

Exploring trends in U.N. Peacekeeping Activities in Korea through Topic Modeling and Social Network Analysis

Donghyeon Jung* · Chansong Kim* · Kangmin Lee* · Soeun Bae** · Yeon Seo*** · Hyeonju Seol****†

*Republic of Korea Navy

**Public Safety Data Science Research Center, Korean National Police University

***Department of Business Administration, Chungnam National University

****School of Integrated National Security, Chungnam National University

토픽모델링과 사회연결망 분석을 통한 우리나라 유엔 평화유지활동 동향 탐색

정동현* · 김찬송* · 이강민* · 배소은** · 서연*** · 설현주****†

*대한민국 해군, **경찰대학 치안데이터과학연구소,

충남대학교 경영학과, *충남대학교 국가안보융합학부

The purpose of this study is to identify the major peacekeeping activities that the Korean armed forces has performed from the past to the present. To do this, we collected 692 press releases from the National Defense Daily over the past 20 years and performed topic modeling and social network analysis. As a result of topic modeling analysis, 112 major keywords and 8 topics were derived, and as a result of examining the Korean armed forces's peacekeeping activities based on the topics, 6 major activities and 2 related matters were identified. The six major activities were 'Northeast Asian defense cooperation', 'multinational force activities', 'civil operations', 'defense diplomacy', 'ceasefire monitoring group', and 'pro-Korean activities', and 'general troop deployment' related to troop deployment in general. Next, social network analysis was performed to examine the relationship between keywords and major keywords related to topic decision, and the keywords 'overseas', 'dispatch', and 'high level' were derived as key words in the network. This study is meaningful in that it first examined the topic of the Korean armed forces's peacekeeping activities over the past 20 years by applying big data techniques based on the National Defense Daily, an unstructured document. In addition, it is expected that the derived topics can be used as a basis for exploring the direction of development of Korea's peacekeeping activities in the future.

Keywords : PKO, Topic Modeling, Social Network Analysis, Trend

1. 서론

한국이 유엔(UN, United Nations)에 공식적으로 가입한

지 30주년이 되는 지금, 한국군은 전 세계 곳곳에서 세계 평화유지를 위해 각자의 자리에서 임무를 수행해 오고 있다. 한국은 '도움받는 나라'에서 '도움주는 나라'로 성장했음을 세계에 알리며 국가 이미지를 제고하고, 국제평화와 국익증진 그리고 국가위상에 걸맞는 역할 실천을 위하여 유엔활동에 참여하고 있다. 한국군은 2021년 10월 기준 부

Received 10 November 2023; Finally Revised 3 December 2023;
Accepted 4 December 2023

† Corresponding Author : hjseol@cnu.ac.kr

대단위로 유엔 PKO활동(PKO, Peace Keeping Operations)에 482명, 다국적군에 303명, 국방협력에 148명이 참가하고 있으며 개인단위로는 유엔 임무단에 23명, 다국적군엔 11명이 참여하는 등 회원국 중에서도 적극적으로 파병활동을 수행하고 있다. 이는 미국, 중국, 일본, 러시아, 인도와 같은 선진국과 비교해도 높은 참여 수준이다[18].

특히 한국군 파병부대들은 성공적인 임무 수행으로 현지인들에게 ‘신이 내린 선물(동명부대)’, ‘레오간의 천사(단비부대)’ 같은 긍정적인 평가를 받고 있으며, 유엔에서도 ‘유엔 메달’을 수여하여 한국군의 우수성을 인정하였다.

이렇듯 파병부대가 성공적으로 임무를 수행해 오고 있음에 따라 파병부대의 성과에 대한 연구들도 확대되고 있다. 그러나 개별 파병부대의 제한적인 성과분석을 중심으로 한 연구가 다수 이루어져, PKO활동에 대한 전반적인 동향 분석에 대한 연구는 부족한 실정이다.

따라서 본 연구에서는 한국군 파병부대의 활동에 대한 동향 분석을 통해 PKO활동의 초기부터 현재까지 한국군 파병의 흐름을 정리하고 그 연관성을 분석하는 것에 목적이 있다. 이를 위하여 한국군 파병부대 및 PKO활동 관련 기사를 대상으로 분석을 수행하였다. 분석에 사용되는 데이터는 국방일보로부터 수집하였는데, 이는 국방일보가 국방부 발행 일간지로 파병활동의 주체인 군에 대하여 전문성을 갖추었기 때문에 분석에 적절한 데이터로 판단하였다.

