

토픽모델링을 활용한 공공갈등 유형 및 경향 분석*

Analysis on Types and Trends of Public Conflicts using Topic Modeling

양 연 희**

Yang, Yeon Hee

■ 목 차 ■

- I. 서론
- II. 이론적 논의 및 선행연구 검토
- III. 분석틀 구성
- IV. 분석결과
- V. 결론

본 연구는 국내 공공갈등 관련 신문기사를 활용하여 국내에서 발생하는 공공갈등의 유형과 경향을 분석하였다. 분석결과, 공공갈등은 총 11개 유형-개발·환경, 선호시설 유치, 교통인프라, 비선호시설 입지, 군사시설, 무상복지, 지역발전사업 유치, 전력시설물, 행정구역 변경, 원전시설물, 공공시설물 명칭-으로 분류되었다. 이는 국내에서 발생한 공공갈등은, 선행연구에서 공공갈등 유형을 갈등 표출여부, 갈등성격, 갈등원인, 갈등주체 등의 다양한 기준으로 분류한 것과는 달리, 갈등을 유발하는 정책대상을 기준으로 분류할 수 있었다. 그리고 비선호시설 입지갈등이 가장 많이 발생하는 것을 확인할 수 있었으며, 그 중에서도 전력, 원전 및 군사시설과 관련된 갈등이 각각의 유형으로 분류될 만큼 두드러지게 발생하는 특징을 보여주었다. 이것은 전력, 원전, 군사시설 관련 갈등은 이해관계가 첨예하게 대립하여 표면적으로 갈등이 표출되거나 장기화되는 경향을 가진다는 것을 말해준다. 따라서 관련 정책 수립 시에는 비선호시설이라도 각 유형별 특징에 맞는 갈등관리체계가 구축되어야 함을 시사한다. 한편 공공갈등의 경향분석을 통해서 대부분의 공공갈등은 증가하는 추세이므로 선제적인 갈등관리시스템 구축의 필요성을 보여준다. 이는 후속연구에서 국내 공공갈등의 유형에 따른 갈등해결요인과 관리방안을 파악하기 위한 자료가 될 수 있다

* 본 연구는 박사학위 논문 「토픽모델링을 활용한 공공갈등 유형화와 해결방안에 관한 연구」(2020) 일부를 요약·수정한 것입니다.

** 행정안전부 사무관, 부산대학교 공공정책학 박사

논문 접수일: 2021. 5. 14. 심사기간: 2021. 5. 14. ~ 2021. 6. 14. 게재확정일: 2021. 6. 14.

는 데에 의의가 있다.

□ 주제어: 공공갈등, 공공갈등관리, 갈등 유형화, 경향분석, 토픽모델링

This study analyzed types and trends of public conflicts in South Korea using news articles on public conflicts in South Korea. The analysis results showed that public conflicts are classified into a total of 11 types; Development/environment, attraction of preferred facilities, transportation infrastructure, locations of non-preferred facilities, military facilities, free welfare, attraction of projects for regional development, electric power facilities, changes of administrative districts, nuclear power facilities, and names of public facilities. Unlike previous studies, where public conflict types were classified by various criteria such as whether conflicts were expressed, conflict nature, conflict causes, and conflict subjects, public conflicts that occurred in South Korea were able to be classified according to policy objects that caused conflicts. In addition, conflicts on locations of non-preferred facilities occurred the most, of which conflicts related to electric power, nuclear power plants, and military facilities occurred prominently enough to be classified into each type. This indicates that conflicts related to electric power, nuclear power plants, and military facilities have a tendency to be expressed to the surface or to be prolonged. Therefore, it suggests that a system for conflict management suitable for each type, should be established when establishing relative policies even for non-preferred facilities. On the other hand, analysis of trends of public conflicts shows that most public conflicts tend to increase, so it is necessary to establish a preemptive conflict management system. This study is meaningful in that it can serve as data to identify conflict resolution factors and management methods according to types of public conflicts in South Korea in subsequent studies.

□ Keywords: Public Conflicts, Public Conflict Management, Conflict Typification, Trend Analysis, Topic Modeling

I. 서론

정부정책 사업으로 인한 공공갈등은 세계적으로 나타나는 보편적이고 필연적인 현상으로 다양한 이해관계자가 존재하는 사회에서는 자연스러운 현상이다. 이와 같이 사회가 복잡해지고 다변화되어감에 따라 갈등해결을 규제나 통제와 같은 기존의 정책수단에 의지하기에는 힘들어졌다. 또한 최근 다원화, 민주화, 지방화, 정보화 등의 현대 행정환경의 변화에 따라 다양한 행위자들의 참여가 확대되는 것도 공공정책의 전반적인 과정에서 갈등이 증폭되는 계기가 되기도 한다(김정인, 2018). 이를 해결하기 위하여 학술적·실무적인 차원에서의 연구가 진행되고 있으며, 지방정부, 지역주민, 시민단체 등의 다양한 이해관계자가 추구하는 복잡한 가치의 차이를 고려한 협력과 참여를 기반으로 한 정책수단이 요구됨에 따라 사회문제가 발생하는 맥락에 대한 정확한 분석이 필요하다(강창민, 2015).

한편 지금까지 선행연구 및 소수 사례분석을 통한 공공갈등 유형을 정리한 연구가 대부분으로 갈등유형을 좀 더 학술적으로 정교화하는 측면(김대중, 2019)과 갈등 사례들을 더 많이 취합하여 우리나라에서 발생하는 공공갈등의 유형과 경향을 파악할 필요성이 제기되어 왔다. 이러한 선행연구의 한계 및 필요성의 일환으로 본 연구는 일정기간동안 국내에서 발생한 공공갈등의 유형과 경향을 파악하기 위하여 공공갈등 관련 신문기사를 활용하여 텍스트 분석을 통하여 국내에서 발생한 공공갈등의 유형과 경향의 특징을 파악하고자 하였다.

따라서 본 연구에서는 공공갈등을 해결하기 위하여 과거 유사한 갈등사례에서 갈등해결의 정책적 함의를 찾는 것에 의미를 두고자 한다. 이를 위하여 지금까지 갈등연구가 갈등 사례별, 연구자의 주관적 인식에 따라 연구범위가 정해지는 점에 착안하여, 우리나라에서 갈등을 공식적인 정부 어젠다(agenda)로 다루기 시작한 「공공기관 갈등 예방과 해결에 관한 규정」이 시행(2007.5.13.)된 이후 지금까지 일어난 모든 국내 갈등 사례를 연구범위로 하여 국내 공공갈등의 유형과 공공갈등의 경향을 탐색하였다.

II. 이론적 논의 및 선행연구 검토

1. 공공갈등 유형분석 및 한계

사회가 다양하고 복잡하게 발달하면서, 정부와 관련된 정책적 사안에 대해서도 다수의 이해관계자들이 복잡하게 얽혀서 대립하는 갈등이 증가하고 있다. 이와 같은 갈등은 사회 전체의 비용을 높게 하는 부정적인 역할을 하지만, 갈등이라는 것이 사회 구성원 간의 관계에서 자연스럽게 이루어지는 것임을 고려한다면 갈등의 발생보다는 진단과 해결에 초점을 두는 것이 더욱 바람직할 것이다.

공공갈등과 관련한 선행연구들을 살펴보면 다양한 시각에서 연구가 진행되고 있는 것을 알 수 있다. 임동진(2012)은 국내의 공공갈등 선행연구를 공공갈등의 개별사례분석을 통한 갈등 해결방안 제시, 다수 갈등사례 분석을 통한 갈등을 유형화, 일반화 및 이론화, 공공갈등을 효과적으로 관리하는 방안에 관한 연구, 공공갈등관리에 대한 법·제도·기구에 대한 연구 등으로 분류하였다. 또한 주상현(2012)은 갈등연구가 갈등원인을 분석하고 그에 따른 효과적인 갈등관리방안을 연구하는 시각과(전주상, 2000; 이민창, 2005), 갈등관리전략에 대한 논의에서 갈등사례를 통해 갈등관리전략의 특성을 분석하는 연구(임정빈, 2007)로 설명하고 있다.

선행연구를 통하여 공공갈등에 관한 주요 선행연구를 다음과 같이 크게 두 가지 측면으로 구분하여 정리하였다. 첫째, 공공갈등 연구에서 가장 많은 비중을 차지하고 있는 연구로서 공공갈등의 유형연구로서, 공공갈등 사례분석을 통해 갈등의 진행과정을 보여주면서 과정 속에서 경험적 시사점을 제시하는 연구이다. 둘째, 공공갈등 해결방안에 관한 연구로서 먼저, 갈등요인에 관한 연구이다. 갈등요인에 따라 갈등해결방안이 다양하게 도출될 수 있으며, 갈등 발생 및 증폭의 요인파악이 가능하다는 의의가 있다. 그리고 공공갈등관리에 관한 법·제도적 연구로서 향후 갈등관리 및 해결기제를 위한 시스템을 제시하는 데 의의가 있다.

한편, 공공갈등의 개념정의를 다양한 것은 공공갈등 자체의 특징과 밀접한 관련이 있다. 공공갈등의 양태가 매우 다양하기 때문에 다양한 공공갈등 전체를 아우르는 종합적인 개념 정의를 이끌어내기는 어렵기 때문이다(심준섭, 2014). 따라서 이러한 한계를 극복하기 위해 국내의 갈등 연구자들은 공공갈등의 유형을 분류하기 위한 노력을 해왔다. 국외연구에서는 갈등원인에 따라 분류하거나(Deustch, 1973; Moor, 2003; Kriesberg, 2003), 개발갈등에 따라 분류하는 경우도 있다(Campbell, 1996). 한편 국내연구를 살펴보면 갈등표출 여부를 기준으로(서울시정개발연구원, 2003) 구분하거나, 갈등의 원인이 된 정책의 특성에 따라 분류하거나(하혜영, 2007), 갈등당사자의 특성에 따라 분류하는 경우도 있다(하혜영, 2007; 정용덕, 2010).

국내의 공공갈등 유형분류의 연구결과를 종합해본 결과, 가장 대표적인 분류방식은 이익갈등, 가치갈등, 권한갈등, 정책갈등과 같이 갈등의 성격 및 원인, 내용을 기준으로 분류하는 방식이다(Deustch, 1973; Kriesberg, 2003; Moor, 2003; 서울시정개발연구원, 2003; 지속가능발전위원회, 2005; 하혜영, 2007; 정용덕, 2010; 국무총리실, 2010; 임동진, 2011).

이상의 공공갈등 유형분류의 선행연구를 정리하면 아래 <표 1>과 같다.

