(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년02월02일
(11) 등록번호 10-2633826
(24) 등록일자 2024년01월31일

(51) 국제특허분류(Int. Cl.)
G06F 21/55 (2013.01) *G06F 40/247* (2020.01)
G06F 40/289 (2020.01) *G06N 3/04* (2023.01)
H04W 4/14 (2018.01)
(52) CPC특허분류
G06F 21/55 (2013.01)
G06F 40/247 (2020.01)
(21) 출원번호 10-2021-0186855
(22) 출원일자 2021년12월24일
심사청구일자 2021년12월24일
(65) 공개번호 10-2023-0097395
(43) 공개일자 2023년07월03일
(56) 선행기술조사문헌
KR1020210114256 A*
방지훈 외, "단어 군집화를 통한 스미싱 탐지 규칙 추천"(2020.11.)*
Junkang Wu et al., "DisenKGAT: Knowledge Graph Embedding with Disentangled Graph Attention Network"(2021.10.)*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이근배
서울특별시 서초구 서운로 221, 103동 1203호(서초동, 래미안서초스위트아파트)
백성빈
경상북도 포항시 남구 효자로77번길 14-1, 205호 (대잠동)
(74) 대리인
특허법인(유한)아이시스

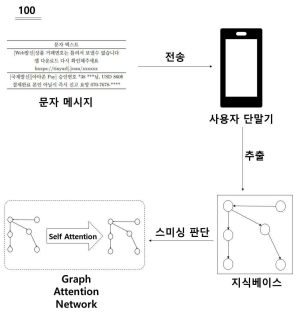
전체 청구항 수 : 총 2 항 심사관 : 정성훈

(54) 발명의 명칭 **지식베이스 데이터 기반 스미싱 탐지 방법 및 장치**

(57) 요 약

개시된 기술은 지식베이스 데이터 기반 스미싱 탐지 방법 및 장치에 관한 것으로, 사용자 단말기의 프로세서가 상기 사용자 단말기에 전송된 문자 메시지를 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성하는 단계; 상기 프로세서가 상기 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력하는 단계; 및 상기 프로세서가 상기 스미싱 탐지 모델의 출력 결과에 따라 상기 문자 메시지에 대한 스미싱 여부를 판단하는 단계;를 포함한다.

대 표 도



(52) CPC특허분류
G06F 40/289 (2020.01)
G06N 3/04 (2023.01)
H04W 4/14 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711125943
과제번호 2019-0-01906-003
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송혁신인재양성(R&D)
연구과제명 인공지능대학원지원(포항공과대학교)
기 여 율 1/2
과제수행기관명 포항공과대학교 산학협력단
연구기간 2021.01.01 ~ 2021.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711138812
과제번호 2021-0-00575-001
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 ICT기반사회문제해결기술개발(R&D)
연구과제명 음성·텍스트 딥러닝 기반 보이스피싱 예방 기술 개발
기 여 율 1/2
과제수행기관명 서울대학교 산학협력단
연구기간 2021.04.01 ~ 2021.12.31
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

사용자 단말기의 프로세서가 상기 사용자 단말기에 전송된 문자 메시지를 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성하는 단계;

상기 프로세서가 상기 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력하는 단계; 및

상기 프로세서가 상기 스미싱 탐지 모델의 출력 결과에 따라 상기 문자 메시지에 대한 스미싱 여부를 판단하는 단계;를 포함하되,

상기 지식베이스 데이터는 상기 문자 메시지에 포함된 복수개의 단어들 중 일부를 노드(Node)로 포함하고, 상기 일부의 단어들 간의 관계를 엣지(Edge)로 포함하는 그래프이고,

상기 지식베이스 데이터는 상기 문자 메시지에 포함된 단어와 유사한 단어로서 통계적으로 자주 사용되는 용어를 더 이용하여 생성되는,

상기 스미싱 탐지 모델은 그래프 어텐션 신경망(Graph Attention Network, GAN) 기반 모델로서, 상기 지식베이스 데이터에 대한 셀프 어텐션(Self Attention) 기반 임베딩을 수행하여 그래프 전체에 대한 벡터를 생성하고, 상기 벡터를 계산하여 스미싱 확률을 출력하는 지식베이스 데이터 기반 스미싱 탐지 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

사용자 단말기로 전송되는 문자 메시지를 수신하는 통신장치;