연구 방법론으로는 토픽모델링(Topic Modeling)과 사회연결망 분석(SNA, Social Network Analysis)을 활용하였다. 토픽모델링은 비구조화된 텍스트 자료들의 집합에서 의미 있는 주제를 추출해 주는 확률적 모델 알고리즘이다. 이러한 토픽모델링 기법의 특성에 따라 PKO활동을 다룬 국방일보 기사에는 어떠한 활동을 주로 다루고 있는지를 살펴보기 위하여 토픽모델링 기법을 활용하였다. 사회연결망 분석은 개체 간의 관계를 노드와 링크로 네트워크화하고 이러한 네트워크의 특징을 계량적으로 살펴보는 방법이다. 본 연구에서는 토픽모델링을 통해 도출한 주제와 관련된 키워드를 노드로 보고 이러한 키워드들의 관계를 네트워크화하여 연관성 및 인과관계를 살펴보기 위하여 활용하였다.

2. 한국군 파병활동

한국군은 1991년 유엔에 가입한 직후 소말리아 지역에 상륙수부대를 파병하는 것을 시작으로 본격적인 파병을 시작하였으며, 이후 다국적군 참여와 우호국 간의 국방교류협력을 확대함에 따라 다양한 유형으로 파병을 진행하게 되었다.

한국군의 파병유형은 크게 3가지 유형으로 구분할 수 있는데, 첫째, 유엔의 요청으로 분쟁지역에 파병되는 ‘PKO 활동’ 유형, 둘째, 특정 국가나 지역기구가 임명한 사령관의 통제하에 국제평화활동을 수행하는 ‘다국적군 활동’ 유형, 셋째, 우호국과 ‘국방교류협력’ 목적으로 파병되는 유형이다[22].

2.1 유엔 PKO활동

유엔 PKO활동은 유엔이 회원국을 대상으로 분쟁지역에 대한 병력파병을 요청함에 따라 정부 및 국회가 동의하여 병력을 편성하고 분쟁지역에 전개하면서 진행된다. 부대 파병의 경우, 자위와 임무를 위한 최소한의 무장을 갖춘 비전투병과 중심 편제로 민사작전과 분쟁 억제를 위한 감시작전 등을 수행하며 개인 파병의 경우, 유엔 임무단의 목표에 따라 읍저버, 참모장교 등의 임무를 수행한다. 한국군은 유엔 가입 이래 수차례의 PKO활동을 부대/개인 파병을 통해 수행하고 있으며, 주요 구간별 활동 및 단위별 파병부대는 <Table 1>과 같다.

<Table 1> Status of UN PKO Participating Individuals/Units[18]

Division	Period	Deployment Area	No. of People	Shift Cycle
Unit	1993-1994 1999-2003	Somalia/East Timor (Sangnoksu Unit)	250/420	8 months
	1994-2006	Western Sahara (Medical Support Group)	20-42	8 months
	1995-1996	Angola (Military Engineer Unit)	198	6 months
	2007-	Lebanon (Dongmyeong Unit)	279	8 months
	2010-2012	Haiti (Danbee Unit)	240	6 months
	2013-	South Sudan (Hanbit Unit)	203	8 months
Personal	1994-	India/Pakistan (Ceasefire Monitoring Group)	8	12 months
	2007-	Lebanon (Peacekeeper)	4	12 months
	2009-	Sudan (Dafur Mission Group)	7	12 months
	2009-	Western Sahara (Election Observation Group)	3	12 months
	2011-	South Sudan (Mission Group)	7	12 months
	2019-	Yemen (Political Mission Group)	-	12 months

2.2 다국적군 활동

다국적군 활동은 유엔이 독자적으로 감당할 역량이나

재원이 부족할 시 지역기구 및 특정 국가에 협력을 촉구함에 따라 이루어지는 파병활동이다[13]. 다국적군은 ‘평화강제’의 성격이 강하며, 유엔이 아닌 각 파견국이 소요경비를 부담한다. 한국군은 미국 주도의 다국적군에 참가하여 이라크, 아프가니스탄에 병력을 보내 임무를 수행하였으며 해상무역로 보호를 위해 청해부대를 파병하는 등 육·해·공 전력을 각지에 파병하였다. 그러나 국민정서 등을 비롯한 여러 문제로 인하여 유엔 PKO와 마찬가지로 전투 임무를 수행했다기 보다는 후방지원 및 재건 활동 등의 민사작전에 전념하였다[26]. 주요 구간별 다국적군 활동은 <Table 2>와 같다.

