



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0014979
(43) 공개일자 2022년02월08일

(51) 국제특허분류(Int. Cl.)
G06Q 30/02 (2012.01) G06F 40/205 (2020.01)
G06F 40/268 (2020.01)
(52) CPC특허분류
G06Q 30/0248 (2013.01)
G06F 40/205 (2020.01)
(21) 출원번호 10-2020-0094899
(22) 출원일자 2020년07월30일
심사청구일자 2020년07월30일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
김영갑
서울특별시 광진구 능동로 209, 세종대학교 701호
오세라
서울특별시 광진구 능동로 209, 세종대학교 732호
(74) 대리인
특허법인이상

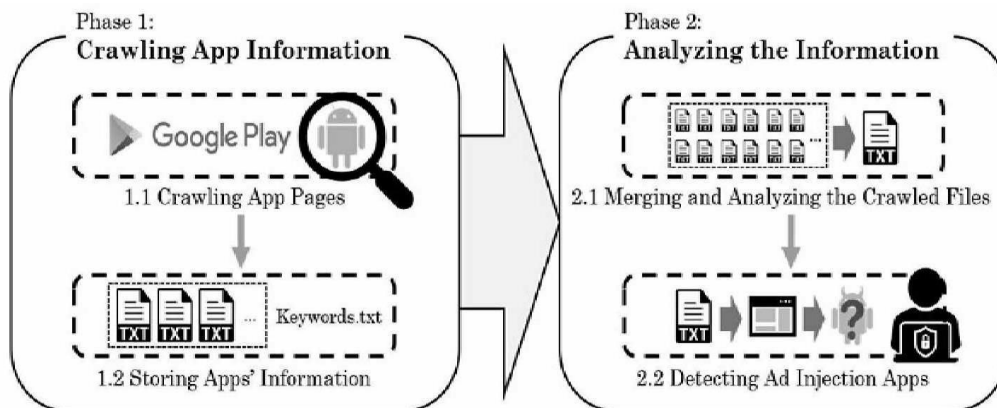
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 광고 인젝션 앱 검출 방법 및 장치

(57) 요약

본 발명은 모바일 앱의 정보를 크롤링하는 단계; 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계; 상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하는 단계; 및 상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하는 단계를 포함하는, 광고 인젝션 앱 검출 방법을 개시한다.

대표도



(52) CPC특허분류

G06F 40/268 (2020.01)

G06F 40/279 (2020.01)

G06Q 30/0278 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103210
과제번호	2019-0-00231-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보보호핵심원천기술개발(R&D)
연구과제명	공공 인프라 안전을 위한 인공지능 기반 영상보안 기술 및 시스템 개발
기 여 율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

모바일 앱의 정보를 크롤링하는 단계;

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계;

상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하는 단계; 및

상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하는 단계를 포함하는, 광고 인젝션 앱 검출 방법.

청구항 2

청구항 1에 있어서,

상기 모바일 앱의 정보를 크롤링하는 단계는,

특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하는 단계를 포함하는, 광고 인젝션 앱 검출 방법.

청구항 3

청구항 1에 있어서,

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는,

정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하는 단계를 포함하는, 광고 인젝션 앱 검출 방법.

청구항 4

청구항 3에 있어서,

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는,

학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하는 단계를 더 포함하는, 광고 인젝션 앱 검출 방법.

청구항 5

청구항 4에 있어서,

상기 학습된 단어 빈도수 테이블은,

상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습되는, 광고 인젝션 앱 검출 방법.

청구항 6

청구항 4에 있어서,

상기 나이트 베이지안 분류기는,

사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류하는, 광고 인젝션 앱 검출 방법.

청구항 7

청구항 4에 있어서,

상기 나이브 베이지안 분류기는,

광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여하는, 광고 인젝션 앱 검출 방법.

청구항 8

청구항 1에 있어서,

상기 광고 인젝션 점수는,

나이브 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출되는, 광고 인젝션 앱 검출 방법.

청구항 9

프로세서; 및

상기 프로세서를 통해 실현되는 적어도 하나의 명령을 저장하는 메모리를 포함하고,

상기 적어도 하나의 명령은,

모바일 앱의 정보를 크롤링하도록 하는 명령;

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령;

상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하도록 하는 명령; 및

상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하도록 하는 명령을 포함하는, 광고 인젝션 앱 검출 장치.

청구항 10

청구항 9에 있어서,

상기 모바일 앱의 정보를 크롤링하도록 하는 명령은,

특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하도록 하는 명령을 포함하는, 광고 인젝션 앱 검출 장치.

청구항 11

청구항 9에 있어서,

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은,

정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하도록 하는 명령을 포함하는, 광고 인젝션 앱 검출 장치.

청구항 12

청구항 11에 있어서,

상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은,

학습된 단어 빈도수 테이블을 활용하는 나이브 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하도록 하는 명령을 더 포함하는, 광고 인젝션 앱 검출 장치.

청구항 13

청구항 12에 있어서,

상기 학습된 단어 빈도수 테이블은,

상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습되는, 광고 인젝션

앱 검출 장치.

청구항 14

청구항 12에 있어서,

상기 나이브 베이지안 분류기는,

사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류하는, 광고 인젝션 앱 검출 장치.

청구항 15

청구항 12에 있어서,

상기 나이브 베이지안 분류기는,

광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여하는, 광고 인젝션 앱 검출 장치.

청구항 16

청구항 9에 있어서,

상기 광고 인젝션 점수는,

나이브 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출되는, 광고 인젝션 앱 검출 장치.

발명의 설명

기술 분야

[0001] 본 발명은 광고 인젝션 앱을 검출하는 방법 및 장치에 관한 것으로, 더욱 상세하게는 사용자 리뷰에 기반하여 광고 인젝션 앱을 검출하는 방법 및 장치에 관한 것이다.

배경 기술

[0002] 모바일 생태계가 성장하면서 PC 위주였던 광고 시장이 모바일로 이동하고 있고, 모바일 환경을 대상으로 하는 보안 위협이 늘어나고 있다. 모바일 광고 인젝션은 일반적으로 정당한 광고주의 광고가 아닌 타 광고를 모바일 앱에 노출 시키는 것을 말하고, 모바일 광고 인젝션에 의해 광고주는 광고에 대한 기대효과를 얻지 못하여 광고 업체에 대해 신뢰를 잃게 되고, 광고를 게시하는 콘텐츠 제공자는 정당한 수입을 얻지 못하는 문제가 발생한다. 또한, 광고 인젝션이 수행되는 모바일 앱을 사용하는 이용자는 무분별한 광고 노출로 불편함을 겪을 수 있다.