지식베이스 추출 모델 및 스미싱 탐지 모델을 저장하는 저장장치; 및

상기 문자 메시지를 상기 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성하고, 상기 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력하여 상기 스미싱 탐지 모델의 출력 결과에 따라 상기 문자 메시지에 대한 스미싱 여부를 판단하는 연산장치;를 포함하되,

상기 지식베이스 데이터는 상기 문자 메시지에 포함된 복수개의 단어들 중 일부를 노드(Node)로 포함하고, 상기 일부의 단어들 간의 관계를 엣지(Edge)로 포함하는 그래프이고,

상기 지식베이스 데이터는 상기 문자 메시지에 포함된 단어와 유사한 단어로서 통계적으로 자주 사용되는 용어를 더 이용하여 생성되는,

상기 스미싱 탐지 모델은 그래프 어텐션 신경망(Graph Attention Network, GAN) 기반 모델로서, 상기 지식베이스 데이터에 대한 셀프 어텐션(Self Attention) 기반 임베딩을 수행하여 그래프 전체에 대한 벡터를 생성하고, 상기 벡터를 계산하여 스미싱 확률을 출력하는 지식베이스 데이터 기반 스미싱 탐지 장치.

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

발명의 설명

기 술 분 야

[0001] 개시된 기술은 문자 메시지에서 추출한 지식베이스 데이터를 기반으로 스미싱을 탐지하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 스미싱은 사용자의 휴대 단말기로 전송되는 문자 메시지를 통해 피해자를 현혹시켜 개인정보나 금전 등을 갈취하는 범죄이다. 발전하는 스미싱 범죄 수법에 대응하기 위해서는 새로운 스미싱 범죄 사례에서 데이터를 추출하고 이를 기존 시스템에 적용하여 추가 피해를 예방해야한다.

[0003] 스미싱 범죄를 예방하기 위해서 다양한 종류의 자연어 처리(Natural Language Processing, NLP) 기술이 연구되고 있으며 높은 정확도로 일반 문자 메시지와 스미싱 문자 메시지를 구별하는 성능을 나타내고 있다. 그러나 종래와 유사한 스미싱 기법에 한해서만 높은 성능을 나타낼 뿐, 기존과는 다른 새로운 형태의 스미싱 기법에는 탐지 정확도가 낮은 문제점을 나타내고 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국 공개특허 제10-2017-0024777호

발명의 내용

해결하려는 과제

[0005] 개시된 기술은 문자 메시지에서 추출한 지식베이스 데이터를 기반으로 스미싱을 탐지하는 방법 및 장치를 제공하는데 있다.

과제의 해결 수단

[0006] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 1 측면은 사용자 단말기의 프로세서가 상기 사용자 단말기에 전송된 문자 메시지를 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성하는 단계, 상기 프로세서가 상기 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력하는 단계 및 상기 프로세서가 상기 스미싱 탐지 모델의 출력 결과에 따라 상기 문자 메시지에 대한 스미싱 여부를 판단하는 단계를 포함하는 지식베이스 데이터 기반 스미싱 탐지 방법을 제공하는데 있다.

[0007] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 2 측면은 사용자 단말기로 전송되는 문자 메시지를 수신하는 통신장치, 지식베이스 추출 모델 및 스미싱 탐지 모델을 저장하는 저장장치 및 상기 문자 메시지를 상기 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성하고, 상기 지식베이스 데이터를 스미싱(Smishing)

탐지 모델에 입력하여 상기 스미싱 탐지 모델의 출력 결과에 따라 상기 문자 메시지에 대한 스미싱 여부를 판단하는 연산장치를 포함하는 지식베이스 데이터 기반 스미싱 탐지 장치를 제공하는데 있다.

발명의 효과

[0008] 개시된 기술의 실시 예들은 다음의 장점들을 포함하는 효과를 가질 수 있다. 다만, 개시된 기술의 실시 예들이 이를 전부 포함하여야 한다는 의미는 아니므로, 개시된 기술의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.

[0009] 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 방법 및 장치는 지식베이스 데이터를 토대로 스미싱을 판별하여 새로운 스미싱 수법에 유연하게 대처하는 효과가 있다.

[0010] 또한, 자동으로 추출되는 지식베이스 데이터를 모델의 학습에 이용하여 스미싱 탐지 성능을 향상시키는 효과가 있다.

도면의 간단한 설명

[0011] 도 1은 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 과정을 나타낸 도면이다.

