인과관계 정보 구성 체계를 활용한 디지털 증거 식별/분류 연구

정종진 [1], 박종빈 [2], 김경원 [3], 이지현 [4]
[1, 2, 3] 한국전자기술연구원 정보미디어연구센터

mozzalt@keti.re.kr, jpark@ keti.re.kr, kwkim@ keti.re.kr, jhlee@ keti.re.kr

Digital Evidence Identification/Classification Study Using Causal Information Organization System

[1, 2, 3]Korea Electronics Technology Institute

요 약

본 논문에서는 디지털증거 분석을 위해 확보한 증거파일 들로부터 범죄 정황에 해당하는 단어 및 어휘를 추출하여 해당 범죄를 인과관계 분석을 하기 위해 핵심 단서와 원인을 효과적으로 파악하기 위해 필요한 인과정보를 제안한다. 이 정보들은 개체명 인식 및 분류를 할 수 있도록 구성되어 범죄 관계인, 관계인간 관계, 범죄 수법과 범죄관련 정보를 추출하고 유형화하여, 향후 해당 범죄에 대한 인과 분석 기법을 활용한 범죄 예방 분석과 수사에 기여할 수 있도록 도움을 준다.

1. 서론

최근 살인강도 등 전통적 개념의 범죄는 감소하고 있으나 비대면경제의 성장과 디지털문화의 확산으로 사이버범죄는 폭증하고 있는 추세이다. 201 년 약 11 만건이던 사이버범죄는 2020 년에 들어서면서 23 만건을 넘었으며, 다크웹, 가상자산, 딥페이크 등 과학기술을 악용하는 신종 수법의 사이버 범죄도 증가하고 있다. 특히 SNS, 금융거래 등 다양한 디지털 증거 분석이 필요한 사이버 사기/금융범죄 등 민생 침해범죄는 2022 년 기준 전년대비 9.3% 증가하기도 했다.습니다. 디지털 소통 채널이라는 공간 특성상 관계인간 짧은 대화, 은어등 검찰/경찰등의 범죄수사기관에 보고되지 않는 범죄 관련 정보가 급격하게 늘어남에 따라, 이를 효과적으로 인식하여 범죄학 측면에서 해석할 필요성이 커지고 있는 상황이다. 또한 대부분 구조화되어 있지 않은 범죄 데이터 특성상 다양한 목적에 의해 파생되는 범죄 예측·분류의 범위와 정확도는 원천 데이터에서 추출 가능한 범죄 정보의범위와 정확도에 의존할 수 밖에 없게 된다.[1] 따라서 본 연구에서는 사이버 공간에서 일어나는 디지털범죄 증거에서 범죄 정황에 해당하는 단어 및 어휘를 추출하여 해당 범죄를 인과관계 분석을 하기 위해 핵심 단서와 원인을 효과적으로 파악하기 위해 필요한 인과정보 구성체계를 제안한다.

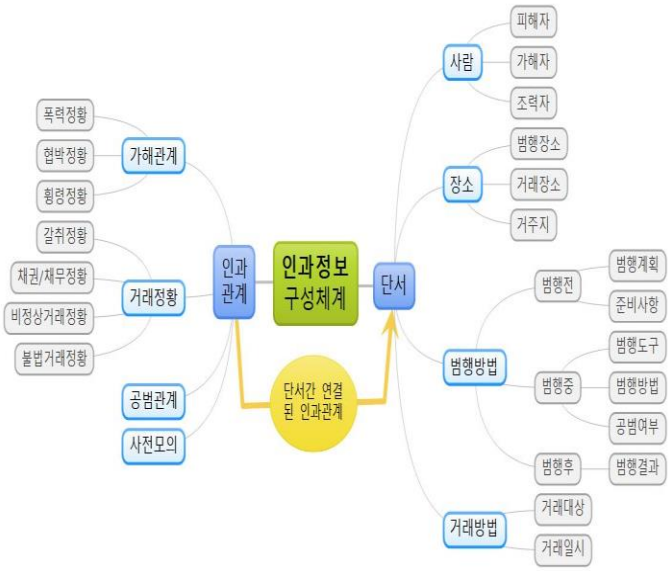
2. 관련연구(디지털 범죄 분석 정보 추출)

