목록을 최신 상태로 유지함과 동시에, 이 과정에서 발생하는 과도하거나 빈번한 크롤링을 감소시킴으로써 크롤링 서버 및 크롤링 대상 시스템의 리소스 부하를 줄일 수 있는 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법, 기록 매체가 제공된다.

[0022] 또한, 본 발명의 실시예에 의하면, 상품의 기본 목록 구성만을 크롤링한 후, 상품의 상세 정보를 요청한 사용자 디바이스를 통해 웹 사이트의 상품 상세 정보를 간접적으로 전달받아 상품 상세 데이터를 수집함으로써 서버의 크롤링 부담을 줄이고 웹 사이트에 의한 크롤링 제한을 방지할 수 있다.

[0023] 또한, 본 발명의 실시예에 의하면, 크롤링 대상 쇼핑몰의 상품 선호도(우선 순위) 기반으로 선택적으로 상품 상세 정보를 수집하여, 일부 상품에 대한 상세 정보만을 수집하여 전체 상품에 대한 상세 정보를 수집한 것에 가까운 효과를 발휘할 수 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치의 구성도이다.

도 2는 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치를 구성하는 크롤링 모듈의 구성도이다.

도 3은 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치를 구성하는 크롤링 대상 분석기의 구성도이다.

도 4는 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법의 순서도이다.

도 5는 도 4의 단계 S110을 나타낸 순서도이다.

도 6은 도 4의 단계 S150을 나타낸 순서도이다.

도 7은 도 4의 단계 S150의 다른 예를 나타낸 순서도이다.

도 8은 도 4의 단계 S150의 또 다른 예를 나타낸 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0025] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0026] 본 명세서에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 본 명세서에서 사용되는 '~부, ~모듈'은 적어도 하나의 기능이나 동작을 처리하는 단위로서, 예를 들어 소프트웨어, FPGA 또는 하드웨어 구성요소를 의미할 수 있다. '~부, ~모듈'에서 제공하는 기능은 복수의 구성요소에 의해 분리되어 수행되거나, 다른 추가적인 구성요소와 통합될 수도 있다. 본 명세서의 '~부, ~모듈'은 반드시 소프트웨어 또는 하드웨어에 한정되지 않으며, 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고, 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 이하에서는 도면을 참조하여 본 발명의 실시예에 대해서 구체적으로 설명하기로 한다.

[0027] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 그 방법은 통합 플랫폼에 상품 기본 정보 목록을 항상 최신 상태로 유지함과 동시에, 그 과정에서 발생하는 과도하거나 빈번한 크롤링을 감소시킴으로써 시스템 리소스 부하를 줄이기 위해 제공된다.

[0028] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치 및 방법은 크롤링 서버에서 다양한 쇼핑몰 사이트에 대해 기본 상품 목록 구성만을 크롤링하도록 하고, 이후 사용자 디바이스가 크롤링 서버에 상품 상세 정보 요청시 해당 사용자 디바이스에서 대상 쇼핑몰 사이트로 접근하게 하여, 사용자 디바이스에서 대상 쇼핑몰 사이트에 조회하여 얻은 데이터를 크롤링 서버로 전송하게 하여 다양한 상품들 중 사용자 디바이스에 의해 요청된 상품에 대한 상세 정보 만을 크롤링할 수 있다.

[0029] 도 1은 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치의 구성도이다. 도 1을 참조하면, 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치(100)는 크롤링 대상 쇼핑몰(10)의 상품의 기본 목록 구성 정보(상품의 상품명과 썸네일 및 가격)를 사용자 대상 사이트에 통합하여 제공하고, 상품들 중 사용자에 의해 요청된 상품이 상세 정보를 사용자 디바이스(20)를 통해 수집하는 크롤링 서버

(crawling server)로 제공될 수 있다.

[0030] 사용자 디바이스(20)는 다양한 쇼핑몰 웹 사이트에서 상품 정보를 조회하여 비교, 분석하고자 하는 사용자의 단말기일 수 있으며, 예를 들어 스마트폰, 스마트패드, 태블릿 PC, 노트북 등의 랩탑(laptop), 데스크탑 컴퓨터 등의 단말기일 수 있으나, 이에 한정되는 것은 아니다.

[0031] 상품의 기본 목록 구성 정보는 쇼핑몰 사이트(12, 14, 16)의 상품에 관한 개략적인 정보를 표시하는 화면 상에 제공되는 웹 페이지에 포함된 정보일 수 있다. 따라서, 상품의 기본 목록 구성 정보는 쇼핑몰 사이트(12, 14, 16) 별로 상이할 수 있다.

[0032] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치(100)는 먼저 각 쇼핑몰 사이트(12, 14, 16)에서 상품의 개략적인 정보를 표시하는 화면에 표시되는 웹 페이지의 데이터에서 설정된 기본 목록 구성에 해당하는 정보를 수집할 수 있다.

[0033] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치(100)는 사용자 대상 사이트를 통해 다수의 사용자에게 상품 데이터를 제공하고, 다수의 사용자 디바이스(20)를 통해 상품의 상세 정보를 수집하여 다수의 쇼핑몰 사이트를 통합한 상품 정보를 제공하는 서버로 제공될 수 있다.

[0034] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치(100)는 크롤링 모듈(110), 에뮬레이터 모듈(120), 사용자 대상 사이트 생성 모듈(130) 및 데이터베이스(140)를 포함할 수 있다.

[0035] 크롤링 모듈(110)은 크롤링 대상 쇼핑몰(10)에 해당하는 다양한 쇼핑몰 사이트(12, 14, 16)에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상품의 상세 정보(예를 들어, 사용자 대상 상품의 상품 상세 이미지, 크기, 중량, 재질, 색상 및 부가 설명 등) 중 기본 목록 구성 정보만 크롤링하도록 구성될 수 있다.

[0036] 크롤링 대상 쇼핑몰(10)은 상품의 기본 목록 구성 정보와 상세 정보를 포함하는 원본 데이터(original data)를 보유하고 있는 쇼핑몰 웹 사이트일 수 있으며, 서버에 지정된 쇼핑몰 웹 사이트 또는 관리자 단말기(도시 생략)에 의해 지정된 쇼핑몰 웹 사이트일 수 있으나, 이에 한정되는 것은 아니다.

[0037] 도 2는 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치를 구성하는 크롤링 모듈의 구성도이다. 도 1 및 도 2를 참조하면, 크롤링 모듈(110)은 크롤링 대상 분석기(112) 및 크롤러(114)를 포함할 수 있다.

[0038] 실시예에서, 크롤링 모듈(110)은 쇼핑몰 사이트 별로 문서 객체 모델의 패턴과, 쇼핑몰 사이트에서 사용자 디바이스(20)에 제공할 크롤링 대상 블록을 분석하여 기본 목록 구성 정보를 크롤링할 수 있다. 일 예로, 크롤링 모듈(110)은 하나 또는 여러 개의 가상 머신으로 구성될 수 있다.

[0039] 크롤링 대상 분석기(112)는 다양한 쇼핑몰 사이트(12, 14, 16) 별로 기본 목록 구성 정보의 배열 패턴을 분석하고, 배열 패턴을 기반으로 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 지정하도록 구성될 수 있다.

