

법률정보시스템을 위한 텍스트 마이닝 적용 방안*

- 명예 훼손 판례를 대상으로 -

Application of Text Mining for Legal Information System: Focusing on Defamation Precedent

김 용 환 (Yong Hwan Kim)**

목 차

- | | |
|-------------------|------------|
| 1. 서론 | 4. 결과 및 분석 |
| 2. 관련 연구 | 5. 결론 |
| 3. 데이터 수집 및 연구 방법 | |

초 록

법률 데이터 중에 판례는 다양한 정보를 포함하고 있는 데이터이다. 본 연구에서는 판례를 대상으로 자동으로 텍스트를 분석하여, 일반인을 위한 법률정보시스템으로 활용할 수 있는 방법을 제안하였다. 명예훼손 판례들을 대상으로 판례의 참조조문, 판시사항, 판결요지, 참조 판례를 활용하여, 분석을 진행한 결과 명예훼손에서 사용되는 법 조항과, 핵심 쟁점, 핵심 판례를 자동으로 추출할 수 있었다. 비록 명예훼손 관련 대법원 판례에만 적용하였지만, 제시한 방법론을 통해 다양한 법률 주제 적용 가능할 것이다.

ABSTRACT

Precedents are data containing various types of information. In this study, I proposed a method to be utilized as legal information system for the public using automatic text analysis performed on precedents. It is carried out to analyze the defamation precedent using reference provision, judgment issues, major points of judgment, and reference precedents. As a result of the analysis, legal provisions used in defamation, key issues covered by defamation, and key cases are extracted. Although only applied to the Supreme Court case regarding defamation, the proposed methodology could be applied to various legal topics.

키워드: 판례, 법률정보서비스, 법률정보시스템, 텍스트마이닝, 명예 훼손

Precedent, Legal Information Service, Legal Information System, Text Mining, Defamation

* 이 논문은 2018~2020학년도에 청주대학교 한국문화연구소가 지원한 학술연구조성비(특별연구과제)에 의해 연구되었음.

** 청주대학교 인문사회대학 인문학부 문헌정보학전공 조교수

(kimyonghwan@cju.ac.kr / ISNI 0000 0004 8306 7840)

논문접수일자: 2020년 1월 17일 최초심사일자: 2020년 2월 4일 게재확정일자: 2020년 2월 14일

한국문헌정보학회지, 54(1): 387-409, 2020. <http://dx.doi.org/10.4275/KSLIS.2020.54.1.387>

1. 서론

법은 개개인의 생활과 밀접한 관계를 맺고 있으나, 법 분야와 관련된 사람들을 제외하고는 법을 잘 파악하고 있는 사람은 드물다. 본인이 법적 분쟁의 당사자가 되더라도 분쟁에 어떠한 법이 적용되는지, 어떠한 쟁점이 있는지를 파악하기 어려워 법률전문가인 변호사에게 위탁하는 경우가 일반적이다. 이렇듯 법은 비전문가인 일반인이 접근하기 어려운 분야로 인식되고 있다.

특정 법률 주제를 파악하기 위해서는 다음과 같은 요소를 충족시켜야 한다. 예를 들어, “사기죄”에 대해 알기를 원한다면, 첫째, 사기죄에 대해서 직접 언급하고 있는 형법 제347조를 알아야 하고, 형법 제347조와 관련된 다른 법 조항도 파악해야 한다. 둘째, 실제 재판에서 다루어지는 사기죄의 쟁점에 대해서 알아야 한다. 그리고 마지막으로 사기죄와 관련된 판례를 통해 사기죄의 쟁점들이 어떠한 판단이 이루어졌는지 파악해야 한다. 따라서, 특정 주제에 대한 법률정보는 최소한 관련 법 조항, 관련 법적 쟁점, 관련 판례라는 세 가지 요소를 충족시켜야 한다.

법률정보서비스를 전문적으로 제공하고 있는 국가법령정보센터(법제처 2020. 2. 7.), 종합 법률정보(대법원 2020. 2. 7.), 상업적인 법률정보서비스인 LAWnB(로앤비 2020. 2. 7.)에서는 법령, 판례, 문헌, 규칙/예규/선례 등의 정보들 또한 제공하고 있다. 그러나 이러한 시스템들은 정보들을 단순 제공하거나, 각각의 유형별로 분류하여 제공하는 단계에서 이루어지고 있다(김기영 2017).

법률전문가들은 법 조항과 법률 쟁점에 대한 기본적인 지식을 보유하기 때문에 현재의 법률정보시스템을 통해 특정 주제에 대한 정보를 쉽게 파악할 수 있다. 그러나 일반인의 경우에는 해당 주제가 어떤 법에 해당하는지부터 파악하기 어렵다는 문제가 있다. 따라서 전문가들이 사용하는 기존의 법률정보시스템과는 차별적인 시스템으로, 특정 주제에 대한 관련 법 조항, 법적 쟁점, 관련 판례를 자동으로 제시해주는 일반인들을 위한 법률정보시스템이 요구된다.

본 연구에서는 일반인들을 위한 법률정보시스템으로 다양한 법률정보원 중에서 판례만을 활용하여 특정 법률 주제의 핵심 정보를 제공할 수 있는 방법을 제시하고자 한다. 판례에는 판결과 관련된 다양한 정보들이 포함되어 있어, 법률 이해를 위한 가장 중요한 데이터라고 할 수 있다. 따라서 다수의 판례를 활용하여 핵심 법률정보를 제공함으로써 비전문가에게도 효과적인 법률정보시스템을 제시하고자 한다.

본 연구에서는 다양한 법률 주제 중에서 “명예훼손”을 대상으로 연구를 진행하고자 한다. 명예훼손에 관한 이슈는 지속적으로 존재하고 있으며, 최근 양육비 미지급 아빠들의 신상을 공개한 ‘Bad Fathers’ 논란(이데일리 2020. 2. 7.)은 명예훼손에 대한 대중의 관심을 증가시키고 있다. 따라서 명예훼손 판례를 중심으로 핵심 법률정보를 제공하는 연구를 수행하고자 한다.

논문의 구성은 다음과 같다. 2장에서는 판례 분석과 관련된 선행연구들을 살펴보고자 한다. 3장에서는 데이터의 수집부터, 데이터 전처리와 네트워크 구축 및 시각화까지의 판례 분석

방법론에 대해서 서술하였다. 4장에서는 방법론을 통해 도출된 결과에 대한 분석에 대해 서술하고, 5장에서 결론 및 향후 연구방향에 대해 기술하고자 한다.

2. 관련 연구

판례 데이터를 분석한 연구는 크게 2가지 영역으로 나누어 살펴볼 수 있다. 하나는 다수의 판례 데이터를 텍스트 마이닝 기법 또는 빅데이터 분석 기법을 적용한 연구 영역이고, 다른 하나는 판례를 수작업으로 살펴보고 분석한 연구 영역이다.

먼저 판례 텍스트를 대상으로 텍스트 마이닝을 적용한 연구들은 국내와 국외로 나누어 살펴볼 수 있다. 국내 연구에서 판례 텍스트를 분석하는 다수의 연구들은 단순 통계를 적용한 연구들이며, 텍스트 마이닝 기법을 적용하는 연구는 소수에 불과하다.

이지희, 이준성, 손정욱(2016)은 R프로그래밍을 이용하여 해외 건설 분쟁 판례 데이터를 대상으로 판례 분석 연구하였다. 단어의 빈도를 기반으로 단어간의 상관관계를 분석하거나 LDA 토픽모델링 기법을 적용하여 데이터를 분석하였으며, 분석 결과 해외 건설 분쟁사건의 5가지 유형을 도출하였다.

심준식, 김형중(2017)은 일반인들도 쉽게 판례를 검색할 수 있는 시스템 설계하였다. 10개의 범죄유형별로 구분되는 1,000개의 판례를 대상으로 LDA 토픽모델링 기법을 적용하여 각각의 판례에 30개의 주제어태그를 부여하였으며, 코사인 유사계수를 이용하여 다수의 판례를 그

룹으로 분류하였다. 그리고 판례 그룹을 대상으로 코사인 유사도 값이 큰 판례를 대상으로 분류하였을 때, 좋은 성능을 나타냈다.

김나리, 김형중(2017)은 '소멸시효'와 '손해배상' 판례의 참조조문 부분을 대상으로 Word2Vec 모형을 적용함으로써 키워드와 관련 높은 연관법령을 도출하는 Law2Vec 모형을 구축하였다. 모델의 성능은 정확율(Precision) 57%, 재현율(Recall) 62%를 나타냈다.

김선우, 지선영, 최성필(2018)은 채권 관련 판례 12,457건을 대상으로 채권회수와 비회수 판례로 구분하는 연구를 수행하였다. 판결문의 주문을 대상으로 Bidirectional LSTM기반의 분류 모델을 구성하였으며, F1 값으로 89.82%의 성능을 보여주었다.

이지훈, 엄태현, 이혁준(2019)은 국가법령정보의 민사, 형사, 행정 문서를 대상으로 TF-IDF 기반의 Convolutional Neural Network를 적용하여 법률문서의 분류 시스템을 구축하는 연구를 수행하였고, 시스템의 성능은 63%의 정확도(Accuracy)를 나타냈다.