디지털 공간에서 벌어지는 범죄 상황들은 다양한 매체,채널에서, 당사자들 만이 알 수 있는 정황이 담긴 짧은 내용으로 이루어지다 보니 국내에서는 이런 정보를 추출하는 선행 연구들이 많지는 않지만 대략은 뉴스·블로그·SNS 등의 텍스트에서 범죄 정보를 추출하여 마약 및 성범죄·사이버범죄 등의 정보를 추적

하는 연구가 진행되었다.[2-5]. 이런 범죄종류의 수사들은 해당 범죄에서 사용되는 특수한 어휘와 용어들을 올바르게 해석하여 인과추론을 해야 하므로 사전에 구축된 어휘사전, 분류 규칙등의 룰(규칙)기반 분석이 주로 연구되어 오다가 최근에는 자연어처리, 사전학습언어 모델 기술이 발달하면서 개체명 분류 모델 등을 활용한 딥러닝 기반의 신경망 분석 연구가 활발히 진행중에 있다.[6]

3. 관련연구(디지털 범죄 분석 정보 추출)

국내 검찰과 경찰에서 디지털범죄 수사를 하면서 디지털 증거들에서 유의미한 범죄 단서를 추출하기 위해 공식적으로 사용하는 인과정보 체계는 존재하지 않는다. 하지만 본 연구에서는 다수의 디지털 범죄 수사 담당자들의 설문을 통해 유의미만 범죄 정황을 찾기 위해 관심있게 살피며 찾는 단서와 그 원인들을 종합 정리하여 그림 1 과 같은 인과관계 구성 정보를 구축하여 인과추론 분석에 활용하였다.



(그림 1) 디지털증거 분석용 인과관계 정보 구성 스키마.

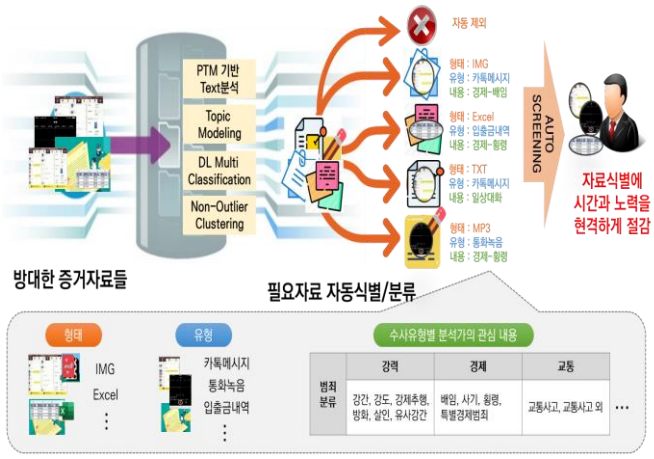
4. 멀티모달 디지털증거 대상 식별 및 분류

주요 범죄 용의자가 소지한 디지털 기기로부터 압수수색을 통해 디지털 증거를 확보할 시점에는 문서, 음성기록 등등 잠재적 범죄혐의가 있을 듯한 가능한 모든 기록을 확보한다. 하지만 확보된 증거들 중 상당수는 수사와 관련성 없는 파일이거나, 하나의 파일안에도 일상 대화 등 범죄혐의와 관련성이 없는 불필요한 부분들이 많다. 따라서 수사와 관련성 없는 정보를 제외하여 사건과 관련 있는 정보들만으로 인과 분석하여 수사역량 집중 유도할 수 있어야 한다.

<표 1> 주요 증거별 주요특징 및 유형.