[0040] 크롤링 모듈(110)은 다양한 쇼핑몰 사이트(12, 14, 16) 별로 문서 객체 모델의 패턴을 인식하여, 쇼핑몰 사이트(12, 14, 16) 별로 인식한 문서 객체 모델 패턴을 기반으로 쇼핑몰 사이트(12, 14, 16)의 웹 페이지에서 필요로 하는 정보를 추출할 수 있다.

[0041] 쇼핑몰 사이트의 문서 객체 모델은 프로그램이나 스크립트가 웹 페이지 내의 구성 요소들에 접근하여 내용이나 스타일 등을 변경할 수 있게 해주는 인터페이스일 수 있다. 브라우저에서 HTML, XML 등의 웹 페이지가 로딩되면 문서 객체 모델은 브라우저 내에 트리(tree) 형태로 문서 내 요소들을 구성하게 된다.

[0042] 크롤러(114)는 다양한 대상 쇼핑몰 사이트(12, 14, 16)로부터 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 기반으로 기본 목록 구성 정보를 크롤링하도록 구성될 수 있다.

[0043] 크롤러(114)는 설정된 주기마다 각 웹 사이트에 접근하여 크롤링을 수행할 수 있다. 크롤러(114)에 의해 크롤링된 데이터는 데이터베이스(140)에 저장될 수 있다.

[0044] 상품의 기본 목록 구성 정보는 크롤러 모듈(110) 외에, 관리자 단말기(도시 생략)에 의해 수집될 수도 있다. 관리자 단말기는 다수개의 쇼핑몰 사이트(12, 14, 16) 별로 마이닝 블록, 속성 및 필드 등을 지정하고, URL 리스트 관리, 에러 로그 관리, 인증 관리 등을 수행할 수 있다.

[0045] 이때, 관리자는 관리자 단말기를 이용하여 쇼핑몰 사이트(12, 14, 16)에서 수집한 웹 페이지들의 구조를 기반으로, 마이닝 블록 지정, 속성(attribute) 지정, 필드(field) 지정, URL 리스트 관리, 에러 로그 관리, 인증 관리 등을 수행할 수 있다.

[0046] 관리자는 관리자 단말기를 이용하여 쇼핑몰 웹 사이트 리스트, 마이닝 영역, 수집 주기, 에러 로그(Error Log) 등을 그래픽 유저 인터페이스(GUI; Graphic User Interface) 형태로 관리할 수 있다. 이러한 관리자 단말기의 기능은 크롤링을 수행하는 서버에 통합될 수도 있다.

[0047] 도 3은 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치를 구성하는 크롤링 대상 분석기의 구성도이다. 도 1 내지 도 3을 참조하면, 크롤링 대상 분석기(112)는 쇼핑몰 사이트에서 웹 크롤링을 위한 크롤링 대상을 분석할 수 있다.

[0048] 크롤링 대상 분석기(112)는 마이닝 블록 지정부(1122), 속성 지정부(1124), 필드 지정부(1126), URL 리스트 관리부(1128), 에러 로그 관리부(1130), 인증 관리부(1132), 패턴 인식기(1134), 스크랩부(1136), 스케쥴러(1138), 에러 핸들러(1140), API 연동 모듈(1142) 및 데이터 저장부(1144)를 포함할 수 있다.

[0049] 마이닝 블록 지정부(1122)는 쇼핑몰 사이트(12, 14, 16)로부터 수집된 데이터에 대해 마이닝 블록(Mining Block)을 지정하기 위한 마이닝 블록 지정 UI(User Interface)를 포함할 수 있다.

[0050] 마이닝 블록 지정부(1122)는 쇼핑몰 사이트(12, 14, 16)의 웹 페이지 내에서 상품의 상품명과 썸네일 및 가격 등의 정보를 나타내는 다양한 블록들의 위치, 크기 등을 설정 또는 지정할 수 있다.

[0051] 속성 지정부(1124)는 쇼핑몰 사이트(12, 14, 16)로부터 수집된 데이터에 대해 속성(Attribution)을 지정하기 위한 속성 지정 UI를 포함할 수 있다. 속성 지정부(1124)는 마이닝 블록 지정부(1122)에 의해 지정된 마이닝 블록들의 속성을 설정 또는 지정할 수 있다.

[0052] 속성 지정부(1124)는 예를 들어, 쇼핑몰 사이트(12, 14, 16)의 웹 사이트 별로 상품의 상품명과 썸네일 및 가격 등의 기본 목록 구성 정보에 관한 속성을 설정 또는 지정할 수 있다.

[0053] 필드 지정부(1126)는 웹 사이트로부터 수집된 데이터에 대해 필드(Field)를 지정하기 위한 필드 지정 UI를 포함할 수 있다. 필드 지정부(1126)는 속성 지정부(1124)에 의해 속성이 지정된 마이닝 블록들의 필드를 지정할 수 있다.

[0054] 필드 지정부(1126)는 예를 들어, 쇼핑몰 웹 사이트 별로 상품의 기본 목록 구성 정보에 해당하는 상품명과 썸네일 및 가격 등에 대해 각각의 필드를 설정 또는 지정할 수 있다.

[0055] URL 리스트 관리부(1128)는 쇼핑몰 웹 사이트 별로 웹 페이지들의 URL(Uniform Resource Locator) 리스트를 관리할 수 있다. 에러 로그 관리부(1130)는 웹 사이트에서 데이터 크롤링 수집시의 오류 정보를 나타내는 에러 로그(Error Log)를 관리할 수 있다.

[0056] 인증 관리부(1132)는 웹 사이트 접속 및 데이터 지정/관리를 위한 인증(Authentication) 정보를 관리할 수 있다. 인증 관리부(1132)에서 관리하는 인증 정보는 관리자 단말기의 인증 정보, 크롤링을 수행하는 서버의 인증 정보 등을 포함할 수 있다.

[0057] 패턴 인식기(1134)는 각 웹 사이트 별로 문서 객체 모델의 패턴을 인식할 수 있다. 패턴 인식기(1134)는 웹 사이트 별로 지정된 마이닝 블록, 속성 및 필드를 기반으로 문서 객체 모델의 패턴을 인식할 수 있다.

[0058] 스크랩부(1136)는 웹 사이트 별로 파악된 문서 객체 모델의 패턴을 기반으로 필요로 하는 정보를 추출할 수 있다. 스크랩부(1136)는 문서 객체 모델의 패턴을 기반으로 웹 사이트의 웹 페이지를 스크래핑하는 스크래핑(scrapping) 모듈과, 웹페이지의 데이터를 파싱(parsing)하는 파서 모듈을 포함할 수 있다.

[0059] 스케쥴러(1138)는 웹 사이트의 데이터를 크롤링하기 위한 쓰레드(Thread) 관리 기능을 수행할 수 있다. 에러 핸들러(1140)는 웹 사이트의 데이터를 크롤링하는 과정에서 오류를 관리하는 에러 핸들링 기능을 수행할 수 있다.