<Table 2> Status of Participating Multinational Forces Units[26]

Areas	Period	Deployment Location	No. of People
Gulf War	1991 (4months)	Saudi Arabia (Medical Support Group)	154
	1991 (3months)	UAE (Beema Unit)	160
Afghanistan War	2001-2003	Diego Garcia (Haeseong Unit)	823/year
	2001-2003	Diego Garcia (Chungma Unit)	1,245
	2002-2007	Afghanistan (Dong-eui Unit)	786
	2003-2007	Afghanistan (Dasan Unit)	1,329
	2010-2014	Afghanistan (Oshno Unit)	1,745
Iraq War	2003-2004	Seohee Unit	956
	2003-2004	Jema Unit	185
	2004-2008	Zaytun Unit	17,708
	2004-2008	Daiman Unit	1,324
Somalia	2009-	Cheonghae Unit	300

다국적군 활동 또한 유엔 PKO활동과 마찬가지로 개인 단위 파병이 존재하며, 연합해군사(CMF), 연합합동기동부대-아프리카의 뿔(CJTF-HOA), 미 중부사 협조단, 미 아프리카사령부 등지에 협조장교 및 참모장교들이 현재 파병되어 있다[17].

2.3 국방교류협력

국방교류협력 활동이란 당해 국가의 요청에 따라 비분쟁지역에 파견되어 교육훈련이나 재난구호 등을 수행하는 것을 말한다[22]. 한국군의 국방교류협력 파병사례는 2011년 UAE에 파병된 아크부대와 2013년 필리핀에 파병된 아라우부대가 있다.

아크부대는 대한민국-UAE 간 ‘전략적 동반자’ 관계의 상징으로 UAE군과 연합훈련을 통해 양국 간의 군사협력을 굳건히 하는 것이 주요 목표이다[8]. 파병은 한국군을 벤치마킹하여 교육훈련의 질을 높이고 국방체계를 선진화하기 위해 UAE가 한국군의 파병을 요청함에 따라 진행되었다[18]. 다른 유형의 파병과 차별화되는 점은 연합훈련 및 상호교류가 주목적이므로 비분쟁지역에 전투병과 중심의 육·해군 특수전 병력이 파병된다는 점이다. 한편, 아크부대 파병은 군사 우호관계의 조성뿐만 아니라 경제적 측면에서 원전 수주를 통해 경제적 성과, 한-UAE 간 방산협력 확대 등으로 이어졌다[25].

아라우부대는 군사협력 및 연합훈련을 목표로 하는 아크부대와 달리 재난구호 목적으로 파병된 부대이다. 슈퍼태풍 하이옌으로 필리핀 레이테 동부지역이 초토화되자 필리핀 정부는 공식적으로 한국군의 파병을 요청하였고, 그에 따라 아라우부대(필리핀 합동지원단)가 파병되었다. 아라우부대는 한국군 최초로 재해 당사국의 요청을 받아들여 진행된 파병이다[1]. 해군 상륙함을 통해 전개되었으며 부대 규모는 520여명으로 파병부대 중에서는 자이툰부대 다음으로 규모가 컸다. 부대는 중장비를 이용한 태풍피해 잔해물 제거, 공공시설물 복구 공사, 오지 주민에 대한 의료 지원, 무료 급식 등의 민사작전을 수행하였으며 2014년 12월 임무를 마치고 철수하였다[24].

