소셜 미디어 상의 마약 범죄 추적을 위한 키워드 수집체계 연구

최민재[O], 이호동[*], 김지연[**], 김창훈[**]
[O]대구대학교 컴퓨터정보공학부(컴퓨터공학전공),
[*]대구대학교 컴퓨터정보공학부(컴퓨터공학전공),
[**]대구대학교 컴퓨터정보공학부
e-mail: dopine@naver.com[O], 2118681@naver.com[*], {jyk, kimch}@daegu.ac.kr[**]

A Study on the Keyword Collection System for Tracking Drug Crimes on Social Media

Min-Jae Choi[O], Ho-Dong Lee[*], Jiyeon Kim[**], Chang-Hoon Kim[**]
[O]Dept. of Computer Engineering, Daegu University,
[*]Dept. of Computer Engineering, Daegu University,
[**]Division of Computer and Information Engineering, Daegu University

● 요 약 ●

소셜미디어를 통한 마약 유통이 증가하면서 젊은 연령층의 마약 문제가 사회 문제로 대두되고 있다. 소셜미디어 상의 마약 유통을 차단하고자 유해 키워드 차단 정책이 마련되고 있지만, 사용자들은 차단 정책을 우회할 수 있는 마약 관련 은어 및 변형어를 사용하여 검색을 시도하기 때문에 다양한 은어 및 변형어를 수집하여 차단 정책에 반영하는 것이 필요하다. 본 논문에서는 마약 유통 및 구매에 빈번하게 사용되는 일반어, 은어, 그리고 일반어와 은어의 각 변형어를 수집하고, 이를 페이스북, 인스타그램, 카카오톡, 트위터, 텔레그램을 통해 검색하여 소셜미디어 유형별로 수집되는 마약 관련 게시글을 정량적으로 분석하였다. 분석 결과, 5종의 소셜미디어 중, 트위터에서 마약 관련 은어 및 변형어를 포함하는 게시글이 가장 많이 검색되는 것을 확인하였다.

키워드: 소셜미디어(Social media), 마약 범죄(Drug Crimes), 사이버 수사(Cyber Investigation), 키워드(Keywords), 변형어(Variational Keywords)

I. 서론

트위터, 페이스북 등 소셜미디어 사용자가 증가하면서 마약 유통에 악용되는 사례가 증가하고 있다. 마약 구매자는 간단한 검색으로 다양한 마약 정보를 수집할 수 있고, 다이렉트 메시지 또는 채팅을 통해 단시간에 마약을 구매할 수 있는 상황이다. 마약은 국민의 건강 및 사회 안전을 심각하게 해치는 범죄이기 때문에 소셜미디어를 통한 마약 유통을 차단하고 예방하기 위한 사이버 수사 기술 개발이 필요하다.

본 논문에서는 국내외 대표적인 소셜미디어에서 발생하는 마약 범죄를 추적하기 위한 키워드 수집체계를 제안하고자 한다. 마약 거래자는 키워드 차단 정책 및 사이버 수사 기술을 우회하기 위해 은어 또는 변형어를 주로 사용하기 때문에 다양한 형태의 마약 관련 키워드를 수집하는 것이 필수적이다. 그러나 기존의 마약 수사 연구에서 은어는 다수 고려되었지만 변형어를 고려한 연구는 미비한 실정이다. 따라서 본 논문에서는 마약 관련 일반 키워드 및 은어 뿐 아니라, 각각의 변형어를 사용하여 소셜미디어 상의 마약관련 게시글을 수집하기 위한 효과적인 수집체계를 제안하고자 한다. 본 논문에서는 소셜미디어 5종(페이스북, 인스타그램, 트위터, 텔레그램, 카카오톡)을 분석 대상으로 선정하고, 제안된 수집체계 기반으로 실제 마약 게시글을 수집하여 소셜미디어 별 마약 추적에 효과적인 변형어를 최종 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구 동향을 살펴보고, 3장에서 마약관련 일반어, 은어, 변형어 정의를 통해 소셜미디어에서 마약 게시글을 수집하기 위한 체계를 제시한다. 4장에서는 제시한 모델을 기반으로 실제 분석을 수행하여 각 소셜미디어에서 수집된 게시글을 정량적으로 비교한다.

II. 관련 연구

기존에 수행된 사이버 수사연구는 검색 엔진에서 범죄 관련 웹 페이지를 수집하는 연구, 소셜미디어에서 범죄 게시글을 수집하는 연구, 다크웹에서 범죄 데이터를 수집하는 연구 등이 존재한다.