국내와는 달리, 판례 텍스트를 대상으로 텍스트 마이닝 기법을 적용한 국외 연구는 많이 있다. 연구들은 판례간의 인용 분석, 판례 요약, 법원 판단 예측 분석, 판례의 영향력 분석과 같은 다양한 영역에서 수행되고 있다. 그중 일부를 살펴보면 다음과 같다.

Bommarito, Katz and Zelnet(2009)는 1791년부터 2005년까지의 미 연방 대법원의 판례를 대상으로 인용 및 시멘틱 네트워크를 구축하는 방법론을 제시하였다. 판례의 인용관계를 추출하여 인용 네트워크를 구축하고, 판례내에 나타나는 각 판사들의 의견들을 추출하여 의견

인용 네트워크를 구축하는 방법뿐만 아니라, 인용 네트워크들을 기반으로 각각의 문서에서 나타나는 단어들간의 유사도를 통해 구축한 시멘틱 네트워크를 구축방법을 제시하였다. 제시한 네트워크들과 인용 네트워크와 시멘틱 네트워크를 혼합하여 분석한다면 더 효과적인 분석이 수행될 수 있음을 설명하였다.

Kumar and Raghuvver(2012)는 인도 민사 법원의 판례를 대상으로 토픽모델링 기법을 적용한 자동 요약 연구를 수행하였다. LDA 기법을 적용하여 판례들에 나오는 문장을 대상으로 특정 토픽에 포함될 확률을 계산하여 각각의 토픽별로 가장 높은 확률을 가지는 2개의 문장을 추출하여 요약문을 구성하는 방식이다. 5개의 영역에 판례 자동 요약 적용한 결과, F 값 평균 0.556의 성능을 보여주었다.

Van Der Haegen(2017)는 2000년부터 2015년까지의 벨기에 대법원의 판례를 대상으로 법률 인용 네트워크를 구축하는 방법에 대해 제안하였다. 구체적인 결과 분석은 이루어지지 않았지만, 텍스트 분석 도구인 iKnow를 사용하여 판례내에서 각각의 개념과 개념간의 관계를 추출하여 네트워크를 구성하는 방법론에 대해서 설명하였고, 이 방법론이 판례 분석을 위해 새로운 접근방법으로 활용될 수 있음을 강조하였다.

Zhong et al.(2018)은 법원의 판단이 하나의 요인이 아닌 다양한 요인에 의해서 이루어진다는 것을 기반으로 topological multi-task learning 프레임워크를 제시하였다. 이 방법은 우선 판례에서 추출한 사실 서술, 법 조항, 혐의, 양형 기간을 이용하여 사실 서술을 제외한 나머지를 하위 요인으로 설정하고, 하위 요인들간의 의

존관계를 설정한다. 이후 사실 서술을 기반으로 법 조항을 예측하고, 사실 서술과 법조항을 기반으로 혐의를 예측하고, 사실 서술, 법 조항, 혐의를 기반으로 양형 기간을 예측하는 방식이다. 이러한 프레임워크를 적용하여 SVM, CNN, 계층적 LSTM과 비교하였을 때, 더 높은 성능을 보여주었다.

Leibon et al.(2018)은 미 연방 대법원의 판례내에 나타나는 의견만을 추출하여 구성된 SCOTUS 코퍼스를 대상으로 인용 및 주제적 유사도를 기반으로 네트워크를 구축하여, 각각의 의견이 향후 어떠한 영향을 미치게 될지를 측정하는 방법론을 제시하였다.

판례 텍스트를 자동으로 분석하는 것과는 다르게 수작업을 통해 분석하는 연구들을 다수가 존재한다. 특히 본 연구에서 다루고 있는 “명예훼손”을 주제로 법적 쟁점을 분석하는 법률 논문은 다수가 존재한다. 그중에서 정진수, 강태경, 김형길(2015)은 2005년에서 2015년까지의 명예훼손죄와 모욕죄에 대한 판례를 전반적으로 분석한 연구를 수행하였다. 명예훼손과 모욕죄 관련 모든 판례를 수작업을 통해 살펴봄으로써 법적 쟁점을 분석하였고, 그 결과 명예훼손죄의 법적 쟁점은 공연성, 사실적시, 허위 사실적시, 공공의 이익을 제시하였다.

판례 텍스트를 대상으로 텍스트 마이닝 기법을 적용한 한 국내 연구들은 주로 자동 분류와 관련된 연구에 속하며, 판례 내용 분석과 관련된 것은 판례의 유형을 도출한 연구만을 살펴볼 수 있었다. 본 연구에서는 국내 판례 텍스트를 대상으로 “명예훼손”과 관련된 법 조항, 법적 쟁점 및 관련 핵심 판례를 추출하는 연구를 수행해 보고자 한다.

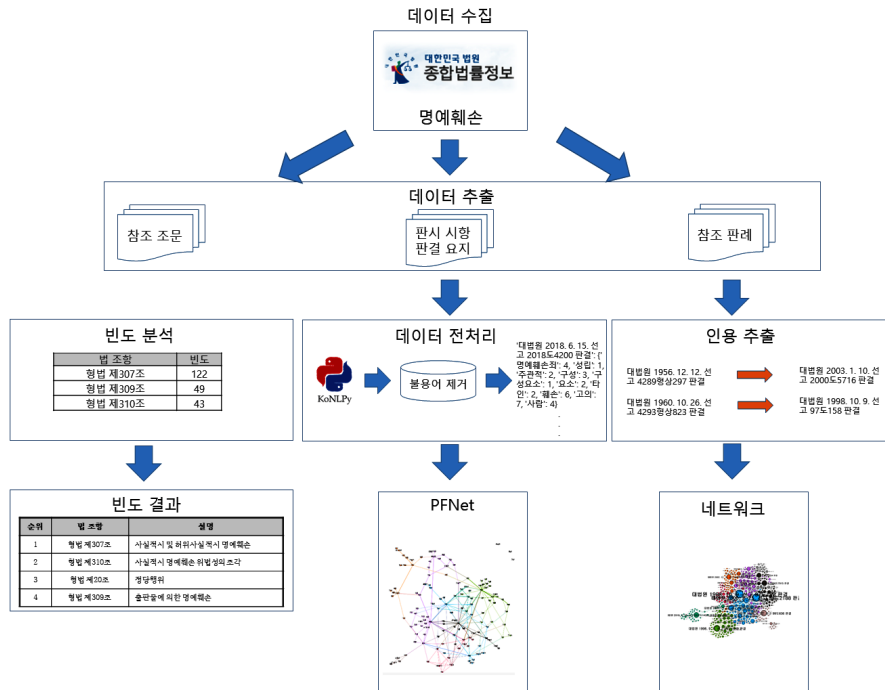
3. 데이터 수집 및 연구 방법

본 연구의 개요는 <그림 1>과 같다. 먼저, 대법원의 종합법률정보를 통해 판례 데이터를 크롤링하여 저장하였다. 그 다음으로, 수집된 데이터에서 판시 사항, 판결 요지, 참조 조문, 참조 판례만을 추출하였다. 이후 추출된 정보에 따라서 각각 다른 방법을 적용하였다. 먼저 참조 조문의 경우에는 빈도 분석을 통해 상위 빈도의 법조문을 분석하였다. 판시사항과 판결요지의 경우에는 형태소 분석과 불용어 작업으로 구성되는 데이터 전처리를 진행한 후, 전처리된 데이터를 기반으로 동시 출현 네트워크를 구축하고 Pathfinder Network(PFNet) 기법을 적용하였다. 참조 판례의 경우에는 판례간

인용 관계를 추출하였고, 이후 네트워크를 구축하여 분석하였다.

3.1 데이터 수집 및 추출

일반인들을 위한 시스템을 구축하기 위해서는 다양한 정보를 포함하고 있는 법률 데이터가 필요하다. 본 논문에서는 판례를 다양한 정보를 포함하고 있는 핵심적인 법률 데이터로 활용하고자 한다. 판례란 법원에서 특정 사건에 대해서 법을 해석하고 적용한 근거와 판단을 문서로 작성한 것을 의미한다. 판결문이라고도 언급되며, 각급 법원에서 생성되는 모든 판결문의 집합을 판례라고 한다. 성문법체제를 유지하는 대륙법계에 속한 한국에서는 실제 재



<그림 1> 전체 방법론 개요

판에서는 법 규정이 모든 요소를 명확하게 규정할 수 없기 때문에, 구체적인 사안에 대해서 판례를 통해 보충하고 있다(홍일표 2007). 또한 판례에는 판결을 위해 적용된 법 조항, 판결에서 주로 다루어진 쟁점, 쟁점에 대한 판단 등이 포함되어 있어, 법률 이해를 위해 중요한 데이터라고 할 수 있다. 따라서 다수의 판례를 통해, 특정 주제에 관한 법 조항, 법률 쟁점, 관련 판례를 파악할 수 있을 것으로 판단하였다.

판례수집은 1998년부터 대법원에서 서비스를 제공하고 있는 종합법률정보(대법원 2020. 2. 7.)를 활용하였다. 상세검색 기능을 이용하여 검색하였다. 키워드는 “명예훼손”으로 하였고, 판결사건종류와 판결법원은 형사와 대법원으로 설정하였고, 사건 종류는 명예훼손이 형법상의 범죄명이기 때문에 형사로 설정하였으며, 법원은 대법원 판례만을 수집하였다. 총 260건

의 판례가 검색되었고, Selenium기반의 자체 제작한 Python 프로그램을 이용하여 데이터를 수집하였다. <그림 2>는 수집된 판례의 예시이다.