<표 1> 국내외 선행연구의 공공갈등의 유형분류

연구자	유형분류 기준		예시 및 유형분류 근거
서울시정개발연구원 (2003)	갈등 표출여부	잠재적 갈등	표출 안 됨
		현재적 갈등	표출됨
	갈등성격	이익갈등	자원의 배분을 놓고 발생하는 갈등
		권한 갈등	법적 권한을 누가 가지는지에 대한 갈등
	갈등내용	지방행정/재정 갈등	지방행정, 지방재정
지역개발갈등		하천, 광역시설, 지역개발 등	
지속가능발전위원회 (2005)	갈등원인	이해갈등	자원의 배분을 놓고 발생하는 이익추구 갈등
		가치갈등	이념, 가치관, 종교적 신념, 지향점 차이에 따른 갈등
하혜영 (2007)	정책특성	선호	핍피
		비선호	넘비
	갈등당사자 간 특성	당사자의 수	갈등당사자 수(이해관계자 모두 포함)
		당사자유형	정부간 갈등, 정부·주민 간 갈등, 시민단체, 환경단체
	갈등이슈	당사자 간 신뢰도	상호 대립 또는 신뢰 수준
가치갈등		이념적 대립, 개발·환경 간 갈등 등	
국무총리실 (2010)·임동진 (2011)	갈등내용	이해갈등	자원배분과 관련된 갈등
		정책 갈등	사례를 바탕으로 귀납적 분류
		이익 갈등	
		입지 갈등	
		노사갈등	
정용덕 (2010)	갈등의 장 · 갈등주체	정부부문	부처간 갈등, 부서간 갈등, 중앙정부·지방정부간 갈등
		사회부문	민간단체 간 갈등, 기업 간 갈등, 계층간 갈등, 세대간 갈등, 노사갈등 등
		정부·사회부문	정부·사회부문간 갈등
	갈등유발 대상	경제계의 경쟁적 배분	빈부 갈등, 노사 갈등, 이익 갈등
		의사결정권의 배분	권력/권한의 배분을 위한 갈등
		가치(이념) 갈등	이념, 문화, 헤게모니(hegemony) 갈등
		정체성 갈등	성, 연령, 종교, 인종, 민족 등 개인이나 집단의 정체성 문제 간 갈등

연구자	유형분류 기준		예시 및 유형분류 근거
Deustch (1973)	갈등원인	자원 갈등	한정된 자원의 배분에 대한 갈등
		선호 갈등	대안의 선호에 대한 갈등
		가치갈등	개발·환경 갈등이 대표적
		신념갈등	이념, 종교적 갈등이 대표적
Campbell (1996)	지속가능한 개발갈등	개발갈등	환경보호와 사회적 평등 관련 갈등
		자원갈등	환경보호와 경제발전 관련 갈등
		자산갈등	사회적 평등과 경제발전 관련 갈등
Kriesberg (2003)	갈등원인	이익갈등	자원의 배분을 놓고 발생하는 이익추구 갈등
		가치갈등	이념, 가치관, 종교적 신념, 지향점 차이에 따른 갈등
Moor (2003)	갈등원인	이해갈등	자원의 배분을 놓고 발생하는 이익추구 갈등
		가치갈등	이념, 가치관, 종교적 신념, 지향점 차이에 따른 갈등
		관계 갈등	강한 거부감, 잘못된 지각, 소통결여, 반복된 부정적 행동 등으로 발생
		데이터 갈등	잘못된 정보, 적합한 자료에 대한 의견 불일치, 정보해석 차이, 상반된 평가절차
		구조적 갈등	제한된 물질적 자원이나 권위 등 갈등상태에 있는 사람 외부에 존재하는 힘에 의한 발생

자료: 선행연구를 근거로 연구자 재구성

이상과 같이 공공갈등 유형에 관한 선행연구를 검토한 결과, 공공갈등 유형은 갈등의 특성을 도출해 낼 수 있는 갈등의 표출여부, 갈등의 성격, 갈등요인, 갈등 이해관계자의 특성, 갈등의 주체 등 다양한 기준을 통하여 분류하였다.

하지만 이와 같은 유형분류가 실제 공공갈등을 완벽하게 분류해 내기에는 한계가 있다. 예를 들어 원자력발전소 입지갈등은 이해갈등, 가치갈등, 관계갈등, 데이터 갈등, 구조적 갈등 등으로 분류할 수 있는데, 이는 갈등원인을 가장 다양하게 분류하고 있는 Moor(2003)의 갈등유형 분류의 모든 범주를 포함하고 있다. 그러나 국무총리실(2010)과 임동진(2011)의 연구에서 사용된 공공갈등 유형분류에 따르면 입지갈등에 해당한다. 반면 정용덕(2010)의 분류에 따르면 정부-사회부문 간 갈등이지만, 갈등원인별 분류로는 명확하게 구분하는 것은 어렵다. 이는 복잡한 가치와 이해관계가 역동적으로 표출되는 공공갈등의 고유한 특성으로 인한 결과로, 본 연구는 이러한 공공갈등의 특성을 인식하고 이러한 갈등 특성을 기준을 종합적으로 접근하여 분류해보고자 하였다.

2. 토픽모델링 분석을 통한 공공갈등 유형연구의 접근

1) 토픽 모델링의 이해

인공지능과 딥러닝(Deep Learning)을 기반으로 한 빅데이터 분석에서 비정형 데이터를 기반으로 하는 모형의 비중이 커지고 있으며 딥러닝의 성장으로 이미지와 텍스트에 대한 분석 요구가 점차 많아지고 있다(박상언 외, 2019). 이와 같이 비정형 데이터에 대한 분석은 기존의 정형 데이터를 활용한 분석에서 발견하지 못하는 새로운 정보 또는 가치를 발견하는 데에 의의가 있다.

빅데이터 분석 기법 중 텍스트 마이닝은 비정형(혹은 무정형)의 특징을 지닌 자연어 텍스트를 특정한 목적에 유용한 정보를 추출하기 위한 분석방법으로(양연희 외, 2019; Inzalkar & Sharma, 2015), 텍스트 데이터에 나타나는 단어를 이용하여 분석을 수행하고, 텍스트 분류, 문서 군집화, 언어 인식, 핵심문구 식별 등 문서의 유사성을 측정하는 방식으로 확대되기도 한다(손복은·고호경, 2018; Khan & Kanth, 2016).

토픽모델링은 구조화되지 않은 방대한 문헌집단에서 일정한 패턴을 발견하여 주제를 찾아내기 위한 알고리즘으로(Blei, 2012; Steyvers & Griffiths, 2007), 기존의 단어 빈도 분석에서 발생하는 희귀성 문제(Sparsity Problem), 다의성(Polysemy), 동의어(Synonymy), 의미계층구조(Semantics Hierarchical Structure) 등의 문제점을 해결하는 대안으로 제시되고 있다. 맥락과 관련된 단서들을 이용해 유사한 의미를 가진 단어들을 클러스터링(clustering)하는 방식으로 주제를 추론하는 방법론으로(Calabrese et al., 2019; Potha & Stamatatos, 2019; Xiong et al., 2019; 신동희, 2016), 본 연구에서는 토픽모델링을 이용하여 신문기사를 통해 토픽을 추출하고 이 토픽들이 시간에 따라 어떻게 변화되는지를 파악해보고자 한다.

토픽 모델링은 일반적으로 두 가지 모형이 있는데, 첫 번째는 분절된 단어들에 벡터 값을 부여하고 차원 축소를 통해 근접한 단어들을 묶어내는 모형(latent semantic analysis: LSA)이다. 두 번째는 확률기법을 활용하여 문서의 토픽 분포를 추론하는 모형(latent dirichlet allocation: LDA)이 있다. 본 연구에서는 LDA 분석을 통해 각 문서를 구성하는 토픽의 비중을 추출하고, 각 토픽을 설명하는 키워드들의 비중과 함께 보여주게 된다¹⁾. 토픽 레이블(labeling)에 있어서는 각 토픽의 구성비중이 높은 키워드를 중심으로 하고, 토픽비중이 높은 문서와 함께 연구자가 종합적인 판단을 하여 결정한다. 데이터를 통해 각 문서를 구성하는 키

1) 최근에는 LSA의 단점을 보완한 LDA 기법이 다량의 문서의 주제를 찾는 연구에서 많이 사용되고 있다(김민관, 이용, 한창희, 2017).

워드들의 확률분포를 구할 수 있으므로 LDA에서는 이를 기반으로 하여 문서를 구성하는 토픽의 확률분포와 토픽을 구성하는 키워드의 확률분포를 확률적으로 추정하게 된다. 따라서 본 연구에서 시기에 따른 토픽 트렌드 분석은 각 논문의 토픽 비중을 추출하여 활용하게 된다 (양연희 외, 2019).

$$\varphi^{(t)} = \frac{\sum_{d=1}^{D_t} \theta^{(d)}}{D_t}$$

* $\varphi^{(t)}$: 기간 t의 토픽 분포 / D_t : 기간 t에 출판된 논문의 수 / $\theta^{(d)}$: 문서 d의 토픽 분포

위의 공식을 이용하여 주어진 기간 동안의 문서들에 대한 토픽분포에 대한 평균을 계산하여 변화추이를 계산할 수 있는데(Kim et al., 2015), 이를 통해 시간이 지날수록 토픽 비중이 증가한다면 주목받는 주제로 해석이 가능하고, 토픽의 비중이 감소한다면 비교적 덜 중요시되거나 관심을 줄어드는 주제라고 볼 수 있다. 이처럼 토픽모델링을 활용하여 문서를 구성하는 토픽들의 비중으로 각 문서의 주요 주제와 시간에 따른 토픽의 변화를 확인할 수 있다 (양연희 외, 2019).

2) 토픽모델링 분석을 통한 공공갈등 유형연구의 차별성

선행연구를 통하여 다양한 분야에서 토픽 모델링이 이용되고 있다는 것을 알 수 있는데, 연구의 경향을 분석하기 위하여 토픽모델링을 활용하는 분석은 2000년대 이후부터 본격화되었다. 국외 연구에서는 Griffiths·Steyvers(2004)는 미국 국립과학원 회보(PNAS)의 초록을 연구대상으로 토픽모델링 분석을 수행하여 의미있는 토픽, 동시에 가장 활발하게 연구되는 주제(Hot Topics)와 점차적으로 사라지는 주제(Cold Topics)를 발견하였다. Wang·McCallum(2006)은 LDA 토픽 모델링을 활용하여 시간의 흐름에 따른 토픽변화를 확인할 수 있는 TOT(Topics Over Time) 모델을 제시하였다. Blei(2012)는 저널「Science」의 논문 17,000편과 저널 「Yale Law」을 대상으로 LDA 토픽 모델링을 소개하였고, 토픽 모델링이 심리학, 정치학, 계량서지분석에도 활용이 가능함을 밝혔다.

국내 연구로서는 허고은(2019)은 과학적 지식에 관한 불확실성의 특징을 파악하기 위하여 생의학 문헌을 대상으로 토픽모델링을 수행하고 과학적 지식의 표현과 불확실성 간의 연관성을 분석하였다. 박상언 외(2019)는 텍스트 마이닝 기법을 활용하여 정부활동 주체인 ‘공무원, 관료, 공직’에 대한 한국 행정학의 연구 경향을 분석하였다. 조재희·조인호(2019)는 2018년 메르

스가 유입되었던 초기의 자료의 데이터를 추출하여 토픽모델링을 수행하여 메르스 유입과 관련 주요 이슈들을 분석하였다. 또한 이준웅·김성희(2018)은 우리나라 중앙일간지의 미세먼지 보도의 특징을 검토하기 위해 분석유목을 정해놓고 뉴스 프레임을 확인하는 전통적인 내용분석 방법이 아닌 ‘자료 기반적’ 방법으로 토픽모델링 중 구조적 주제모형(STM)을 적용해서 뉴스 프레임을 추출한 결과와 비교분석하였다. 그 외 서병조·신선영(2017)은 1987년부터 2017년까지의 논문 215편을 대상으로 토픽모델링을 활용하여 플랫폼 정부의 연구동향을 분석하였다.

