

친밀한 관계 내 폭력에 효과적으로 대응하기 위한 반복사건 추론 기술 연구

홍세은

한국여성정책연구원

seeunhong.korea@gmail.com

A Study on the Inference Technology of Recurrent Events for Effective Field Response to Crimes committed in Intimate Relationships

Hong Seeun

Korean Women's Development Institute

요 약

본 논문에서는 인공지능 기술을 기반으로 가정폭력, 데이트폭력, 아동학대와 같이 친밀한 관계 내에서 발생하는 범죄 관련 반복사건을 추론하여 해당 신고에 보다 적극적이고 효과적으로 대응할 수 있도록 지원하는 기술을 마련하고자 하였다. 이를 위해 신고의 지리적 인접성, 신고내용의 유사성 등의 정보를 통해 반복사건을 추론하는 기술을 설계하였다. 이러한 반복사건 추론 기술을 활용하는 경우 친밀한 관계 내에서, 또한 가정 내에서 반복적으로 발생하는 가정폭력, 데이트폭력, 아동학대를 인지할 수 있기 때문에 보다 효과적인 치안활동을 통해 국민의 생명을 보호할 수 있다.

I. 서론

코로나19가 확산됨에 따라 전반적인 범죄의 발생이 감소하였다는 연구 결과가 국내외에서 보고되었다.[1] 그러나 다른 연구들은 가정폭력이나 데이트폭력, 아동학대와 같이 친밀한 관계 내에서 발생하는 폭력이 코로나19 확산 이후 증가하였다는 결과를 보고하였다.[2][3] 또한 한 연구는 코로나19 이후 국내에서 폭력범죄는 감소하였지만 아동학대는 급증하였다는 결과를 보고하기도 하였다.[4]

이처럼 친밀한 관계 내 폭력이 증가하고 있음에도 여전히 가정폭력과 같은 친밀한 관계 내 폭력범죄는 개인적 영역에서 발생한다는 인식 때문에 외부에서 적극적으로 조치를 가하지 않는 경우가 많고 가해자와 피해자의 분리가 확실히 이루어지지 않기 때문에 심각한 경우에는 피해자가 반복된 폭력에 노출되다가 부상을 입거나 사망에 이르는 등 심각한 결과가 초래될 수 있다는 특성이 있는 범죄이다. 따라서 이러한 측면을 보완할 수 있는 기술 개발이 필요하다.

친밀한 관계 내에서 발생하는 범죄는 대개 반복적으로 발생하며 점차 그 심각성이 높아지는 특성이 있기 때문에[5] 신고가 들어왔을 때 해당 신고가 기 신고 건과 동일한 가해자에 대한 신고라는 것을 추론해낼 수 있는 기술이 필요하다. 특히 친밀한 관계 내 폭력의 특성 상 피해자가 신고하지 않고 이웃이나 다른 가족 등 제3자가 신고하는 경우가 있기 때문에 신고 내용, 지리적 인접성 등 다양한 정보를 통해 반복사건을 추론해낼 수 있는 기술의 개발이 필요하다.

따라서 본 논문에서는 가정폭력, 데이트폭력, 아동학대 등에 관한 긴급신고 접수 시 해당 사건이 기존에 신고된 사건과 동일한 반복사건인지 추론하여 반복사건인 경우에는 현장에 출동한 경찰관이 보다 적극적이고 효과적으로 사건처리를 할 수 있도록 지원하는 기술을 설계하고자 한다. 이를 위해 긴급신고 접수 정보, 즉 신고 발생 장소, 신고 내용 등의 정보가 입력되면 인공지능 기반의 유사성 추론 모델을 통해 반복사건 여부를 추론할 수 있는 기술을 제안하고자 한다.

II. 반복사건 추론 기술의 설계

긴급신고 전화가 걸려오면, 접수 일시와 접수 시 사용된 전화번호는 자동으로 기록되고 접수요원은 신고자와 통화하며 사건의 내용과 유형, 사건 발생 장소의 주소 등을 입력하게 된다. 그리고 이러한 정보를 통해 해당 신고가 기존에 신고된 내용과 반복되는 사건인지 추론하는 기술을 설계할 수 있으며 이는 그림 1과 같이 두 가지 단계로 구성될 수 있다.

먼저 접수요원이 신규 접수된 긴급신고의 유형을 '가정폭력', '아동학대', '데이트폭력', '스토킹' 등으로 분류하는 경우, 해당 신고의 전화번호와 일치하는 기존 신고 전화번호를 데이터베이스의 내에서 찾아 해당하는 신고 내용이 자동으로 시스템을 통해 제시되어 접수요원이 확인하는 것이 가능하도록 구축할 수 있다. 또한 신규 접수된 긴급신고의 사건발생 장소의 주소 또는 위경도 정보를 활용하여 기존 신고 데이터베이스 내 사건발생 장소와 정확히 일치하는 기존 신고를 찾아 그 신고내용이 시스템을 통해 접수요원이 확인할 수 있는 형태로 제시되도록 구축할 수 있다. 이때 검색되는 신고의 유형은 '가정폭력', '데이트폭력', '스토킹', '아동학대', '폭력' 등으로 제한하여 보다 명확한 반복사건의 목록을 제시해줄 수 있다.