판례 데이터는 선고번호, 판시 사항, 판결 요지, 참조 조문, 참조 판례 및 전문으로 구성된다. 선고번호는 사법부에서 선고되는 모든 재판 결과를 식별할 수 있는 번호에 해당한다. 판시사항은 <그림 2>에서 “명예훼손죄의 성립요건”, “명예훼손죄의 구성요건 중 ‘공연성’의 의미”와 같이 판결문의 주된 쟁점이 되는 부분을 의미한다. 판결요지는 판결문을 요약한 내용에 해당한다. 참조 조문은 해당 판결을 위해서 참고한 법률 조항을 의미하며, 참조 판례는 판결문을 위해서 직·간접적으로 참고한 이전 판례를 의미한다. 전문은 해당 판결에 대한 직접적인 내용으로, 피고인과 상고인, 원심판결, 주문, 이유 등으로 구성된다. 여기서 주문은 “상고를



<그림 2> 판례 데이터의 예시

기각한다”와 같이 최종적인 판결 내용을 의미하며, 이유는 주문을 도출하기 위한 근거를 서술한 내용이다.

본 연구에서는 판례 데이터 중에서 선고번호, 판시 사항, 판결 요지, 참조 조문, 참조 판례만을 추출하여 분석을 수행하였다. 선고번호는 판례의 ID로 사용하였고, 판시 사항과 판결 요지는 핵심 쟁점의 도출, 참조 조문은 관련 법조항의 도출 참조 판례는 관련 핵심 판례를 도출을 위해 사용되었다. 전문을 이용한 내용 분석 또한 가능하나 전문에는 분쟁 당사자의 정보뿐만 아니라 각각의 분쟁에 관한 사실관계가 존재하기 때문에 이를 대상으로 분석을 수행하면 오히려 잡음(Noise)가 발생할 위험이 있기 때문에 분석에서 제외하였다.

추가적으로 본 연구에서는 명예훼손의 인정 여부에 따른 쟁점의 차이를 살펴보고자 하였다. 따라서 각각의 판례에 대해서 수작업을 통해 명예훼손의 인정여부를 분류하였다. 전문에서 “명예훼손의 위법성이 있다” 또는 “명예훼손의 성립을 인정한 법리오해의 잘못이 없다”라는 언급이 있거나, 명예훼손의 성립을 인정하는 언급이 있다면, 명예훼손을 인정한 판결로 분류하였고, “명예훼손의 위법성이 있다는 원심의 판단은 잘못이라 할 것이다.” 또는 “명예훼손죄 또는 그 위법성 조각사유에 관한 법리오해 등의 위법이 있다고 볼 수 없다.”라는 언급 및 명예훼손의 불성립을 언급하는 문장이 있다면, 명예훼손을 부정한 판결로 분류하였다. 다만, 재판에서 명예훼손에 대한 언급은 있으나, 소송과정의 문제로 인하여 이에 대한 판결이 진행되지 않은 경우와 명예훼손죄의 당사자들 일부는 인정되고, 일부는 부정되는 경우는

데이터에서 제외하였다. 따라서 수집된 260건의 판례 데이터 중에서 7건의 데이터를 제외하고 최종적으로 총 253건의 데이터를 대상으로 분석을 수행하였다.

3.2 참고 조문 분석 방법

참고 조문은 판례마다 약간씩 다르게 표현되고 있다. 대부분 ‘법률명’, ‘조’, ‘항’까지 상세하게 서술되어 있으나, 일부는 ‘법률명’, ‘조’까지만 서술되어 있다. 법 조문에서 ‘항’은 세부적으로 다른 내용을 포함하고 있으나, 같은 ‘조’에 포함되었다는 것은 유사한 법 내용이라고 볼 수 있기 때문에 이 연구에서는 ‘항’ 부분을 제외하고 ‘법률명’, ‘조’ 까지만 나타냈다.

또한 판례에서 동일한 ‘법률명’이 연속적으로 나타나면, 뒤에 나타나는 ‘법률명’을 ‘동법’으로 작성하기 때문에, ‘동법’의 경우에는 앞서 출현한 ‘법률명’으로 변경하여 분석하였다.

3.3 판시사항 및 판결 요지 분석

3.3.1 전처리 단계

전처리과정은 한글 형태소 분석과 불용어의 제거, 명사추출의 과정으로 진행되었다.

형태소 분석을 위해서 꼬꼬마 형태소 분석기와 코모란 형태소 분석기를 혼합하여 사용하였다. 코모란 형태소 분석기의 경우에는 3.x 버전으로 업그레이드되면서, 문장 구분 기능이 사라졌고, 꼬꼬마 형태소 분석기는 이용자 사전 구축이 복잡하게 되어 있어, 단일 형태소 분석기를 통해서서는 적합한 분석이 어렵기 때문에, 두 가지 형태소 분석기를 혼합하여 사용하였다.

본 연구에서는 꼬꼬마 형태소 분석기를 통해 문장 구분을 한 다음, 코모란 형태소 분석기를 통해 문장의 형태소 분석하는 방식으로 수행하였다. 형태소 분석기의 적용을 위해서 Konlpy 라이브러리(박은정, 조성준 2014)를 활용하여 Python 프로그램 자체 제작하였다. Konlpy는 다양한 언어로 개발된 한국어 형태소 분석기 Python 기반으로 사용할 수 있도록 지원하고 있으며, 꼬꼬마 형태소 분석기와 코모란 형태소 분석기 모두 Konlpy 기반으로 사용 가능하다.

코모란 형태소 분석기를 통한 형태소 분석을 위해 이용자 사전을 추가하였다. 판례 텍스트는 일반 텍스트와는 다른 어휘를 많이 사용하기 때문에 이를 위하여 법률 용어를 이용자 사전을 통해 추가하였다. <표 1>은 이용자 사전의 예시이다.

<표 1> 이용자 사전의 예시

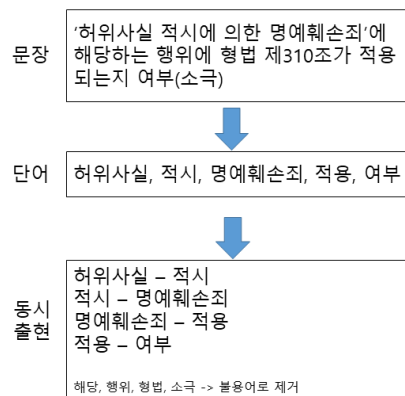
단어	품사
정보통신망 이용 촉진 및 정보보호 등에 관한 법률	NNP
출판물에 의한 명예훼손죄	NNP
사회통념	NNP
사회 상규	NNP
미필적 고의	NNP
공공의 이익	NNP
...	...

코모란 형태소 분석 결과를 기반으로 명사만을 추출하였다. 추출된 단어 중에서는 ‘등’, ‘개’와 같은 주제적 의미가 없는 일반명사들이 다수 포함되어, 불용어으로써 이를 제거하였다.

3.3.2 네트워크 구축 및 PFNet

형태소 분석을 통해 추출된 단어를 기반으로

동시 출현 네트워크를 구축하였다. 판례 데이터는 단문보다는 장문으로 구성되어 있기 때문에, 하나의 문장안에 다수의 접속사와 구가 포함되어 있다는 특징을 가진다. 따라서, 문장 단위로 동시 출현 단어쌍을 추출하게 되면, 동시 출현 단어들 사이에 아무런 관련성이 없더라도 관련 있는 단어로 연결될 수 있으므로, 본 연구에서는 문장에서 단어의 출현 순서를 기반으로 각 단어의 앞뒤 단어만을 이용하여 동시 출현 단어쌍을 추출하였다. 그리고 가중치는 문헌빈도(Document Frequency)로 하였다. <그림 3>은 동시 출현 단어 추출 과정 예시이다.



<그림 3> 동시 출현 단어 추출 과정 예시

추출된 동시 출현 단어로 네트워크를 구축하기 위해서, ‘명예훼손’ 단어를 기준으로 연결 단계가 2단계인 단어만을 활용하였다. ‘명예훼손’에서 나타나는 쟁점을 파악하고자 하였기 때문에 네트워크에서 ‘명예훼손’과 직접 연결되거나, 매개단어를 통해 간접적으로 연결되는 단어들이 쟁점을 더욱 잘 나타낼 수 있을 것이라 판단하였다.

우선, ‘명예훼손’ 단어를 기준으로 직접 연결

되는 단어 쌍을 추출하였고, 가중치가 1인 단어 쌍들은 제외하였다. 그 다음 단계로 '명예훼손' 단어와 직접적으로 연결되는 단어들과 직접적으로 연결되는 단어쌍을 추출하고, 가중치가 1인 데이터들은 모두 제거하였다. <그림 4>는 네트워크를 구축하는 예시이다.