도 2는 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 방법에 대한 순서도이다.

도 3은 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 장치에 대한 블록도이다.

도 4는 지식베이스 데이터에 대한 예시를 나타낸 도면이다.

도 5는 스미싱 탐지 모델의 동작을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0012] 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.

[0013] 제 1 , 제 2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0014] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 한다. 그리고 "포함한다" 등의 용어는 실시된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.

[0015] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다.

[0016] 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다. 따라서, 본 명세서를 통해 설명되는 각 구성부들의 존재 여부는 기능적으로 해석되어야 할 것이다.

[0017] 도 1은 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 과정을 나타낸 도면이다. 도 1을 참조하면 문자 메시지는 사용자 단말기의 SMS(Short Message Service)를 통해 타 단말기로부터 전송될 수 있다. 사용자 단말기는 사용자가 휴대하는 스마트폰이나 태블릿과 같은 휴대 단말기일 수도 있고 PC와 같은 고정된 단말기일 수도 있다.

[0018] 한편, 사용자 단말기로 문자 메시지가 전송되면 사용자 단말기의 프로세서는 문자 메시지를 지식베이스 추출 모

델에 입력한다. 지식베이스 추출 모델은 입력된 문자 메시지에 포함된 복수개의 단어들을 객체로 추출하고 특정 객체와 다른 객체 간의 관계를 추출하여 지식베이스 데이터를 생성하는 모델을 의미한다. 지식베이스 추출 모델은 객체 간 의미적 연관관계를 자동으로 추출하는 OpenIE6를 기반으로 구축된 모델일 수 있다. 지식베이스 추출 모델은 문자 메시지와 같이 복수개의 단어들을 포함하는 정형화되지 않은 텍스트 데이터를 이용하여 단어, 문장 또는 구문의 유사도, 관계, 시맨틱 트리 등을 분류할 수 있다. 지식베이스 추출 모델은 문자 메시지 원문에 포함된 복수개의 단어들 중 일부를 추출하거나 문자 메시지에 포함되지 않은 유사 단어를 이용하여 지식베이스 데이터를 생성할 수 있다. 가령 통계적으로 자주 사용되는 용어를 이용하여 지식베이스 데이터를 생성할 수 있다.

[0019] 한편, 지식베이스 추출 모델에서 생성한 지식베이스 데이터는 문자 메시지에 포함된 복수개의 객체들 중 일부를 노드(Node)로 포함하고, 상기 일부의 객체들 간의 관계를 엣지(Edge)로 포함하는 그래프일 수 있다. 각각의 객체와 객체들 간의 관계를 설명하는 엣지를 포함하고 있으므로 그 자체로 사람이 이해할 수 있다. 다만, 사람이 직접 보고 스미싱인지 판단하는 것이 아니라 아래의 스미싱 탐지 모델에 지식베이스 데이터를 입력하여 자동으로 스미싱 여부를 판단할 수 있다.

[0020] 스미싱(Smishing) 탐지 모델은 그래프 어텐션 신경망(Graph Attention Network, GAN)을 기반으로 구축된 모델일 수 있다. 스미싱 탐지 모델은 그래프 형태로 생성된 지식베이스 데이터를 입력값으로 하여 해당 지식베이스 데이터가 스미싱 문자로부터 생성된 것인지 확률을 계산할 수 있다.

[0021] 스미싱 탐지 모델은 기본적으로 복수의 임베딩 레이어(Embedding Layer)와 복수의 덴스 레이어(Dense Layer)들을 포함한다. 스미싱 탐지 모델은 복수의 임베딩 레이어들을 이용하여 지식베이스 데이터에 대한 셀프 어텐션(Self Attention) 기반 임베딩을 수행한다. 이 과정에서 그래프 전체에 대한 벡터를 생성할 수 있다. 그리고, 덴스 레이어를 통해 전체 벡터를 계산하여 스미싱 확률을 출력할 수 있다. 물론 학습 과정에서도 상술한 과정과 동일한 프로세스가 수행되어 자연어의 맥락을 이해하도록 학습될 수 있다. 사용자 단말기의 프로세서는 스미싱 탐지 모델의 출력 결과에 따라 문자 메시지에 대한 스미싱 여부를 판단할 수 있다. 예컨대, 스미싱 탐지 모델에서 지식베이스 데이터가 스미싱일 확률이 높은 것으로 출력되면 문자 메시지가 스미싱인 것으로 판단하고 스미싱일 확률이 낮은 것으로 출력되면 문자 메시지가 정상적으로 수신된 메시지라고 판단할 수 있다. 사용자 단말기의 프로세서가 이러한 과정을 사용자의 별도의 입력 없이 자동으로 수행하여 스미싱 피해를 방지할 수 있다.