[0060] API 연동 모듈(1142)은 웹 사이트의 데이터를 크롤링하기 위한 외부 API 연동 기능을 수행할 수 있다. 서버는 웹 사이트에서 응용 프로그램 인터페이스(API) 제공시 API 연동 모듈(1142)에 의해 API를 통한 데이터 조회를 수행할 수 있다.

[0061] 데이터 저장부(1144)는 크롤링 모듈(110)에 의해 크롤링된 데이터(상품 기본 목록 구성 정보), 에뮬레이터 모듈(120)에 의해 수집된 데이터(상품 상세 정보)를 데이터베이스(140)에 저장할 수 있다.

[0062] 다시 도 1을 참조하면, 데이터베이스(140)는 다양한 쇼핑몰 사이트(12, 14, 16)로부터 다양한 상품에 대해 각각 기본 목록 구성 정보를 크롤링하여 수집된 다양한 상품에 관한 상품 데이터를 저장하도록 구성될 수 있다.

[0063] 사용자 대상 사이트 생성 모듈(130)은 데이터베이스(140)에 저장된 상품 데이터로부터 다양한 쇼핑몰 사이트(12, 14, 16)가 통합된 사용자 대상 사이트를 생성하고, 사용자 디바이스(20)로부터 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청을 입력받도록 구성될 수 있다.

[0064] 사용자 대상 사이트 생성 모듈(130)은 다양한 대상 쇼핑몰 사이트(12, 14, 16)로부터 수집되는 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 사용자 대상 사이트에 표시하도록 구성될 수 있다.

[0065] 사용자 대상 사이트 생성 모듈(130)은 다양한 쇼핑몰 사이트(12, 14, 16)로부터 수집한 데이터를 통합하여 사용자 디바이스(20)에 일괄하여 제공하는 서비스를 제공할 수 있다. 이에 따라 사용자는 다양한 쇼핑몰 웹 사이트의 정보를 보다 빠르고 효율적으로 파악하여 상호 간에 비교, 분석을 수행할 수 있다.

[0066] 실시예에서, 사용자 대상 사이트 생성 모듈(130)은 다양한 웹 사이트의 관련된 필드들(동일하거나 유사한 필드들)을 하나의 필드명으로 통일하여 정보를 제공할 수 있다. 이때, 관련된 필드들은 다양한 웹 사이트의 필드들 간의 단어 유사도, 유의어 사전 등을 기반으로 분류될 수 있다.

[0067] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치는 다양한 웹 사이트의 데이터를 한번에 사용자에게 반환할 수 있다. 이에 따라 사용자들은 여러 사이트를 이동하지 않고 한번에 간편하게 데이터를 조회하여 다양한 웹 사이트의 데이터를 비교할 수 있으며, 원하는 데이터를 빠르고 편리하게 찾을 수 있다.

[0068] 또한, 사용자들이 크롤링 서버에 자체적으로 데이터를 업로드할 수 있도록 구성되는 경우, 웹 사이트의 데이터뿐 아니라 크롤링 서버에 자체적으로 업로드되는 데이터까지 한번에 보여줄 수도 있다.

[0069] 에뮬레이터 모듈(120)은 사용자 대상 사이트에 생성된 사용자 대상 상품에 대해 사용자 디바이스(20)로부터 상기 상세 정보 요청이 입력되면, 사용자 디바이스(20)를 통해 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 사용자 대상 상품의 상세 정보를 수집하여 데이터베이스(140)에 저장시키도록 구성될 수 있다.

[0070] 실시예에서, 에뮬레이터 모듈(120)은 사용자 디바이스(20)로 하여금 사용자 대상 상품에 해당하는 상품 데이터를 수집한 대상 쇼핑몰 사이트에 상세 정보를 조회하여 열람하게 할 수 있다.

[0071] 에뮬레이터 모듈(120)은 사용자 디바이스(20)에 응용 프로그램(application program)을 실행시켜 대상 쇼핑몰 사이트에 상품 상세 정보를 호출, 조회하게 하고, 사용자 디바이스(20)가 대상 쇼핑몰 사이트로부터 전달받은 상품 상세 정보를 수집하는 방법으로 대상 쇼핑몰 사이트의 상품 상세 정보를 크롤링할 수 있다.

[0072] 또한, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 대상 쇼핑몰 사이트를 열람한 사용자 대상 상품의 상세 정보를 사용자 디바이스(20)로부터 전달받을 수 있다.

[0073] 에뮬레이터 모듈(120)은 사용자 디바이스(20)로부터 전달받은 상품의 상세 정보를 사용자 대상 상품에 관한 상품 데이터에 포함시켜 데이터베이스(140)에 저장시키도록 구성될 수 있다.

[0074] 에뮬레이터 모듈(120)은 사용자 디바이스로부터 사용자 대상 상품에 대해 상세 정보 요청이 입력되면, 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있는지 판단할 수 있다.

[0075] 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있는 경우, 에뮬레이터 모듈(120)은 데이터베이스(140)에 저장된 상세 정보를 사용자 대상 사이트를 통해 사용자 디바이스(20)로 제공할 수 있다.

[0076] 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있지 않은 경우, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 응용 프로그램 어플리케이션을 실행하여 대상 쇼핑몰 사이트에 접근하게 할 수 있다.

[0077] 그리고 나서, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 실행되는 응용 프로그램 어플리케이션에 의해 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 데이터베이스(140)에 저장하도록 구성될 수 있다.

[0078] 에뮬레이터 모듈(120)은 다수의 가상머신을 통한 병렬 처리를 통해 상품 상세 정보를 수집할 수 있으며 크롤링 처리 속도를 높일 수 있다. 또한, 사용자 디바이스와 유사한 환경인 가벼운 모바일 OS를 가상환경으로 병렬 구동시켜 빠른 속도를 확보할 수 있으며 비용 절감이 가능하다.

[0079] 또한, 에뮬레이터 모듈(120)은 병렬 처리로 분산 호출(call)이 이루어지므로 크롤링 속도 측면에서 우수하고, API를 호출하는 형태를 유지하여 크롤링 대상 서비스 시스템 측에 일반 유저가 상세 페이지를 조회하는 것과 큰

차이 없는 수준으로 부하를 경감시킬 수 있다.

[0080] 실시예에서, 에뮬레이터 모듈(120)은 대상 쇼핑몰 사이트(12, 14, 16)에 포함된 상품들 중, 상세 정보 요청이 입력된 사용자 대상 상품의 선호도 순위(우선 순위)를 판단하고, 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단할 수 있다.

[0081] 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는 것으로 판단되면, 에뮬레이터 모듈(120)은 사용자 디바이스(20)를 통해 대상 쇼핑몰 사이트(12, 14, 16)에서 사용자 대상 상품의 상세 정보를 수집할 수 있다.

[0082] 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 에뮬레이터 모듈(120)은 사용자 디바이스(20)를 통해 사용자 대상 상품의 상세 정보를 수집하지 않도록 구성될 수 있다.