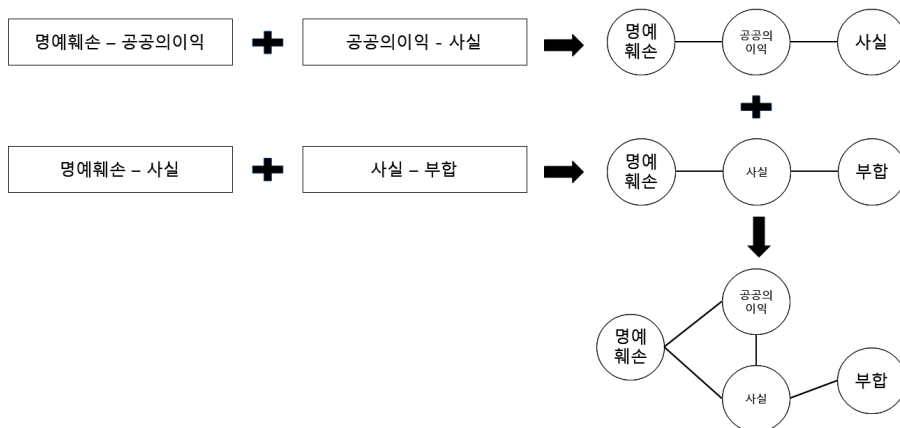
구축된 네트워크를 기반으로 Pathfinder Network (PFNet) 알고리즘을 적용하였다. PFNet은 네트워크에서 상대적으로 중요한 링크만을 선정하여 네트워크를 구성하는 알고리즘 중의 하나이다. Schvaneveldt(1990)에 의해서 개발된 이 알고리즘은 삼각부등식을 위반하는 경로를 제거하는 방식으로 네트워크를 구성한다. 이 알고리즘은 r과 q라는 2개의 파라미터를 설정해주어야 하는데, $r = \infty$ 이고 $q = n - 1$ 인 경우 가장 엄격하게 주요 링크만을 선정하기 때문에 이를 이용하여 구축하였다(이재운 2006). PFNet의 구축을 위해서 WNET(이재운 2013) Tool을 이용하였다.

이후, PFNet을 통해 구축된 네트워크는 오픈 소스 Tool인 Gephi(Bastian, Heymann and

Jacomy 2009)를 통해 시각화하였다. Gephi에서는 네트워크의 시각화 뿐만 아니라, 다양한 네트워크 분석 기법을 제공하고 있다. Gephi는 Centrality와 같은 핵심노드를 구분하는 기법과 네트워크에 나타나는 커뮤니티를 식별함으로써 네트워크의 구조를 파악하는 Modularity (Blondel et al. 2008) 기법 등을 제공하고 있어, 이를 활용한 네트워크 분석을 수행하였다.

3.4 참고 판례 분석

참고 판례는 두 가지 형태로 표현되고 있다. 하나는 판결문 내의 참고 판례 영역에 표현되어 있고, 다른 하나는 '이유' 영역 내에 괄호를 이용하여 표현되어 있다. 일부 판례에서는 참고 판례 부분에 모든 참고 판례가 포함되어 있지 않기 때문에, 정확한 분석을 위해서는 '이유' 영역내에 있는 참고 판례를 추출해야 한다. 정규식을 기반으로 '이유' 영역내의 참고 판례를 수집하였고, 참고 판례 영역의 판례와 중복되는 경우에는 제거하였다.



<그림 4> 동시 출현 네트워크 구축 예시

그 다음으로, 판례의 선고번호와 참고 판례의 인용 관계를 추출하여 인용 네트워크를 구축하였다. 네트워크 분석기법중 하나인 In-Degree Centrality를 기준으로 가장 많이 인용된 핵심적인 판례 상위 5개를 추출하였다. 판례 추출시 In-Degree Centrality가 동일한 경우에는 최신의 판례를 우선 추출하였다.

한가지 고려할 점은, In-Degree Centrality를 이용하는 경우 오래된 판례일수록 값이 높아질 수 밖에 없다는 점이다. 판례의 경우에는 사회의 변화에 따라서 기존의 판례와는 다르게 나타날 수 있기 때문에 상대적으로 최근의 판례가 더 중요하게 여겨진다. 따라서 추가적으로 최근 10년간의 판례를 대상으로 In-Degree Centrality 상위 5개를 추출하였다.

4. 결과 및 분석

판례 분석을 통해 총 3가지의 결과를 도출하였다. 첫째는 참고 조문으로 명예훼손과 관련된 법 조항을 추출하였다. 둘째는 판시사항과 판결 요지를 통해 수행된 내용분석 결과이다. 그 결과 '명예훼손'의 핵심 쟁점을 도출하였고, 시각화를 통해 나타냈다. 셋째는 관련 판례를 제시하였다. 인용 네트워크 분석을 통해 명예훼손과 관련된 핵심 판례를 제시하도록 하였다.

4.1 참고 조문 분석

판례 텍스트의 참고 조문을 기반으로 명예훼손과 관련된 법 조문에 대한 정보를 찾아 볼 수

있다. 그러나 다수의 판례에서 참고 조문을 수집하면 <표 2>와 같은 정보를 얻을 수 있다. <표 2>는 명예훼손 관련 판례에서 나타난 조항을 빈도 기준으로 순서대로 나열한 것이다. 단, 형사 소송법의 경우에는 소송과정과 관련된 법률이기 때문에 형사소송법을 제외한 나머지 법 조항만을 나타낸 것이다.

관련 조항을 살펴보면, 대부분 명예훼손과 직접 또는 간접적으로 관련된 법 조항으로 구성되어 있다. 우선 형법 제307조, 제308조, 제309조, 제310조, 제312조, 공직선거및선거부정방지법 제251조, 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조는 명예훼손죄를 규정하는 것으로 명예훼손과 직접적인 영향을 미치는 법 조항에 해당한다. 각각의 조항을 살펴보면, 명예훼손의 범위는 사실적시에 의한 명예훼손, 허위사실 적시에 의한 명예훼손, 사자에 대한 명예훼손, 출판물에 의한 명예훼손, 공직선거후보자 명예훼손, 정보통신망을 통한 명예훼손이 있고 명예훼손죄를 저지르더라도 공공의 이익을 위한다는 특별한 사유가 있다면 위법성을 제한하는 위법성조각사유와 친고규정 등이 해당한다.

다른 법 조항 중의 일부는 명예훼손과 간접적으로 관련 있는 법 조항이라 볼 수 있다. 형법 제20조, 제21조, 제30조, 제40조, 제156조, 제311조, 헌법 제21조, 제37조는 직접적으로 명예훼손을 언급하고 있지 않지만, 명예훼손과 유사하거나 명예훼손에서 다룰 수 있는 쟁점을 간접적으로 포함하고 있는 법 조항이기 때문이다.

기타 다른 법 조항들이 등장하는 이유는 판례에 특징으로 인해 나타난 현상이라 볼 수 있다. 판례는 하나 이상의 주제가 다루어 질 수

〈표 2〉 판례에서 수집한 법 조항의 순위 및 설명

순위	법 조항	설명	빈도
1	형법 제307조	사실적시 및 허위사실 적시 명예훼손	124
2	형법 제309조	출판물에 의한 명예훼손	49
3	형법 제310조	사실적시 명예훼손 위법성의 조각	43
4	형법 제20조	정당행위	19
5	구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조	정보통신망을 통한 명예 훼손(현재 법 조항이 제70조로 변경)	15
6	형법 제314조	업무방해	13
7	형법 제355조	횡령과 배임	11
8	형법 제311조	모욕	10
9	형법 제156조	무고	8
10	형법 제356조	업무상 횡령과 배임	8
11	공직선거및선거부정방지법 제251조	후보자비방죄	7
12	헌법 제21조	언론, 출판의 자유	4
13	형법 제308조	사자 명예훼손 관련 조항	4
14	형법 제347조	사기	4
15	헌법 제37조	자유와 권리 보장 및 제한	3
16	형법 제21조	정당방위	3
17	형법 제30조	공동정범 관련 조항	3
18	형법 제312조	고소와 피해자의 의사	3
19	형법 제327조	강제집행면탈	3
20	형법 제40조	상상적 경합 관련 조항	3

있고, 각각의 주제에는 하나 이상의 쟁점이 포함될 수 있다. 대법원에서도 이러한 현상은 동일하게 나타나기 때문에 명예훼손 관련 쟁점을 다루는 판례라 할지라도 다른 주제의 쟁점이 포함되어 있을 수 있다. “대법원 2006. 4. 27. 선고 2003도4735 판결”을 보면, 피고인은 횡령배임죄에 대한 상고를 하였고, 검사는 명예훼손죄에 대한 상고를 서로 진행하였기 때문에 하나의 판례에 두 가지 주제에 관한 쟁점이 동시에 나타나고 있음을 살펴볼 수 있다. 여러 쟁점중 형법 제355조 배임과 횡령 그리고 형법 제356조 업무상 배임과 횡령이 명예훼손과 같이 자주 등장하는 다른 주제의 쟁점이라고 볼 수 있다.