[0003] 광고 인젝션은 모바일 생태계를 지지하는 광고 시장을 직접적으로 위협하는 공격이지만 아직까지 모바일 광고 인젝션에 대해 수행된 연구가 부족하며, 현재 광고 인젝션 모바일 앱을 검출하기 위한 방법이나 도구의 개발도 연구가 미진하다.

발명의 내용

해결하려는 과제

[0004] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은 광고 인젝션 앱을 검출하는 방법을 제공하는 데 있다.

[0005] 상기와 같은 문제점을 해결하기 위한 본 발명의 다른 목적은 모바일 앱의 사용자 리뷰에 기반하여 광고 인젝션 앱을 검출하는 방법을 제공하는 데 있다.

과제의 해결 수단

[0006] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법은, 모바일 앱의 정보를 크롤링하는 단계; 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계; 상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하는 단계; 및 상기 광고 인젝션 점수를 토대로

광고 인젝션 앱을 검출하는 단계를 포함할 수 있다.

- [0007] 여기서, 상기 모바일 앱의 정보를 크롤링하는 단계는, 특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하는 단계를 포함할 수 있다.
- [0008] 한편, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는, 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하는 단계를 추출하는 단계를 포함할 수 있다.
- [0009] 여기서, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는, 학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하는 단계를 더 포함할 수 있다.
- [0010] 한편, 상기 학습된 단어 빈도수 테이블은, 상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습될 수 있다.
- [0011] 또한, 상기 나이트 베이지안 분류기는, 사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류할 수 있다.
- [0012] 또한, 상기 나이트 베이지안 분류기는, 광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여할 수 있다.
- [0013] 또한, 상기 광고 인젝션 점수는, 나이트 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출될 수 있다.
- [0015] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 장치는, 프로세서; 및 상기 프로세서를 통해 실행되는 적어도 하나의 명령을 저장하는 메모리를 포함하고, 상기 적어도 하나의 명령은, 모바일 앱의 정보를 크롤링하도록 하는 명령; 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령; 상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하도록 하는 명령; 및 상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하도록 하는 명령을 포함할 수 있다.
- [0016] 여기서, 상기 모바일 앱의 정보를 크롤링하도록 하는 명령은, 특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하도록 하는 명령을 포함할 수 있다.
- [0017] 한편, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은, 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하도록 하는 명령을 포함할 수 있다.
- [0018] 여기서, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은, 학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하도록 하는 명령을 더 포함할 수 있다.
- [0019] 한편, 상기 학습된 단어 빈도수 테이블은, 상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습될 수 있다.
- [0020] 또한, 상기 나이트 베이지안 분류기는, 사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류할 수 있다.
- [0021] 또한, 상기 나이트 베이지안 분류기는, 광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여할 수 있다.
- [0022] 또한, 상기 광고 인젝션 점수는, 나이트 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출될 수 있다.

발명의 효과

- [0023] 본 발명의 일 실시예에 따르면, 광고 인젝션 점수를 통해 광고 인젝션 앱을 검출하는 장점을 가진다.
- [0024] 본 발명의 일 실시예에 따르면, 나이트 베이지안 분류기를 통해 모바일 앱의 사용자 리뷰를 분석하는 장점을 가진다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법의 개념도이다.
- 도 2는 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법의 동작 순서도이다.
- 도 3은 모바일 앱의 정보를 크롤링하는 방법의 개념도이다.
- 도 4는 단어 빈도수 테이블을 학습시키는 방법의 개념도이다.
- 도 5는 단어 빈도수 테이블을 학습시키기 위한 사용자 리뷰의 예시도이다.
- 도 6a는 학습된 단어 빈도수 테이블의 제 1 예시도이다.
- 도 6b는 학습된 단어 빈도수 테이블의 제 2 예시도이다.
- 도 7은 나이트 베이지안 분류기를 통해 사용자 리뷰를 분석한 결과의 예시도이다.
- 도 8은 광고 인젝션 점수가 부여된 모바일 앱들의 예시도이다.
- 도 9는 모바일 앱 상세 정보의 예시도이다.
- 도 10은 APK(Android Application Package) 분석창의 예시도이다.
- 도 11은 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 장치의 블록 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 본 발명은 다양한 변경을 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 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [0027] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는 데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는"이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0028] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0029] 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.
- [0030] 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 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0032] 이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0033] 도 1은 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법의 개념도이다.
- [0034] 광고 인젝션 앱을 검출하기 위해서는 많은 양의 학습 데이터(예를 들어, 실제 광고 인젝션을 수행하는 앱)가 필

요하나, 현재 광고 인젝션 앱을 탐지하는 분야에서는 발견된 광고 인젝션 앱의 개수가 적으므로 딥 러닝과 같이 성능이 뛰어나지만 많은 학습 데이터를 요구하는 기계 학습을 적용하여 광고 인젝션 앱을 탐지할 수 없다. 다만, 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법을 따르면, 모바일 앱에서 광고 인젝션과 관련된 사용자 리뷰를 학습 데이터로 확보하여 광고 인젝션 검출에 활용할 수 있다. 따라서, 본 발명에 따라 학습 데이터를 확보하여 기계 학습을 적용하고, 학습된 단어 빈도수 테이블에 기반하여 앱 스토어에 등록된 모바일 앱에서 광고 인젝션 앱을 탐지할 수 있다. 학습 데이터를 확보하는 방법에 대해서는 후술한다.

[0035] 도 1을 참조하면, 본 발명의 일 실시예에 따른 광고 인젝션 앱을 검출하는 방법은 모바일 앱의 정보를 크롤링하는 단계, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계, 상기 사용자 리뷰에 기반하여 광고 인젝션 점수를 산출하는 단계 및 상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하는 구성을 포함할 수 있다. 광고 인젝션 앱을 검출하는 방법의 구체적인 내용은 후술한다.

[0036] 본 발명의 광고 인젝션 앱을 검출하는 방법은 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 것이 특징이다. 사용자는 일반적으로 모바일 앱에 대한 평가를 리뷰로써 남기며, 평가는 모바일 앱의 기능(동작)에 대해서 이루어진다. 즉, 모바일 앱 스토어에 등록된 특정 광고 인젝션 앱이 주기적으로 다수의 광고를 노출시키면 사용자는 이에 대해서 부정적인 리뷰를 남기고, 사용자의 부정적 리뷰는 광고 인젝션 검출을 위해 분석해야 하는 동적 분석 보고서가 될 수 있다.