[0083] 사용자 대상 상품의 선호도 순위는 예를 들어, 각 쇼핑몰 사이트의 상품 인기 순위, 상품 노출 순서, 크롤링 서버에서 산출하거나 예측한 상품 인기 순위 등을 기반으로 산출될 수 있으나, 이에 한정되는 것은 아니다.

[0084] 일반적으로 웹 사이트에 불특정 상대가 웹 페이지를 스크랩하지 못하게 막는 기능이 구비되어 있는 경우, 관리자 단말기 또는 서버가 웹 사이트에서 데이터를 수집하기 위해 웹 사이트에 반복적으로 접근 시도시 웹 사이트로의 접근이 쉽게 차단되어 웹 사이트의 데이터를 크롤링하지 못하게 될 수 있다.

[0085] 그러나, 본 발명의 실시예에 의하면, 서버가 사용자 디바이스로 하여금 웹 사이트에서 데이터를 호출하게 하여, 웹 사이트의 데이터를 간접적으로 수집하여 크롤링하도록 구성되므로, 웹 사이트에 의해 서버의 크롤링이 제한되는 것을 방지할 수 있다.

[0086] 즉, 본 발명의 실시예는 크롤링 대상에 대한 정보를 수집할 때, 사용자가 조회할 지 여부를 알 수 없는 상세 정보 대신 기본 목록 구성에 필요한 정보만 크롤링하여 신규 목록을 유지하면서, 많이 조회되는 상품의 특성을 분석하여 우선 순위가 높다고 판단되는 상품 정보를 중심으로 상세 정보를 수집한다.

[0087] 따라서, 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치에 의하면, 사용자 상품 접근 데이터를 활용하여 상품 선호도(우선 순위) 기반으로 효율적으로 상품 상세 정보를 수집함으로써 일부 데이터 수집으로 전체 데이터 수집에 가까운 효과를 발휘할 수 있다.

[0088] 또한, 본 발명의 실시예에 의하면, 다양한 상품들의 기본 목록 구성을 유지하여 사용자의 선택의 폭을 넓히면서도, 서버의 크롤링 부담을 줄일 수 있고, 서버의 잦은 접근에 따라 쇼핑몰 사이트로부터 웹 페이지 접근이 차단되는 것을 방지할 수 있다.

[0089] 따라서, 본 발명의 실시예에 의하면, 사용자가 상품의 상세 정보에 접근했을 때 사용자 디바이스를 통해 타겟 서비스 사이트의 추가 정보를 수집하여 크롤링 주체 시스템에 전달하고 상품 접근시 활용함으로써, 사용자 디바이스 자원을 사용하여 중앙 시스템의 부하를 감소하고, 타겟 서비스 시스템 부하 또한 줄일 수 있다.

[0090] 또한, 본 발명의 실시예에 의하면, 사용자 디바이스(20)의 요청에 따라 상품 상세 정보를 수집하므로, 이 과정에서 사용자 디바이스(20)의 상품 상세 정보 요청 히스토리를 수집할 수 있다.

[0091] 따라서, 이러한 사용자 디바이스(20)의 상품 상세 정보 요청 히스토리를 기반으로 사용자의 상품 선택 성향 데이터를 분석하여 사용자 대상 사이트로 피드백함으로써 사용자에 적합한 상품 정보를 우선적으로 추천하는 것도 가능하다.

[0092] 도 4는 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법의 순서도이다. 이하에서 도 1 및 도 4를 참조하여 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법에 대해 설명하다.

[0093] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법은 크롤링 대상 쇼핑몰(10)의 상품의 기본 목록 구성 정보를 사용자 대상 사이트에 통합하여 제공하고, 상품들 중 사용자에 의해 요청된 상품이 상세 정보를 사용자 디바이스(20)를 통해 수집하기 위한 것이다.

[0094] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법은 사용자 대상 사이트를 통해 다수의 사용자에게 상품 데이터를 제공하고, 다수의 사용자 디바이스(20)를 통해 상품의 상세 정보를 수집할 수 있다.

[0095] 크롤링 모듈(110)은 크롤링 대상 쇼핑몰(10)에 해당하는 다양한 쇼핑몰 사이트(12, 14, 16)에서 상품의 상품명과 썸네일 및 가격을 포함하는 기본 목록 구성 정보와, 상품의 상세 정보(예를 들어, 사용자 대상 상품의 상품 상세 이미지, 크기, 중량, 재질, 색상 및 부가 설명 등) 중 기본 목록 구성 정보만 크롤링할 수 있다(S110).

[0096] 이때, 크롤링 모듈(110)은 쇼핑몰 사이트 별로 문서 객체 모델의 패턴과, 쇼핑몰 사이트에서 사용자 디바이스(20)에 제공할 크롤링 대상 블록을 분석하여 기본 목록 구성 정보를 크롤링할 수 있다.

[0097] 도 5는 도 4의 단계 S110을 나타낸 순서도이다. 도 1, 도 4 및 도 5를 참조하면, 크롤링 모듈(110)은 다양한 쇼핑몰 사이트(12, 14, 16) 별로 기본 목록 구성 정보의 배열 패턴을 분석하고(S112), 배열 패턴을 기반으로 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 지정할 수 있다(S114).

[0098] 크롤링 모듈(110)은 다양한 쇼핑몰 사이트(12, 14, 16) 별로 문서 객체 모델의 패턴을 인식하여, 쇼핑몰 사이트(12, 14, 16) 별로 인식한 문서 객체 모델 패턴을 기반으로 쇼핑몰 사이트(12, 14, 16)의 웹 페이지에서 필요로 하는 정보를 추출할 수 있다. 크롤링 모듈(110)은 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 기반으로 기본 목록 구성 정보를 크롤링할 수 있다(S116).

[0099] 다시 도 1 및 도 4를 참조하면, 다양한 쇼핑몰 사이트(12, 14, 16)로부터 다양한 상품에 대해 각각 기본 목록 구성 정보를 크롤링하여 수집된 다양한 상품에 관한 상품 데이터는 데이터베이스(140)에 저장될 수 있다(S120).

[0100] 사용자 대상 사이트 생성 모듈(130)은 데이터베이스(140)에 저장된 상품 데이터로부터 다양한 쇼핑몰 사이트(12, 14, 16)가 통합된 사용자 대상 사이트를 생성할 수 있다(S130).

[0101] 사용자 대상 사이트 생성 모듈(130)은 다양한 대상 쇼핑몰 사이트(12, 14, 16)로부터 수집되는 다양한 상품에 대한 상품 데이터를 지정된 객체 패턴으로 통합하여 사용자 대상 사이트에 표시하도록 구성될 수 있다.

[0102] 사용자 대상 사이트 생성 모듈(130)은 다양한 쇼핑몰 사이트(12, 14, 16)로부터 수집한 데이터를 통합하여 사용자 디바이스(20)에 일괄하여 제공하는 서비스를 제공할 수 있다. 이에 따라 사용자는 다양한 쇼핑몰 웹 사이트의 정보를 보다 빠르고 효율적으로 파악하여 상호 간에 비교, 분석을 수행할 수 있다.