[0022] 도 2는 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 방법에 대한 순서도이다. 도 2를 참조하면 지식베이스 데이터 기반 스미싱 탐지 방법(200)은 210 내지 230 단계를 포함한다. 지식베이스 데이터 기반 스미싱 탐지 방법(200)은 사용자 단말기의 프로세서를 통해 자동으로 수행될 수 있다.

[0023] 210 단계에서 사용자 단말기의 프로세서는 사용자 단말기에 전송된 문자 메시지를 지식베이스 추출 모델에 입력한다. 지식베이스 추출 모델은 입력된 문자 메시지로부터 지식베이스 데이터를 생성한다. 지식베이스는 복수개의 객체를 노드로 하고 각 노드들 간의 관계를 엣지로 포함하는 그래프 형태의 데이터이다. 즉, 일종의 이미지에 해당하므로 딥러닝 모델의 입력값으로 활용이 가능하다.

[0024] 220 단계에서 프로세서는 지식베이스 추출 모델이 생성한 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력한다. 스미싱 탐지 모델은 그래프 어텐션 신경망을 기반으로 하는 딥러닝 모델에 해당한다. 따라서, 지식베이스 데이터를 입력값으로 이용하여 해당 지식베이스 데이터가 정상적인 문자 메시지의 특징을 갖는 것인지 아니면 스미싱 메시지의 특징을 갖는 것인지 판단할 수 있다.

[0025] 230 단계에서 프로세서는 스미싱 탐지 모델의 출력 결과에 따라 문자 메시지에 대한 스미싱 여부를 판단한다. 만약 문자 메시지가 스미싱 메시지인 것으로 판단하면 자동으로 문자 메시지를 삭제하거나 사용자에게 해당 문자 메시지가 스미싱 메시지일 확률이 높다는 것을 알릴 수 있다.

[0026] 도 3은 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 장치에 대한 블록도이다. 도 3을 참조하면 지식베이스 데이터 기반 스미싱 탐지 장치(300)는 통신장치(310), 저장장치(320) 및 연산장치(330)를 포함한다.

[0027] 통신장치(310)는 사용자 단말기로 전송되는 문자 메시지를 수신한다. 통신장치(310)는 사용자 단말기의 데이터 통신을 수행하는 통신모듈로 구현될 수 있다. 통신장치(310)는 와이파이, 4G, 5G 등 여러가지 통신기술을 지원할 수 있다.

[0028] 저장장치(320)는 지식베이스 추출 모델 및 스미싱 탐지 모델을 저장한다. 저장장치(320)는 데이터 저장이 가능한 용량을 가진 메모리로 구현될 수 있다. 저장장치(320)는 사용자 단말기와 일체형으로 구비될 수도 있고, 별도로 구비된 연결 단자를 통해 외부에서 연결될 수 있다. 저장장치(320)가 저장하는 2개의 모델들은 각각 학습

데이터를 이용하여 사전에 학습될 수 있다.

[0029] 연산장치(330)는 문자 메시지를 지식베이스 추출 모델에 입력하여 지식베이스 데이터를 생성한다. 그리고, 생성된 지식베이스 데이터를 스미싱(Smishing) 탐지 모델에 입력한다. 그리고, 스미싱 탐지 모델의 출력 결과에 따라 문자 메시지에 대한 스미싱 여부를 판단한다.

[0030] 한편, 상술한 지식베이스 데이터 기반 스미싱 탐지 장치(300)는 컴퓨터에서 실행될 수 있는 실행 가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수도 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.

[0031] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.

[0032] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.