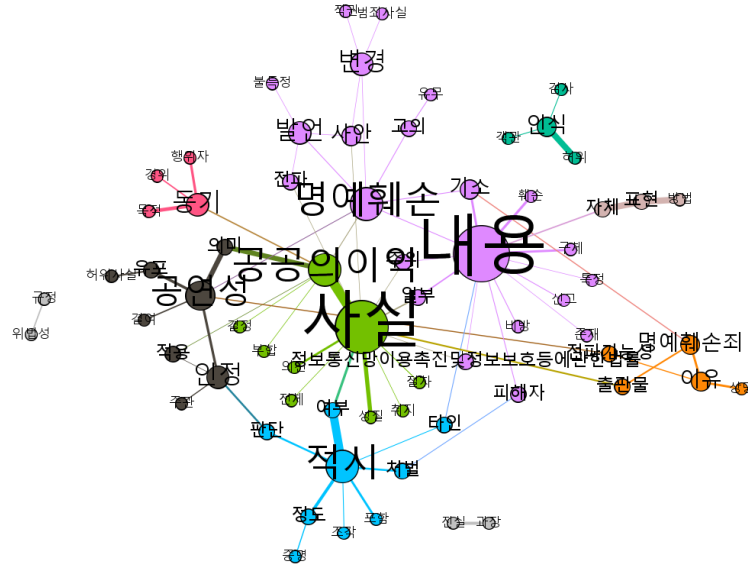
4.2 판시사항 및 판결요지 분석

4.2.1 판례 쟁점 분석

명예훼손 판례 데이터로부터 〈그림 5〉와 같이 PFNet을 시각화하였다. 네트워크에 크게 나타나는 노드 중에서 사실, 적시, 내용, 공공의 이익, 공연성이 명예훼손의 가장 핵심적인 단어라는 것을 확인할 수 있다.

보다 상세한 분석을 위해, Modularity기법을 적용하여 분석을 수행한 결과, 총 10개의 커뮤니티가 나타났다. 〈표 3〉은 명예훼손 판례 PFNet에서 출현하는 커뮤니티를 나타낸 것이다.

〈표 3〉에서의 주제는 각각의 커뮤니티에 출



〈그림 5〉 명예훼손 판례의 PFNet

〈표 3〉 명예훼손 판례의 PFNet에서 출현하는 커뮤니티

번호	주제	출현 단어
1	명예훼손의 내용	내용, 명예훼손, 발언, 변경, 사안, 기소, 고의, 불특정 등
2	공공의 이익 부합 여부	사실, 공공의 이익, 부합, 성질, 의견, 전체, 절차, 취지 등
3	공연성 인정 여부	공연성, 인정, 동기, 유포
4	과장 여부	과장, 진실
5	동기 및 목적	동기, 경위, 목적
6	위법성 규정	위법성, 규정
7	출판물 전과가능성	명예훼손죄, 상당, 이유, 전과가능성, 출판물
8	적시 정도 여부	적시, 여부, 정도, 처벌, 타인, 판단, 조각, 증명, 포함 등
9	표현 방법	방법, 자체, 표현
10	객관적 허위 사실 인식	객관, 검사, 허위, 인식

현하는 단어를 기반으로 구성한 것이다. 먼저 1번 커뮤니티의 주제는 ‘명예훼손의 내용’으로 어떠한 내용의 명예훼손이 성립되는지 살펴보는 쟁점이라고 할 수 있다. 사실의 내용을 서술하였는지, 또는 허위 사실을 서술하였는지, 명예훼손의 내용이 특정 대상을 지정하였는지 또는 구체적인 내용을 서술하였는지 등은 명예훼손

손죄의 구성 요건중의 하나에 해당한다. 예로써, “대법원 1994. 4. 12. 선고 93도3535 판결”은 일부 알려진 사실의 내용을 서술하였을 때에도 명예 훼손죄의 성립되는지를 다루고 있다. 또한 “대법원 1987. 5. 12. 선고 87도739판결”은 명예훼손죄의 대상에 대해 어떤 특정한 사람 또는 인격을 보유하는 단체에 대하여만 성

립된다는 것을 쟁점으로 다루고 있다.

2번 커뮤니티의 주제인 “공공의 이익 부합 여부”는 위법성 조각 사유에 해당하는 것으로 공공의 이익은 명예훼손죄의 성립의 중요한 쟁점이 된다. “대법원 1995. 11. 10. 선고 94도1942 판결”, “대법원 2011. 9. 2. 선고 2010도17237 판결” 등은 형법 제310조에서 규정하는 위법성 조각사유에 대한 판단 기준으로 피고인의 행위가 공공의 이익에 부합한지를 판단하여 명예훼손죄의 성립 여부를 다루는 판례에 해당한다.

3번 커뮤니티의 주제인 “공연성 인정 여부”는 명예훼손의 죄를 구성하는 구성요건중의 하나로, 행위자의 명예훼손의 내용을 다른 사람에게 전파되었는지를 나타내는 것에 해당한다. “대법원 2010. 10. 28. 선고 2010도2877 판결”은 한 사람에게 명예훼손의 내용을 유포한 것이 공연성을 충족하는 것인지를 쟁점으로 다루고 있다.

4번 커뮤니티의 주제인 “과장 여부”는 과장한 것이 허위사실로 인정될 것인지를 보는 것으로 허위사실에 적시에 의한 명예훼손죄가 성립되는지를 판단하는 요소 중의 하나이다. “대법원 2005. 4. 29. 선고 2003도2137 판결”은 객관적 사실과 맞는 경우, 세부적인 내용에 있어서 과장이 있다고 하더라도 이를 허위의 사실로 볼 수 있을 것인지를 쟁점으로 다루고 있다.

5번 커뮤니티의 “동기와 목적”은 명예훼손의 동기와 목적으로 명예훼손의 고의를 판단하는 기준이 된다. 이와 관련하여 “대법원 2008. 10. 23. 선고 2008도6515 판결”에서는 명예훼손 내용을 발설하였는지에 관한 질문에 대답하는 과정에서 명예훼손 내용을 발설한 경우에 명예훼손이 성립되는지를 다루고 있다.

6번 커뮤니티의 “위법성 규정”은 형법 제 310조의 내용이 위법성 조각 사유에 관한 내용으로 여기에서는 사실적시에 의한 명예훼손의 경우에만 공공의 이익이 있을 때, 위법성 조각사유로 인정하고 있다. 이 경우, 위법성 조각사유 규정에 적합한지에 대한 판단 역시 중요한 쟁점이 될 수 있다. “대법원 2000. 5. 12. 선고 99도5734 판결”의 경우에는 명예훼손의 내용이 객관적인 사실인지, 공공의 이익에 관한 것인지와 기타 제반사항을 고려하여 명예훼손죄가 성립되는지를 다루고 있다.

7번 커뮤니티는 “출판물 전파가능성”으로 출판물에 의한 명예훼손의 경우에는 출판물이 유포 여부가 구성요건에 해당한다. “대법원 2000. 5. 16. 선고 99도5622 판결”에서는 기자에게 사실을 이야기하였지만, 그 기자가 기사화하지 않은 경우, 출판물에 의한 명예훼손이 성립되는지를 다루고 있다.

8번 커뮤니티의 “적시 정도 여부”는 명예훼손죄가 성립되기 위해서는 어느 정도의 명예훼손이 이루어져야 하는지에 관한 것이다. “대법원 2011. 8. 18. 선고 2011도6904 판결”에서는 명예훼손이 되기 위해서는 특정인의 사회적 가치가 침해될 정도의 내용이 서술되어야 한다고 언급하고 있다.

9번 커뮤니티의 “표현방법”은 표현방법은 명예훼손 내용을 어떠한 방식으로 표현하였는가를 의미한다. 이는 명예훼손의 고의와 관련된 것으로 표현방법에 따라서 고의의 여부를 고려할 수 있음을 의미한다. “대법원 2002. 8. 23. 선고 2000도329 판결”은 이러한 부분에 대해서 다루고 있다.

10번 커뮤니티의 “객관적 허위 사실 인식”은

명예훼손의 행위자가 명예훼손 행위를 진행하였는데 해당 내용이 허위사실임을 인지하였는지 여부가 논점이 된다. 만약, 허위사실을 인지하였다면 허위 사실 적시에 의한 명예훼손에 해당하고, 허위사실임을 인지하지 못하였을 때에는 사실적시에 의한 명예훼손에 해당되게 되어 있다. “대법원 2011. 6. 24. 선고 2011도3824 판결” 등은 허위성의 인식 유무에 대해 다루고 있다.

위와 같이 판례에 대한 PFNet 분석을 통해 판례에서 다루고 있는 명예훼손 관련 쟁점들이 도출되었다고 볼 수 있다.

추가적으로 분석 과정과 관련하여 고려할 점은 법률 텍스트 전문 형태소 분석기의 필요성이다. 판시사항 및 판결요지에 대한 형태소 분석을 위해서 오픈 소스 형태소 분석기를 활용하였고, 보다 정확한 분석을 위해서 이용자 사전을 다수 사용하였다. 그러나 결과에서도 나타나듯이 “조각”, “상당”과 같이 다른 단어들의 도움없이 해석하기에는 어려운 단어들도 여전히 남아있음을 확인할 수 있다. 이러한 이유는 판례가 권위적인 표현, 한문 표현, 일본식 어휘, 비정상적인 장문들로 구성되기 때문이다(장소원 2009). 이는 일반 텍스트들과는 다른 특징을 가진다. 따라서 판례를 분석하기 위해서는 판례를 전문적으로 다루는 형태소 분석기의 개발이 필수적이다. 형태소 분석기는 텍스트 분석의 가장 기본이므로, 이 형태소 분석기의 개발이 이루어진다면, 향후 판례에 대한 내용 분석 연구가 더 활발히 이루어 질 수 있을 것이다.