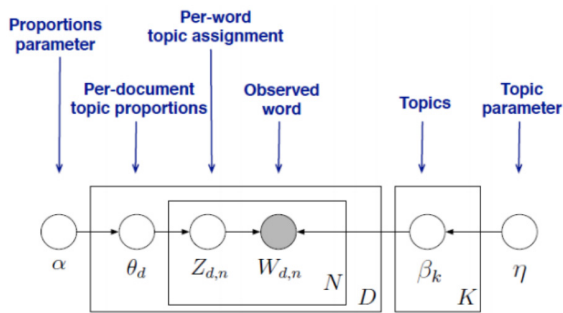
3. 분석방법론

3.1 토픽모델링

토픽모델링(Topic Modeling)은 텍스트 전체의 내용을 대상으로 의미 있는 주제를 추출하고, 그 경향성을 파악할 수 있다는 점에서 각 분야에서 활발하게 사용되고 있다. 토픽모델링을 활용하면 신문이나 소설, 학술논문, 역사적 자료 등의 비구조화된 방대한 자료 가운데 얽힌 맥락과 관련된 단서를 이용하여 해석 가능성이 높은 주제가 추출된다. 이처럼 단어의 종류와 빈도를 분석하여 문서에 등장하는 핵심적인 주제를 탐색하고, 문서의 지식 구조 체계를 밝혀냄으로써 문서 집합 속에 잠재적인 규칙을 파악할 수 있다[3]. 이러한 이점을 활용하여 토픽모델링은 학문분야의 연구 동향을 파악하거나, 매체에 따른 보도의 입장을 분석하거나, 사회역사적인 사건에 따라 구간별 주제 변화를 해석하는데 적극적으로 활용되고 있다[12].

토픽모델링의 대표적인 알고리즘으로는 LSA(Latent Semantic Analysis), pLSA(probabilistic Latent Semantic Analysis), LDA(Latent Dirichlet Allocation)가 있다. 이러한 알고리즘에서 가장 많이 활용되는 것은 LDA로 이는 문서

의 주제를 알지 못하는 상황에서도 단어의 패턴을 바탕으로 어떤 주제에 단어가 포함될 확률을 모형화하는 장점이 있다. 즉, LDA는 방대한 문헌 집단에서 잠재적으로 의미 있는 주제를 찾아내기 위한 절차적 확률 분포 모델이다[30]. LDA는 단어 간의 관계에 의존하여 서사나 구문, 단어 간의 동시 발생을 측정함으로써 각 문서에 어떤 토픽이 존재하는지 확률적으로 판단하여 토픽을 추출하는 것으로, 단어 배열의 순서는 고려하지 않고 주위의 어떤 단어가 있는지에 따라 군집이 형성되는 원리이다[11]. 이에 디리클레(Dirichlet) 확률 분포를 활용하여 문서가 일정한 토픽에 속할 잠재 확률을 할당하고, 토픽과 토픽 분포가 출현할 확률을 계산한다. LDA 알고리즘의 시각적 모델은 <Figure 1>과 같다.



<Figure 1> LDA Algorithm

이 계산에 MCMC(Markov Chain Monte Carlo) 기법의 일종인 깁스 샘플링(Gibbs Sampling) 알고리즘을 사용하여 복잡하게 주어진 파라미터 사후 분포 함수 속성을 도출한다[3]. LDA에서 중요하게 여겨지는 α 는 각 문서별 토픽 분포에 대한 디리클레 분포의 설정값으로, 각 문서-토픽 간 밀도를 나타내어 α 값이 커질수록 문서의 집합이 더 많은 수의 토픽으로 구성되었음을 의미한다. β 는 토픽별 단어 분포에 대한 디리클레 분포의 설정값으로, 문서와 토픽 간 밀도를 의미하며 β 값이 커질수록 토픽이 더 많은 단어로 구성되어 있다고 이해할 수 있다. 이처럼, LDA는 사전에 계산된 확률 분포를 바탕으로 단어가 어떠한 주제에 포함되는가를 예측하는 모델이며, 이 과정에서 α 값과 β 값은 결과에 많은 영향을 주기 때문에 최적의 값을 찾아내는 것이 중요하다[25].

3.2 사회연결망 분석

사회연결망 분석(SNA, Social Network Analysis)은 개인, 집단, 사물 등 개체 간의 관계를 노드 간의 링크로 간주하고, 그 링크의 구조가 결정하는 네트워크의 특징을 계량적으로 연구하는 방법론으로, 보이지 않는 복수의 링크

들이 형성하는 패턴, 패턴 상의 나타나는 노드의 위치, 패턴 형성의 인과관계 등을 밝혀내는 것이 가능하다[5].

이러한 SNA 분석 중 키워드를 활용한 키워드 네트워크 분석은 데이터에서 추출한 단어를 각각의 노드로 하며, 단어들의 관계를 네트워크로 형성해 계량적으로 분석하여 텍스트의 의미구조와 맥락 등을 나타낼 수 있다. 키워드 네트워크 분석은 키워드 개념들의 관계 구조를 공간적으로 나타내 줌으로써 텍스트에 대한 새로운 해석 가능성을 열어주기 때문에 최근에는 연구 동향 뿐만 아니라 다양한 학문 분야에서 시도되고 있다[23, 33].