증거 유형	주요 특징	정보 유형
메신저	많게는 1 인당 100 만개 이상 대화내역 존재	Text(대화,줄임말 등)
이메일	수신, 참조, 숨은참조 등 중복 이메일 다수 존재	Text
문서	사건과 관련없는 파일 등이 대부분, 분류 및 발췌에 너무 많은 시간이 소요됨	Text, Image
음성 파일	화자식별 및 확인에 많은 시간 필요	Audio
캡쳐된 웹/문서	자동화된 분석 방법이 없어 사람이 직접 확인	Text 포함된 Image

담당 수사와 관련성 높은 증거 및 핵심내용 분류/식별은 그림 2 와 같이 수사 유형별 관련 정보분류/정제에 대한 사전지식/모델을 통해 수행된다.

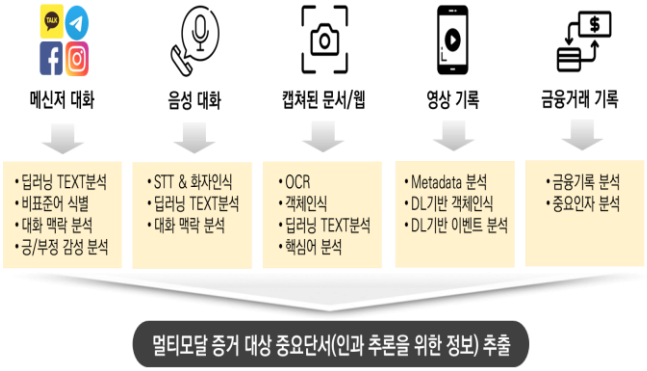


(그림 2) 디지털증거 자동 분류 및 식별 과정.

5. 식별된 멀티모달증거들로부터 인과정보 추출

디지털매체와 디지털 소통 수단이 많아 지면서, 범죄 정황이 담긴 증거 유형은 오디오, 비디오 형태로 다양하게 존재한다. 수사현장에 근무하는

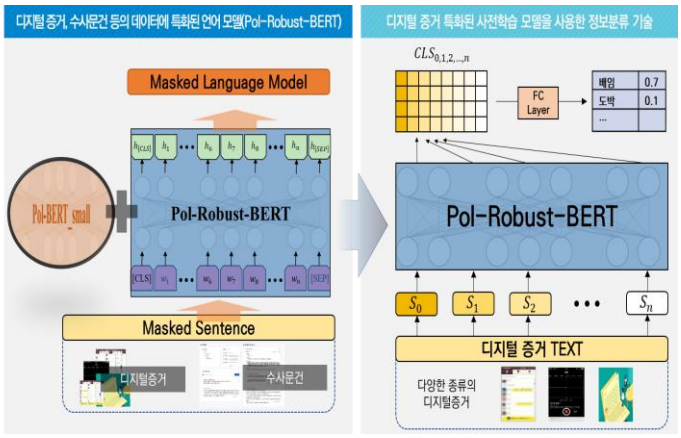
분석가들이 이야기하는 것도 문서, 메신저 대화등 Text 형태로 존재하는 증거 뿐만 아니라 용의자간 음성대화녹음, 캡쳐된 웹자료, CCTV 영상등에 범죄혐의가 담긴 내용들이 많이 존재하고 이들을 모두 자동 분석하여 인과관계 자동 연결 분석에 대한 수요가 많다. 그림 3 은 다양한 형태의 증거 유형으로부터 중요 인과정보를 추출하기 위해 구조화된 Text 정보를 추출하고, 추출된 단어와 문장들 로부터 그림 1 에서 정의한 인과관계 구성정보에 맞춰 유의미한 범죄 해석 정보로 처리하는 과정을 개념적으로 설명한다.



(그림 3) 멀티모달 증거 대상 구조화된 인과정보 추출 개념도.