4.2.2 명예훼손 인정 여부에 따른 판례 쟁점 분석

수작업을 통해 명예훼손의 인정여부를 구분

하여, 명예훼손을 인정하는 판례와 불인정하는 판례 사이의 논점의 차이가 있는지 살펴보았다. <그림 6>은 명예훼손을 인정하는 판례를 기준으로 PFNet을 구축한 것이다.

<그림 6>을 보면 명예훼손을 인정받은 판결에서, 위법성, 진실, 표현이 핵심적인 단어로 도출되고 있다. 해당 내용을 기반으로 명예훼손의 내용이 공공의 이익에 부합하는지 여부, 출판물에 의한 명예훼손 인정 여부, 명예훼손의 내용이 진실인지 여부, 표현 방법에 대한 것이 쟁점으로 나타났다고 볼 수 있다.



<그림 6> 명예훼손 인정 판례의 PFNet

반면에 <그림 7>은 명예훼손을 불인정한 판례를 기반으로 PFNet을 구축한 것이다. 명예훼손이 불인정되는 경우에는 출판물, 내용, 사실, 처벌, 위법성, 공연성 적시 등이 핵심적인 단어로 나타나는 것을 살펴 볼 수 있다. 해당 내용을 기반으로 출판물의 유통 여부, 내용의 진실 여부, 공연성의 인정 여부, 공공의 이익을 위한 사실 여부, 허위 사실의 적시 여부, 처벌의 의사 여부 등이 쟁점으로 나타나고 있다.

명예훼손을 인정한 판례 또는 불인정한 판례에서 도출되었다고 해당 쟁점이 명예훼손을 인

정하는 쟁점 또는 불인정하는 쟁점으로 간주하기는 어렵다. <그림 6>과 <그림 7>의 네트워크에 등장하는 모든 쟁점들은 명예훼손이 인정되는 경우와 부정되는 경우에서 모두 나타날 수 있다는 의미이다. 다만, 인정 또는 불인정 되는 판례에서 상대적으로 더 나타나는 쟁점들이라는 의미로 생각할 수 있다. 그렇지만, 상대적으로 인정 또는 불인정 되는 판례에서 주로 나타나는 핵심 쟁점을 사전에 파악하고 있다면, 재판에 위한 전략을 찾을 때 활용할 수 있을 것이다.

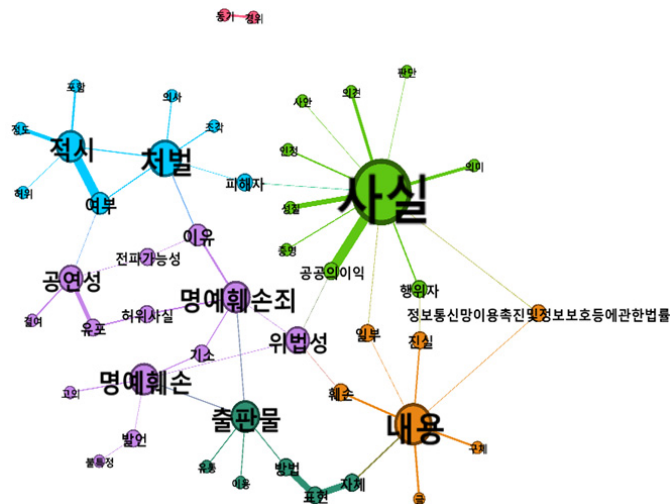
분석 과정과 관련해서 고려할 점은 명예훼손 인정 판례와 불인정 판례의 차이를 분석하기 위하여 수작업이 포함되었다는 점이다. 판례의 자동 분석을 통해 일반인에게 핵심 정보를 제공하는 시스템을 위해서는 수작업의 영역이 최소화되어야 한다. 쟁점의 인정 여부를 고려하지 않고 정보를 제공하는 경우에는 현재의 과정으로도 제공할 수 있으나, 쟁점의 인정 여부를 자동으로 분류할 수 있게 된다면, 법률정보

시스템을 자동화된 시스템으로써 활용할 수 있을 것이고, 다양한 법률 주제에 효과적으로 적용할 가능할 것이다.

4.3 참조 판례 분석

판례에도 이전의 판례가 있는 경우에는 이를 참고하여 작성한다. 즉, 판례간의 인용관계가 성립되기 때문에, 핵심 판례를 도출하기 위해서 인용 네트워크를 통한 분석을 수행하였다. <그림 8>은 판례 인용 네트워크를 구축한 것을 시각화한 것이다.

중앙에 다수의 노드를 포함하고 있는 큰 네트워크 하나와 함께 소수의 노드를 포함하는 작은 네트워크 여러 개를 형성하고 있는 것을 볼 수 있다. 중앙의 큰 네트워크는 전체적으로 명예훼손과 관련된 네트워크로 볼 수 있으며, 다른 네트워크들은 주로 명예훼손이 아닌 다른 주제의 판례에 대한 인용 네트워크에 해당한다.



<그림 7> 명예훼손 불인정 판례의 PFNet



〈그림 8〉 판례 인용 네트워크

〈표 4〉는 Degree Centrality 상위 5개 판례를 나타낸 것이다.

핵심 판례를 살펴보면, 사실 유포 및 공연성의 의미를 쟁점으로 다루는 “대법원 1981. 10. 27. 선고 81도1023 판결”을 제외하고, 모든 판례에서 공공의 이익에 해당하는지가 쟁점으로 등장하는 것을 확인할 수 있다. 그 외에는 특정 행위가 명예훼손을 구성하는 요건에 해당하는지 여부를 다루는 것이 쟁점으로 나타나는 것을 볼 수 있다.

핵심 판례의 최신성을 위해서 최근 10년간의 데이터만을 활용하여 인용네트워크를 구성하였다. 〈그림 9〉를 살펴보면, 전체 데이터를 사용한 것과는 동일하게, 중앙의 큰 네트워크는 전체적으로 명예훼손과 관련된 네트워크로 볼

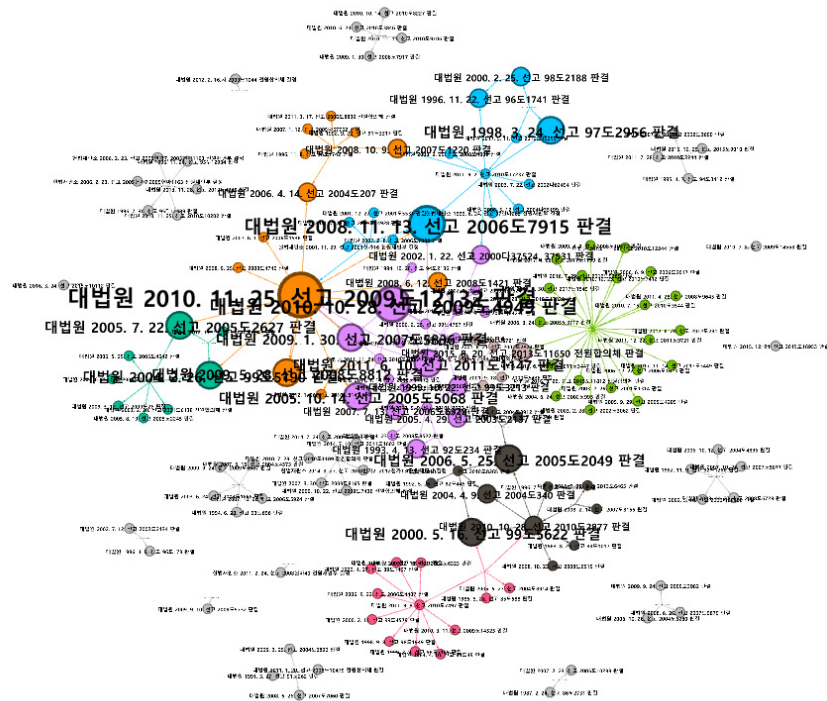
수 있으며, 소수의 노드만 포함되어 있는 다른 네트워크들은 주로 명예훼손이 아닌 다른 주제의 판례를 표현하고 있다.

〈표 5〉의 In-Degree Centrality 상위 5개의 핵심 판례들에는 허위 사실에 의한 명예훼손과 관련된 판시 사항과 정보통신망 관련된 명예훼손에 관한 판시 사항이 주로 나타나고 있음을 확인할 수 있다. 인터넷의 발전으로 인하여 정보통신망 관련 판결 사항이 주로 나타나고 있음을 확인할 수 있다.