[0038] 도 2는 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법의 동작 순서도이다.

[0039] 도 2를 참조하면, 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법은, 모바일 앱의 정보를 크롤링하는 단계(S210)를 포함할 수 있다,

[0040] 여기서, 상기 모바일 앱의 정보를 크롤링하는 단계는, 특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하는 단계를 포함할 수 있다.

[0041] 한편, 본 발명은 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계(S220)를 포함할 수 있다.

[0042] 여기서, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는, 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하는 단계를 포함할 수 있다.

[0043] 또한, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하는 단계는, 학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하는 단계를 더 포함할 수 있다.

[0044] 한편, 상기 학습된 단어 빈도수 테이블은, 상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습될 수 있다.

[0045] 또한, 상기 나이트 베이지안 분류기는, 사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류할 수 있다.

[0046] 또한, 상기 나이트 베이지안 분류기는, 광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여할 수 있다.

[0047] 또한, 본 발명은 상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하는 단계(S230)를 포함할 수 있다.

[0048] 여기서, 상기 광고 인젝션 점수는, 나이트 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출될 수 있다.

[0049] 한편, 본 발명은 상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하는 단계(S240)를 포함할 수 있다.

[0051] 도 3은 모바일 앱의 정보를 크롤링하는 방법의 개념도이다.

[0052] 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 방법은, 모바일 앱의 정보를 크롤링하는 단계(S210)를 포함할 수 있다. 여기서, 상기 모바일 앱의 정보를 크롤링하는 단계는, 특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하는 단계를 포함할 수 있다.

- [0053] 도 3을 참조하면, 모바일 앱의 정보를 크롤링하는 방법은 특정 키워드를 통해 URL을 구성하여 앱 스토어에 요청하는 단계를 포함할 수 있다. 예를 들어, "https://play.google.com/store/search?q=게임&c=apps"을 통해 구글 플레이 스토어에 등록된 키워드 "게임"과 관련된 모바일 앱들을 검색할 수 있으며, HTML 파싱을 통해 검색된 모바일 앱들의 아이디(ID)를 얻으면, 이에 기반하여 모바일 앱 페이지에 접속하여 특정 모바일 앱에 관련된 정보들을 크롤링할 수 있다.
- [0054] 본 발명은 모바일 앱의 사용자 리뷰에 기반하여 키워드(특정 단어) "광고"가 포함되어 있는 모바일 앱의 정보를 크롤링할 수 있다. 예를 들어, 모바일 앱의 사용자 리뷰에 키워드 "광고"가 포함되어 있지 않은 경우 해당 모바일 앱은 광고 인젝션 앱과 관련이 없는 모바일 앱으로 판단하여 크롤링하지 않을 수 있다. 모바일 앱 페이지에서 크롤링할 수 있는 모바일 앱의 정보는 앱 이름, 버전, 앱 업데이트 날짜, 필요 권한, 리뷰, 권한의 목록 및 사용자 리뷰 중 적어도 하나의 정보를 크롤링할 수 있다.
- [0055] 한편, 광고 인젝션 앱을 검출하기 위한 가장 중요한 정보는 모바일 앱이 필요로 하는 권한의 목록 및 사용자 리뷰일 수 있다. 예를 들어, 광고 인젝션 앱은 다른 모바일 앱이나 홈 화면 등에 광고를 출력하기 위해 주로 "다른 앱 위에 그리기" 권한을 사용하므로 권한의 목록 정보에 대한 크롤링이 필요할 수 있고, 특정 모바일 앱이 광고 인젝션 앱과 얼마나 관련이 있는지에 대한 분석은 사용자의 리뷰를 기반으로 분석할 수 있으므로 사용자의 리뷰 정보에 대한 크롤링이 필요할 수 있다.
- [0057] 도 4는 단어 빈도수 테이블을 학습시키는 방법의 개념도이고, 도 5는 단어 빈도수 테이블을 학습시키기 위한 사용자 리뷰의 예시도이다.
- [0058] 본 발명의 학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하기 위해서는 단어 빈도수 테이블들이 사전에 생성되어야 한다. 즉, 본 발명의 나이트 베이지안 분류기는 사전에 생성된 단어 빈도수 테이블에 기반하여 사용자의 리뷰를 광고 인젝션 또는 일반 리뷰로 분류할 수 있다.
- [0059] 도 4를 참조하면, 단어 빈도수 테이블을 학습시키는 방법은 학습 데이터 파일에서 사용자의 리뷰와 리뷰의 카테고리어를 읽어오는 단계(S410), 학습 데이터 파일에서 읽어온 사용자의 리뷰를 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 전처리하는 단계(S420), 전처리된 사용자의 리뷰에 기반하여 단어 빈도수 테이블을 생성하는 단계(S430)를 포함할 수 있다. 여기서, 학습 데이터 파일은 사전에 저장되었던 사용자의 리뷰와 리뷰의 카테고리뿐만 아니라 본 발명의 광고 인젝션 앱을 검출하는 과정에서 사용되는 모바일 앱의 사용자 리뷰도 포함될 수 있다.
- [0060] 예를 들어, 학습 데이터 파일에서 광고 인젝션 앱으로 분류된 사용자 리뷰 "이 앱을 설치한 후에 광고가 홈 화면에 나와요"가 있는 경우, 상기 사용자의 리뷰를 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 "이Noun", "앱Noun", "설치Noun", "후Noun", "광고", "홈Noun", "화면Noun", "나오다Verb"로 분류한 후 단어 빈도수 테이블에 반영할 수 있다. 즉, 상기 사용자의 리뷰는 광고 인젝션 앱으로 분류된 모바일 앱에서 사용자의 리뷰를 읽어와서 단어 빈도수 테이블을 학습시키는데 사용될 수 있다.
- [0061] 도 5를 참조하면, 학습된 단어 빈도수 테이블에 기반하여 사용자의 리뷰를 나이트 베이지안 분류기를 통해 분석할 수 있다. 여기서, 나이트 베이지안 분류기는 광고 인젝션 리뷰에는 1의 값을 부여하고, 일반 리뷰에는 0의 값을 부여할 수 있다. 예를 들어, "걸리지도 않은 바이러스 걸렸다고 앱 설치 하라는 광고 뜨고 최악"이라는 사용자의 리뷰는 광고 인젝션 앱과 관련이 있으므로 나이트 베이지안 분류기를 통해 1의 값이 부여될 수 있다. 한편, 사용자의 리뷰 좌측에는 사용자가 모바일 앱을 평가한 별점이 표시될 수 있다.
- [0063] 도 6a는 학습된 단어 빈도수 테이블의 제 1 예시도이고, 도 6b는 학습된 단어 빈도수 테이블의 제 2 예시도이다.
- [0064] 도 6a 및 도 6b를 참조하면, 도 6a의 학습된 단어 빈도수 테이블은 "많다", "짜증나다"라는 단어의 빈도수가 높은 것으로 보아 광고 인젝션 앱을 포함하는 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하여 학습된 단어 빈도수 테이블을 의미할 수 있고, 도 6b는 도 6a와 달리 "많다", "짜증나다"라는 단어의 빈도수가 상대적으로 낮은 것으로 보아 광고 인젝션 앱을 포함하지 않는 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하여 학습된 단어 빈도수 테이블을 의미할 수 있다.