검색 엔진을 활용하는 경우는 불법 공유사이트, 유해사이트 등을 크롤링을 통해 수집하는 연구를 진행하였고[1], 소셜미디어를 활용하는 연구는 성범죄자 증거수집 연구, 한국형 프로파일링 기술개발 연구가 존재한다[2-3]. 다크웹에서 범죄를 추적하는 연구로는 PGP(Pretty Good Privacy) 공개키 기반으로 범죄단서 간의 연관성 분석을 수행하는 연구가 존재하고[4], 마약 수사 연구 방법으로는 텔레그램 상의 범죄집단을 규제하기 위한 방법론, 마약류 관리법 정비방안에 대한 제안 연구 등도 수행되었다. 효과적인 마약 수사를 위해서는 정보 수집을 위한 유효 키워드를 다수 확보하는 것이 필수적이며 본 논문에서는 소셜미디어 상의 마약 추적을 위한 다양한 일반어, 은어, 변형어를 정의하고, 대표적인 소셜미디어에서의 사용정도를 분석하고자 한다.

III. 소셜미디어 마약 게시글 키워드 수집체계

본 논문에서는 그림 1과 같은 과정으로 소셜미디어상의 마약 거래 게시물을 수집하기 위한 모델을 제시한다. 첫 번째 단계에서는 소셜미디어 게시물을 검색하기 위한 키워드를 정의하는 단계로서 마약과 관련된 일반어, 은어, 그리고 각 변형어를 정의한다. 일반어에 해당하는 '마약'은 소셜미디어 키워드 차단 정책을 우회하기 위해 '아이스', '작대기' 등의 은어를 사용하기도 하지만, 'ㅁr약'과 같은 변형어로 검색하는 사례도 빈번하다. 또한, 은어 또한 'ㅇroi스' 또는 '작ㄷㅐ기'와 같이 변형된 키워드가 존재하기 때문에 다양한 변형어를 확보하여 마약 게시글을 수집하는 것이 필요하다.

Fig. 1. 마약 관련 검색 키워드 수집체계

다음 단계는 소셜미디어에서 데이터를 수집하는 단계로서 본 논문에서는 인스타그램, 카카오톡, 페이스북, 트위터, 텔레그램에서 전 단계에서 정의한 키워드를 사용하여 게시글을 수집하였다.

IV. 소셜미디어 마약 게시글 검색 및 결과 분석

본 논문에서는 표 1부터 표 4와 같이 소셜미디어 5종에서 일반어 10종, 일반어의 변형어 12종, 은어 14종, 은어의 변형어 12종을 사용하여 마약 관련 게시글을 수집하였다.

일반어의 변형어 경우도 같은 환경에서 검색하였고 "ㅁr약", "LS디", "엘s디", "대ㅁr", "ㄷHㅁr" 등 총 10종의 일반어의 변형어를 검색하여 수집한 게시글의 수는 표 1과 같다. 소셜미디어별 일반어 검색 게시글 수는 트위터와 텔레그램이 각각 182건 이상, 165건 이상으로 가장 많이 검색되었다. 게시글 수가 일정 건수 이상인 경우, 정확한 수치가 확인되지 않기 때문에 '+(기준수치 제공)' 및 '∞(기준수치 미제공)' 기호로서 추가 게시글이 존재함을 표현하였다.

한가지 일반어에 집중된 인스타그램에 비해 트위터와 텔레그램에서는 다양한 마약 일반어로 게시글이 검색되고 있음을 알 수 있다.

Table 1. 소셜미디어 유형별 일반어 검색 게시물 수 (FB:페이스북, Insta:인스타그램, Kakao:카카오톡, Twitter:트위터, Tele:텔레그램)

키워드	소셜미디어				
	FB	Insta	Kakao	Twitter	Tele
마약	0	0	0	5	20+
LSD	0	0	0	10	20+
GHB	0	0	0	∞	20+
대마초	0	0	0	14	20+
마리화나	2	0	0	34	20+
필로폰	0	0	0	32	20+
메스암페타민	0	0	0	0	10
엑스터시	5	0	0	63	20+
케타민	0	166	0	3	20+
펜타닐	0	0	0	4	12
계	7	166	0	165+	182+

표 2는 일반어를 변형한 형태의 키워드로 마약 관련 게시글을 검색한 결과를 보여준다. 트위터를 제외한 모든 소셜미디어에서 정의된 일반어의 변형어가 발견되지 않았으며 트위터에서도 '대ㅁr'를 포함하는 게시글이 총 5건 검색된 것을 확인할 수 있다.

Table 2. 소셜미디어 유형별 일반어의 변형어 검색 게시물 수 (FB:페이스북, Insta:인스타그램, Kakao:카카오톡, Twitter:트위터, Tele:텔레그램)

키워드	소셜미디어				
	FB	Insta	Kakao	Twitter	Tele
ㅁr약	0	0	0	0	0
LS디	0	0	0	0	0
엘s디	0	0	0	0	0
대ㅁr	0	0	0	5	0
ㄷH마초	0	0	0	0	0
대ㅁr초	0	0	0	0	0
malihwana	0	0	0	0	0
필lo폰	0	0	0	0	0
math암페타민	0	0	0	0	0
x터시	0	0	0	0	0
케ㅌr민	0	0	0	0	0
펜ㅌr닐	0	0	0	0	0
계	0	0	0	5	0

표 3은 정의된 은어를 사용하여 마약 관련 게시글을 검색한 결과를 보여준다. 소셜미디어별 은어 검색 게시글 수는 트위터가 가장 많았고, 인스타그램도 총 2,466건 검색된 것을 볼 수 있다.