본 연구에서는 “명예훼손”을 주제로 판례를 활용하여 수작업 분석이 아닌 자동 분석을 통해 법 조항, 법적 쟁점, 핵심 판례를 추출하였다. 추출된 정보를 확인한 결과, 실제 명예훼손과의 관련성을 확인할 수 있었으며, 명예훼손에 관한

〈표 4〉 참조 판례를 기반을 추출한 핵심 판례

판례	In-Degree Centrality	명예훼손 관련 판시 사항
대법원 1998. 10. 9. 선고 97도158 판결	25	1) 형법 제310조 소정의 '진실한 사실'의 의미 2) 형법 제309조 제1항 소정의 '기타 출판물'의 의미 3) 형법 제309조 제1항과 제310조와의 관계 4) 형법 제310조 소정의 '오로지 공공의 이익에 관한 때'의 의미 및 그 판단 기준
대법원 2000. 2. 25. 선고 98도2188 판결	22	1) 명예훼손죄 성립에 필요한 사실 적시의 정도 2) 언론매체의 표현행위가 명예훼손죄의 사실 적시에 해당하는지 여부에 대한 판단 기준 3) 형법 제310조의 위법성조각사유에 해당하기 위한 요건인 '진실한 사실', '공공의 이익'의 의미 4) 적시된 사실이 공공의 이익에 관한 것인 경우, 형법 제309조 소정의 '사람을 비방할 목적'이 부인되는지 여부(적극)
대법원 2003. 11. 13. 선고 2003도3606 판결	11	1) 형법 제310조의 위법성조각사유에 해당하기 위한 요건인 '진실한 사실', '공공의 이익'의 의미와 그 판단 기준 및 형법 제309조 소정의 '사람을 비방할 목적'의 의미 2) 공직선거및선거부정방지법 제251조 단서 소정의 '진실한 사실로서 공공의 이익에 관한 때'의 의미
대법원 1981. 10. 27. 선고 81도1023 판결	10	1) 특정한 한 사람에 대한 사실의 유포와 명예훼손죄에 있어서의 공연성
대법원 2003. 10. 26. 선고 2003도6036 판결	9	1) 형법 제309조 제1항, 제2항 소정의 '사람을 비방할 목적'의 의미 및 그 판단 방법 2) 형법 제309조 제1항과 제310조와의 관계 3) 출판물에 의한 명예훼손죄에 있어서 비방의 목적이 인정되지 않는다고 한 사례



〈그림 9〉 최근 10년의 판례 인용 네트워크

〈표 5〉 참조판례를 기반을 추출한 최근 10년간의 핵심 판례

판례	In-Degree Centrality	명예훼손 관련 판시 사항
대법원 2010. 11. 25. 선고 2009도12132 판결	5	1) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제70조 제2항의 허위사실 적시 정보통신망을 통한 명예훼손죄에서 '허위의 인식' 등에 관한 증명책임의 소재 및 그 증명 여부의 판단 기준 2) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제70조에서 정한 '사람을 비방할 목적'이 있는지 여부의 판단 기준 및 '공공의 이익'과의 관계
대법원 2010. 10. 28. 선고 2009도4949 판	4	1) 허위사실 적시에 의한 명예훼손죄의 성립요건인 '허위의 인식'과 이에 대한 증명책임의 소재 및 허위사실 유포 기타 위계에 의한 업무방해죄에도 동일한 법리가 적용되는지 여부(적극) 2) 甲 회사와 乙의 공유인 특허발명에 대해 특허심판원의 무효심결이 내려진 후 확정되기 전에 甲 회사의 대표인 피고인이 '丙이 생산·판매한 제품은 위 특허권을 침해한 제품이다'라는 사실을 인터넷을 통하여 적시하고, 또한 丙의 거래처들에 같은 내용의 내용증명을 발송하였다는 내용으로 기소된 사안에서, 피고인에게 위와 같이 적시된 사실이 '허위'라는 인식이 있었다고 보기 어렵다고 한 사례
대법원 2008. 11. 13. 선고 2006도7915 판결	4	1) 명예훼손죄에서 '사실의 적시'의 의미와 그 정도 2) 형법 제309조 제2항의 허위사실적시 출판물에 의한 명예훼손죄에서 공표사실의 허위성에 관한 증명책임 및 증명의 정도 3) 형법 제309조 제2항의 허위사실적시 출판물에 의한 명예훼손의 공소사실에 대하여 공소장변경 절차 없이 같은 조 제1항의 사실적시 출판물에 의한 명예훼손으로 처벌할 수 있는지 여부(적극) 4) 형법 제309조에 정한 '사람을 비방할 목적'의 의미 및 판단 방법 5) 감사원 소속 공무원이 재벌그룹의 콘도미니엄 사업승인과 관련한 특혜의혹사건에 관하여 기자들에게 "양심선언"이란 제목 아래 감사원 국장이 외부의 압력을 받아 감사를 이유 없이 중단시켰다는 내용의 유인물을 배포한 사안에서, 비방의 목적이나 허위라는 인식이 없으므로 출판물에 의한 명예훼손죄가 성립하지 않는다고 한 사례
대법원 2011. 6. 10. 선고 2011도1147 판결	3	1) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조 제2항의 명예훼손죄 성립요건인 '허위의 인식'과 이에 대한 증명책임의 소재 2) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조 제2항의 명예훼손죄에서 '사실의 적시'의 의미 및 적시한 사실이 '거짓의 사실'인지 판단하는 기준 3) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조 제2항의 명예훼손죄에서 '사람을 비방할 목적'의 의미 및 적시한 사실이 공공의 이익을 위한 것인지 판단하는 기준
대법원 2009. 5. 28. 선고 2008도8812 판결	3	1) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조 제1항에 정한 '사실의 적시'의 의미 및 그 판단 방법 2) 구 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제61조 제1항에 정한 '사람을 비방할 목적'이 있는지 여부의 판단 방법 및 공공의 이익에 관한 것일 경우와의 관계 3) 인터넷 포털사이트의 지식검색 질문·답변 게시판에 성형시술 결과가 만족스럽지 못하다는 주관적인 평가를 주된 내용으로 하는 한 줄의 댓글을 게시한 사안에서, '사실을 적시한 것은 맞지만 '비방할 목적'이 있었다고 보기 어렵다고 한 사례

사전 법률지식이 없더라도, 제시된 정보를 통해 명예훼손에 관한 기본적인 배경 지식을 파악할 수 있었다. 이 방법을 법률정보시스템에 적용하면 일반인들을 위한 시스템으로 충분히 활용할 수 있을 뿐만 아니라 법률 교육에서도 활용될 수 있을 것이다. 대부분의 법학전문대학원의 교육과정에는 사서들에 의해 진행되는 법률정보조사론과 같은 정보시스템 활용 교육이 포함되어 있다(김성은, 정종기 2011). 본 연구에서 제안한 시스템은 이러한 과목에서 법학전문대학원 학생들을 위한 정보 서비스가 될 뿐만 아니라 교육적으로도 활용할 수 있을 것으로 본다. 뿐만 아니라, 법학 관련 전문 지식이 필수적인 법학전문사서를 양성을 위한 교육(오일석, 기성훈 2014)에서도 활용될 수 있을 것이다.

5. 결론

본 연구는 일반인을 위한 법률정보시스템을 구축하기 위해 판례를 대상으로 텍스트 마이닝 기법을 적용하는 방안을 제안하였다. 명예훼손 판례를 대상으로 참조조문, 참조 판례, 판시사항, 판결요지를 활용하여 텍스트 마이닝 기법을 적용한 결과를 제시 하였으며, 그 결과 명예훼손에서 사용되는 법 조항과, 핵심 쟁점, 그리고 핵심 판례를 추출하였고, 실제 명예훼손과 관련 있는 주제들이 등장하는 것을 확인할 수 있었다.

이는 일반인에게도 특정 주제에 관한 핵심적인 정보를 제공해 줄 수 있는 방법이며, 정보서비스뿐만 아니라 교육적인 측면에서도 활용 가

능하다는 점에서 그 의의를 가진다. 비록 명예훼손 관련 대법원 판례인 253개의 판례만을 활용하였고, 일부 수작업이 포함되어 있지만, 일부 수작업을 통해 진행된 과정을 극복하게 된다면, 대량의 판례에도 적용 가능하며, 다양한 법률 주제에도 적용 가능할 것으로 본다.

본 연구와 관련하여 향후 연구 방향으로 몇 가지를 제안하고자 한다. 우선 대법원 판례뿐만 아니라 모든 법원의 판례를 활용한 시스템의 구축이다. 김지현, 서은경(2015)은 변호사들의 법률정보 데이터베이스 개선사항으로 하급심 판례 검색의 확대를 요구하고 있다. 법원에서 공개하고 있는 판례는 주로 대법원 판례이며, 하급심의 판례의 경우에는 일부만 공개하고 있어, 이러한 부분에서 개선이 된다면, 시스템에 활용되는 판례의 범위를 확대하여 보다 상세한 정보를 제공할 수 있을 것이다.

둘째, 판례 텍스트만을 전문적으로 분석하는 형태소 분석기를 구축하는 것이다. 앞서 서술하였듯이 판례 텍스트의 특이성으로 인하여 형태소 분석에 어려움이 있다. 이러한 형태소 분석기를 구축함으로써 보다 상세한 결과를 도출할 수 있을 것이다.

셋째, 판례에서 나타나는 법률 쟁점의 여부를 자동으로 분류하는 방법론에 관한 연구가 필요하다. 판례에는 여러 주제가 혼합되어 있는 것이 일반적이다. 그 중 일부인 특정 쟁점에 대한 인정 여부를 판정하기 위해서는 판례 내에서 특정 쟁점에 관한 텍스트만을 추출하는 방법과 쟁점의 해당 쟁점의 인정 여부를 분류하는 방법이 필요하다. 이러한 연구들은 현재도 지속적으로 이루어지고 있으나 판례 데이터를 대상으로 소수에 불과하다. 따라서 보다 상

제한 정보를 제공하는 법률정보시스템을 제안하기 위해서는 판례내에 나타나는 법적 쟁점의 인정여부를 자동으로 판단해주는 연구가 선행적으로 필요하다.