[0065] 본 발명에서는 사용자의 리뷰를 정규화(normalization), 어간 추출(stemming), 형태소 분석(morphological analysis) 및 품사 태깅(part of speech tagging) 중 적어도 하나를 사용하여 사용자의 리뷰를 분석할 수 있다. 예를 들어, "광고가 너무 많아서 짜증나요"라는 사용자의 리뷰가 있는 경우, 상기 사용자의 리뷰를 "광고 Noun", "너무Adverb", "짜증나다Verb"로 분류할 수 있다.

[0066] 이어서, 본 발명은 분류된 사용자의 리뷰에서 단어의 발생 빈도를 추출할 수 있고, 단어의 발생 빈도에 기반해 학습된 단어 빈도수 테이블을 활용하는 나이브 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류할 수 있다. 예를 들어, "광고가 너무 많아서 짜증나요"라는 사용자의 리뷰는 도 6a 및 도 6b를 참조하면, "짜증나다"라는 단어를 포함하고 있으므로 광고 인젝션 앱을 포함하는 모바일 앱에 대한 사용자의 리뷰로 분류될 수 있다. 또한, 상기 사용자의 리뷰는 나이브 베이지안 분류기를 통해 1의 값이 부여될 수 있다.

[0067] 한편, 광고 인젝션 앱을 검출하는 과정에서 분류된 광고 인젝션 리뷰 또는 일반 리뷰는 본 발명의 단어 빈도수 테이블을 재학습시키는 데 사용될 수 있다.

[0069] 도 7은 나이브 베이지안 분류기를 통해 사용자 리뷰를 분석한 결과의 예시도이다.

[0070] 나이브 베이지안 분류기를 통해 모바일 앱에 포함된 사용자의 리뷰를 분석할 경우 0과 1의 값이 부여되므로 분석된 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로서 분류하여 학습 데이터로서 활용하거나 광고 인젝션 앱을 검출하는데 사용할 수 있다. 예를 들어, 도 7을 참조하면, "광고 좀 없애줘라", "다 좋은데 광고가 너무 많음"과 같은 사용자의 리뷰는 나이브 베이지안 분류기를 통해 광고 인젝션 리뷰로 분류될 수 있고, "광고에서 재미있어 보여서 깔았는데 아직 모름"과 같은 사용자의 리뷰는 일반 리뷰로 분류될 수 있다.

[0072] 도 8은 광고 인젝션 점수가 부여된 모바일 앱들의 예시도이다.

[0073] 본 발명은 나이브 베이지안 분류기에 의해 광고 인젝션 리뷰로 분류된 사용자의 리뷰에 기반하여 광고 인젝션 점수를 부여할 수 있다. 예를 들어, 광고 인젝션 점수(Advertisement Injection Score; AIS)는 수학적 식 1에 따라 산출될 수 있다.

수학적 식 1

$$AIS = \sum_{i=1}^{i=n} ((ClassIndex * C_{weight}) * ((6 - SR) * r_{weight}))$$

[0075] 여기서, ClassIndex는 나이브 베이지안 분류기에 의해 부여된 값을 의미할 수 있고, SR은 사용자의 리뷰에 포함된 별점(Star Rating; SR)을 의미할 수 있고, C_{weight} 는 ClassIndex에 부여된 가중치를 의미할 수 있고, r_{weight} 는 별점에 부여된 가중치를 의미할 수 있다. 여기서, 가중치는 사용자의 설정에 의해 미리 설정될 수 있다. 따라서, 광고 인젝션 점수는 모바일 앱에 등록된 사용자의 리뷰가 광고 인젝션 리뷰로 분류된 것이 많을수록 높고, 사용자의 리뷰에 포함된 별점이 낮을수록 높을 수 있다.

[0076] 또한, 본 발명은 광고 인젝션 앱을 검출하기 위해서 GUI(Graphical User Interface)를 활용할 수 있다. GUI는 광고 인젝션 앱 검출 장치에 적용될 수 있으나 이에 제한되지 않는다.

[0077] 도 8을 참조하면, GUI가 적용된 화면의 예로서 광고 인젝션 점수가 부여된 모바일 앱들을 표시하고 있음을 알 수 있다. GUI가 적용된 화면을 살펴보면 광고 인젝션 앱들이 주로 사용하는 권한들을 가진 특정 앱들만 필터링할 수 있음을 알 수 있다.

[0078] 즉, 본 발명은 수학적 식 1에 따라 산출된 광고 인젝션 점수에 따라 광고 인젝션 앱으로 의심되는 모바일 앱을 우선적으로 필터링하여 광고 인젝션 앱을 효과적으로 검출할 수 있다.

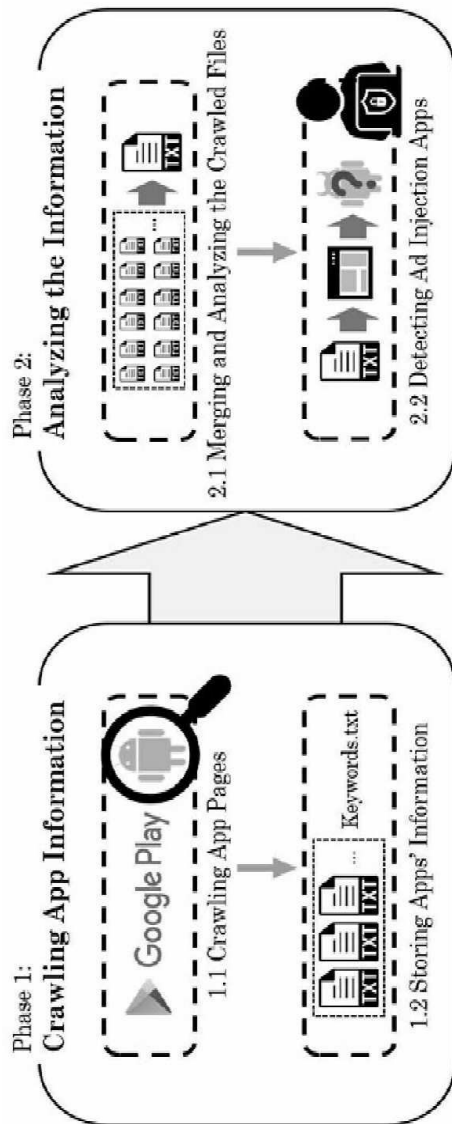
[0080] 도 9는 모바일 앱 상세 정보의 예시도이고, 도 10은 APK(Android Application Package) 분석창의 예시도이다.

