

# 인스타그램 프로필 사진 유사도 분석을 통한 마약 유통 계정 추적 기술

박은영<sup>○</sup>, 조경현<sup>\*</sup>, 김지연<sup>\*\*</sup>, 김창훈<sup>\*\*</sup>

<sup>○</sup>대구대학교 컴퓨터정보공학부(컴퓨터공학전공),

<sup>\*</sup>대구대학교 일반대학원 컴퓨터정보공학과,

<sup>\*\*</sup>대구대학교 컴퓨터정보공학부

e-mail: pey6693@daegu.ac.kr<sup>○</sup>, jgh1118@daegu.ac.kr<sup>\*</sup>, {jyk, kimch}@daegu.ac.kr<sup>\*\*</sup>

## Tracking Drug Distribution Accounts Through Similarity Analysis of Instagram Profile Photos

Eun-Young Park<sup>○</sup>, Kyeong-Hyun Cho<sup>\*</sup>, Jiyeon Kim<sup>\*\*</sup>, Chang-Hoon Kim<sup>\*\*</sup>

<sup>○</sup>Dept. of Computer Engineering, Daegu University,

<sup>\*</sup>Dept. of Computer and Information Engineering, Graduate School, Daegu University,

<sup>\*\*</sup>Division of Computer and Information Engineering, Daegu University

### ● 요약 ●

국내·외 소셜미디어 사용자가 증가하면서 마약 유통, 불법 촬영물 유포 등과 같은 사이버 범죄가 소셜미디어를 통하여 발생하고 있다. 사이버 범죄 수사를 위해 소셜미디어 크롤링 연구가 진행되고 있지만, 주로 ‘N번방’ 등 불법 촬영물 및 성 착취물 유포와 같은 성범죄 수사를 대상으로 한다. 그러나 최근에는 성범죄뿐만 아니라, 소셜미디어를 통한 마약 유통이 급격히 증가하고 있으므로 소셜미디어 크롤링을 통한 마약 수사 기술 개발이 필요하다. 본 논문에서는 소셜미디어 중, 인스타그램의 마약 유통을 추적하기 위해 실제 사용되는 마약 은어를 정의하고, 정의된 은어를 검색 키워드로 입력하여 사용자 계정을 수집하였다. 또한, 수집된 사용자 계정의 프로필 사진을 추출하고 유사도 분석을 수행하여 동일 마약 유통자 식별에 프로필 사진이 효과적임을 검증하였다.

**키워드:** 사이버 수사, 마약, 인스타그램(Instagram), 크롤링(Crawling), 소셜미디어(Social Media)

## 1. 서론

전 세계 소셜미디어(Social Media) 사용자가 증가하면서 마약 유통, 불법 촬영물 유포 등과 같은 사이버 범죄에 악용되고 있다. 국내의 경우에도 소셜미디어를 사용한 10대·20대 마약 사범이 급격히 늘어나고 있으며 전체 마약 사범은 전년 대비 약 29.7% 증가한 것으로 조사되었다[1]. 소셜미디어는 접근성이 매우 높으므로 마약 사범의 평균 연령이 지속적으로 낮아지고 있는 추세이지만, 소셜미디어에서 행해지는 마약 유통 행위를 제재하는 것이 현실적으로 어려우므로 범죄자들은 소셜미디어의 게시물 또는 다이렉트 메시지(Direct Messenger, 이하 DM)를 통해 마약 복용자와 쉽게 소통하는 것이 가능하다.

소셜미디어에서 생성된 방대한 양의 데이터를 자동으로 크롤링하여 마케팅 또는 사회현상을 분석하기 위한 목적으로 수행된 연구들은 다수 존재하지만, 사이버 범죄를 추적하기 위한 소셜미디어 연구는 상대적으로 부족한 실정이다. 특히, 사회적으로 큰 이슈가 되었던 ‘N번방’과 같은 불법 촬영물 추적 연구는 최근에 다수 수행되었지만,

최근 급격히 증가하고 있는 마약 유통을 추적하기 위한 연구 생태는 미비한 실정이다. 본 논문에서는 대표적인 소셜네트워크 서비스인 인스타그램(Instagram)에서 사용자 계정 분석을 통해 마약 유통자를 추적하는 크롤링 기술을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 다양한 분야에 활용되는 소셜미디어 크롤링 연구 동향 및 사이버 범죄 수사를 위해 기존에 수행된 관련 연구를 살펴본다. 3장에서는 인스타그램에서 마약 판매 계정을 수집하기 위하여 검색 키워드로 사용할 은어를 선정하고, 선정된 키워드를 사용하여 인스타그램에서 마약 계정을 확보하기 위한 마약 유통 계정 추적 모델을 설계한다. 4장에서는 설계된 추적 모델을 기반으로 실제 인스타그램 크롤러를 개발하여 마약 계정을 수집하고, 마약 계정 간의 유사도를 판단하여 동일 마약 유통 계정 식별 결과를 분석한다. 마지막으로 5장에서는 결론 및 향후 연구를 제시한다.