마지막으로, 법률정보의 네트워크 구축이다. 변호사들은 법률 각 조항들이 연관 정보와 함께

검색되기를 바라고 있다(김지현, 서은경 2015). 실제 법 조항과 판례, 주석서 등은 서로 연관성을 가지고 있다. 이러한 연관성을 바탕으로 정보 네트워크로 구축하여 제시한다면 보다 효과적인 시스템으로 제안될 수 있다고 생각한다.

참 고 문 헌

[1] 김기영. 2017. 인공지능 활용 법률정보서비스의 발전과 관련 쟁점. 『문화·미디어·엔터테인먼트법』, 11(1): 95-113.

[2] 김나리, 김형중. 2017. 연관법령검색을 위한 워드 임베딩기반 Law2Vec 모형 연구. 『한국디지털콘텐츠학회 논문지』, 18(7): 1419-1425.

[3] 김선우, 지선영, 최성필. 2018. 심층 학습 기반의 채권 회수 판례 분석. 『한국정보과학회 언어공학연구회』, 373-377, 학술대회논문집.

[4] 김성은, 정종기. 2011. 우리나라 법학전문대학원 교육과정 실행에 나타난 법률 정보활용교육 실태 분석. 『한국문헌정보학회지』, 45(4): 103-122.

[5] 대법원. “종합법률정보.” [online] [cited 2020. 2. 5.] <<https://glaw.scourt.go.kr/wsjo/intsrch/sjo022.do>>

[6] 로앤비. “LAWnB.” [online] [cited 2020. 2. 5.] <<https://www.lawnb.com/>>

[7] 박은정, 조성준. 2014. KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지. 『제26회 한글 및 한국어 정보처리 학술대회 논문집』, 2014년 10월 10일, 춘천: 강원대학교 춘천캠퍼스: 133-136.

[8] 법제처. “국가법령정보센터.” [online] [cited 2020. 2. 5.] <<http://www.law.go.kr/LSW/main.html>>

[9] 심준식, 김형중. 2017. LDA 토픽 모델링을 활용한 판례 검색 및 분류 방법. 『전자공학회논문지』, 54(9): 67-75.

[10] 오일석, 김성훈. 2014. 법학전문사서의 교육과 육성에 관한 연구. 『서울법학』, 21(3): 483-526.

[11] 이테일리. “양육비 미지급 부모 신상 공개...공익 활동 vs 명예훼손.” [online] [cited 2020. 2. 5.] <<https://www.edaily.co.kr/news/read?newsId=01413686622686376&mediaCodeNo=257&OutLnkChk=Y>>

[12] 이재윤. 2006. 지적 구조의 규명을 위한 네트워크 형성 방식에 관한 연구. 『한국문헌정보학회지』,

- 40(2): 333-355.
- [13] 이재윤. 2013. tnet과 WNET의 가중 네트워크 중심성 지수 비교 연구. 『정보관리학회지』, 30(4): 241-264.
- [14] 이지훈, 엄태현, 이혁준. 2019. TF-IDF 와 CNN 을 사용한 법률문서 분류 시스템에 관한 연구. 『한국통신학회 학술대회논문집』, 982-983.
- [15] 이지희, 이준성, 손정욱. 2016. R 프로그래밍 기반의 비정형 건설 데이터 분석: 해외건설 분쟁 판례 데이터를 대상으로. 『대한건축학회 논문집-구조계』, 32(5): 37-44.
- [16] 장소원. 2009. 법률 텍스트 문장의 문법성. 『텍스트언어학』, 27: 1-29.
- [17] 정진수, 강태경, 김형길. 2015. 명예훼손죄·모욕죄에 대한 판례의 판단 기준 연구: 최근 10년간 (2005~2015)의 판례를 중심으로. 『형사정책연구원 연구총서』, 1-109
- [18] 홍일표. 2007. 판례의 형성과 구속력의 범위. 『일감법학』, 12: 1-41.
- [19] Bastian M., Heymann S. and Jacomy M. 2009. "Gephi: an open source software for exploring and manipulating networks." *International AAAI Conference on Weblogs and Social Media*.
- [20] Bommarito, M. J., Katz, D. M. and Zelner, J. 2009. "Law as a Seamless Web? Comparison of Various Network Representations of the United States Supreme Court Corpus (1791-2005)." *Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL 2009)*, 234-235.
- [21] Blondel, V. D., Guillaume, J. L., Lambiotte, R. and Lefebvre, E. 2008. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment*, 2008(10): P10008.
- [22] Kumar, R. and Raghuvver, K. 2012. "Legal document summarization using latent dirichlet allocation." *International Journal of Computer Science and Telecommunications*, 3: 114-117.
- [23] Leibon, Greg, Michael Livermore, Reed Harder, Allen Riddell and Dan Rockmore. 2018. "Bending the law: geometric tools for quantifying influence in the multinet network of legal opinions." *Artificial Intelligence and Law*, 26(2): 145-167.
- [24] Schvaneveldt, R. W. 1990. *Pathfinder associative networks: Studies in knowledge organization*. New York: Ablex Publishing.
- [25] Van Der Haegen, M. 2017. "Building a Legal Citation Network: The Influence of the Court of Cassation on the Lower Judiciary." *Utrecht Law Review*, 13(3): 65-76.
- [26] Zhong, Haoxi, Zhipeng Guo, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu and Maosong Sun. 2018. "Legal judgment prediction via topological learning." *In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3540-3549.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kim, K. Y. 2017. "Development of legal information service utilized AI and related issues." *The institute for Culture, Media, and Entertainment Laws*, 11(1): 95-113.
- [2] Kim, N. and Kim, H. J. 2017. "A Study on the Law2Vec Model for Searching Related Law." *The Journal of Digital Contents Society*, 18(7): 1419-1425.
- [3] Kim, S. W., Ji, S. Y. and Choi, S. P. 2018. "Analysis of Judicial Precedent Information related to Debt Recovery based on Deep-Learning." *In Annual Conference on Human and Language Technology*, 373-377. Human and Language Technology.
- [4] Kim, S. E. and Jung, J. K. 2011. "An Analysis of Information Literacy Education Based on the Operation of Korean Law Schools Education Curriculum." *Journal of the Korean Society for Library and Information Science*, 45(4): 103-122.
- [5] Supreme Court of Korea, *Comprehensive law information*. [online] [cited 2020. 2. 5.] <<https://glaw.scourt.go.kr/wsjo/intsrch/sjo022.do>>
- [6] LAWnB, *LAWnB*. [online] [cited 2020. 2. 5.] <<https://www.lawnb.com/>>
- [7] Park, Eunjeong L. and Cho, Sungzoon. 2014. "KoNLPy: Korean Natural Language Processing in Python." *In Proceedings of the 26th Annual Conference on Human and Cognitive Language Technology*, October 10, 2014, Chuncheon: Kangwon National University Chuncheon Campus, 133-136.
- [8] Korea Ministry of Government Legislation, *National Law Information Center*. [online] [cited 2020. 2. 5.] <<http://www.law.go.kr/LSW/main.html>>
- [9] Sim, J. S. and Kim, H. J. 2017. "A Searching Method for Legal Case Using LDA Topic Modeling." *Journal of the Institute of Electronics and Information Engineers*, 54(9): 67-75.
- [10] Oh, I. S. and Kim, S. H. 2014. A Study on Educating and Training Law Librarians. *Seoul Law Review*, 21(3): 483-526.
- [11] edaily. *The disclosure of parents who unpaid child support. Public interest activities vs defamation*. [online] [cited 2020. 2. 5.] <<https://www.edaily.co.kr/news/read?newsId=01413686622686376&mediaCodeNo=257&OutLnkChk=Y>>
- [12] Lee, J. Y. 2006. "A Study on the Network Generation Methods for Examining the Intellectual Structure of Knowledge Domains." *Journal of the Korean Society for Library and Information Science*, 40(2): 333-355

- [13] Lee, J. Y. 2013. "A comparison study on the weighted network centrality measures of tnet and WNET." *Journal of the Korean Society for information Management*, 30(4): 241-264.
- [14] Lee, J. H., Eom, T. H. and Lee, H. J. 2019. A Study on Legal Document Classification Systems Using TF-IDF and CNNs. *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, 982-983.
- [15] Lee, J. H., Yi, J. S. and Son, J. W. 2016. "Unstructured Construction Data Analytics Using R Programming - Focused on Overseas Construction Adjudication Cases." *Journal of architectural institute of Korea structure & Construction*, 32(5): 37-44
- [16] Chang, S. 2009. "The Grammaticality of Legal Texts." *Text linguistics*, 27: 1-29.
- [17] Chung, J. S., Gahng, T. Y. and Kim, H. K. 2015. "A Study on the Korean Courts Standards for Defamation and Insult: Focused on Defamation and Insult Decisions during the Last Ten Years (2005-2015)." *Korean Institute of Criminology*, 1-109.
- [18] Hong, I. p. 2007. "Judicial Precedent and the Scope of Its Binding Force." *Ilkam Law Review*, 12, 1-41.