〈표 2〉 토픽모델링 분석의 특징

구분	내용	분석효과	비고
토픽 모델링 분석	<ul style="list-style-type: none"> 구조화되지 않은 방대한 문헌 집단에서 주제를 찾아내기 위한 알고리즘 유사한 의미를 가진 단어들을 그룹화 하는 방식으로 주제를 추론하는 모델 	<ul style="list-style-type: none"> 자료의 주제와 각 자료별 주제 비중, 각 단어들이 각 주제에 포함될 확률 등을 파악할 수 있음 	<ul style="list-style-type: none"> Griffiths·Steyvers(2004) Wang·McCallum(2006) Blei(2012) 서병조·신선영 (2017) 이준웅·김성희(2018) 박상언 외(2019), 허고은(2019) 조재희·조인호(2019)

자료: 선행연구를 근거로 연구자 구성

이상과 같이 본 연구에서는 텍스트 분석을 활용하여 ‘대한민국’의 특징적인 공공갈등 유형을 도출하는 것을 목적으로 한다. 이에 본 연구에서는 약 12년간(2007년 5월~2018년 12월)의 공공갈등 관련 신문기사 데이터를 기반으로 토픽모델링 분석을 통하여 공공갈등을 유형화하였다.

Ⅲ. 분석틀 구상

1. 연구문제 도출

본 연구는 국내에서 발생한 ‘공공갈등’의 경향과 유형을 밝히기 위하여 아래와 같은 연구문제를 도출하였다. 이러한 연구문제는 기존에 진행된 연구 흐름의 전환, 그리고 공공갈등의 예방 및 해결을 위한 방향을 제시하는 것을 목적으로 한다.

구체적으로는 국내에서 발생한 공공갈등의 유형은 어떠한가, 시기별로 갈등 유형의 비중변화를 통해 공공갈등 트렌드가 보여주는 시사점을 알아보고자 한다.

- 연구문제 1: 국내에서 발생한 공공갈등은 어떤 유형으로 분포되어 있는가?
- 연구문제 2: 시기별로 공공갈등 유형의 비중변화는 어떻게 변화되었는가?

2. 분석대상 및 자료수집

1) 분석대상 범위

본 연구에서는 대표적인 비정형 데이터인 신문기사의 텍스트, 즉 질적 자료를 연구대상으로, 공공갈등을 종합적인 관점에서 유형을 분류하기 위하여 양적 방법인 토픽모델링 기법을 활용하고, 그 결과에 대해 질적인 해석을 시도하였다(Shim et al., 2015; 임도빈, 2009; 심준섭, 2010).

본 연구는 공공갈등 관련하여 국내 신문사 중 중앙지에 실린 기사를 활용하여, 기술통계 분석과 토픽 모델링분석을 진행하였다²⁾. 기술통계 분석을 통해서는 분석에 활용된 기사의 분포 결과와 이후 토픽 모델링 분석을 통해 도출된 공공갈등 유형의 비중변화를 파악하였다. 그리고 텍스트 마이닝 기법 중 토픽 모델링 분석을 이용하여 국내 공공갈등의 유형과 전체 맥락을 파악하고, 연도별 공공갈등이 갖는 이슈 및 시사점의 변화를 분석하였다.

연구의 분석대상의 범위는 크게 공간·대상적 범위, 시간적 범위로 구분할 수 있다. 첫째, 공간적·대상적 범위는 국내에서 발생한 갈등사례 중에서 정부의 정책이나 사업을 수행하는 과정에서 정부와 정부, 정부와 민간 사이에서 발생하는 공공갈등 관련 신문기사 내용을 대상으로 하였다. 이때 민간과 민간 간의 갈등은 대규모 노사갈등과 같이 그 규모가 크거나 사회적 영향이 큰 경우도 공공갈등으로 보는 견해가 있으나, 본 연구의 범위에서는 제외하고, 정부가 주요 갈등의 주체로 개입되는 사례만을 분석대상으로 한정하였다. 또한 해당 기간 내 발생 또는 진행, 완료된 갈등을 대상으로 하였다. 이때 완료라는 것은 갈등의 완벽한 해소가 아닌 잠정적으로 일단락된 경우를 포함하며, 그 중에서 직접적으로 외부로 표출되어 사회적으로 이슈가 된 갈등위주로 선정하여 분석하였다. 이때 외부 표출의 기준은 국내 신문사 중 중앙지에 보도된 갈등으로 선정하였다.

둘째, 연구의 시간적 범위로는 대통령령인 「공공기관의 갈등 예방과 해결에 관한 규정」이

2) 기존에는 공공갈등사례를 통해 이슈를 발견하고 갈등관리방안을 도출하기 위하여 문헌 조사 또는 전문가 인터뷰 등과 같은 정성적 연구방법이 주로 사용되어 왔다. 그러나 이 방법으로 대량의 정보로부터 결과를 도출해기에는 많은 시간과 비용이 소요될 뿐만 아니라 연구자의 주관적 가치가 반영될 가능성이 크다. 이와 같은 한계점을 보완하고자 텍스트 마이닝과 같은 정량적 연구방법을 보완적으로 활용하는 방식으로 트렌드를 분석하는 연구 방법론의 패러다임 변화가 시도되고 있다(김상겸 외, 2016; 김성근 외, 2016).

시행된 2007년 5월 13일부터 2018년 12월 31일까지 약 12년 동안 국내에서 발생한 공공갈등 사례를 모집단으로 한다. 「공공기관의 갈등 예방과 해결에 관한 규정」의 제정은 중앙행정기관의 갈등 예방과 해결에 관한 역할과 책무 및 절차 등을 규정하고 중앙행정기관의 갈등 예방과 해결 능력을 향상시킴으로써 사회적 합의를 통한 정책수용성 제고와 사회통합을 목적으로 하고 있다.

본 규정을 통해 중앙행정기관과 국민을 포함한 이해관계집단 간에 대화와 타협, 자율과 신뢰, 참여와 협력에 기반한 갈등해결 문화를 형성하고, 중앙행정기관의 갈등관리기제의 발전에 기여할 것으로 기대된다. 또한 해당 규정을 기점으로 정부가 자발적·적극적으로 갈등관리의 의사를 표명한 것으로 볼 수 있으며, 중앙행정기관뿐만 아니라 지자체 및 그 밖의 공공기관의 갈등관리제도 운용의 근간이 되었다는 점에서 의의가 있다. 따라서 본 연구의 시간적 범위를 해당 법령의 시행을 기점으로 설정한다.

2) 분석대상 수집

갈등사례분석의 연구대상인 신문기사 수집은 신문검색 종합사이트인 ‘빅카인즈’³⁾를 사용한 다. 갈등이 전국적으로 표출된 사례로 선정하기 위해 신문사는 중앙지와 주요 경제지로 선정한다. 신문검색에서 키워드는 ‘갈등’으로만 설정하고, 2007년 5월 15일부터 2018년 12월까지 기간에 발생 또는 진행, 완료된 기사를 검색하였다. 그 결과로 얻은 기사에서 연구자가 읽고 본 연구와 부합되는 갈등사례일 경우를 분석하였다. 빅카인즈는 한국언론진흥재단에서 운영하는 신문·방송 등 각종 매체의 27년 동안의 5천만 건이 축적된 뉴스 DB로서 빅데이터 기술을 접목해 만든 뉴스 분석서비스이다. 데이터 축적에 있어 형평성과 공공성을 가지므로 연구논문 및 정책보고서의 언론동향 및 데이터 수집에 활용되고 있다. 연구목적에 맞게 수집한 신문기사의 전문을 분석하며, 신문기사와 같이 비정형 데이터의 키워드는 넷마이너(Netminer) 프로그램을 사용하여 추출한다.

3. 토픽모델링 분석과정

1) 토픽모델링 분석의 목적

대부분의 공공갈등은 반복적으로 나타난다. 각 유형별 공공갈등은 유사하거나 동질적인 특

3) 빅카인즈(BIGKINDS) 홈페이지 : www.kinds.or.kr

성을 보유하고 있고, 공공갈등을 유발하는 핵심적인 원인 역시 유형별로 유사할 가능성이 높다. 이런 점을 고려하면 기존의 공공갈등에 대한 누적적 분석을 통해서 공공갈등의 핵심원인을 분석하고 체계적이고 지속적으로 관리한다면 향후 공공갈등의 예방과 해결의 중요한 자산이 될 수 있다(이주형 외, 2014).

한편, 공공갈등 유형에 관한 선행연구가 많이 축적되어 있으나, 공공갈등 전체를 아우르는 종합적인 개념 정의를 이끌어내기 어렵기 때문에 공공갈등의 유형 또한 다양한 기준으로 분류되고 있다⁴⁾. 따라서 토픽모델링을 활용하는 목적은 첫째, 국내에서 발생한 공공갈등의 전 범위사례(본 연구에서는 신문기사의 텍스트를 분석함)를 대상으로 대한민국의 공공갈등은 어떤 유형을 나타내고 있는지 보고자 하였다는 것이다. 그리고 둘째, 선행연구에서 많이 이루어진 공공갈등의 유형분류가 각각 다양한 기준으로 이루어졌다면, 본 연구에서는 종합적인 관점에서 공공갈등의 유형을 분류하고자 하였다. 이 두 가지 목적을 달성하기 위하여 계량적인 방법으로 유형을 분류함으로써 신뢰성과 객관성을 모두 확보할 수 있는 방법인 ‘토픽모델링’을 활용하였다.

2) 데이터 전처리과정

토픽모델링은 단순한 키워드 분석의 한계를 극복하고, 잠재된 토픽 계층을 설정하여 통계적으로 대량의 자료를 효율적으로 종합할 수 있는 효율적인 분석방법으로서 국내 공공갈등분야에서 적용한 사례가 거의 없다.

데이터 준비를 위한 텍스트 수집 단계에서는 빅데이터와 언론사 홈페이지를 활용하여 텍스트를 수집하고 자연어 처리를 수행하였다. 텍스트 전처리 단계에서는 수집된 비정형 텍스트를 교정 작업, 통제 작업 및 제거 작업을 수행하였다(이수상, 2013). 교정 작업에서는 추출한 텍스트에서 약어, 띄어쓰기, 단·복수 변경과 같은 클렌징(cleansing) 작업을 하였다. 통제 작업에서는 형태소를 분석하고 유사 단어를 통일하여 정제하였다(최수정, 2018). 마지막으로 제거 작업에서는 문장의 의미를 해치지 않는 ‘있’, ‘없’ 등의 동사를 제거하여 명사 및 형용사 형태의 단어만을 사용하였다. 정제된 텍스트의 출현빈도 및 TF-IDF 지수, 중심성 지수를 기준으로 공공갈등의 주요 요인 및 환경에 포함될 단어 목록을 구축하고, 상위 빈도를 나타내는 단어들의 속성을 분석하여 분류하였다(안효선, 2017).

4) 선행연구의 분석으로 종합해본 결과 가장 대표적인 분류방식은 갈등의 성격 및 원인 또는 이슈를 기준으로 분류하는 방식이다(Deustch, 1973; Kriesberg, 2003; Moor, 2003; 서울시정개발연구원, 2003; 지속가능발전위원회, 2005; 하혜영, 2007; 국무총리실, 2010; 정용덕, 2010; 임동진, 2011).

3) 토픽모델링 분석방법

(1) 할당값 부여

토픽모델링은 각 단어들이 통계적으로 특정 토픽에 포함될 확률을 파악하여 문서의 주제를 추정하는 기법으로, 텍스트 분석에서 널리 활용되는 방법이다. 토픽모델링 기법 중 하나인 LDA는 단순하다는 특징과 함께 데이터의 차원을 축소하는 데 유용하다. 또한 의미적으로 일관성이 있는 주제들을 추출하는 데 장점을 가지기 때문에 텍스트 분석에 있어서 각광받고 있다(Mimno & McCallum, 2012).