[0033] 도 4는 지식베이스 데이터를 나타낸 도면이다. 도 4와 같이 지식베이스 데이터는 두 객체와 그 객체들 간의 관계를 나타내는 3-tuple의 집합들로 구성된다. 지식베이스 데이터에서 객체는 노드(Node)로 표현되며 객체 간의 관계는 엣지(Edge)로 표현된다. 일 실시예로, 도 4와 같이 객체가 ‘언어’이고 연결된 다른 객체가 ‘영어’라면 두 객체를 연결하는 엣지는 ‘하위어’로 정의될 수 있다. 이와 같이 지식베이스 데이터는 비정형 문자 메시지로부터 정형화된 자료구조를 추출한 것이므로 그 자체로도 사람이 이해할 수 있으며 복수개의 노드와 엣지들로 이루어진 그래프 형태의 데이터이므로 딥러닝 모델의 입력값으로 활용할 수 있다. 지식베이스 데이터는 스미싱 탐지 모델의 입력값 뿐만 아니라 문장을 이해하여 사용자와 대화하는 챗봇과 같은 시스템에서도 모델을 학습시키는 데이터로도 활용될 수 있다.

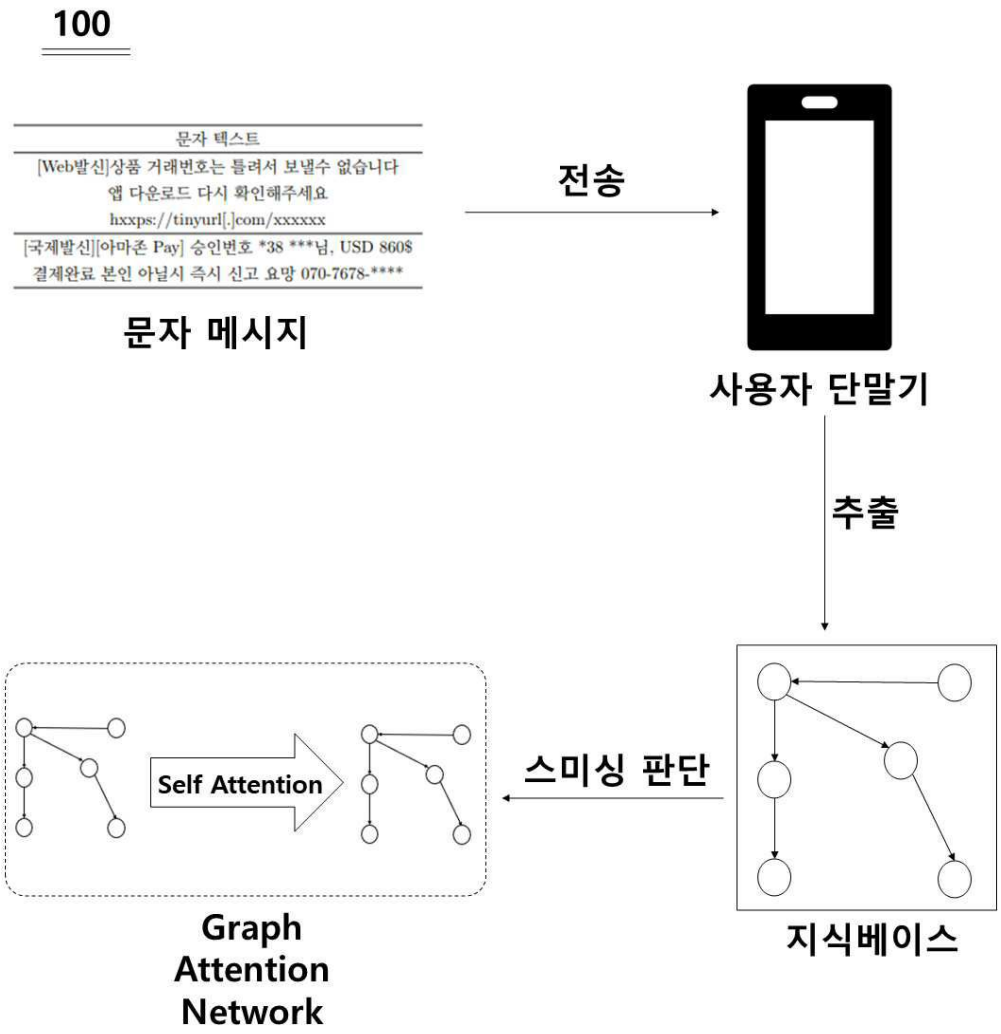
[0034] 도 5는 스미싱 탐지 모델의 동작을 나타낸 도면이다. 스미싱을 탐지하기 위해서는 단순히 문자 메시지에 포함된 단어나 키워드 뿐만 아니라 각 단어들 사이의 관계, 맥락 등을 이해해야한다. 따라서, 그래프 전체에 대한 셀프 임베딩을 수행하여 각각에 대한 가중치를 계산한다. 도 5에 도시된 바와 같이 스미싱 탐지 모델은 지식베이스 데이터가 입력되면 특정 노드를 중심으로 이웃하는 노드들과의 관계를 학습한다. 셀프 어텐션을 통해 지식베이스 데이터에 포함된 노드와 엣지들을 벡터로 임베딩할 수 있으며 임베딩된 벡터들을 계산하여 스미싱 확률을 출력할 수 있다.