[0103] 실시예에서, 사용자 대상 사이트 생성 모듈(130)은 다양한 웹 사이트의 관련된 필드들(동일하거나 유사한 필드들)을 하나의 필드명으로 통일하여 정보를 제공할 수 있다. 이때, 관련된 필드들은 다양한 웹 사이트의 필드들 간의 단어 유사도, 유의어 사전 등을 기반으로 분류될 수 있다.

[0104] 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 장치는 다양한 웹 사이트의 데이터를 한번에 사용자에게 반환할 수 있으며, 사용자들은 여러 사이트들을 이동하지 않고 한번에 데이터를 간편하게 조회하여 여러 웹사이트의 데이터를 비교할 수 있으며, 원하는 데이터를 빠르고 편리하게 찾을 수 있다.

[0105] 또한, 사용자들이 서버에 자체적으로 데이터를 업로드할 수 있도록 구성되는 경우, 웹 사이트의 데이터 뿐 아니라 서버에 자체적으로 업로드되는 데이터까지 한번에 보여줄 수도 있다.

[0106] 사용자 디바이스(20)로부터 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청이 입력되면(S140), 에뮬레이터 모듈(120)은 사용자 디바이스(20)를 통해 대상 쇼핑몰 사이트에서 사용자 대상 상품의 상세 정보를 수집하여 데이터베이스(140)에 저장할 수 있다(S150).

[0107] 도 6은 도 4의 단계 S150을 나타낸 순서도이다. 도 1, 도 4 및 도 6을 참조하면, 에뮬레이터 모듈(120)은 사용자 디바이스(20)로 하여금 사용자 대상 상품에 해당하는 상품 데이터를 수집한 대상 쇼핑몰 사이트에 상세 정보를 열람하게 할 수 있다(S152).

[0108] 에뮬레이터 모듈(120)은 사용자 디바이스(20)에 응용 프로그램(application program)을 실행시켜 대상 쇼핑몰 사이트에 상품 상세 정보를 호출, 조회하게 하고, 사용자 디바이스(20)가 대상 쇼핑몰 사이트로부터 전달받은 상품 상세 정보를 수집하는 방법으로 대상 쇼핑몰 사이트의 상품 상세 정보를 크롤링할 수 있다.

[0109] 또한, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 대상 쇼핑몰 사이트를 열람한 사용자 대상 상품의 상세 정보를 사용자 디바이스(20)로부터 전달받을 수 있다(S154).

[0110] 에뮬레이터 모듈(120)은 사용자 디바이스(20)로부터 전달받은 상품의 상세 정보를 사용자 대상 상품에 관한 상품 데이터에 포함시켜 데이터베이스(140)에 저장할 수 있다(S156).

[0111] 도 7은 도 4의 단계 S150의 다른 예를 나타낸 순서도이다. 도 1, 도 4 및 도 7을 참조하면, 에뮬레이터 모듈(120)은 사용자 디바이스로부터 사용자 대상 상품에 대해 상세 정보 요청이 입력되면, 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있는지 판단할 수 있다(S162). 이때 서버는 URL 정보를 통해 사용자 디바이스가 요청한 정보가 서버에 저장되어 있는지 여부를 판단할 수 있다.

[0112] 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있는 경우, 에뮬레이터 모듈(120)은 데이터베이

스(140)에 저장된 상세 정보를 사용자 대상 사이트를 통해 사용자 디바이스(20)로 제공할 수 있다(S164, S166).

[0113] 사용자 대상 상품의 상세 정보가 데이터베이스(140)에 저장되어 있지 않은 경우, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 응용 프로그램 어플리케이션을 실행하여 대상 쇼핑몰 사이트에 접근하게 할 수 있다(S164, S168).

[0114] 그리고 나서, 에뮬레이터 모듈(120)은 사용자 디바이스(20)에서 실행되는 응용 프로그램 어플리케이션에 의해 대상 쇼핑몰 사이트로부터 상세 정보를 수집하여 데이터베이스(140)에 저장할 수 있다(S170).

[0115] 도 8은 도 4의 단계 S150의 또 다른 예를 나타낸 순서도이다. 도 1, 도 4 및 도 8을 참조하면, 에뮬레이터 모듈(120)은 대상 쇼핑몰 사이트(12, 14, 16)에 포함된 상품들 중, 상세 정보 요청이 입력된 사용자 대상 상품의 선호도 순위(우선 순위)를 판단하고(S172), 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는지 판단할 수 있다(S174).

[0116] 사용자 대상 상품의 선호도 순위가 기 설정된 순위 조건을 만족하는 것으로 판단되면, 에뮬레이터 모듈(120)은 사용자 디바이스(20)를 통해 대상 쇼핑몰 사이트(12, 14, 16)에서 사용자 대상 상품의 상세 정보를 수집할 수 있다(S174, S176).

[0117] 사용자 대상 상품의 선호도 순위가 상기 기 설정된 순위 조건을 만족하지 않으면, 에뮬레이터 모듈(120)은 사용자 디바이스(20)를 통해 사용자 대상 상품의 상세 정보를 수집하지 않도록 구성될 수 있다(S174, S178).

[0118] 본 발명의 실시예에 의하면, 서버가 사용자 디바이스로 하여금 웹 사이트에서 데이터를 호출하게 하여, 웹 사이트의 데이터를 간접적으로 수집하여 크롤링하도록 구성되므로, 웹 사이트에 의해 서버의 크롤링이 제한되는 것을 방지할 수 있다.

[0119] 본 발명의 실시예에 따른 상품 데이터 크롤링 방법은 크롤링 대상에 대한 정보를 수집할 때, 사용자가 조회할지 여부를 알 수 없는 상세 정보 대신 기본 목록 구성에 필요한 정보만 크롤링하여 신규 목록을 유지하면서, 많이 조회되는 상품의 특성을 분석하여 우선 순위가 높다고 판단되는 상품 정보를 중심으로 상세 정보를 수집한다.

[0120] 따라서, 본 발명의 실시예에 따른 에뮬레이터를 활용한 상품 데이터 크롤링 방법에 의하면, 사용자 상품 접근 데이터를 활용하여 상품 선호도(우선 순위) 기반으로 효율적으로 상품 상세 정보를 수집함으로써 일부 데이터 수집으로 전체 데이터 수집에 가까운 효과를 발휘할 수 있다.

[0121] 또한, 본 발명의 실시예에 의하면, 다양한 상품들의 기본 목록 구성을 유지하여 사용자의 선택의 폭을 넓히면서도, 서버의 크롤링 부담을 줄일 수 있고, 서버의 잦은 접근에 따라 쇼핑몰 사이트로부터 웹 페이지 접근이 차단되는 것을 방지할 수 있다.

[0122] 따라서, 본 발명의 실시예에 의하면, 사용자가 상품의 상세 정보에 접근했을 때 사용자 디바이스를 통해 타겟 서비스 사이트의 추가 정보를 수집하여 크롤링 주체 시스템에 전달하고 상품 접근시 활용함으로써, 사용자 디바이스 자원을 사용하여 중앙 시스템의 부하를 감소하고, 타겟 서비스 시스템 부하 또한 줄일 수 있다.

[0123] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/ 또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPGA(Field Programmable Gate Array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다.