- [0081] 도 9는 광고 인젝션 앱으로 의심되는 앱에 대한 상세 정보를 제공하는 GUI가 적용된 화면이다. 즉, 상세 정보에는 해당 모바일 앱의 패키지명, 개발자, 앱 버전, 사용자 리뷰, 권한의 목록 등이 표시될 수 있고, 상세 정보를 참고하여 광고 인젝션 앱을 검출할 수 있다.
- [0082] 또한, 사용자는 APK(Android Application Package) 다운로드 기능과 APK 분석 기능을 활용하여 광고 인젝션 앱을 검출할 수 있다. 예를 들어, 도 10을 참조하면, APK 분석창에 나타나는 앱이 사용하는 광고 네트워크(ad-network), 권한, 액티비티, 서비스, 리시버 등과 함께 AndroidManifest.xml, Main 함수의 코드를 통해 광고 의심 앱을 검출할 수 있다.
- [0084] 도 11은 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 장치의 블록 구성도이다.
- [0085] 도 11을 참조하면, 본 발명의 일 실시예에 따른 광고 인젝션 앱 검출 장치(1100)는 프로세서(1110) 및 프로세서를 통해 실행되는 적어도 하나의 명령 및 명령 수행의 결과를 저장하는 메모리(1120) 및 네트워크와 연결되어 통신을 수행하는 송수신 장치(1130)를 포함할 수 있다.
- [0086] 광고 인젝션 앱 검출 장치(1100)는 또한, 입력 인터페이스 장치(1140), 출력 인터페이스 장치(1150), 저장 장치(1160) 등을 더 포함할 수 있다. 광고 인젝션 앱 검출 장치(1100)에 포함된 각각의 구성 요소들은 버스(Bus)(1170)에 의해 연결되어 서로 통신을 수행할 수 있다.
- [0087] 프로세서(1110)는 메모리(1120) 및 저장 장치(1160) 중에서 적어도 하나에 저장된 프로그램 명령(program command)을 실행할 수 있다. 프로세서(1110)는 중앙 처리 장치(central processing unit, CPU), 그래픽 처리 장치(graphics processing unit, GPU), 또는 본 발명의 실시예에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다. 메모리(1120) 및 저장 장치(1160) 각각은 휘발성 저장 매체 및 비휘발성 저장 매체 중에서 적어도 하나로 구성될 수 있다. 예를 들어, 메모리(1120)는 읽기 전용 메모리(read only memory, ROM) 및 랜덤 액세스 메모리(random access memory, RAM) 중에서 적어도 하나로 구성될 수 있다.
- [0088] 저장 장치(1160)는 또한, 기존 학습 데이터 및 본 발명의 일 실시예에 따라 광고 인젝션 앱으로 분류된 모바일 앱의 사용자 리뷰로서 분석된 사용자의 리뷰, 크롤링된 모바일 앱에 부여된 광고 인젝션 점수, 나이트 베이지안 분류기에 의해 부여된 값을 갖는 모바일 앱의 목록 등을 저장할 수 있다.
- [0089] 여기서, 적어도 하나의 명령은, 모바일 앱의 정보를 크롤링하도록 하는 명령; 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령; 상기 사용자 리뷰에 기반하여 광고 인젝션 점수(Advertisement Injection Score; AIS)를 산출하도록 하는 명령; 및 상기 광고 인젝션 점수를 토대로 광고 인젝션 앱을 검출하도록 하는 명령을 포함할 수 있다.
- [0090] 여기서, 상기 모바일 앱의 정보를 크롤링하도록 하는 명령은, 특정 단어를 포함하는 사용자 리뷰에 기반하여 모바일 앱의 정보를 크롤링하도록 하는 명령을 포함할 수 있다.
- [0091] 한편, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은, 정규화, 어간 추출, 형태소 분석 및 품사 태깅 중 적어도 하나를 사용하여 사용자 리뷰를 단어별로 추출하도록 하는 명령을 포함할 수 있다.
- [0092] 여기서, 상기 모바일 앱의 정보에 포함된 사용자 리뷰를 분석하도록 하는 명령은, 학습된 단어 빈도수 테이블을 활용하는 나이트 베이지안 분류기를 통해 사용자의 리뷰를 광고 인젝션 리뷰 또는 일반 리뷰로 분류하도록 하는 명령을 더 포함할 수 있다.
- [0093] 한편, 상기 학습된 단어 빈도수 테이블은, 상기 광고 인젝션 앱에 포함된 사용자의 리뷰에서 추출한 단어의 빈도수를 활용하여 재학습될 수 있다.
- [0094] 또한, 상기 나이트 베이지안 분류기는, 사용자의 리뷰를 학습된 단어 빈도수 테이블의 특정 단어와 비교하여 광고 인젝션 리뷰 또는 일반 리뷰로 분류할 수 있다.
- [0095] 여기서, 상기 나이트 베이지안 분류기는, 광고 인젝션 리뷰에는 1의 값을 부여하고 일반 리뷰에는 0의 값을 부여할 수 있다.
- [0096] 또한, 상기 광고 인젝션 점수는, 나이트 베이지안 분류기를 통해 사용자의 리뷰에 부여된 값과 사용자의 리뷰에 포함된 별점을 활용하여 산출될 수 있다.