<Table 3> Keyword Network Centrality Analysis Indicator[22]

Type	Indicator	Definition	Meaning
Centrality	Degree Centrality	<ul style="list-style-type: none"> Sum of the number of neighboring nodes connected to each other The larger the number of connected nodes, the larger the value 	Content-oriented keywords
	Eigenvector Centrality	<ul style="list-style-type: none"> Node connectivity/ importance measurement 	Most influential keywords

키워드 네트워크 분석을 위한 대표적인 지표 중 중심성(Centrality) 분석은 전체 또는 시계열적 변화에 따른 동향 및 핵심 키워드 등을 도출해 낼 수 있다. 본 연구에서 사용된 키워드 네트워크 중심성 분석지표에 대한 개념은 <Table 3>과 같다.

중심성 분석의 지표에는 연결중심성(Degree Centrality)과 위세중심성(Eigenvector Centrality)이 대표적이며 키워드 네트워크 분석에서 연결중심성과 위세중심성이 높은 키워드일수록 해당 분야에서 내용 구성에 중요한 영향을 끼친 대상이라 할 수 있으며, 이는 곧 해당 분야를 논할 때 필수불가결한 영역이라고 할 수 있다[20].

노드와 링크 수가 많은 조밀한 네트워크에서 노드들을 그대로 유지하면서 중요한 링크만을 남겨놓는 방식인 PFnet(Path Finder Network)은 정제된 네트워크를 통해 네트워크를 시각화할 때 직관적인 네트워크 지도를 형성할 수 있다는 장점이 있다.

네트워크 분석에 대한 결과값을 시각화하는 대표적인 방법으로 Spring Map(스프링형 맵)이 사용된다. Spring Map은 기존 노드 사이의 실제 경로 거리를 반영하는 MDS와 달리 그래프의 좋은 특징을 극대화되는 상태로 그래프를 그리며, 이로 인해 노드 사이를 밀고 당기면서 연결 구조를 가장 잘 보이는 그래프를 만들어 내기 때문에 노드의 좌표는 특정한 의미를 갖지 않는다는 특징이 있다[10]. Spring Map의 가장 대표적인 알고리즘으로는 그래프 상 노드 간의 거리가 경로거리를 반영하도록 하는 Kamada & Kawaii가

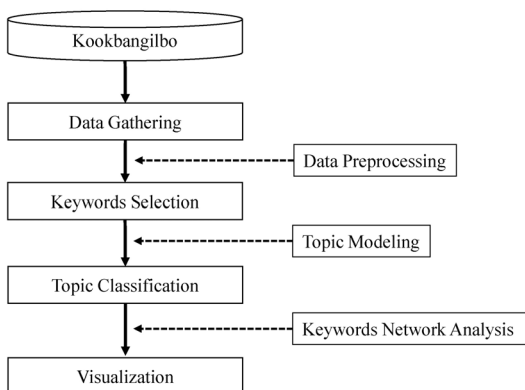
있으며, 직접 연결된 노드는 서로 당기고, 직접 연결되지 않은 노드는 서로 밀어내는 배치를 통해 응집성의 구분이 좀 더 뚜렷하게 드러나는 Fruchterman & Reingold 알고리즘도 많이 사용된다.

4. 연구방법

4.1 연구 프로세스

본 연구에서는 국방일보에서 보도한 PKO활동과 관련된 보도 자료에 토픽모델링 분석을 적용하여 전반적인 활동 동향을 살펴보고, 도출된 토픽과 이에 대한 키워드를 중심으로 연관성을 분석하고자 한다. <Figure 2>는 이를 위한 전체적인 연구 흐름도이다. 먼저, 국방일보에 게시된 자료 중 홈페이지 자체 검색 기능에서 PKO를 검색하여 2001년부터 2020년까지 보도된 자료를 수집하였다. 수집한 데이터에서는 기사의 본문 텍스트를 바탕으로 키워드를 구축하였다. 이후 확률적 토픽모델링 기법인 LDA를 활용하여 토픽 및 키워드를 도출한 뒤 토픽을 정의하였다. 정의한 토픽과 키워드를 바탕으로 시계열적 변화에 따른 PKO활동의 동향을 분석하였다. 상위 키워드를 중심으로 키워드 연관 네트워크 분석을 위해 동시출현 네트워크를 구축하였고 이를 PFnet을 통해 정제하였다.