한편 첫 번째 단계와 같은 방법으로는 이웃이나 지인, 다른 가족 구성원 등, 제3자가 신고한 건은 반복사건일 가능성이 있음에도 누락될 수 있다는 한계가 있다. 예를 들어 신규 접수된 신고사건과, 인근에서 발생했던 신고 중 '소란'으로 분류된 사건이, 실제로는 반복된 가정폭력일 가능성이 있다. 따라서 두 번째 단계로 인공지능 기술을 활용한 자연어처리 기법을 활용하여 반복사건일 가능성이 높은 기존 신고를 알려주는 기술을 설계하여 적용할 필요가 있다.

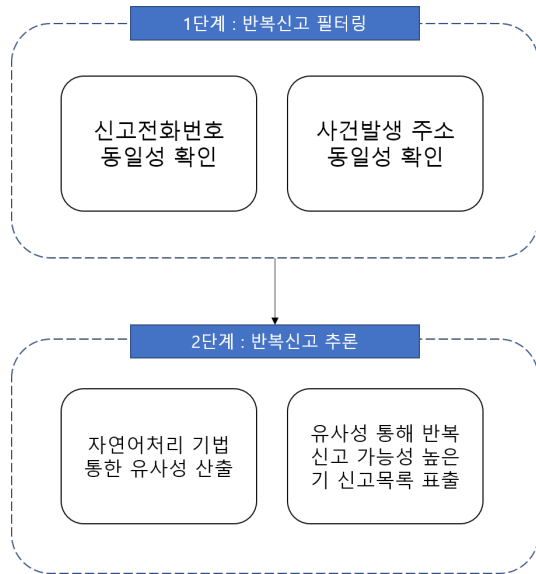


그림 1. 반복사건 추론 단계 예시

인공지능 기반의 자연어처리 기법을 활용하여 반복사건을 추론하는 기술은 법률 관련 자연어처리 연구 분야 중 ‘유사 케이스 매칭’ 연구에 착안하여 구성할 수 있다.[6] 유사 케이스를 찾는 연구는 전통적으로는 단어-단어 유사도를 계산하는 데 초점을 맞추어 이루어졌으나 최근에는 단어-단어 유사도뿐만 아니라 문서의 의미론적(semantic) 유사도를 고려하는 방향으로 이루어졌고, 일부에서는 행렬분해 등의 머신러닝 기법이나[7] CNN, RNN과 같은 딥러닝 기법을 적용하여 문서 유사도를 계산하고자 하였다.[8][9][10] 긴급신고 내 반복사건의 추론 또한 인근에서 신고되었던 사건 중 신규 접수된 사건의 내용과 벡터값이 유사한 신고를 반복사건으로 추론하는 방식을 적용할 수 있다. 즉 문서 임베딩 기반 유사도 측정 모델을 기반으로 근접성 등을 함께 고려하여 신고의 유사성을 산출하고 반복사건 가능성을 추론할 수 있는 기술을 그림 2와 같이 설계할 수 있다.

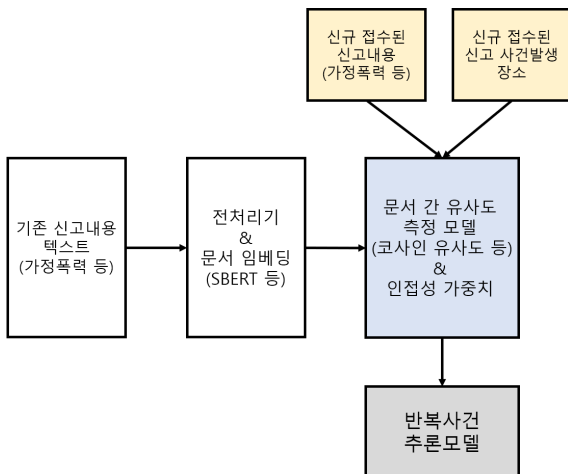


그림 2. 반복사건 추론기 구성 예시

반복사건 추론기의 구성을 살펴보면 먼저 기존의 긴급신고 내용 텍스트 데이터를 전처리에 투입하여 불필요한 단어나 특수기호 등을 삭제하여 정제하고 이어서 정제된 텍스트 데이터는 SBERT 모델과 같은 문장 임베딩 모델을 적용하여 문서 임베딩을 수행하고 벡터값을 산출한다. 그런 뒤 문서 간 유사도 측정 모델을 구성하여 코사인 유사도 등을 통해 신고내용의 벡터값 사이의 유사도를 측정할 수 있는 모델을 구성한다.