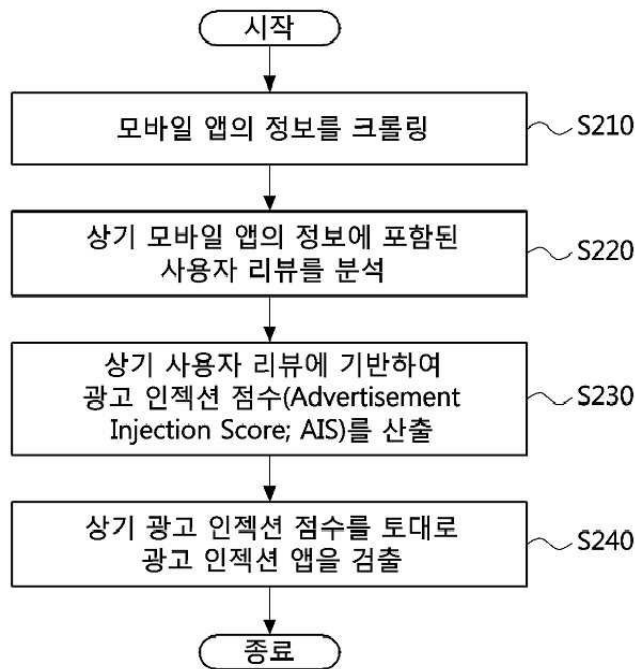
- [0098] 본 발명의 실시예에 따른 방법의 동작은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 프로그램 또는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의해 읽혀질 수 있는 정보가 저장되는 모든 종류의 기록장치를 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산 방식으로 컴퓨터로 읽을 수 있는 프로그램 또는 코드가 저장되고 실행될 수 있다.
- [0099] 또한, 컴퓨터가 읽을 수 있는 기록매체는 롬(rom), 램(ram), 플래시 메모리(flash memory) 등과 같이 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함할 수 있다. 프로그램 명령은 컴파일러(compiler)에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터(interpreter) 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함할 수 있다.
- [0100] 본 발명의 일부 측면들은 장치의 문맥에서 설명되었으나, 그것은 상응하는 방법에 따른 설명 또한 나타낼 수 있고, 여기서 블록 또는 장치는 방법 단계 또는 방법 단계의 특징에 상응한다. 유사하게, 방법의 문맥에서 설명된 측면들은 또한 상응하는 블록 또는 아이템 또는 상응하는 장치의 특징으로 나타낼 수 있다. 방법 단계들의 몇몇 또는 전부는 예를 들어, 마이크로프로세서, 프로그램 가능한 컴퓨터 또는 전자 회로와 같은 하드웨어 장치에 의해(또는 이용하여) 수행될 수 있다. 몇몇의 실시예에서, 가장 중요한 방법 단계들의 하나 이상은 이와 같은 장치에 의해 수행될 수 있다.
- [0101] 실시예들에서, 프로그램 가능한 로직 장치(예를 들어, 필드 프로그램블 게이트 어레이)가 여기서 설명된 방법들의 기능의 일부 또는 전부를 수행하기 위해 사용될 수 있다. 실시예들에서, 필드 프로그램블 게이트 어레이는 여기서 설명된 방법들 중 하나를 수행하기 위한 마이크로프로세서와 함께 작동할 수 있다. 일반적으로, 방법들은 어떤 하드웨어 장치에 의해 수행되는 것이 바람직하다.
- [0102] 이상 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 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.