[0035] 종래 그래프를 입력받아 클래스 분류를 수행하는 그래프 신경망(Graph Neural Network, GNN)의 경우 이웃하는 노드들에 대한 정보를 동일한 가중치로 평균을 내는 방식을 이용하였다. 즉, 노드들 간의 연결 자체에만 의미를 가졌을 뿐, 가중치를 판단하지 않았다. 그러나 실제 자연어의 맬락을 이해하는데 있어서 모든 맥락이 동일한 가중치를 나타내는 것은 아니다. 예컨대, 도 5에서도 노드 1이 중심이라 가정하였을 때, 노드 1과 노드 2의 관계인 ‘a 2,1’는 노드 1과 노드 3의 관계인 ‘a 1,3’은 서로 동일하지 않은 가중치일 수 있다. 그러므로 그래프 어텐션 신경망의 셀프 어텐션을 통해 보다 중요도가 높은 가중치가 무엇인지 학습하고 이를 토대로 문자 메시지에 포함된 자연어를 보다 깊이 이해하는 것이 가능하다. 이러한 학습 과정에 따라 스미싱 탐지 모델은 단순 키워드 추출 기반의 스미싱 탐지가 아닌 입력된 문자 메시지 전체에 대한 이해를 기반으로 하는 스미싱 탐지를 수행하는 것이 가능하다. 따라서, 종래와 다른 새로운 기법의 스미싱 문자가 수신되더라도 자동으로 스미싱 여부를 탐지할 수 있다.

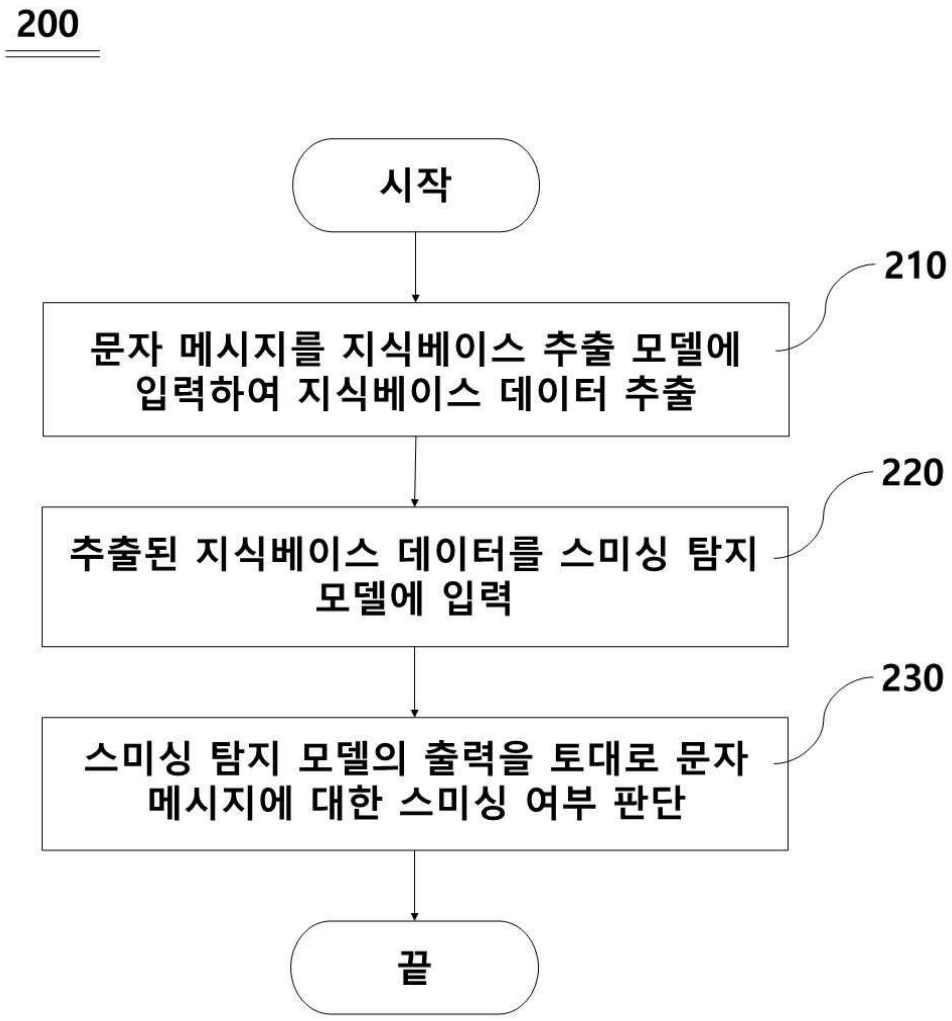
[0036] 개시된 기술의 일 실시예에 따른 지식베이스 데이터 기반 스미싱 탐지 방법 및 장치는 이해를 돕기 위하여 도면에 도시된 실시 예를 참고로 설명되었으나, 이는 예시적인 것에 불과하며, 당해 분야에서 통상적 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 개시된 기술의 진정한 기술적 보호범위는 첨부된 특허청구범위에 의해 정해져야 할 것이다.