## II. 관련 연구

본 장에서는 기존에 수행된 소셜미디어 크롤링 연구 분야 및 기술 동향을 살펴본다. 인스타그램 크롤링 연구는 주로 비즈니스 계정 중심의 마케팅, 인플루언서에 대한 연구와 사회적 문제를 분석하는 연구가 진행되었다. 비즈니스 마케팅에는 브랜드 시장의 계층적 군집 분석을 통해 공통 키워드를 도출하고, 도출한 키워드에 대한 연관 규칙 분석(Association rule analysis), 감성 분석(Sentiment analysis) 및 중심성 분석(Centrality analysis)을 수행하여 마케팅 효과 및 산업 기업에 시사점을 제공하는 연구가 수행되었다[2]. 사회적 문제를 다루는 연구로는 자살 자해 키워드 크롤링을 수행하고 수집된 데이터를 텍스트 마이닝 기반으로 분석하여 부정적인 사회 문제를 해결하고자 하였다[3]. 또한, 소셜미디어 크롤링을 통해 디지털 성범죄, 보이스 피싱(Voice Phishing) 등에 사용되는 범죄 조직의 은어를 활용하여 크롤링을 수행하고, 범죄 피해 예방을 위한 특징을 제시하였다[4].

## III. 마약 유통 계정 추적 모델 설계

인스타그램에서 마약을 유통하는 동일 계정을 추적하기 위하여 본 논문에서는 Fig. 1과 같이 4단계에 걸쳐 데이터 수집 및 분석을 수행한다.

검색에 사용할 키워드를 정의하는 1단계에서는 ‘마약과 같은 표준어 대신 소셜미디어에서 마약 거래를 위해 일반적으로 사용하는 은어를 선정한다. 소셜미디어에서는 유해 키워드 차단 정책 또는 사이버 수사망을 우회하기 위하여 주로 은어를 사용하여 마약 거래가 이루어지고 있으므로, 실제 거래에 사용되는 은어를 검색 키워드로 사용할 필요성이 있다

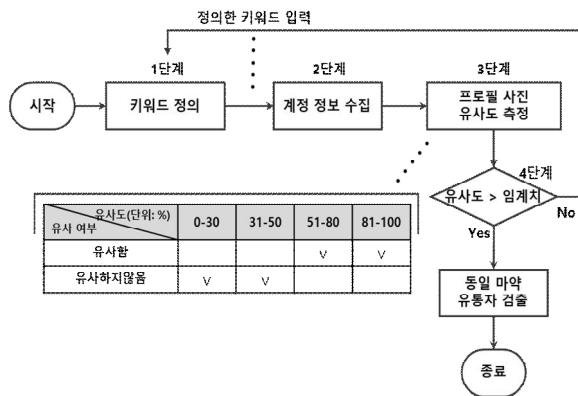


Fig. 1. 마약 유통 계정 추적 모델

본 논문에서는 ‘마약의 대표적인 은어로서 Table 1과 같은 ‘작대기’, ‘차가운 술’, ‘히로뽕’을 선정하였으며 해당 은어를 활용하여 인스타그램 계정을 수집하기 위한 검색어로 사용한다.

Table 1. 마약 유통 계정 추적을 위한 검색 키워드 정의

키워드	의미
작대기	주사기를 통해 투약하는 필로폰
차가운 술	북한산 필로폰
히로뽕	메스암페타민

2단계에서는 상기 3개의 은어를 검색 키워드로 사용하여 인스타그램의 사용자 계정을 크롤링하고, 키워드에 따른 계정 정보를 수집한다. 3단계와 4단계에서는 계정 정보 수집 결과 중 프로필 사진을 유사도 측정을 위한 계정 요소로 선정하고, 유사도를 백분율로 표현하여 유사도가 높은 경우 동일 마약 유통자라고 판별한다.

## IV. 마약 유통 계정 크롤링 및 결과 분석

본 논문에서는 인스타그램 크롤러를 직접 개발하여 3장에서 정의한 마약 관련 검색 키워드로 사용자 계정을 총 91건을 수집하였다. 이 중 마약 판매 계정과 관련된 총 17건의 계정을 키워드별로 수집하였을 때, 사용자 계정 현황은 Table 2와 같다.

Table 2. 검색 키워드별 마약 유통 계정 수집 결과

키워드	총 건수	동일 유통자 수
작대기	8	2
차가운 술	3	1
히로뽕	6	2

본 논문에서는 ‘차가운 술’ 키워드에 대한 예시를 토대로 결과를 분석한다. Fig. 2는 본 연구에서 직접 개발한 크롤러를 사용하여 수집한 마약 유통 계정 목록을 보여준다.

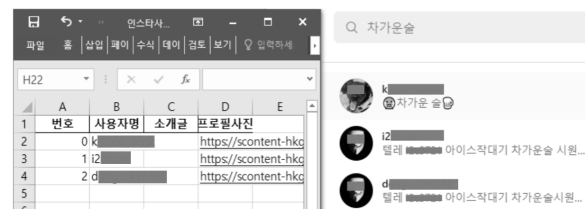


Fig. 2. ‘차가운 술’ 키워드 기반 크롤링 검색 결과

Fig. 3은 수집된 사용자 계정에서 추출한 프로필 사진을 보여주며 본 논문에서는 동일 마약 계정을 판별하기 위한 이미지 유사성 분석 알고리즘으로 절대 오차 알고리즘(Absolute Error Algorithm)을 사용한다. 절대 오차 알고리즘은 결과 데이터 이미지의 픽셀값 절대 차이를 측정하여 두 사진 간의 유사도를 분석한다.