도면

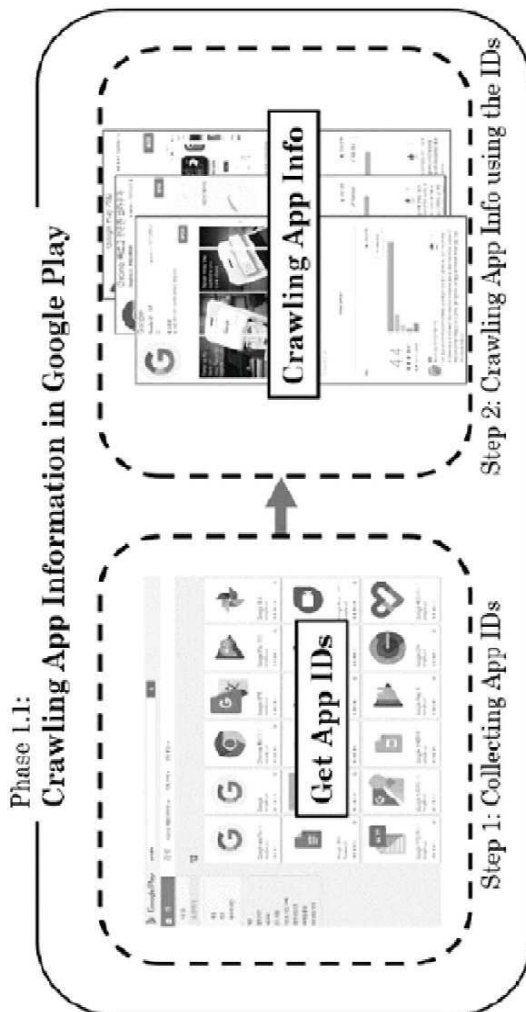
도면1



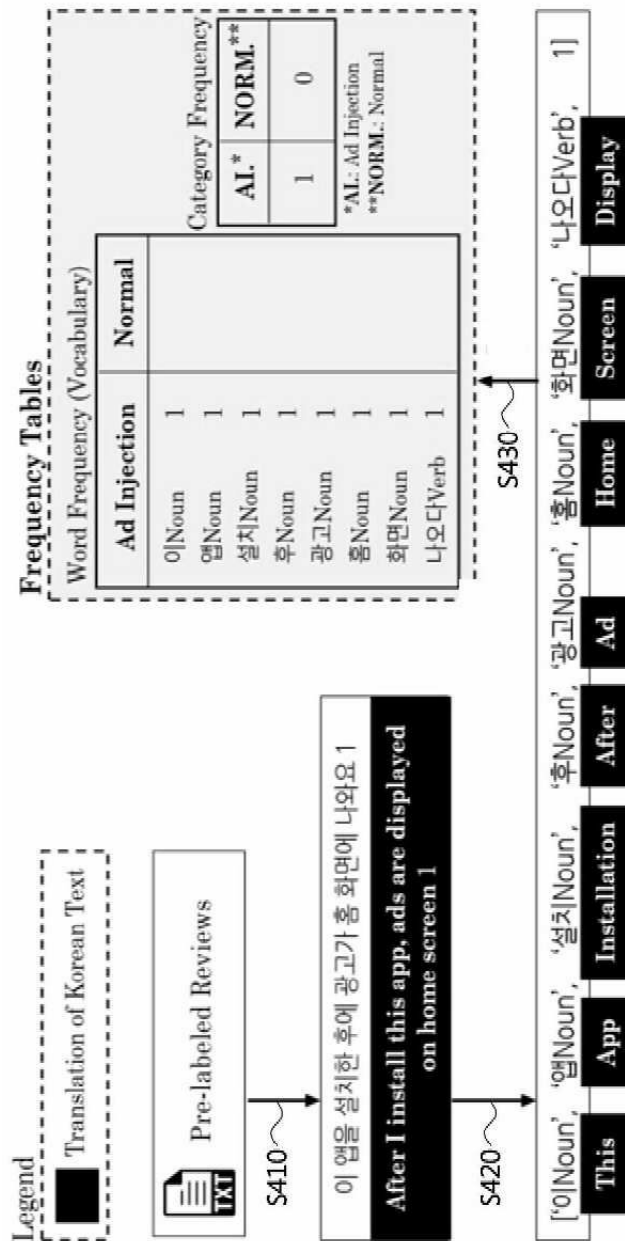
도면2



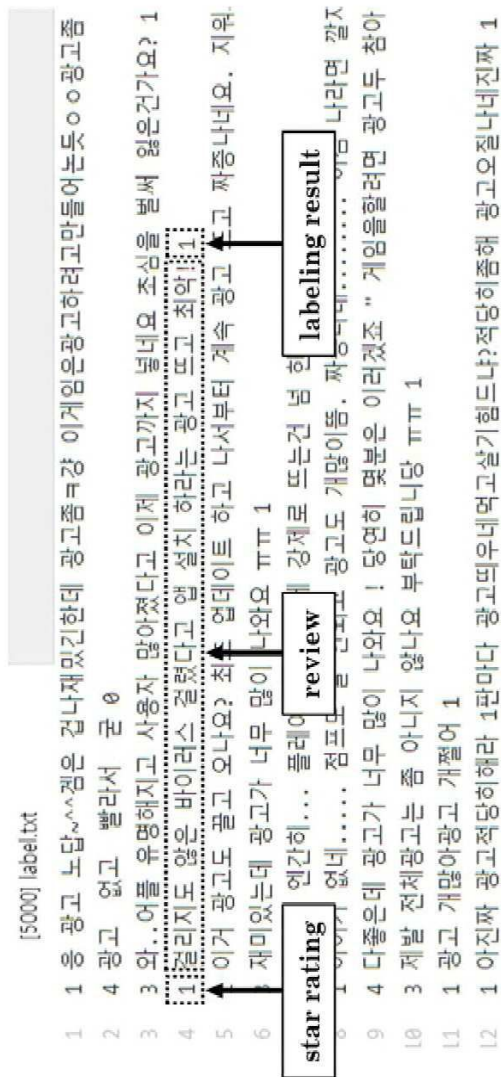
도면3



도면4



도면5



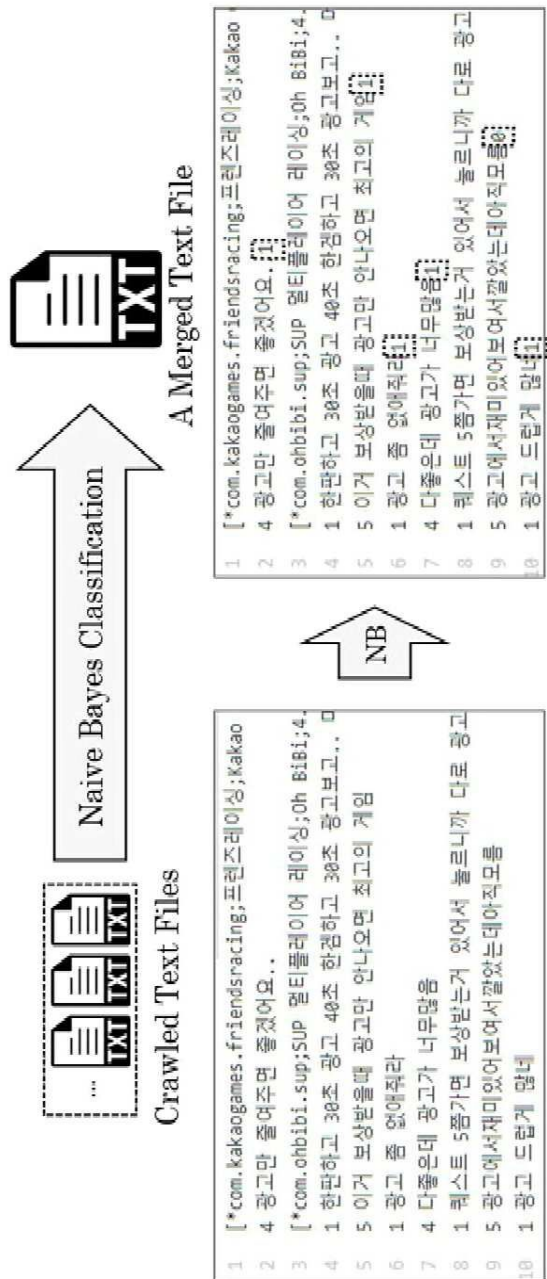