이후 친밀한 관계 내 폭력에 관련된 새로운 긴급신고가 접수되는 경우, 그 신고내용의 텍스트가 전처리와 임베딩 과정을 거쳐 문서 간 유사도 측정 모델에 투입되고, 해당 모델을 통해 반복사건일 가능성이 높은 기존 신고 건의 목록이 추출된다. 이때 신규 접수된 신고 사건발생 장소와 각 반복사건의 발생장소 간 인접성을 산출하여 가중치로 적용하게 된다. 즉 신고내용의 벡터값이 유사한 신고 중에서도 발생장소가 인접한 곳을 우선적으로 추천하여 보다 반복사건일 가능성이 높은 사건부터 제시될 수 있도록 설계하였다. 이때 반복사건일 확률을 의미하는 R 은 수식 1과 같이 계산할 수 있다.

$$R = f(X_n, X_e) + d \quad (1)$$

수식 1에서 $f(X_n, X_e)$ 는 문서 간 유사도 측정 함수이며 X_n 과 X_e 는 각각 신규 접수된 신고내용과 기존에 접수된 신고내용을 의미한다. 이때 X_e 는 기존 신고내용들을 의미하기 때문에 복수의 데이터가 될 수 있다. 즉 $f(X_n, X_e)$ 은 신규 접수된 신고내용을 기존 신고내용 데이터들과 비교하여 유사도를 산출하는 함수로 이해할 수 있다. 한편 d 는 인접성에 대한 가중치를 의미하며 신규 접수된 사건의 발생장소와 해당 기존 접수 사건의 발생장소 사이의 거리를 통해 결정된다. 즉 d 는 장소 간 거리가 가까울수록, 다시 말해 인접한 장소에서 발생하였을수록 큰 값을 나타내게 된다. 이와 같이 구축한 반복사건 추론기를 활용하면 신규 접수된 친밀한 관계 내 폭력 사건과 유사성이 높은 기존 신고 건을 추론하여 현장대응 시 해당 정보를 참고할 수 있게 된다.

이처럼 본 연구에서 제안한 반복사건 추론기를 통해 신규 접수된 사건이 파생되었을 가능성이 높은 기존 반복사건의 신고내용과 종결내용 등을 추출할 수 있으며, 이 기술을 활용하여 가정폭력이나 아동학대 등 친밀한 관계 내 폭력 사건이 접수되었을 때 접수요원과 현장출동 경찰관이 해당 사건에 대한 심각성을 보다 정확하게 인지하고 적극적으로 대응하고 사건처리를 할 수 있도록 지원할 수 있다.

III. 결론

본 연구를 통해 제안한 반복사건 추론 기술은 기존에 접수된 친밀한 관계 내 폭력에 대한 긴급신고의 정보, 즉 신고내용과 발생 장소, 관계성 정보를 활용하여 인공지능을 기반으로 학습을 수행하고 이를 통해 새롭게 접수되는 긴급 사건과 동일하거나 유사한 과거의 신고를 추론하는 기술이다. 본 연구에서 제안하는 기술을 활용하면 신고 현장에 출동한 경찰이 사건의 반복성과 심각성을 인지하고 적극적으로 대응할 수 있기 때문에 보다 효과적인 현장대응이 가능해질 것이며, 이를 통해 국민의 생명을 보호하는 데 기여할 수 있을 것이라고 기대한다.

참고 문헌

- [1] Nivette, A.E., Zahnow, R., Aguilar, R. et al., “A global analysis of the impact of COVID-19 stay-at-home restrictions on crime”, Nature Human Behavior, 5, 868-877, 2021.
- [2] Mohler, G., Bertozzi, A. L., Carter, J., Short, M. B., Sledge, D., Tita, G. E., ... & Brantingham, P. J. “Impact of social distancing during COVID-19 pandemic on crime in Los Angeles and Indianapolis”, Journal of criminal justice, 68, 101692, 2020.
- [3] Honey-Roses, Jordi, et al. “The Impact of COVID-19 on Public

Space: A Review of the Emerging Questions.” OSF Preprints, 21 Apr. 2020. Web.

- [4] 조상현, 정병수, “코로나19 시대 범죄 관련 이슈와 시사점”, 한국치안 행정논집, 18(3), 207-220, 2021.
- [5] Rand, M. R., & Saltzman, L. E., “The nature and extent of recurring intimate partner violence against women in the United States”, *Journal of Comparative Family Studies*, 34(1), 137-149, 2003.
- [6] Kang, Y. J., Kang, H. R., Park, S. Y., Jang, Y. J., & Kim, H. S. “A Study on the building Dataset of Similar Case Matching in Legal Domain using Deep Learning Algorithm”, In *Annual Conference on Human and Language Technology*, 72-76, 2021.
- [7] Kabbur, Santosh, Xia Ning, and George Karypis. "Fism: factored item similarity models for top-n recommender systems." *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2013.
- [8] Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., & Sun, M. “How does NLP benefit legal system: A summary of legal artificial intelligence”, *arXiv preprint arXiv:2004.12158*, 2020.
- [9] Qiu, Xipeng, and Xuanjing Huang. "Convolutional neural tensor network architecture for community-based question answering." *Twenty-Fourth international joint conference on artificial intelligence*. 2015.
- [10] Palangi, Hamid, et al. "Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24.4 (2016): 694-707.