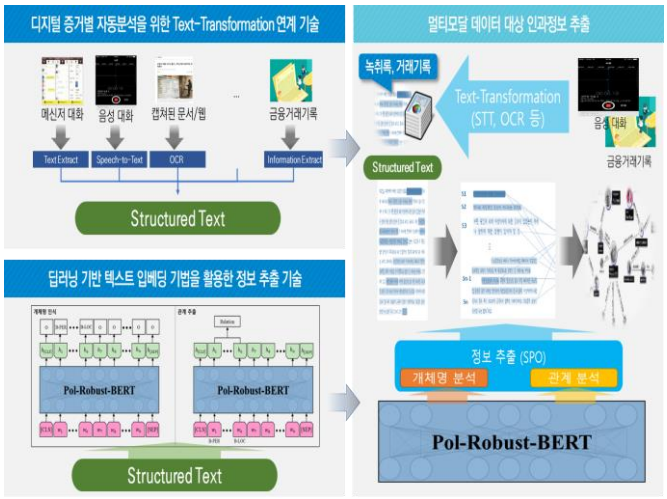
6. 인과정보 추출에 필요한 사전학습언어모델

그림 2 처럼 확보한 디지털증거들을 수사중인 사건특성 관련성 높은 것들만 식별/분류를 하기 위해서는 치안 도메인에 특화된 일반 범죄 관련 사전학습 언어모델인 Pol-BERT_small 에 디지털 증거, 수사문건/자료 데이터를 추가학습하여 Fine-tunning 하여 특화된 언어모델(Pol-Robust-BERT)을 필요하다. 이 2 개의 모델을 활용하여 Document-to-Sequence 를 이용한 핵심어/주제분석을 통해 개별 증거의 내용을 파악 한 뒤 수사 특성에 맞춰 디지털 증거들을 선별할 수 있다.



(그림 4) 인과정보 추출 위한 사전학습 언어모델.

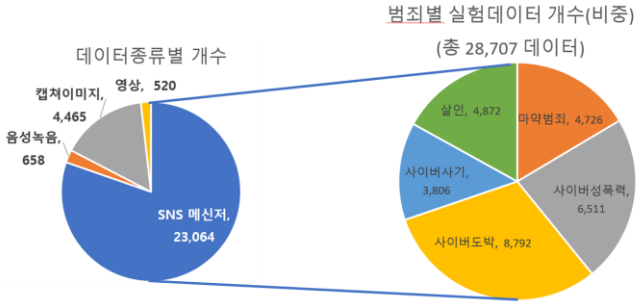
그림 4 에 설명된 2 개의 BERT 모델을 활용하여 선별된 증거 들에는 메신저 대화 캡처 화면 음성대화 녹음, 금융거래 기록 사진 등의 멀티모달 데이터로 이루어져 이를 분석하기 위해서는 그림 5 에서처럼 구조화된 텍스트로 변환된다. 구조화된 텍스트는 심층학습 기반의 개체명/관계 분석 등의 NLP 기술을 활용하여 지식그래프에서 사용될 트리플(SPO) 형태의 정보로 추출된다.



(그림 2) 사전학습언어모델을 활용한 멀티모달 증거대상 인과정보 추출과정.

7. 구현 및 실험

본 연구는 디지털범죄 수사 관련하여 그림 6 의 데이터 분포와 같이 범죄 죄종 4 개별를 대상으로 SNS 메신저, 영상, 음성녹음, 캡처된 이미지 4 종에 대해 총 28,707 개의 학습데이터셋을 구축하였다.



(그림 6) 인과추론 정보 추출 학습모델 구축에 활용한 데이터셋 규모 및 분포

7.1 인과관계 구성 정보 추출 결과

준비된 실험 데이터를 현재 수사중인 범죄 유형에 맞춰 식별하려면 각 데이터셋에 어떤 내용이 담겨있는지를 파악해야 하므로 앞서 그림 1 정의된 인과관계 구성정보로 개체명 분석이 선행되어야 한다. 개체명 분석은 6 장에서 제시된 디지털증거분석에 특화된 사전학습언어모델 (Pol-Robust-BERT)를 활용하여 전체데이터대상 80%를 학습용으로 20%를 평가용으로 구성하여 4 가지 정확도 측정지표 (precision, Recall, Accuracy, F1-Score)를 표 1 과 같이 측정하였다.