LDA(Latent Dirichlet Allocation) 알고리즘은 대량의 문서를 분석하는 방법으로써 가장 유명한 모델 중의 하나이다. 이 알고리즘의 기본 아이디어는 문서들을 각 토픽의 분포로 표현하는 것이다. 이때 토픽은 각 키워드 분포 특징으로 묶이게 된다. $p(z|d_i)$, $p(w|z_i, j)$ 를 각각 문서 i 에 대한 토픽 확률 밀도 함수와 문서 i 의 j 번째 단어에 할당된 토픽에 대한 단어 확률 밀도 함수라고 표기한다. 이러한 분포가 주어졌을 때 LDA는 다음 생성 프로세스를 통하여 새로운 문서를 생성한다.

LDA는 문서 생성 프로세스를 모델링하도록 설계되지만, 이것은 메인 노드와 하위 노드 집합으로 이루어진 2-모드 네트워크를 생성 프로세스 모델링에 사용할 수 있다. LDA는 또한 행렬 인수분해(Matrix Factorization)로 해석되는 경우가 있다(Steyvers & Griffiths, 2007). LDA 분석은 단어-문서 데이터에서 토픽이라는 새로운 레이어를 추출하는 분석이기 때문에, 연구자가 직접 지정해야하는 토픽의 수, 알파(alpha), 베타(beta) 값, 반복 횟수(iteration)의 하이퍼파라미터(hyperparameter) 값들이 있다.

먼저 알파 값은 문서별 토픽의 할당(Distribution)에 직접적으로 영향을 주는 값으로, 알파 값이 크면 한 문서에 여러 토픽이 비슷한 확률로 할당된다. 즉, 한 문서를 한 개의 토픽에 치우치지 않게 하여 한 문서에 여러 토픽이 골고루 분포하게 하는 것이다. 베타 값은 토픽별 단어의 할당에 직접적으로 영향을 주는 값으로, 예를 들어 1번 토픽에 모든 단어를 할당시키는데 만약 베타 값이 크면 모든 단어가 비슷한 확률로 1번 토픽에 할당되게 된다. 즉, 모든 단어가 비슷한 확률로 1번 토픽에 할당될 수 있게 되어 모든 단어가 1번 토픽과 연관성이 있다는 것처럼 분석되는 오류가 발생할 수 있다.

따라서 알파와 베타 값을 크게 하면 모든 문서, 모든 단어가 모든 토픽에 비슷한 확률로 할당되기 때문에, 작은 값을 쓸수록 소수의 토픽이 문서에 할당되고, 소수의 단어가 토픽에 할당된다. 본 연구에서는 Griffiths & Steyvers(2004)의 알파 값은 50/토픽수, 베타 값은 0.1을 사용하였고, 이후 베타 값이 0.01이 최적값이라고 언급한 연구를 토대로 알파, 베타 값을

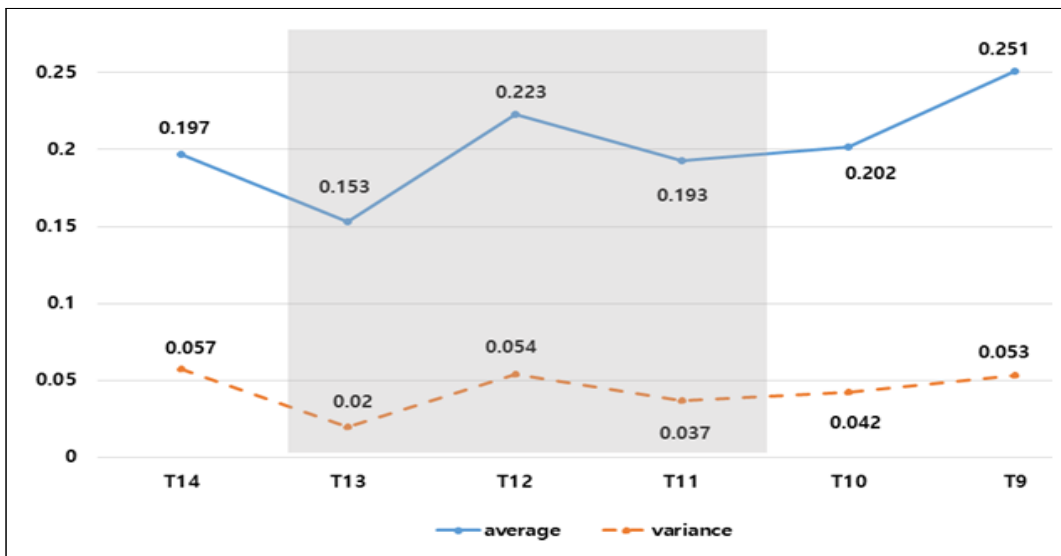
설정하여 진행하였다. 다음으로 반복횟수(iteration)는 가장 적합한 토픽수를 도출하기 위한 반복횟수로 1000번을 적용하였다(Kyrölä, A).

(2) 최적의 토픽수 선정

마지막으로 최적 토픽수는 선정에 있어서는 혼잡도(Perplexity)를 통해 선정하는 방법이 있다. 혼잡도는 특정 확률모델이 실제로 관측되는 값을 얼마나 잘 예측하는지 평가할 때 사용하는 방법으로 낮을수록 모델의 성능이 좋다는 것을 의미한다. 하지만 텍스트 분석에서 해당 방법을 적용하는 것이 정확한 결과도출이 어렵다는 견해(넷마이너, 2019)와 이수상(2016)에서 토픽수를 결정하는 방법이 분석자가 보기에 범주화가 잘 되었다고 판단되는 것과 토픽 확률분포의 퍼플렉서티(perplexity)값을 사용하는 방법, 주제의 평균 코사인 거리가 최소에 도달 할 때 LDA 모델이 가장 잘 수행된다(Cao et al., 2009)는 선행연구를 종합하여 본 연구에서는 토픽 간 유사도 값과 범주화를 통해 토픽 수를 선정하였다.

최적의 토픽 수를 구하기 위하여 토픽 간 토픽의 코사인(cosine) 유사도 값의 평균 및 분산이 가장 낮은 구간을 도출하였다. 이를 위하여 토픽의 수를 21개에서 한 개씩 줄여가며 토픽 간 유사도를 분석하였으며, 그 결과 아래 <그림 1>과 같이 토픽의 수가 11개에서 13개 구간에서 토픽 간 유사도가 가장 낮게 나타났다. 그리고 토픽의 주제를 선정하는 데 결정적인 역할을 하는 핵심 키워드의 중복이 나타나지 않는 토픽의 수인 11개로 선정하였다.

<그림 1> 토픽 간 코사인 유사도



토픽 간의 유사도가 낮게 나타났다는 것은 일차적으로 토픽 간에 의미적으로 주제가 잘 구분되었다는 것을 나타낸다. 따라서 이차적으로 기계적으로 분석된 적정 구간의 범주화가 적절하게 이루어졌는지를 연구자가 판단하게 된다(이수상, 2016). 이 때, 토픽의 주제를 선정하는 데 결정적인 역할을 하는 핵심 키워드의 중복이 나타나지 않는 토픽의 수를 선정한다. 예를 들어, 공공갈등에 있어서 주요 이해관계자인 ‘주민’, ‘지자체’와 같은 경우는 대부분의 토픽에서 핵심 키워드로 등장하고, 토픽의 주제를 정할 때에도 해당 키워드가 노출되지 않는다. 따라서 토픽의 주제를 정함에 있어서는 이러한 키워드들의 중복여부에 의미를 두기보다는 ‘원자력’, ‘전력시설물’, ‘군사시설’, ‘환경’, ‘교통’ 등과 같이 주제의 범주화에 영향을 미치는 키워드의 중복여부에 의미를 두며 판단하게 된다.

IV. 분석결과

1. 자료수집 결과

국내 공공갈등 현황을 분석하기 위하여 국내 언론보도자료 DB를 대상으로 약 12년간의 신문기사의 세부정보인 제목과 본문을 수집하여 분석하였다. 신문기사 수집은 빅카인즈를 활용하였고, 분석은 텍스트 마이닝 패키지인 Netminer 4.0 프로그램을 사용하였다. 분석대상이 되는 공공갈등 관련 신문기사의 대부분이 ‘공공갈등’이라는 키워드를 사용하지 않기 때문에, 검색 키워드를 ‘갈등’으로 입력하였고, 그 결과로 얻은 기사에서 분석대상 범위에 맞게 연구자가 읽고 본 연구와 부합되는 국내에서 발생한 갈등사례 중에서 정부의 정책이나 사업을 수행하는 과정에서 정부와 정부, 정부와 민간 사이에서 발생하는 공공갈등 관련 사례를 추출하였다.

검색 언론사는 국내 언론사 중에서 중앙지 및 주요 경제지로 선정하였으며, 신문기사 수집 사이트인 빅카인즈에서 검색이 가능한 언론사를 중심으로 하였다. 중앙지로는 경향신문, 국민일보, 내일신문, 동아일보, 문화일보, 서울신문, 세계일보, 조선일보, 중앙일보, 한겨레, 한국일보로 총 11개, 경제지로는 매일경제, 머니투데이, 서울경제, 아시아경제, 아주경제, 파이낸셜뉴스, 한국경제, 헤럴드경제 총 8개의 언론사의 기사를 검색하였다. 이 때 ‘조선일보’, ‘중앙일보’, ‘동아일보’는 최근에 검색기능에 추가되었으므로 검색되지 않는 그 이전의 기사는 각 언론사 홈페이지) 및 네이버 신문기사(news.naver.com) Tool을 활용하여 기사내용을 추출하였다.

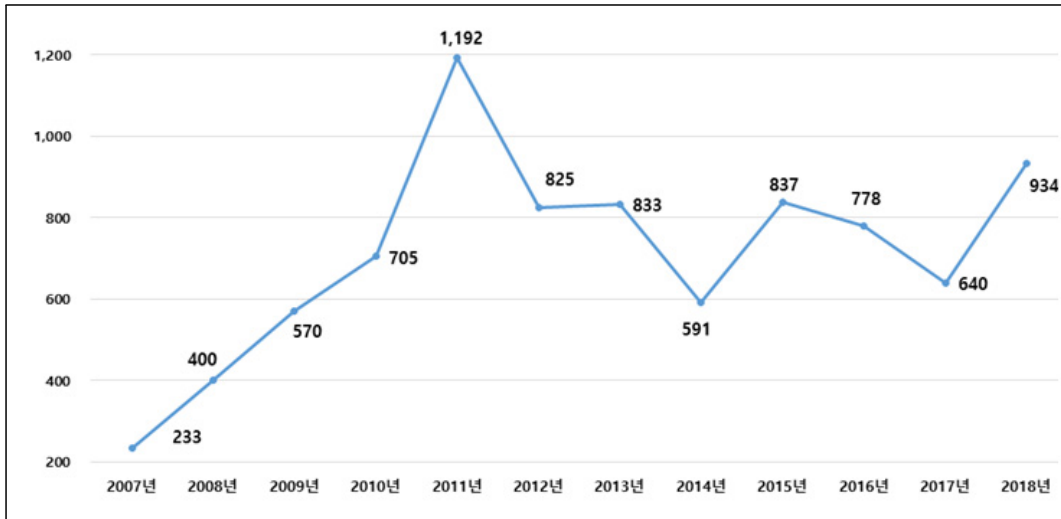
〈표 3〉 공공갈등 현황 분석대상 자료수집 결과

기사 수	검색기간	검색키워드(제목, 본문)	검색수준
8,538편	2007.05.13.~2018.12.31.	갈등	국내 언론사, 중앙지 및 주요 경제지

2007년 5월 13일부터 2018년 12월 31일까지의 공공갈등 분야 기사 중 기사를 검토한 결과, 중복되거나 관련이 없는 기사를 제외하여 최종 분석대상은 8,538편이다.