[0124] 처리 장치는 운영 체제 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리요소를 포함할 수 있음을 이해할 것이다.

[0125] 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor) 와 같은, 다른 처리 구성(Processing configuration)도 가능하다. 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의

조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다.

[0126] 소프트웨어 및/ 또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody) 될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0127] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.

[0128] 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDROM, DVD와 같은 광기록 매체(optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0129] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 그러므로, 다른 구현들, 다른 실시예들 및 청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

부호의 설명

[0130] 10: 크롤링 대상 쇼핑몰

12, 14, 16: 쇼핑몰 사이트

20: 사용자 디바이스

100: 에뮬레이터를 활용한 상품 데이터 크롤링 장치

110: 크롤링 모듈

112: 크롤링 대상 분석기

114: 크롤러

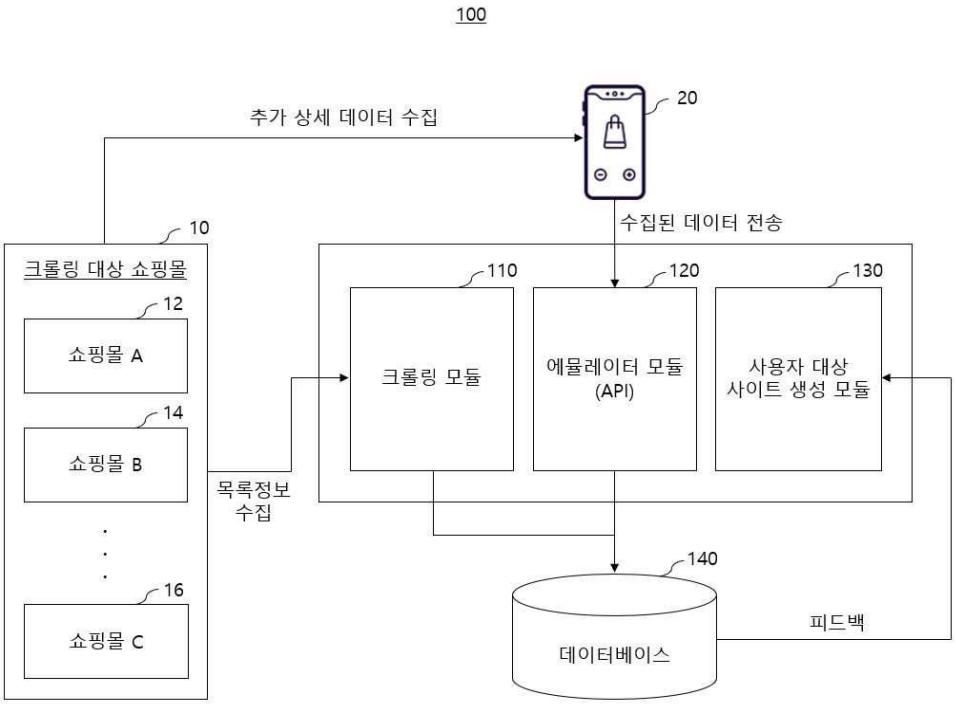
120: 에뮬레이터 모듈

130: 사용자 대상 사이트 생성 모듈

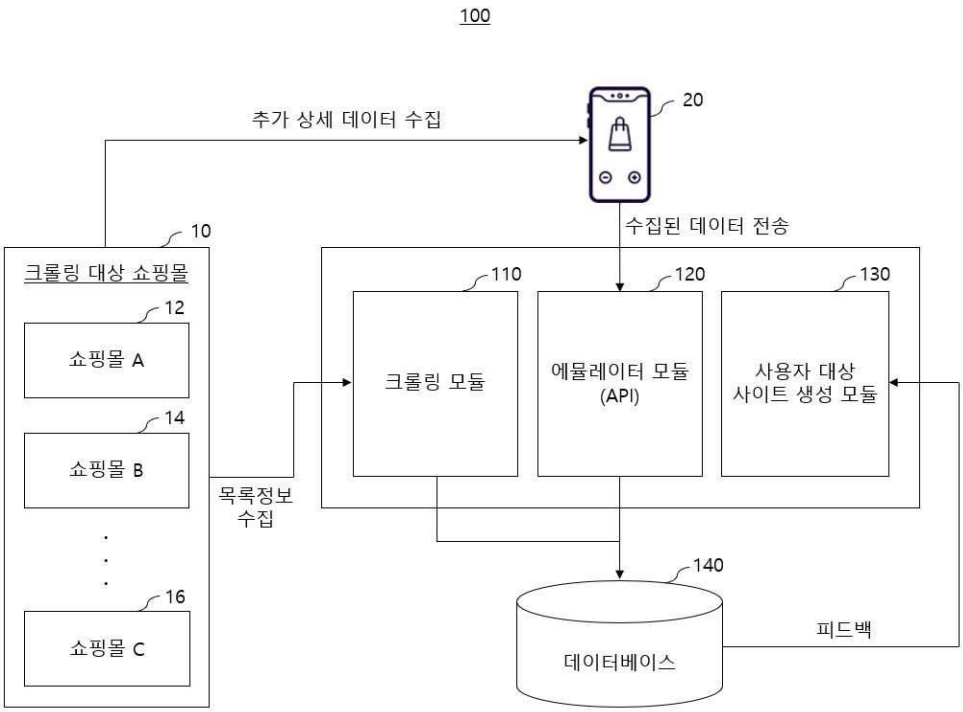
140: 데이터베이스

도면

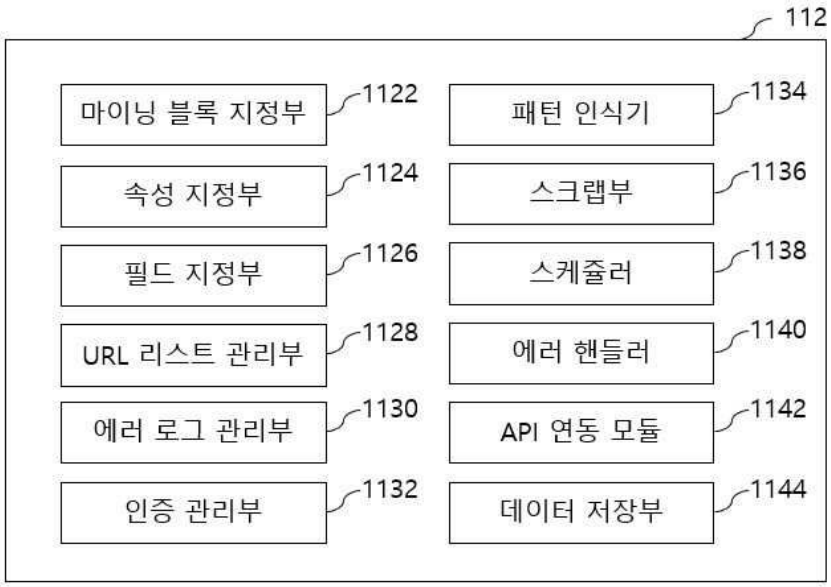
도면1


100
추가 상세 데이터 수집
20
수집된 데이터 전송
10
크롤링 대상 쇼핑몰
12
쇼핑몰 A
14
쇼핑몰 B
16
쇼핑몰 C
목록정보 수집
110
크롤링 모듈
120
에뮬레이터 모듈 (API)
130
사용자 대상 사이트 생성 모듈
140
데이터베이스
피드백