정제된 네트워크를 5년 단위로 구분하여 동시출현 네트워크를 활용한 중심성 분석 및 시각화를 통해 핵심 키워드 트렌드를 도출하였으며, 시계열적 변화에 따른 PKO활동의 동향을 분석하였다.



<Figure 2> Research Process

4.2 분석 대상 자료 및 분석 절차

본 연구에서는 2001년부터 2020년까지의 국방일보 보

도 자료 중 PKO활동을 주제로 한 보도 자료를 활용하였다. 본 연구의 목적이 국방일보의 보도 자료를 통하여 PKO활동의 동향을 분석하기 위함이기 때문에 자료 수집을 위한 단어로는 광의적인 단어인 ‘PKO’로 선택하여 데이터를 수집하였다. 그 결과 총 1,030개의 보도 자료 중 2021년 자료 91건, 대분류 항목 중 포토갤러리 항목 등에 포함된 44건의 데이터와 동향 분석에 불필요한 자료 203건을 제외하여 총 692개의 기사를 최종 분석 자료로 선정하였다.

데이터 수집을 마친 후 토픽모델링 분석을 수행하기 위해 데이터 전처리를 수행하였다. 데이터 전처리는 비정형 텍스트 자료인 보도 자료를 데이터 분석에 가장 적합한 형태로 만드는 과정으로 분석에 있어 가장 중요한 단계이다[30]. 본 연구의 데이터 전처리는 크게 불용어 제거, 빈도수 산출, 사용자 사전 구성 순으로 진행하였다.

먼저 불용어 제거 단계에서는 명사형 단어를 제외한 동사, 형용사, 접속사 등과 더불어 수사 및 숫자, 단위 등을 제거하여 분석에 용이한 명사형 키워드만을 추출한다. 이후 전체 키워드 중 최하 빈도수를 시작으로 분석에 큰 의미가 없는 키워드 및 사용에 부적절한 키워드 등을 제거하였다. 본 연구에서는 전체 키워드에서 빈도 5 이하의 단어는 분석에 의미가 없다고 판단하여 모두 불용어로 처리하였다.

사용자 사전 구성을 통해 본격적으로 분석에 필요한 단어와 불필요한 단어 등을 정리한다. 이 과정에서 유의어, 지정어, 제외어를 선정하며, 이 과정은 결과의 정확도를 높이는데 중요한 과정이다[4].

유의어 사전은 띄어쓰기가 다른 단어나 같은 의미를 갖는 유사 단어 등을 하나의 대표 단어로 지정하는 사전으로 본 연구에서는 대표성을 갖는 단어를 중심으로 유사한 의미의 단어를 하나로 묶을 수 있도록 하였다. 예를 들어 장관, 참모부, 사령부, 대통령, 간부 등의 단어는 ‘고위급’ 단어가 대표할 수 있도록 하였다.

지정어 사전은 의미 및 형태소 단위로 분리된 단어를 다시 하나의 단어로 처리하기 위한 사전으로, 본 연구에서는 ‘집단적 자위권’과 같이 하나의 키워드 형태를 가졌을 때 큰 의미를 갖는 단어 등을 중점으로 처리하였으며, 고유 명사인 국가, 군함, 부대의 명칭 등 또한 지정어 사전을 통해 하나의 단어로 처리하였다.

제외어 사전은 분석 키워드에서 제외할 단어의 목록으로 앞선 과정에서 처리한 불용어와 더불어 문서에서 많은 빈도수를 차지하지만 큰 의미를 가지고 있지 않으며, 토픽 키워드로 사용하기에는 부적절한 단어를 제외어 사전에 적용하였다. 그 결과 <Table 4>와 같이 분석을 위한 사전을 구성하였다.

〈Table 4〉 Dictionary Composition Status

Division	Thesaurus	Defined Words	Stopword
Word Count	535	54	10,494

본 연구에서 토픽모델링은 데이터 분석 툴인 (주)사이람의 NetMiner 4.4.3 프로그램을 활용하여 진행하였다. NetMiner 프로그램은 GUI 기반의 사용자 친화적 인터페이스 및 Python 기반의 스크립트를 활용하여 네트워크 분석과 심리학, 교육학, 경영학, 정보처리학 등의 다양한 분야에서 사용되고 있다.