도면

도면1


100
문자 텍스트
[Web발신]상품 거래번호는 틀려서 보낼수 없습니다
앱 다운로드 다시 확인해주세요
hxxps://tinyurl[.]com/xxxxxx
[국제발신][아마존 Pay] 승인번호 *38 ***님, USD 860$
결제완료 본인 아닐시 즉시 신고 요망 070-7678-****
문자 메시지
전송
사용자 단말기
추출
지식베이스
스미싱 판단
Self Attention
Graph
Attention
Network

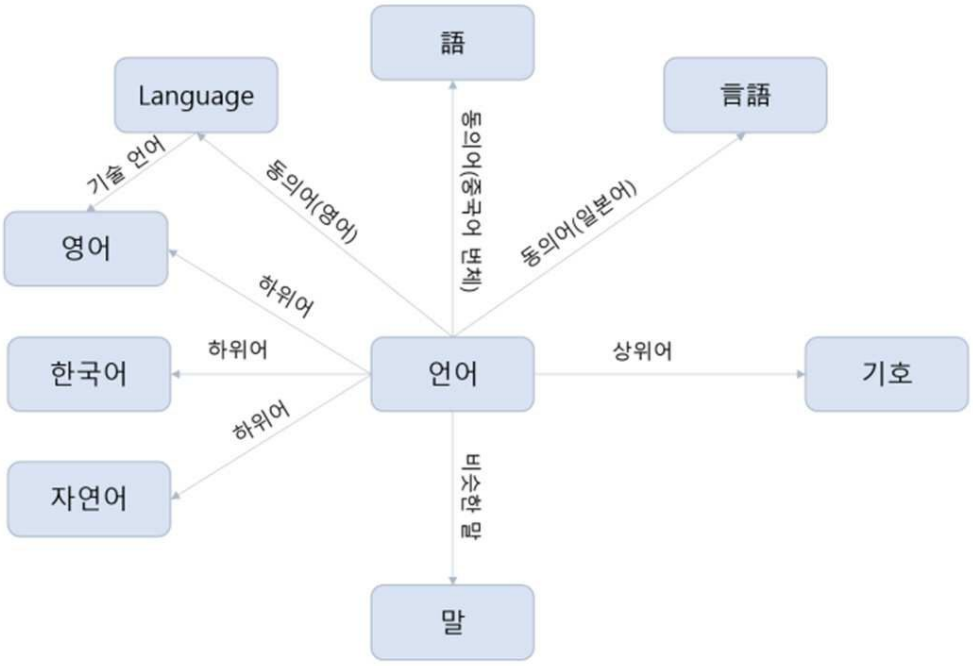
도면2


200
시작
210
문자 메시지를 지식베이스 추출 모델에 입력하여 지식베이스 데이터 추출
220
추출된 지식베이스 데이터를 스미싱 탐지 모델에 입력
230
스미싱 탐지 모델의 출력을 토대로 문자 메시지에 대한 스미싱 여부 판단
끝

도면3

도면4

400


語
Language
言語
기술 언어
동의어(영어)
동의어(중국어 번체)
동의어(일본어)
영어
하위어
한국어
하위어
언어
상위어
기호
하위어
자연어
비슷한 말
말

도면5