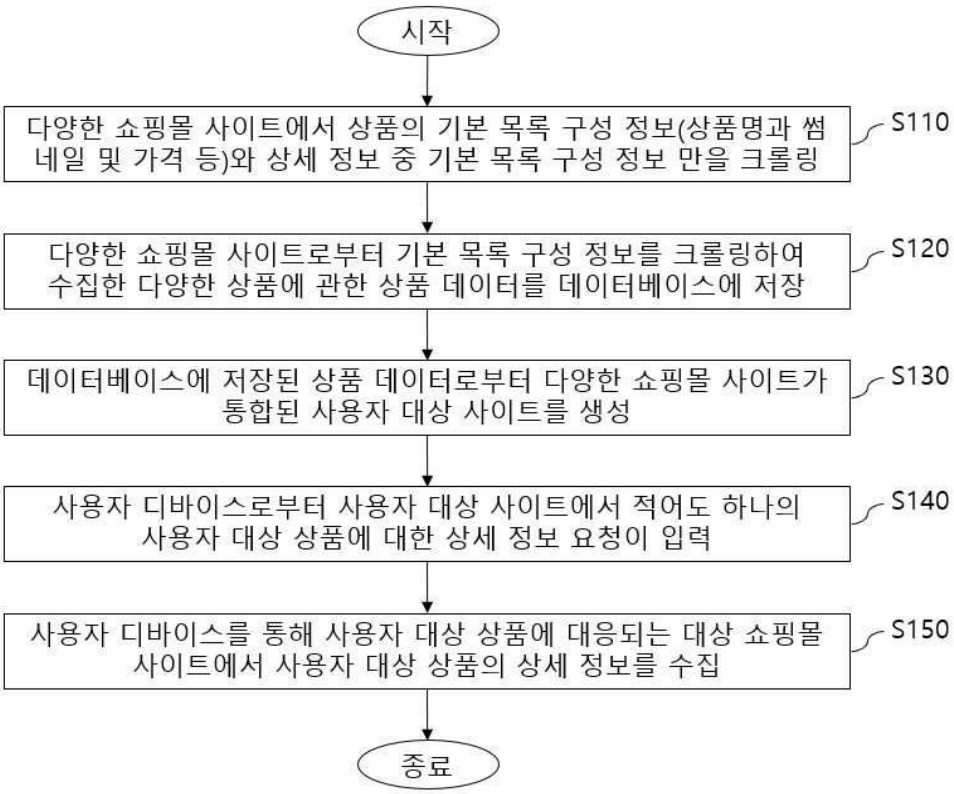
도면2


100
추가 상세 데이터 수집
20
수집된 데이터 전송
10
크롤링 대상 쇼핑몰
12
쇼핑몰 A
14
쇼핑몰 B
16
쇼핑몰 C
목록정보 수집
110
크롤링 모듈
120
에뮬레이터 모듈 (API)
130
사용자 대상 사이트 생성 모듈
140
데이터베이스
피드백

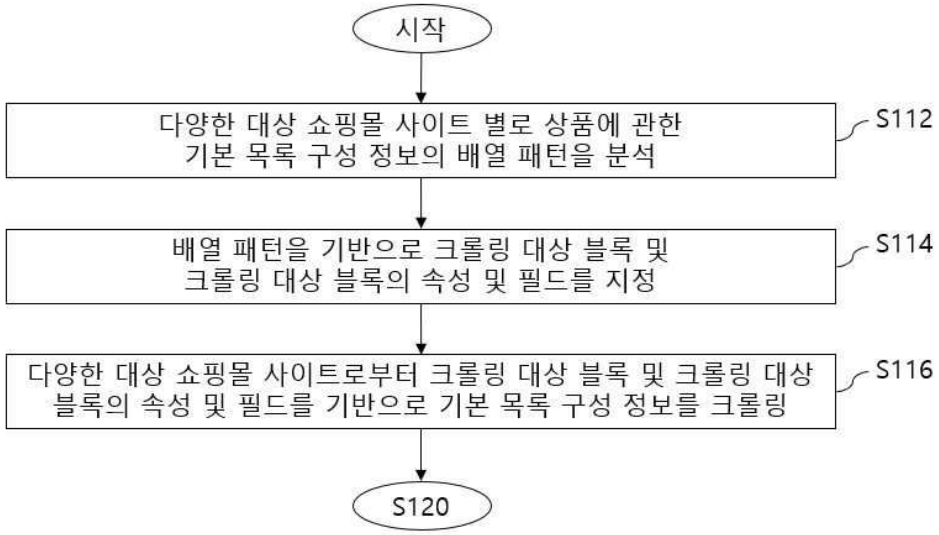
도면3


112
마이닝 블록 지정부
1122
속성 지정부
1124
필드 지정부
1126
URL 리스트 관리부
1128
에러 로그 관리부
1130
인증 관리부
1132
패턴 인식기
1134
스크랩부
1136
스케줄러
1138
에러 핸들러
1140
API 연동 모듈
1142
데이터 저장부
1144

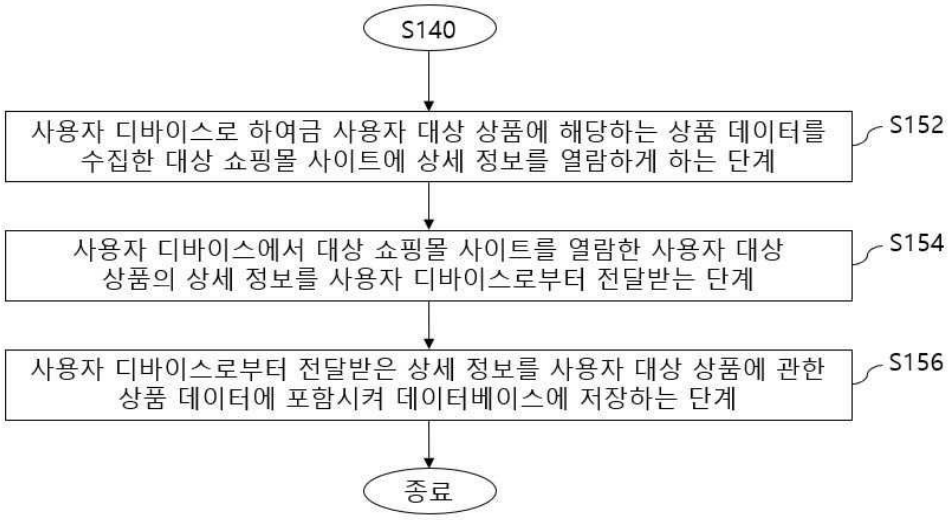
도면4


시작
다양한 쇼핑몰 사이트에서 상품의 기본 목록 구성 정보(상품명과 썸네일 및 가격 등)와 상세 정보 중 기본 목록 구성 정보 만을 크롤링
S110
다양한 쇼핑몰 사이트로부터 기본 목록 구성 정보를 크롤링하여 수집한 다양한 상품에 관한 상품 데이터를 데이터베이스에 저장
S120
데이터베이스에 저장된 상품 데이터로부터 다양한 쇼핑몰 사이트가 통합된 사용자 대상 사이트를 생성
S130
사용자 디바이스로부터 사용자 대상 사이트에서 적어도 하나의 사용자 대상 상품에 대한 상세 정보 요청이 입력
S140
사용자 디바이스를 통해 사용자 대상 상품에 대응되는 대상 쇼핑몰 사이트에서 사용자 대상 상품의 상세 정보를 수집
S150
종료