Fig. 3. 마약 유통 계정에서 추출한 프로필 사진

본 논문에서는 픽셀 간의 차이(민감도)를 상(10)·중(5)·하(0) 단계로 설정하여 알고리즘을 실행하였으며 절대 오차 알고리즘을 활용하여 사진간 유사성을 분석하였다.

Table 3은 Fig. 3에서 제시한 세 개의 이미지 데이터를 {이미지 1, 이미지 2}, {이미지 2, 이미지 3}, {이미지 1, 이미지 3}의 세 개의 그룹으로 분류하여 비교하였을 때 민감도에 따른 유사도와 오차 범위를 나타낸다. {이미지 1, 이미지 2}를 비교한 경우, 민감도가 상(10)일 때의 유사도는 60.3%로 민감도를 상승시켰음에도 다른 데이터인 {이미지 2, 이미지 3} 및 {이미지 1, 이미지 3}의 비교 결과보다 오차 범위가 적은 것을 확인할 수 있다. 또한, 민감도를 각 중(5), 하(0)로 설정하였을 때는 유사도가 99.91%, 100%로, 두 이미지 데이터의 오차범위가 매우 낮고 유사도가 높다는 것을 알 수 있다.

Table 3. 민감도에 따른 프로필 이미지 유사도 분석

비교		프로필 이미지								
		1	2	3	1	2	3	1	2	3
민감도 상(10)	오차 범위									
	유사도(%)	60.36	0	0						
민감도 중(5)	오차 범위									
	유사도(%)	99.91	2.33	2.38						
민감도 하(0)	오차 범위									
	유사도(%)	100	7.45	7.47						

또한, {이미지 2, 이미지 3}의 민감도를 상(10)·중(5)·하(0)로 나누어 결과를 확인하였을 때, 각각의 유사도가 0%, 2.33%, 7.45%로 매우 낮은 값이 추출되며 {이미지 1, 이미지 3}의 민감도를 구분하여 유사도 분석을 한 결과 또한, 0%, 2.38%, 7.47%로 {이미지 2, 이미지 3}과 유사하게 낮은 값을 확인할 수 있다.

## V. 결론 및 향후 연구

소셜미디어를 통한 사이버 범죄가 확산됨에 따라 범죄 예방 및 대응을 위한 사이버 수사 기술 개발이 필요하다. 본 논문에서는 소셜미디어상의 마약 범죄를 추적하기 위한 기술로서 인스타그램 크롤링을 통해 마약 관련 계정을 수집하고, 사용자 계정의 프로필 사진 유사도 분석을 통해 동일 마약 유통자를 판별하는 수사 기술을 개발하였다. 본 논문에서 개발한 모델을 활용하여 마약 유통 계정으로 의심되는 프로필 사진을 수집하고, 프로필 사진 간의 유사도를 분석한 결과, 동일 마약 유통 계정의 경우에는 프로필 사진의 유사도가 높은 것을 확인할 수 있었다.

본 연구는 마약뿐 아니라, 소셜미디어를 통한 다양한 사이버 범죄를 추적하는 수사 기술로 활용될 수 있으며 향후에는 프로필 사진 외에도 사용자 소개 글, ID, 게시물 등의 요소를 추가로 수집하여 고성능으로 마약 유통자를 추적하는 크롤링 기술을 개발할 계획이다.

## ACKNOWLEDGEMENT

이 논문은 과학기술정보통신부·경찰청이 공동지원한 ‘폴리스랩2.0 사업(www.kipot.or.kr)’의 지원을 받아 수행된 연구 결과입니다. [과제명: 다크웹 범죄 예방을 위한 능동형 다크웹 정보 수집 및 분석·추적 기술 개발 / 과제번호 : RS-2023-00244362]

## REFERENCES

- [1] THE WOMEN'S NEWS, <https://www.womennews.co.kr/news/articleView.html?idxno=237655>, 2023.
- [2] Y.J Choi and S.C Lee, "Investigating of user opinions on virtual influencer based on instagram post, comment", Proceedings of the 2022 Autumn Conference on Korean Institute of Industrial Engineers, 2073-2078, 2022.
- [3] S.H Park and K.L Yu, "Analysis of instagram Posts Related to Self-Injury and Suicide Using Text Mining", Journal of the Korea Counseling And Psychotherapy 33.3, 1429-1455, 2021.
- [4] H.J Moon, H.M Park and K.B Kim, "A Study of Slang of Voice Phishing Criminal Organization and Practical Use for Investigation", The journal of Police Science, 21(4), 137-160, 2021
- [5] J.H Lee, "Building an SNS Crawling System Using Python", Journal of the Korea Society Industrial Information System 23.5, 61-76, 2018.