토픽모델링을 위한 알고리즘으로는 LDA모형을 적용하였다. LDA모형은 문서의 주제를 알지 못하는 상황에서 단어의 패턴을 바탕으로 어떠한 주제에 단어들이 포함될 확률을 모형화하는 비지도 학습 방법의 확률 모형이다[29]. 본 연구에서는 LDA 파라미터 추정 방법 중 표집 기반 방법 MCMC 알고리즘의 한 종류인 깁스 샘플링(Gibbs Sampling) 알고리즘을 채택하였다.

LDA 분석을 위해 디리클레 파라미터 값(α , β), 토픽 수(K)와 샘플링 반복 횟수를 결정한다. 적절한 토픽의 개수 및 샘플링 반복 횟수는 연구자가 결과를 가장 효과적으로 해석할 수 있는 수준에서 결정할 수 있으며, 이는 토픽모델링을 활용하여 연구하는 분야에서는 통상적이라고 할 수 있다[15, 29]. 따라서 본 연구에서는 최적의 토픽 및 토픽 키워드를 도출하기 위해 디리클레 파라미터 값 중 α 는 문서 집합에 의해 정해진 기본값 0.1을 사용하였고, β 는 분석 패키지의 0.01을 사용하였다. 토픽 수는 6, 7, 8, 9개로 변화시키며 수행하였으며, 샘플링 반복 횟수 또한 3,000회, 5,000회, 6,000회, 8,000회, 9,000회로 변화시키며 수행하였다. 분석 결과 최종적으로 토픽 수 8개, 샘플링 반복 횟수는 8,000회로 설정하였을 때 가장 적절하다고 판단하여 이를 최종적인 토픽을 파악하는 데 적용하였다.

토픽의 명칭은 각 토픽별로 토픽 내 확률값이 높은 유의미한 단어를 바탕으로 정의하였으며, 특히 토픽 내 순위가 높고 해당 토픽에만 나타난 특징적인 단어를 참고하여 정의하였다.

토픽의 명칭을 정의한 후에는 각 기사별로 확률값이 가장 높은 토픽을 제1토픽으로 할당하였으며, 이를 연도별로 정렬하여 연도별 토픽의 동향을 도출하였다. 연도별 동향을 보다 직관적으로 파악하기 위해 2001~2020년까지를 5년 단위로 묶어 시계열적 변화에 따른 토픽의 흐름을 살펴보고, 전반적인 PKO활동 동향을 분석하였다.

다음 단계로 핵심 키워드를 중심으로 연관성을 도출하기 위해 네트워크를 분석하고, 이를 시각화하여 PKO활동 동향 분석을 수행하였다. 이를 위한 핵심 키워드는 8개 토픽

픽에서 출현된 각 상위 20개에서 중복을 제외한 키워드 총 112개의 단어를 추출하여 선정하였다. 이후 키워드 연관 네트워크 분석을 위해 단어-논문 간 네트워크를 단어-단어 간 네트워크로 변환하여 단어-단어 간 동시출현 네트워크를 형성하였다.

도출한 키워드 동시출현 네트워크를 기반으로 핵심 키워드 분석을 위한 중심성 분석을 진행하였다. 본 연구에서는 네트워크 분석의 지표로 연결중심성(Degree Centrality)과 위세중심성(Eigenvector Centrality)을 선정하였다. 연결중심성이 높은 단어일수록 대상 논의의 중심이 되는 키워드가 될 가능성이 크며, 이 값이 크게 나온 키워드들은 연결된 대상들이 많기 때문에 이 키워드를 제거하면 내용의 구성이 어렵게 된다. 위세중심성은 노드의 연결성과 영향력까지 고려한 값이다. 위세중심성이 큰 단어일수록 내용 구성에 중요한 영향력을 끼쳤다고 볼 수 있다[21].

본 연구에서 사용된 키워드들의 네트워크는 총 12,321개를 형성하고 있으며, 이러한 조밀한 네트워크에서 중요한 네트워크만을 남겨두는 처리 과정으로 PFnet을 사용하였다. 그 결과 64개의 네트워크를 제외한 나머지 링크들은 시각화에서 제외하여 전반적인 동향 분석에 핵심적인 네트워크만을 활용한 용이한 시각화 모델을 얻을 수 있었다. 분석한 결과를 직관적으로 파악하기 위해 노드 간 거리에 많은 영향을 주는 Fruchterman & Reingold 알고리즘의 Spring Map의 형태로 시각화하여 해석하고 결과를 도출하였다.