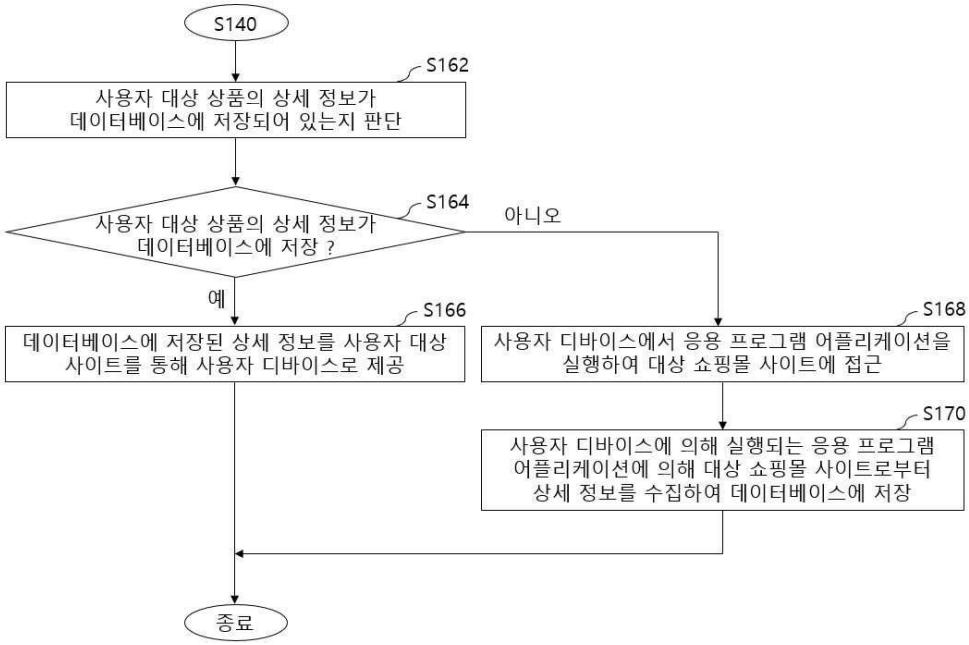
도면5


시작
다양한 대상 쇼핑몰 사이트 별로 상품에 관한 기본 목록 구성 정보의 배열 패턴을 분석
S112
배열 패턴을 기반으로 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 지정
S114
다양한 대상 쇼핑몰 사이트로부터 크롤링 대상 블록 및 크롤링 대상 블록의 속성 및 필드를 기반으로 기본 목록 구성 정보를 크롤링
S116
S120

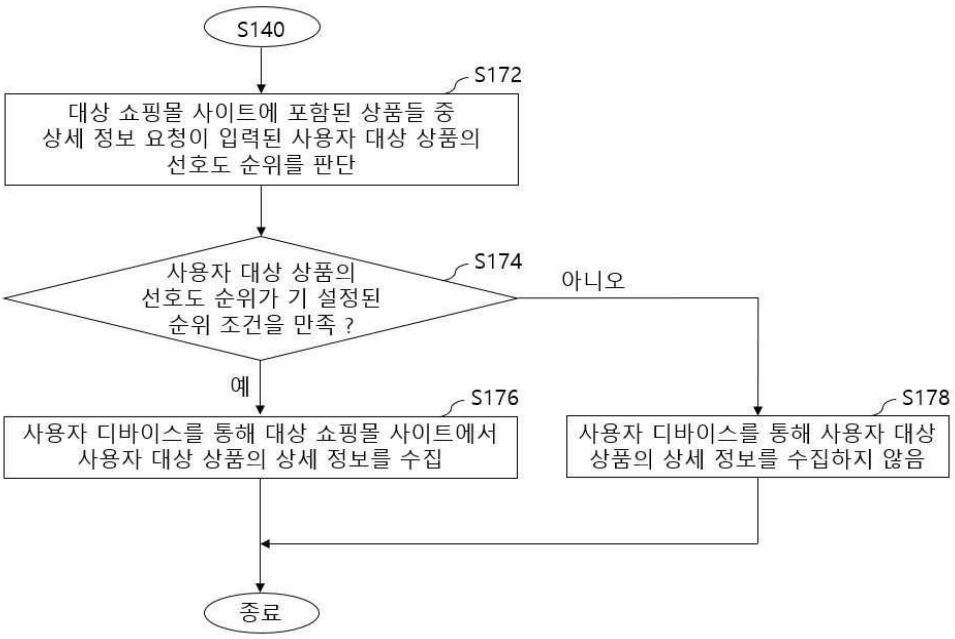
도면6


S140
사용자 디바이스로 하여금 사용자 대상 상품에 해당하는 상품 데이터를 수집한 대상 쇼핑몰 사이트에 상세 정보를 열람하게 하는 단계
S152
사용자 디바이스에서 대상 쇼핑몰 사이트를 열람한 사용자 대상 상품의 상세 정보를 사용자 디바이스로부터 전달받는 단계
S154
사용자 디바이스로부터 전달받은 상세 정보를 사용자 대상 상품에 관한 상품 데이터에 포함시켜 데이터베이스에 저장하는 단계
S156
종료

도면7


S140
S162
사용자 대상 상품의 상세 정보가
데이터베이스에 저장되어 있는지 판단
S164
사용자 대상 상품의 상세 정보가
데이터베이스에 저장 ?
아니오
예
S166
데이터베이스에 저장된 상세 정보를 사용자 대상
사이트를 통해 사용자 디바이스로 제공
S168
사용자 디바이스에서 응용 프로그램 어플리케이션을
실행하여 대상 쇼핑몰 사이트에 접근
S170
사용자 디바이스에 의해 실행되는 응용 프로그램
어플리케이션에 의해 대상 쇼핑몰 사이트로부터
상세 정보를 수집하여 데이터베이스에 저장
종료

도면8


S140
S172
대상 쇼핑몰 사이트에 포함된 상품들 중
상세 정보 요청이 입력된 사용자 대상 상품의
선호도 순위를 판단
S174
사용자 대상 상품의
선호도 순위가 기 설정된
순위 조건을 만족 ?
아니오
예
S176
사용자 디바이스를 통해 대상 쇼핑몰 사이트에서
사용자 대상 상품의 상세 정보를 수집
S178
사용자 디바이스를 통해 사용자 대상
상품의 상세 정보를 수집하지 않음
종료