도면6a

단어	빈도수
많다	10
짜증나다	43

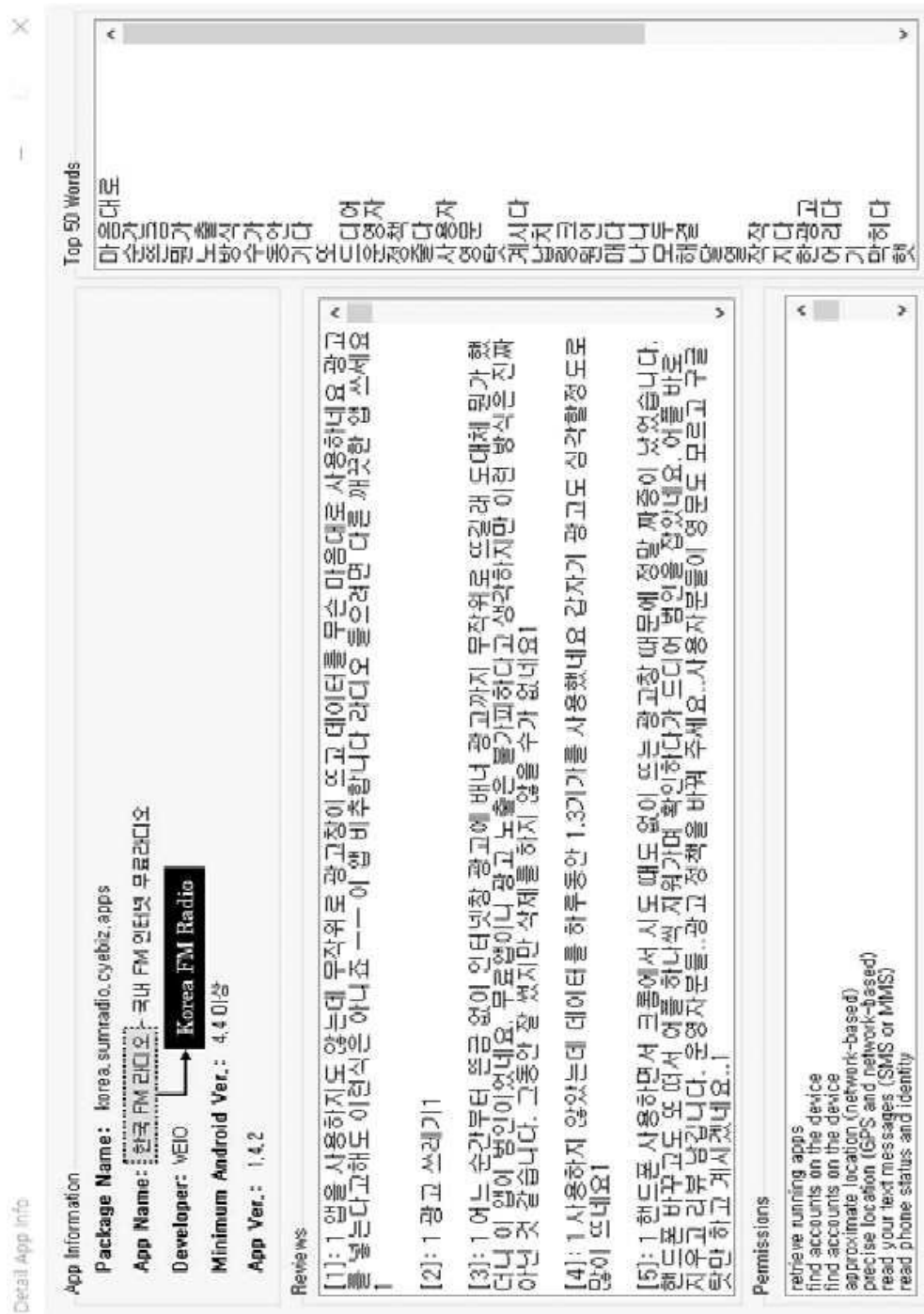
도면6b

단어	빈도수
많다	5
짜증나다	6

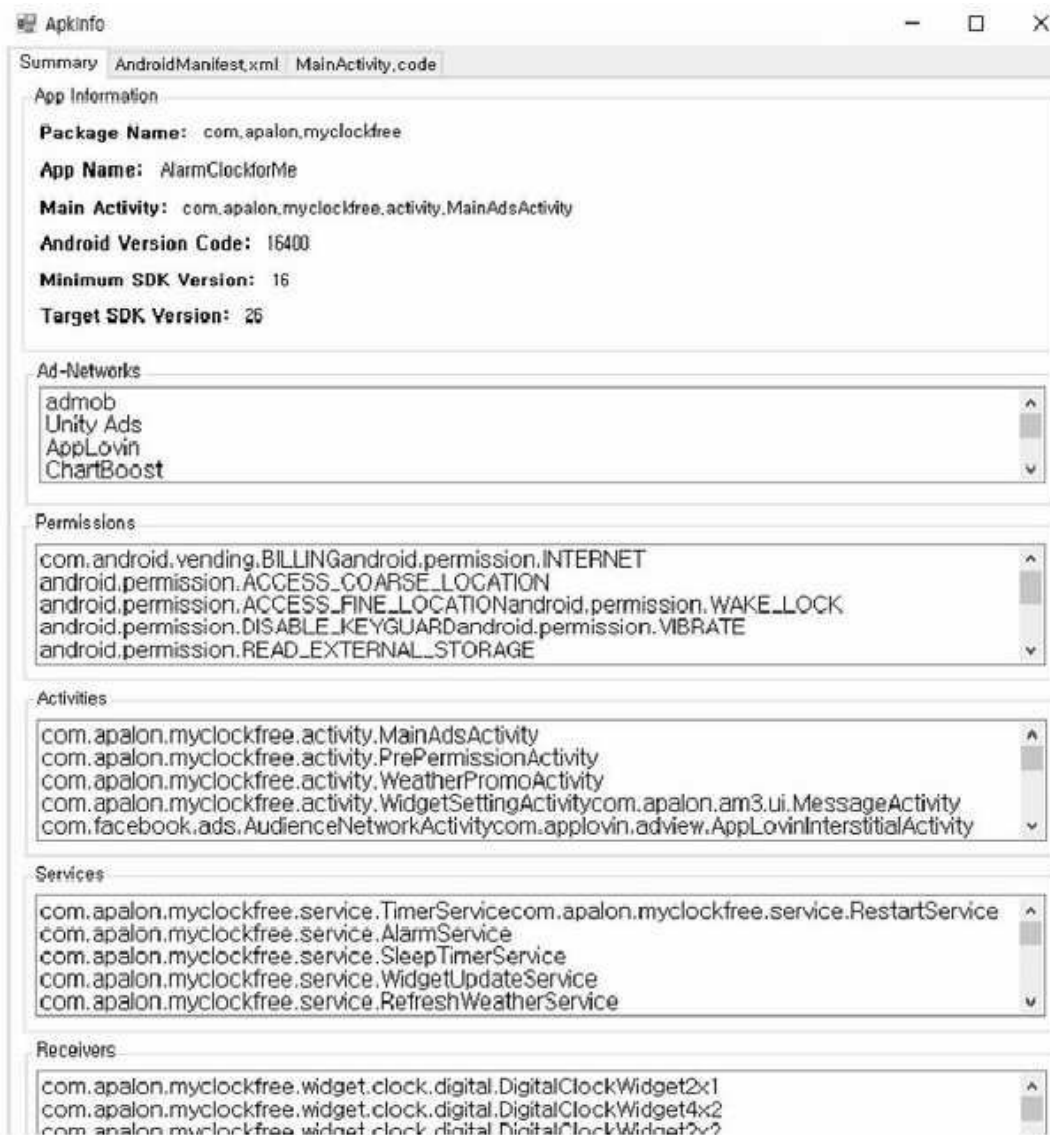
도면7



도면9



도면10



도면11