〈그림 2〉 공공갈등 관련 기사 데이터 현황

(단위: 건, 총 8,538 건)



위의 그래프에서 확인할 수 있듯이 공공갈등 관련 언론보도 노출빈도는 전체적으로는 증가하는 추세를 보이고 있다. 그 중에서 2011년에 급격하게 증가하는 것을 볼 수 있는데, 2011년은 제주해군기지 건설갈등과 동남권 신공항 유치 갈등으로 인한 신문기사가 급격하게 증가한 시기이다. 반면, 증가추세에 있는 공공갈등 언론보도 노출빈도가 2014년, 2017년에는 떨어지는 것을 볼 수 있다. 2014년은 세월호 참사, 2017년은 박근혜 대통령 탄핵 등의 이슈로 인해 상대적으로 공공갈등에 대한 보도가 적게 나타난 것을 알 수 있다⁵⁾.

5) 조선일보 www.chosun.com 중앙일보 joongang.joins.com 동아일보 www.donga.com

6) 세월호 참사와 대통령 탄핵은 본 연구에서 다루고자 하는 공공갈등의 범위에 벗어나므로 제외하였다.

2. 토픽모델링 분석결과

1) 국내 공공갈등 유형 분류

본 연구는 토픽모델링 분석을 통하여 공공갈등 관련 신문기사에 내포된 주제들과 의미의 맥락구조 및 분포도 등을 파악하여, 공공갈등의 유형에 따른 특성을 분석하고자 하였다. 이는 궁극적으로 공공갈등유형에 적합한 갈등해결기제를 탐색하는데 의의가 있을 것으로 기대한다.

국내 공공갈등 관련 신문기사 총 8,538건의 제목과 본문을 대상으로 토픽모델링을 활용하여 11개의 주요 토픽과 각 토픽별 핵심 키워드를 도출하였다. 또한 각 토픽의 키워드와 해당 토픽에 속하는 신문기사를 바탕으로 특정 토픽이 나타내는 갈등의 유형을 파악할 수 있었다. 이때 갈등의 토픽당 추출된 키워드의 확률이 나타내는 의미는, 하나의 신문기사 당 키워드들의 동시출현⁷⁾ 확률로써 특정 토픽으로 서로 연결되어 있는 정도를 나타낸다. 따라서 확률 값이 클수록 해당 토픽에서 주요 키워드이며, 다른 키워드들과의 연관성이 높음을 의미한다(정원준, 2018; 양연희 외, 2019).

한편, 각 토픽의 주제를 명명하기 위하여 토픽의 주요 키워드, 즉 해당 토픽으로 연결확률이 상위에 링크되는 키워드를 중심으로 토픽 주제를 추론하게 된다. 상위 키워드들을 중심으로 주제의 의미를 파악해 보면, ‘토픽 1’의 핵심 키워드는 사업, 개발, 지자체, 추진, 계획, 구역, 환경, 생태, 케이블카, 골프장 등이 등장하고 있다. 해당 토픽의 비중이 높은 기사의 원문을 탐색해 본 결과, 개발로 인한 환경갈등을 주제로 하고 있음을 알 수 있었다. 따라서 ‘토픽 1’은 ‘개발·환경 갈등’으로 명명할 수 있으며, 전체 언론보도 중에서 11.97%(1,022건)을 차지하여 국내에서 발생한 공공갈등 중에서 해당 갈등이 가장 비중이 크다는 것을 확인할 수 있었다. ‘토픽 2’의 주요 키워드로는 지자체, 공항, 이전, 정부, 유치, 입지, 선정, 공약 등이 나타났다. 해당 토픽의 비중이 높은 기사의 원문을 살펴본 결과 선호시설 유치로 인한 경쟁갈등임을 확인할 수 있었다. 따라서 ‘토픽 2’는 ‘선호시설 유치갈등’으로 나타낼 수 있다. 총 900건의 기사가 해당 갈등에 해당하여 전체 분석대상의 10.54%를 차지하여 두 번째로 많은 비중을 차지하였다.

‘토픽 3’의 상위권으로 나타난 지자체, 도로, 노선, 공사, 교통, 구간, 철도, 버스 등의 키워드와 해당 토픽의 비중이 높은 기사의 원문을 종합하며 살펴본 결과 교통체계 및 인프라와 관련되었다. 따라서 ‘토픽 3’은 ‘교통인프라 갈등’으로 명명할 수 있으며 총 889건의 기사로 전체 데이터의 10.41%를 차지하여 세 번째로 많은 비중을 차지하였다. ‘토픽 4’의 상위권으로

7) 본 연구에서 동시출현의 단위를 paragraph을 기준으로 하였는데, 이는 문서(신문기사) 단위로 분석하는 것보다 신문기사가 가지는 맥락구조를 세부적으로 파악할 수 있다.

시설, 주민, 부지, 건립, 처리, 아파트, 공장, 고충민원, 쓰레기 등의 키워드가 나타났다. 관련 기사의 원문을 살펴본 결과 쓰레기 처리장, 장례추모시설 등의 혐오 및 기피시설과 관련된 주제임을 알 수 있어 '비선호시설 입지갈등'으로 명명할 수 있다. 총 810건의 기사가 해당하였으며 전체 데이터의 9.49%의 비중을 차지하였다.

그리고 '토픽 5'는 주민, 제주해군기지, 제주강정마을, 지자체, 공사, 반대, 평화, 철거 등의 키워드가 상위로 나타났다. 해당 갈등으로 분류된 관련 기사의 원문을 통해 제주해군기지와 같은 군사시설과 관련된 갈등임을 알 수 있었는데, 특히 '제주해군기지 건설갈등'이 장기화되고 갈등이 첨예하였던 대표적인 사례라는 점에서 핵심 키워드에 특정 사례에 해당하는 키워드가 두드러졌다. 따라서 '토픽 5'는 '군사시설 갈등'으로 정하였다. 전체 데이터에서 총 797건의 기사(9.33%)에 해당하였다.

한편 '토픽 6'은 지자체, 예산, 교육청, 학교, 무상급식, 학생, 교육감, 복지, 재정, 누리과정⁸⁾ 등이 주요 키워드로 등장하였으며, 관련 기사의 원문을 살펴본 결과 무상급식 및 무상보육과 같은 무상복지와 관련된 내용임을 알 수 있었다. 따라서 '토픽 6'은 '무상복지 갈등'으로 명명하였으며, 해당 갈등은 총 791건의 기사로 전체 데이터의 9.26%의 비중을 차지하였다. '토픽 7'은 핵심 키워드로 지자체, 문화, 도시, 산업, 경제, 공동, 발전, 유치, 기업, 개최, 상생 등이 나타났으며, 관련 기사의 내용을 살펴본 결과 국제 행사 및 지역 행사 등의 사업개최로 인한 지역발전을 꾀하는 사업유치 갈등임을 알 수 있었다. 따라서 '토픽 7'은 '지역발전사업 유치갈등'으로 명명할 수 있으며, 전체 기사의 9.21%(총 786건)에 해당한다.

'토픽 8'의 중요 키워드는 주민, 건설, 마을, 발전, 전력시설물, 발전소, 한국전력공사, 보상 등의 키워드가 등장하였으며, 해당 토픽의 비중이 높은 신문기사의 원문을 살펴본 결과 전력시설물 설치로 인한 갈등을 나타내고 있었다. 따라서 '토픽 8'은 '전력시설물 갈등'으로 설정하였다. 해당 갈등은 총 728건의 기사로 전체 데이터의 8.53%로 나타났다. '토픽 9'는 통합, 의회, 지자체, 행정, 조례, 공무원, 지자체장 등의 키워드가 상위로 나타났다. 비중이 높은 관련 기사의 원문을 살펴본 결과 대표적으로 시, 군의 통합과 같은 행정구역 변경과 관련된 갈등임을 알 수 있었다. 따라서 '토픽 9'는 '행정구역 변경갈등'으로 정하였다. 총 655건의 기사로 전체 데이터의 7.67%에 해당하였다.

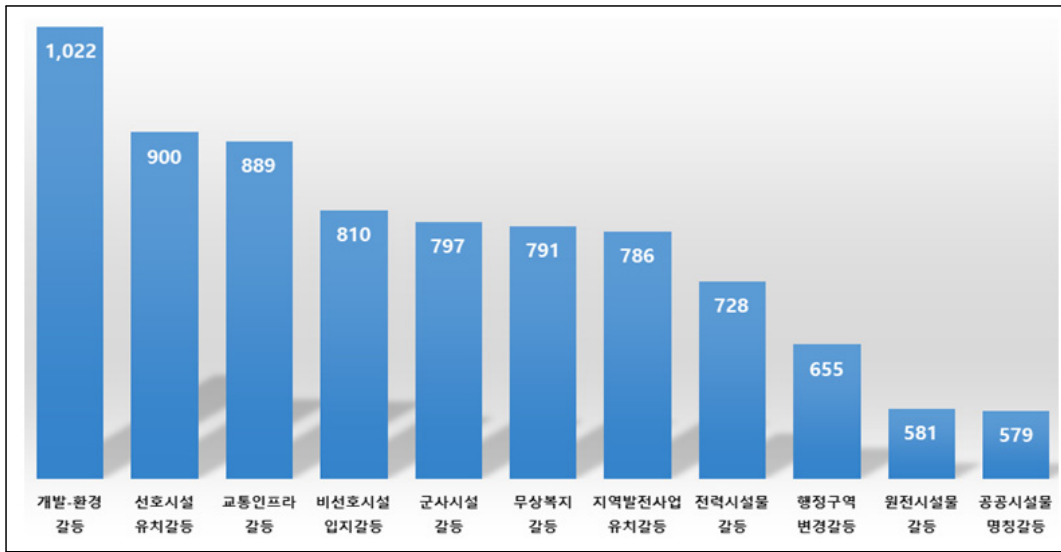
'토픽 10'은 주민, 원자력발전소, 참여, 반대, 과정, 갈등해결, 주민소환제도, 여론 등의 키워드가 상위로 나타났으며, 해당 토픽의 비중이 높은 기사의 원문을 함께 살펴보면 원자력발전소 및 방폐장과 같은 원자력 관련 시설로 인한 갈등이었다. 따라서 '토픽 10'은 '원전시설물 갈등'으로 명명하였다. 총 581건의 기사로 전체 데이터의 6.80%의 비중을 나타냈다. 마지

8) 만 3~5세 유아에게 공통적으로 제공하는 교육·보육 과정. 2012년 3월 5세를 대상으로 시행에 들어가 2013년에 3~4세까지 확대되었다.

막으로 ‘토픽 11’은 지자체, 주민, 명칭, 변경, 청사, 이전 등의 키워드가 상위로 등장하였으며, 해당 비중이 높은 기사의 원문을 살펴본 결과, 고속철도역사, 교량 등 공공시설물 명칭과 관련한 갈등임을 알 수 있었다. 따라서 ‘토픽 11’은 ‘공공시설물 명칭갈등’으로 정하였다. 해당 갈등은 총 579건의 기사로 전체 데이터의 6.78%로 가장 낮은 비중을 차지하였다.