Table 3. 소셜미디어 유형별 은어 검색 게시물 수 (FB:페이스북, Insta:인스타그램, Kakao:카카오톡, Twitter:트위터, Tele:텔레그램)

키워드	소셜미디어				
	FB	Insta	Kakao	Twitter	Tele
아이스	0	283	0	∞	20+
시원한술	21	283	0	∞	20+
차가운술	22	283	0	∞	20+
작대기	13	283	0	∞	20+
콕	0	166	0	0	16
허브	0	166	0	∞	20+
떨	0	283	0	∞	20+
얼음	0	283	0	∞	20+
빙두	28	166	0	40	20+
캔디	1	266	0	1	20+
히로뽕	0	0	0	3	9
가루	0	0	0	0	11
밀가루	0	0	0	0	1
물뽕	3	4	0	1	20+
계	88	2466	0	45+	237+

표 4는 은어를 변형한 형태의 키워드로 마약 관련 게시글을 검색한 결과를 보여준다. 일반어의 변형어 검색 결과와 유사하게 트위터를 제외한 소셜미디어에서는 해당 키워드를 포함한 게시글이 검색되지 않았고, 게시글에서 검색된 은어의 변형어는 “ㅇrㅇI스”, “ㅇrㅇIs”, “짝ㄷH긔”, “떠ㄹ” 총 4개 종류이다.

Table 4. 소셜미디어 유형별 은어의 변형어 검색 게시글 수 (FB:페이스북, Insta:인스타그램, Kakao:카카오톡, Twitter:트위터, Tele:텔레그램)

키워드	소셜미디어				
	FB	Insta	Kakao	Twitter	Tele
아e스	0	0	0	0	0
ㅇre스	0	0	0	0	0
ㅇrㅇ이스	0	0	0	1	0
아ㅇls	0	0	0	1	0
작ㄷH기	0	0	0	0	0
짝ㄷH긔	0	0	0	1	0
캔ㄷI	0	0	0	0	0
크ㄹ이스탈	0	0	0	0	0
ㅎ이로뽕	0	0	0	0	0
떠ㄹ	0	0	0	1	0
ga루	0	0	0	0	0
가ru	0	0	0	0	0
계	0	0	0	4	0

Ⅴ. 결론 및 향후연구

본 논문에서는 마약 거래에 사용되는 키워드를 정의하고, 소셜미디어 검색을 통해 마약 게시글을 수집하는 연구를 수행하였다.

소셜미디어 5종에 본 논문에서 수집한 마약관련 일반어, 은어, 변형어를 검색한 결과, 은어가 마약 유통 및 구매를 위해 가장 많이 사용되는 것을 확인하였고, 트위터에서 가장 많은 마약 거래 게시글이 검색되는 것을 알 수 있었다.

본 연구결과는 소셜미디어의 마약 예방 정책을 마련하는 데에 활용될 수 있으며 향후에는 능동적으로 은어 및 변형어를 수집하여 마약 유통 정보를 수집 및 추적하는 고성능의 사이버 수사 크롤링 기술을 개발할 계획이다.

ACKNOWLEDGEMENT

이 논문은 과학기술정보통신부·경찰청이 공동 지원한 ‘폴리스랩 2.0사업(www.kipot.or.kr)’의 지원을 받아 수행된 연구결과입니다. [과제명: 다크웹 범죄 예방을 위한 능동형 다크웹 정보 수집 및 분석·추적 기술 개발 / 과제번호 : RS-2023-00244362]

REFERENCES

[1] Seungyong Choo, Yeseong Hwang, Sangjin Lee.(2021). Methods for Collecting Harmful Websites Using Web Crawling.Journal of Digital Forensics ,15(3),127-138.

[2] Hyun-Jeong Kim, Ka-Eun Park, Ye-Won Sohn, Joongjin Kook.(2020).Digital Sexual Crime Evidence Collection

System Using SNS Crawl. KEEE Workshop,123-125.
[3] SuGyeong Lee, Choi Eun-jung, JIYEON KIM, Insoo Lee, Seunghoon Lee, Kim Myuhng Joo.(2022).A Study on the Expansion of Workflow for the Collection of Surface Web-based OSINT(Open Source Intelligence). Journal of Digital Convergence, 20(4),367-376.
[4] Jae-Jin Kim.(2019).Dark Web Crime Investigation Using PGP Public Keys Cues, Journal of Digital Forensics, 13(4),219-230.