(표 1) Pol-Robust BERT 를 활용한 인과관계정보 자동추출 성능

사용모델	Precision	Recall	Accuracy	F1-Score
Pol-Robust-BERT	0.76	0.81	0.87	0.85

7.2 디지털 증거 식별 정확도

실험 7.1 의 수행결과로 디지털 증거들을 대상으로 분석되어 추출된 인과관계 구성 정보들을 활용하여 자동 증거 식별/분류를 수행하여 분류 정확도를 실험하였다. 각 4 개 죄종별로 학습데이터셋과 테스트셋을 8:2 비율로 구성하여 실험 7.1 과 같은 방법으로 표 2 와 같이 도출하였다.

(표 2) 죄종별 증거자동 분류 모델 성능

범죄유형	Precision	Recall	Accuracy	F1-Score
마약범죄	0.87	0.91	0.94	0.95
살인사건	0.84	0.87	0.90	0.91
사이버사기	0.72	0.79	0.81	0.82
사이버도박	0.81	0.77	0.84	0.84
사이버성폭력	0.91	0.92	0.92	0.96
계/평균	0.83	0.852	0.882	0.896

8. 결론

본 연구에서는 범죄 수사에서 확보한 방대하고 다양한 종류의 디지털 증거를 제한된 수사기간내에 적은 수사인력이 효과적으로 범죄 내용을 파악할 수 있도록 딥러닝 기반 자연어 처리 기술 및 사전학습언어모델과 디지털유형별 적합한 내용 추출 분석 방법론을 병행 응용하여 유의미한 인과추론 분석을 할 수 있는 시스템을 제안하였다. 디지털 증거들로부터 중요하게 해석해야 하는 정보를 체계화하여 인과관계 구성정보 스키마 구조를 제시함으로써 규격화된 정보 추출 및 정보 처리를 가능하게 하였고, 이 정보들을 종합 인과분석 하여 담당 수사에 적합한 증거들로 위주로 식별 및 분류가 가능하여 적은 수사인력만으로도 보다 신속하고 정확한 수사가 가능하도록 가능성을 보여 주었다. 7 장 실험 결과는 통상 확보 가능한 증거 유형들에 대해 주요 사이버 범죄를 대상으로 80% 중반대의 모델 정확도 성능을 갖춰 현장 가능성을 기대할 수 있었다. 하지만 사이버사기의 실험예처럼 학습데이터셋이 적거나, 다른 사건들과 추출되는 정보가 유사한 경우 정확도가 다소 낮게 나오는 것을 알수 있었다. 이 결과는 보다 많은 현실의 데이터로 학습하고, 정보체계를 좀더 세밀하게 구체화한다면 성능향상이 가능하며, 이를 통해 다양한 수사관점에 따라 범죄를 면밀히 분석할 수 있는 시스템 개선이 될 것으로 기대된다.

Acknowledgement

이 논문은 2023 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. RS-2023-00225661, 디지털 증거의 증명력 제고를 위한 인과관계 추론 및 표현 기술 개발)

참고문헌

[1] 김희두, 임희석,"사전학습 언어모델을 활용한 범죄 수사 도메인 개체명 인식", 한국융합학회논문지 제 13 권 제 2 호, pp. 13-20, 2022.

[2] K. R. Rahem & N. Omar. (2014). Drug-related crime information extraction and analysis. Proceedings of the 6th International Conference on Information Technology and Multimedia, pp.250-254. DOI : 10.1109/ICIMU.2014.7066639

[3] A. Alkaff & M. Mohd. (2013). Extraction of nationality from crime news.Journal of Theoretical and Applied Information Technology, 54, 304-312.

[4] S. Sathyadevan, M. S. Devan & S. S. Gangadharan (2014). Crime analysis and prediction using data mining. 2014 First International Conference on Networks & Soft Computing (ICNSC2014),406-412.DOI : 10.1109/CNSC.2014.6906719.

[5] M. Asharef, N. Omar & M. Albared. (2012). Arabic named entity recognition in crime documents. Journal of Theoretical and Applied Information Technology, 44(1), 1-6.

[6] P Gohel. (2016) Crime information extraction from news articles. M Tech Dissertations. Dhirubhai Ambani Institute of Information and Communication Technology. Gandhinagar.