〈그림 3〉 토픽별 기사 수

(단위: 건)



위의 〈그림 3〉은 토픽별 기사의 수를 도식화한 것이고, 〈표 4〉는 각 토픽의 명명과 주요상위 키워드, 그리고 전체 분석대상(신문기사) 대비 비중을 나타낸 것이다.

〈표 4〉 국내 공공갈등의 유형분류

Topic	Keyword	기사수 (비중 ⁹⁾)
개발·환경 갈등	사업, 개발, 광역지자체, 추진, 계획, 구역, 환경, 지정, 지구, 지역, 공원, 조성, 정비, 관리, 생태, 설치, 공급, 케이블카, 방식, 골프장	1,022 (11.97%)
선호시설 유치갈등	광역지자체, 공항, 이전, 정부, 지역, 유치, 도지사, 입지, 건설, 선정, 추진, 시장, 발표, 후보, 용역, 후보지, 영남, 주장, 공약, 국가	900 (10.54%)
교통인프라 갈등	광역지자체, 도시, 도로, 노선, 계획, 공사, 한국토지주택공사, 교통, 구간, 사업, 혁신, 철도, 관계자, 요구, 운행, 국토교통부, 방식, 버스, 신설, 공공기관	889 (10.41%)
비선호시설 입지갈등	시설, 주민, 부지, 건립, 계획, 관리, 사용, 조성, 처리, 아파트, 업체, 공원, 규모, 공장, 고층민원, 쓰레기, 발생, 운영, 설치, 건물	810 (9.49%)

Topic	Keyword	기사수 (비중 ⁹⁾)
군사시설 갈등	주민, 제주해군기지, 제주강정마을, 광역지자체, 공사, 반대, 건설, 해군, 정부, 경찰, 평화, 갈등해결, 기지, 소송, 마을, 병원, 중단, 청구, 현장, 철거	797 (9.33%)
무상복지 갈등	광역지자체, 예산, 지원, 교육청, 학교, 교육, 무상급식, 부담, 감사, 학생, 교육감, 편성, 복지, 의회, 재정, 설립, 계획, 누리과정, 도지사, 학부모	791 (9.26%)
지역발전사업 유치갈등	광역지자체, 문화, 도시, 지역, 사업, 산업, 시장, 경제, 공동, 운영, 협력, 센터, 발전, 유치, 기업, 개최, 세계, 투자, 상생, 조성	786 (9.21%)
전력시설물 갈등	주민, 건설, 마을, 발전, 지역, 전력시설물(송전), 발전소, 한국전력공사, 공사, 기초지자체, 반대, 보상, 사업, 에너지, 피해, 태양광, 설치, 대책위원회, 계획, 환경	728 (8.53%)
행정구역 변경갈등	통합, 의회, 의원, 광역지자체, 행정, 추진, 조례, 지방, 자치, 위원회, 단체, 지역, 공무원, 회의, 논의, 도지사, 제출, 시장, 군수	655 (7.67%)
원전시설물 갈등	주민, 조사, 원자력발전소, 의견, 참여, 구성, 반대, 결정, 위원회, 정책, 과정, 방식, 갈등해결, 추진, 주민소환제도, 투표, 여론, 찬반, 시장, 실시	581 (6.80%)
공공시설물 명칭갈등	기초지자체, 지역, 주민, 명칭, 지방자치단체, 변경, 주장, 반대, 입장, 결정, 청사, 추진, 시장, 수원, 시청, 단체, 위원회, 이전, 요구, 위치	579 (6.78%)

* 위의 토픽 번호는 비중 순으로 부여하였으며 중요도와는 무관하다.

** 표에 제시된 비중은 전체 기사 수(100%)에 대한 각 토픽의 기사 수의 비율을 나타낸다.

*** 키워드는 비중이 높은 순으로 상위 20개를 정렬하였다.

분석결과 중에서 ‘전력시설물 갈등’, ‘원전시설물 갈등’, ‘군사시설 갈등’, ‘비선호시설 입지 갈등’에 해당하는 4개의 갈등은 유형의 범위를 넓게 보면 비선호시설로 인식되는 갈등이다. 이는 전체 데이터의 34.15%를 차지하는 것으로 국내에서 발생하는 공공갈등 중 비선호시설 입지갈등에 따른 신문기사가 가장 많다는 것을 알 수 있다. 이와 같은 결과는 공공갈등 연구 경향 분석에서 최근 20년간의 국내에서 이루어진 연구가 전체 756편의 논문 중에 56.61%인 428편이 비선호시설 관련 연구¹⁰⁾라는 점이 공공갈등에서 비선호시설 관련된 갈등이 가장 많은 비중을 차지하기 때문에 그와 관련한 연구가 가장 많이 이루어졌다(양연희 외, 2019)는 것을 뒷받침해 주는 결과라고 볼 수 있다.

한편 전력시설물과 원전시설물, 군사시설과 관련된 갈등이 선행연구에서는 비선호시설로 분류되었으나 본 연구에서는 따로 구분되어 분류되었다. 이는 각각의 갈등 비중은 전체 공공갈등 분야에서는 낮은 순위(군사시설 5번째, 전력시설물 8번째, 원전시설물 10번째)를 차지하였으나 비선호시설과 구분되어 분류되었다는 것은 그만큼 국내에서 두드러지게 발생하는 특징이라는 점을 알 수 있다. ‘교통인프라 갈등’은 선호시설 유치갈등과 개발·환경 갈등이 함

9) 전체 기사수 대비 해당 토픽 기사수의 비중

10) 전체 756편의 연구 중, 비선호시설 인식프레임 연구(14.2%), 비선호시설 입지갈등 및 갈등해결 요인 연구(16.5%), 비선호시설 입지관련 주민지원 및 참여를 보장하는 법·제도적 연구(10.7%), 개발갈등 관련 이해관계자 연구(15.2%)가 이루어졌다(양연희 외, 2019).

게 나타나는 특징을 가지고 있는 유형이다¹¹⁾. 갈등주체를 살펴보면 크게는 경제적 이익과 손실에 따른 정부 간 갈등과 좁게는 지역주민과의 갈등이 나타나며, 교통 인프라 건설에 따른 환경가치 갈등도 나타난다.

2) 공공갈등의 유형별 트렌드 변화

한편 신문기사의 발행일을 활용하여 공공갈등 주요 이슈의 트렌드 변화를 분석해보았다. 이 분석을 통해서 시간의 흐름에 따라 증가하게 되는 공공갈등의 이슈를 파악해 볼 수 있고, 역으로 공공갈등이 줄어드는 분야 및 이러한 변화의 특성을 파악해 볼 수 있다.

트렌드 변화는 정권 변화를 기준으로 시기를 구분하였다. 총 분석대상이 되는 언론보도 발행기간을 2007년 5월부터 2008년 2월까지를 Period 0는 1년도 채 되지 않기 때문에 그 이후 시기와 비교하기에는 짧지만 전체 연구의 기준점으로 분석하기 위하여 독립적 시기로 구분하였다. 따라서 아래 <표 5>와 같이 Period 0은 2007년 5월부터 2008년 2월, Period 1은 2008년 3월부터 2013년 2월, Period 2¹²⁾는 2013년 3월부터 2017년 4월¹³⁾, Period 3은 2017년 5월부터 2018년 12월까지로 하여, 총4구간으로 구분하여 분석하였다.

<표 5> 공공갈등 관련 신문기사의 연도별 현황

(단위 : 건)

구분	기간	신문기사 수
Period 0	2007. 05. 13. - 2008. 02. 28	276
Period 1	2008. 03. 01. - 2013. 02. 28	3,788
Period 2	2013. 03. 01 - 2017. 04. 30	3,094
Period 3	2017. 05. 01 - 2018. 12. 31	1,380
총 합계		8,538

Period 1에는 3,788건으로 기사 수가 가장 많았으며, Period 2에는 3,094건으로 다소 감소하였다. 한편 Period 3은 1,380건으로 앞의 두 기간과 비교하여 짧은 기간임에도 불구하고 기간 대비 기사수가 상대적으로 많은 것을 통해 공공갈등의 이슈가 증가하는 추세라는 것

11) 연구자는 해당 갈등유형을 선호-비선호의 개념을 부각하여 두 개의 유형으로 구분하는 것을 시도하였으나 분석결과의 타당도가 떨어지는 것으로 나타났다. 따라서 분석결과의 타당도를 손상시키지 않는 방법으로 복합적인 성격을 가지는 유형으로 정리하였다.

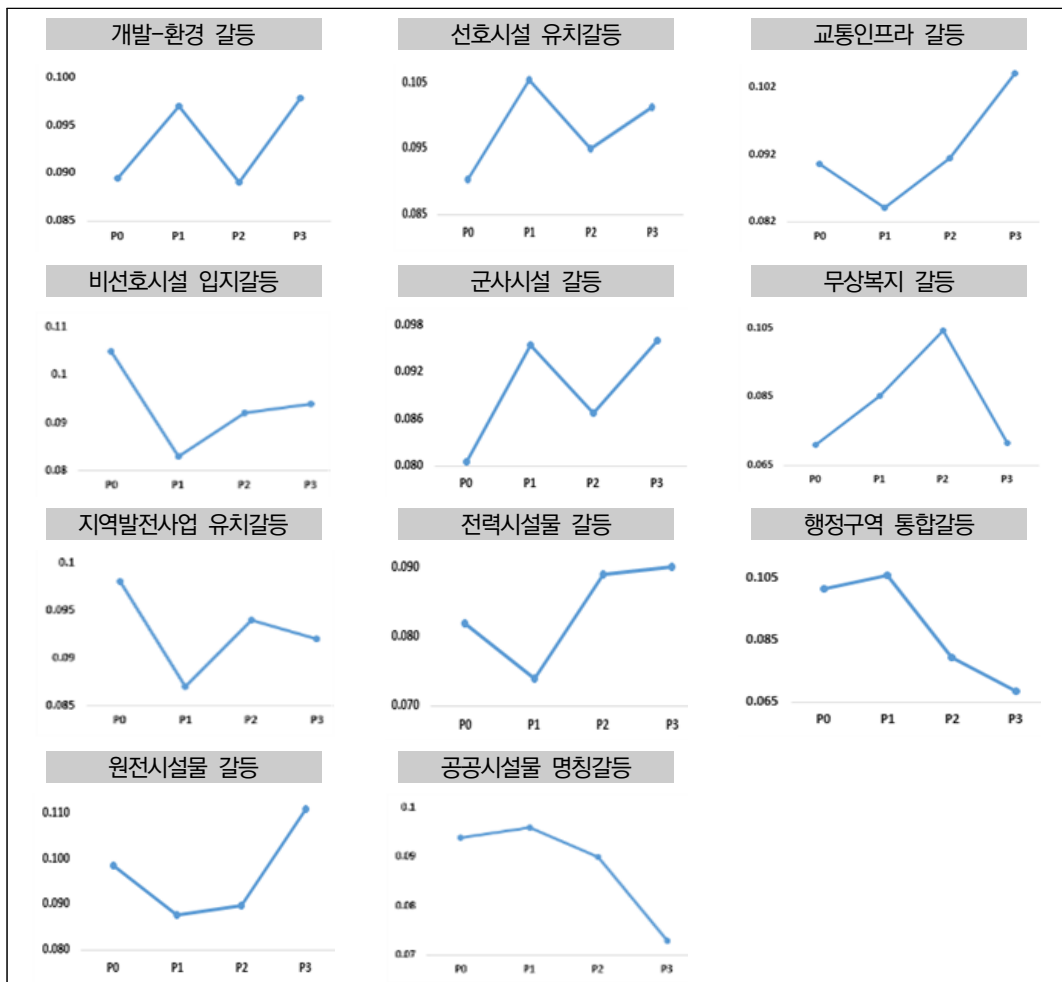
12) 박근혜 대통령이 2017년 2월 탄핵되었고 문재인 정부로 교체되기까지의 약 2개월간은 정권교체시기로 박근혜 정부기간으로 포함하였다.

13) 시기의 구분은 편의상 정권교체시기별로 구분하였으며, 대통령(박근혜) 탄핵이 있는 정권교체시기도 이전 대통령 집권기간으로 설정하였다.

을 확인할 수 있다.

한편 이와 같이 기간구분에 따른 일수의 차이가 분명하므로 기간별 기사수의 빈도보다는 기간에 따른 토픽의 비중변화를 중심으로 분석하였다¹⁴⁾. 아래 <그림 4>는 시간의 흐름에 따라 토픽의 비중변화를 구체적으로 살펴보기 위하여, 각 문서에 대한 각각의 토픽의 구성 비중을 추출하여 그 평균값을 기간별로 비교분석하였다.

<그림 4> 전체 공공갈등 언론보도 기간별 토픽비중



14) 분석대상인 공공갈등 언론보도는 노무현 정권은 2007년 5월부터 수집되었고 문재인 정권은 2017년 5월부터 수집되어 다른 정권과 비교하여 시간적 범위에서 차이가 나므로 발행수가 상대적으로 적다. 하지만 본 연구의 시기 구분을 정권에 따른 것은 편의적 구분이지 정권별 성격에 따른 차이를 보고자 하는 것이 아니기 때문에 분석에서는 발행수가 아닌 토픽의 비중변화를 통해 분석결과를 도출하였다.

기존의 사회는 대규모 국책사업의 경우 국민이 정책대상자로서 소극적인 주체로 정책수립 및 집행을 수용하는 입장이었다면, 한국 사회가 민주화, 다원화, 탈권위주의화 되었으며 시민 사회 영역이 빠르게 성장하면서 시민의 권리의식과 정책 참여에 대한 욕구, 민주적 절차에 대한 요구가 크게 향상되었다(김강민 외, 2018). 이는 분석결과를 통해서도 확인할 수 있는데 전반적인 갈등의 시간에 따른 추이를 살펴보면 국민이 정책주체자로 정책수립 및 집행에 있어 적극적으로 목소리를 냄으로써 전반적인 공공갈등이 증가하는 것을 볼 수 있다. 특히 비선호시설의 입지갈등이 두드러지는 것을 확인할 수 있다.

유형별로 살펴보면, 군사시설 갈등은 2008년~2013년 초에는 갈등이 급격하게 증가하는 것을 볼 수 있다. 군사시설의 경우 일반적으로 거주민이 많지 않은 지역에 설치하는 경우가 대부분이나, 도시화로 인해 지역개발이 활발하게 이루어지면서 지역주민의 거주범위도 넓어지게 되면서 갈등이 증가하는 배경으로 설명할 수 있다. 또한 2007년 제주해군기지건설 찬반갈등이 첨예화되면서 관련 갈등이슈들이 증가하게 되었다. 2013년~2017년에는 감소하는 추세였으나 최근까지 완료되지 않은 갈등으로 지속적으로 높은 비중을 차지하고 있다. 개발·환경갈등 또한 2000년 초반까지 대규모 국책사업의 경우 국민이 정책대상자로서 소극적인 주체로 정책수립 및 집행을 수용하는 입장이었다면, 최근으로 올수록 정책주체자로 정책수립 및 집행에 있어서 적극적으로 목소리를 냄으로써 관련 갈등이 두드러지는 것을 알 수 있다.

한편 무상복지 갈등은 이명박 정권 때 경기도를 중심으로 서울시 「서울시 친환경 무상급식 등 지원에 대한 조례」 제정에 이르기까지 논란이 점화되었다. 이후 2013~2017년에 와서는 심각한 예산부족으로 정책 실행에 곤란을 겪으면서 큰 논란과 이슈를 가져오면서 갈등이 확대되는 경향을 보이면서 갈등이슈가 증폭되었다.

그리고 전력시설물 갈등은 현재까지도 높은 비중으로 유지되는 이슈로 전력공급을 둘러싼 송전탑, 변전소 및 송전선로 설치를 둘러싼 갈등은 매년 꾸준히 발생하고 있으며 정부 통계에 의하면 연도별로 증가하는 추세이다. 특히 2012년 밀양 송전선로 갈등 과정에서 주민의 분신 시도로 갈등이 전국화 되고 증폭 및 심화되면서 급격하게 증가하였다. 이처럼 송·변전 설비건설 관련 갈등이 빈번하게 발생하는 이유는 국가차원의 안정적인 전력공급과 송·변전 설비가 설치되는 지역주민의 생활권(재산권 및 건강권 등)이 상호 충족되지 못하고, 관련 법·제도의 미비점 등 복합적 요인으로 인해 이해관계가 충돌하기 때문이다(은재호, 2011). 밀양송전탑 갈등 발생 이후 송전선로 건설에 따른 갈등은 공공갈등의 대표적인 유형으로 인식되고 있다. 해당 갈등은 송전선로 경과지를 따라 여러 지역에 걸쳐 연쇄적으로 발생한다는 특징과 함께 원전지역과 인접할 경우에는 반핵운동과도 접목된다는 등의 독특한 양상을 보인다(심준섭, 2014). 반면 원전시설물 갈등은 상대적으로 2017년~2018년에서 큰 폭으로 증가한 것을 알 수 있다. 이는 신고리 5·6호기 공론화 과정을 거치면서 국민의 관심이 집중되었고, 이러한 경

향을 언론에서도 그대로 반영한 것도 큰 영향을 미쳤다.

그 외 비선호시설 건립갈등은 2007년~2008년 초와 비교하여 2008년~2013년 초에 급격하게 감소하였으나, 그 이후 조금씩 증가하는 것을 볼 수 있다. 일반적으로 국민의 소득수준과 의식수준이 높아질수록 국가에 대한 사회서비스 요구는 증가하게 되고, 이와 함께 비선호시설이 증가는 불가피하게 된다. 따라서 해당 갈등은 지속적으로 증가할 것으로 보인다¹⁵⁾.

지역발전사업 유치갈등은 2013~2017년 상반기 동안에 국제행사를 개최하면서 갈등이 다른 시기보다 상대적으로 증가하였으며, 교통인프라 갈등과 선호시설 유치갈등은 2008년~2013년 초부터 지속적으로 증가하는 것을 볼 수 있다.

행정구역 변경갈등 및 공공시설물 명칭갈등은 지방자치제가 시작되고 지역개발사업과 지방행정구역의 통합 등의 행정체제 변화로 볼 수 있다. 이것은 2000년대 초반까지에 해당하는 2008년~2013년 초, 2013~2017년 상반기에 빈번하게 이루어진 사업으로 인한 갈등으로, 이후 2013~2017년부터는 현재에 와서는 갈등의 다양성으로 인해 상대적으로 큰 폭으로 감소하는 추세라고 해석할 수 있다.

한편 선호시설 및 교통인프라는 지역 내 혹은 지역 간 긍정의 외부효과(positive externalities¹⁶⁾)를 발생시켜 지역 이미지 제고, 지역 및 개인 자산가치 증가, 관광효과, 고용창출 등 지역경제 활성화에 이바지하는 것으로 여겨짐으로써 유치경쟁이 치열하다(김강민 외, 2018). 산업화, 도시화가 진행되고 소득수준이 증가하면서 산업시설 및 생활기반시설과 같은 선호시설 확대의 필요성 및 요구 또한 커지게 되고, 그 과정에서 발생하는 공공갈등의 발생빈도 또한 높아짐을 보여준다.

V. 결론

본 연구는 선행연구 및 소수 사례분석을 통한 공공갈등 유형을 정리한 연구가 대부분이라는 한계에서 시작하여, 실제 우리나라에서 발생하는 공공갈등의 유형과 경향을 파악하였다.

15) 비선호시설의 필요성은 인지하나 내 지역에는 설치하면 안 된다는 의식을 변화시키고, 절차적 민주성을 견지한 갈등관리시스템을 정착시킴으로써 갈등에 대비하여야 한다.

16) 외부효과란 경제행위자가 어떤 특정한 경제적 행위를 할 때, 의도하지 않게 제3자에게 이익 또는 손해를 가져다주며 이에 대한 대가나 벌칙을 받지 않는 경우를 말한다. 사회 전체적으로 보았을 때 긍정적인 측면을 '외부경제', 부정적인 측면을 '외부불경제'라고 한다(Holtermann, 1972; 김강민 외, 2018 재인용)

따라서 이를 위하여 2007년 5월부터 2018년 12월까지 공공갈등 관련 신문기사를 활용하여 국내에서 발생한 공공갈등의 유형과 경향의 특징을 분석하였다.

분석결과, 국내에서 발생하는 공공갈등은 개발·환경 갈등, 선호시설 유치갈등, 교통인프라 갈등, 비선호시설 입지갈등, 군사시설 갈등, 무상복지 갈등, 지역발전사업 유치갈등, 전력시설물 갈등, 행정구역 변경갈등, 원전시설물 갈등, 공공시설물 명칭갈등으로 총 11개의 유형으로 분류되었다. 이는 선행연구에서 공공갈등 유형을 갈등의 표출여부, 갈등의 성격, 갈등요인, 갈등 이해관계자의 특성, 갈등의 주체 등 다양한 기준을 통하여 분류하였지만, 정책수립 시에 적용하기에는 기준이 중복되거나 명확하게 구분하는 것이 어렵다. 이러한 한계로 공공갈등에 관한 많은 연구가 이루어졌음에도 불구하고 최근 갈등이 증가하고 장기화되고 갈등원인이 다르더라도 갈등전개양상이 유사하게 전개되는 특징을 가진다는 것은 체계적인 갈등관리 시스템 구축이 필요하다는 것을 도출할 수 있다. 따라서 본 연구의 분석결과와 같이 선행연구와는 달리 갈등을 유발하는 정책대상을 기준으로 분류하여, 국내에서 두드러지게 발생한 공공갈등의 특징을 반영한 갈등 유형분류를 통해 향후 정책대상별로 정책 수립 시에 적용 가능한 갈등 관리기제를 도출하고, 갈등관리 정책효과를 파악하는 데에도 의의가 있다.

그리고 분석결과 중에서 전력시설물 갈등, 원전시설물 갈등, 군사시설 갈등, 비선호시설 입지갈등에 해당하는 4개의 갈등은 유형의 범위를 넓게 보면 비선호시설로 인식되는 갈등이다. 이는 전체 데이터의 34.15%를 차지하는 것으로 국내에서 발생하는 공공갈등 중 비선호시설 입지갈등에 따른 신문기사가 가장 많다는 것을 알 수 있었다. 이와 같은 결과는 공공갈등 연구경향 분석에서 최근 20년간의 국내에서 이루어진 연구가 전체 756편의 논문 중에 56.61%인 428편이 비선호시설 관련 연구라는 점이 공공갈등에서 비선호시설 관련된 갈등이 가장 많은 비중을 차지하기 때문에 그와 관련한 연구가 가장 많이 이루어졌다는 것으로 확인할 수 있다(양연희 외, 2019).

또한 전력시설물과 원전시설물, 군사시설과 관련된 갈등이 공공갈등 유형으로 분류될 만큼 두드러지게 발생하는 특징이라는 점을 알 수 있었다. 이것은 전력시설물과 원전시설물, 군사시설과 관련한 갈등관계는 이해관계가 첨예하게 대립하여 표면적으로 갈등이 표출되거나 장기화되는 경향을 가진다는 것을 보여준다. 따라서 관련하여 정책 수립 시에는 유형별 특징에 맞는 갈등관리체계가 구축되어야 할 필요가 있다는 점을 말해준다.

공공갈등의 유형별 트렌드 변화를 통해서도 공공갈등 이슈의 흐름과 특징을 파악해 볼 수 있었다. 무상복지, 행정구역 통합, 공공시설물 명칭 갈등은 감소하는 반면 그 외 대부분 유형의 갈등은 증가하는 추세라는 것을 확인할 수 있었다. 이와 같은 결과는 일반적으로 국민의 소득수준과 의식수준이 높아질수록 국가에 대한 사회서비스 요구는 증가하게 되면서 비선호 시설이 증가하게 됨에 따라 해당 갈등은 지속적으로 증가할 것으로 보인다. 이는 절차적 민주

성을 견지한 갈등관리시스템을 구축함으로써 갈등에 대비해야 함을 보여준다.

본 연구는 주관적인 해결방안을 제시한 지금까지의 선행연구와는 달리, 빅데이터 분석을 통하여 공공갈등의 유형을 객관화하였다는 데에 의의가 있다. 이것은 선행연구를 통해서 제기한 문제인식을 토대로 국내에서 발생한 갈등사례를 수집하여 단순 빈도분석을 통한 공공갈등 현황을 확인하는 데에 그쳤던 한계를 극복하기 위하여, 최근 다양한 학문분야에서 적극적으로 활용되기 시작한 '텍스트 마이닝' 기법을 활용하여 최대한 많은 공공갈등 사례정보를 사용하였으며, '토픽모델링' 분석이라는 갈등에 대한 객관적인 분석을 통하여 유형을 분류하고 경향성을 찾고자 하였다. 본 연구를 통해 국내에서 진행되고 있거나 향후 발생할 수 있는 갈등 해결에 적실성 있는 시사점을 도출할 수 있을 것으로 기대한다.

이는 후속연구에서 국내 공공갈등의 유형에 따른 갈등해결요인과 이에 맞는 갈등해결 방안을 파악하기 위한 자료가 될 수 있다는 데에 의의가 있다.

【참고문헌】

- 국무총리실. (2010). 「'09년도 부처 갈등관리실태 점검·평가 결과」.
- 김강민·김학린·임재형·전형준. (2018). 「갈등관리와 협상」. 단국대학교 분쟁해결연구센터 총서2, nosvos.
- 김민관·이용·한창희. (2017). Analysis of Consulting Research Trends Using Topic Modeling. 「산업경영시스템학회지」, 40: 46-54.
- 김정인. (2018). 「참여형 공공갈등관리의 이해」. 서울: 박영사.
- 강창민. (2015). 공공갈등관리의 국내외 동향과 제주지역 시사점. 「정책이슈브리프」, 237: 1-19.
- 김대중. (2019). 사례분석을 통한 사회갈등의 유형별 해소방안에 관한 탐색적 연구. 「협상연구」, 22(2): 69-102.
- 넷마이너. (2019). 「텍스트 네트워크 분석교육 교재」. 서울: 사이람.
- 대통령자문지속가능발전위원회. (2005). 「공공갈등관리의 이론과 기법(상)」. 서울: 논형.
- 박상언·이병량·정민경. (2019). 텍스트 마이닝 기법을 활용한 한국의 정부 활동 주체에 관한 연구 경향 분석. 「정부학연구」, 25(2): 159-190.
- 서병조·신선영. (2017). 토픽 모델링을 활용한 한국의 플랫폼정부 연구동향 분석. 「정보화정책」, 24(3): 3-26.
- 서울시정개발연구원. (2003). 「청계천 복원사업 타당성 조사 및 기본계획: 사회영향 부문」.
- 손복은·고호경. (2018). 수학 담화에서 나타나는 교사의 감성적 언어 빈도 분석. 「E-수학교육 논문집」, 32(4): 555-573.
- 신동희. (2016). 「컴퓨터이셔널 저널리즘」. 서울: 커뮤니케이션북스.
- 심준섭. (2010). 질적 연구에 대한 소고: 질적 연구에서 '질적'의 의미. 「한국행정학회 학술발표논문집」, 674-683.
- 심준섭. (2014). 「갈등관리 Role Model 확산을 위한 연구」. 한국사회갈등해소센터.
- 안효선. (2017). 「빅데이터를 활용한 패션디자인 감성분석 연구」. 이화여자대학교 박사학위 논문.
- 양연희·권영주·이상철. (2019). 토픽모델링과 네트워크 분석을 활용한 공공갈등 연구경향 분석. 「지방정부연구」, 23(3): 427-450.
- 이민창. (2005). 정책갈등현상의 제도론적 해석-NIMBY 사례를 중심으로. 「한국정책학회보」, 14(1): 71-98.
- 이수상. (2013). 연구자 네트워크의 중심성과 연구성과의 연관성 분석: 국내 기록관리학 분야 학술 논문을 중심으로. 「한국도서관·정보학회지」, 44(3): 405-428.
- 이수상. (2016). 독후감 텍스트의 토픽모델링 적용에 관한 탐색적 연구. 「한국도서관·정보학회지」,

- 47(4): 1-18.
- 이주형·가상준·임재형·김강민·김재신. (2014). 「공공갈등관리 사례분석과 외국의 공공갈등관리제도 조사」. 국회예산정책처, 1-163.
- 이준용·김성희. (2018). 미세먼지 재해 보도의 프레임 분석: 구조적 주제모형(Structural Topic Modeling)의 적용. 「한국언론학보」, 62(4): 125-158.
- 임동진. (2011). 공공갈등관리의 실태 및 갈등해결 요인분석. 「한국행정학보」, 45(2): 291-318.
- 임도빈. (2009). 질적 연구 방법의 내용과 적용전략: 양적인 질적 연구와 질적인 질적 연구. 「정부학연구」, 15(1): 155-187.
- 임동진. (2012). 「대안적 갈등해결방식(ADR)의 운영실태 및 개선방안 연구」. 기본연구과제, 2012: 1-613.
- 임정빈. (2007). 남비와 핼피사례의 비교분석을 통한 지방정부간 갈등관리전략. 「지방정부연구」, 11(3): 155-179.
- 주상현. (2012). 지방행정체제 개편논의의 과정분석과 합의-전주·완주 통합논의 사례를 중심으로. 「한국비교정부학보」, 16(2): 255-286.
- 전주상. (2000). 비선호시설 입지갈등요인에 관한 연구: 노원·목동·강남 쓰레기소각장 건설사례의 비교분석. 「한국사회와 행정연구」, 11(2): 275-295.
- 정용덕. (2010). 공공갈등과 정책조정 연구. 「행정논총」, 48(4): 1-30.
- 정원준. (2018). 사드(THAAD) 이슈를 둘러싼 한국과 중국 간 갈등 쟁점의 변화 추이 연구: 빅데이터를 이용한 시계열 토픽 모델링과 언어 네트워크 분석 기법으로. 「한국광고홍보학보」, 20(3): 143-196.
- 조재희·조인호. (2019). 2018 메르스 해외 재유입에 대한 주요 온라인 이슈 탐색: 토픽모델링 분석과 감성 분석을 중심으로. 「한국디지털콘텐츠학회 논문지」, 20(5): 1051-1060.
- 최수정. (2018). 「빅데이터 감성분석을 통한 프로농구 발전 방안 모색」. 숙명여자대학교 박사학위논문
- 하혜영. (2007). 공공갈등 해결에 미치는 영향요인 분석. 「한국행정학보」, 41(3): 273-296.
- 허고은. (2019). 토픽 모델링 기반 과학적 지식의 불확실성의 흐름에 관한 연구. 「정보관리학회지」, 36(1): 191-213.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4): 77-84.
- Campbell, P. (1996). 'The history of the user movement in the United Kingdom', in Heller, T., Reynolds, J., Gomm, R., Muston, R. and Pattison, S. (eds.), *Mental Health Matters*, Basingstoke, Macmillan.
- Calabrese, C., Ding, J., Millam, B. & Barnett, G. A. (2019). The Uproar Over Gene-Edited Babies: A Semantic Network Analysis of CRISPR on Twitter. *Environmental*

- Communication*, 1-17.
- Cao, J., Xia, T., Li, J., Zhang, Y. & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, 72(7-9): 1775-1781.
- Deustch, M. (1973). *The resolution of conflict: constructive and destructive processes*.
- Griffiths, T. L. & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1): 5228-5235.
- Inzalkar, S., & Sharma, J. (2015). A survey on text mining-techniques and application. *International Journal of Research In Science & Engineerin.*, 24: 1-14.
- Khan, R. A., & Kanth, S. (2016). Text Mining: Knowledge Discovery from Unstructured Data. *Artificial Intelligent Systems and Machine Learning*, 8(2): 71-77.
- Kim, W., Won, J. H., Park, S. & Kang, J. (2015). Demand forecasting models for medicines through wireless sensor networks data and topic trend analysis. *International Journal of Distributed Sensor Networks*, 11(9): 907169.
- Kriesberg, L. (2003). *Factors shaping the course of intractable conflict. Beyond Intractability*. http://www.beyondintractability.org/m/factors_shaping_intractable_conflict.jsp.
- Mimno, D. & McCallum, A. (2012). *Topic models conditioned on arbitrary features with dirichlet-multinomial regression*. arXiv preprint arXiv :1206.3278.
- Moor, E. (2003). Branded spaces: the scope of 'new marketing'. *Journal of Consumer Cultur*, 3(1): 39-60.
- Potha, N. & Stamatatos, E. (2019). Improving author verification based on topic modeling. *Journal of the Association for Information Science and Technology*, 70(10): 1074-1088.
- Shim, J., Park, C. & Wilding, M. (2015). Identifying policy frames through semantic network analysis: an examination of nuclear energy policy across six countries, *Policy Sciences*, 48(1): 51-83.
- Steyvers, M. & Griffiths, T. (2007). Probabilistic topic models. *Handbook of latent semantic analysis*, 427(7): 424-440.
- Wang, X. & McCallum, A. (2006). *Topics over time: a non-Markov continuous-time model of topical trends*. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining(pp. 424-433). ACM.
- Xiong, Y., Cho, M. & Boatwright, B. (2019). Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the# MeToo movement. *Public relations review*, 45(1): 10-23.

한국언론진흥재단 빅카인즈 www.kinds.or.kr

조선일보 www.chosun.com

중앙일보 joongang.joins.com

동아일보 www.donga.com

양 연 희: 부산대학교에서 공공정책학 박사학위를 취득하고(토픽모델링을 활용한 공공갈등 유형화와 해결방안에 관한 연구, 2020), 현재 행정안전부에서 사무관으로 재직하고 있다. 주요 관심 분야는 공공정책, 공공갈등관리, 조직이론 등이다. 최근 논문으로는 “토픽모델링과 네트워크 분석을 활용한 공공갈등 연구경향 분석(2019)”, “공공기관 조직구성원의 회복탄력성이 직무만족과 조직몰입에 미치는 영향에 관한 연구(2018)” 등이 있다(himawari135@naver.com).