5. 실증분석

5.1 파병부대의 PKO활동 동향 분석

5.1.1 PKO활동 토픽 도출 및 정의

국방일보 PKO 관련 기사 692건으로부터 수집한 기사 본문을 대상으로 전처리를 수행하였고, 그 결과 3,104개의 단어가 분석대상으로 추출되었다. 추출한 단어를 대상으로 하여 LDA 알고리즘에 기반한 토픽모델링 분석을 수행하였다. 디리클레 파라미터 α 는 기본값 0.1을 사용하였고, β 는 패키지의 0.01을 사용하였다. 깁스 샘플링은 8,000회 반복 시행하였으며, 토픽별 단어 분포를 토대로 토픽 8개와 상위 키워드 20개를 각각 도출하고 도출된 토픽에 대하여 명칭을 부여하였다. 토픽의 단어와 분포만을 토대로 토픽을 추론하는 것은 어렵기 때문에 토픽별 단어를 포함하고 있는 기사 본문을 참고하여 토픽의 명칭을 부여하였다. 도출된 토픽과 부여된 명칭은 <Table 5A>, <Table 5B>와 같다.

<Table 5A> Results of PKO Activity Topic

NO	Topic1	Topic2	Topic3	Topic4
1	Deployment	High Level	Japan	Deployment
2	Overseas	Discussion	Self-Defense Forces	Medical Care
3	High Level	Collaboration	Information	South Sudan
4	Congress	Overseas	Navy	Soldier
5	Soldier	National Defense	Establishment	Reconstruction
6	Government	Security	Power	Afghanistan
7	Collaboration	Korean Peninsula	Defense	Citizen
8	Iraq	Exchange	Air force	Military Engineer
9	Law	Relationship	Battle	Vietnam
10	Discussion	Japan	Landing	War
11	Request	Joint	Weapon	Haiti
12	Union	China	Warship	East Timor
13	Pirate	Enforce	System	Iraq
14	Somalia	Way	Business	Africa
15	Scale	North Korea	Base	Sahara
16	Evaluation	Strategy	Sea	Medical Treatment
17	Nation	Defense Industry	Harbor	Build
18	Special Warfare	Education Training	Marine	Somalia
19	Relationship	Summit	Capability	Overseas
20	Status	Loan	Mountainous	Headquarter
Topic	General Deployment	Northeast Asia Defense Cooperation	Multinational Force Activities	Civil Operation

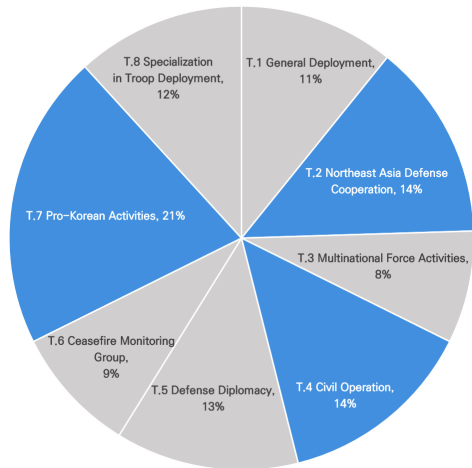
<Table 5B> Results of PKO Activity Topic(cont.)

NO	Topic5	Topic6	Topic7	Topic8
1	Overseas	Overseas	Soldier	Education Training
2	Security	Dispute	Lebanon	High Level
3	Diplomacy	High Level	Deployment	Soldier
4	National Defense	Observer Group	Overseas	Deployment
5	Development	Security Council	Delegation	Capability
6	Government	India	High level	Special Warfare
7	Policy	Ceasefire	Citizen	Overseas
8	Economy	USA	Education Training	Army
9	Society	Subscriber	East Timor	Procedure
10	North Korea	Society	Special Warfare	Organization
11	Enforce	Pakistan	Peacekeeper	Plan

NO	Topic5	Topic6	Topic7	Topic8
12	Strategy	War	Japan	Consolidation
13	Research	Discussion	Culture	Command
14	Nation	Japan	Observer Group	Union
15	Relationship	Start	Taekwondo	Specialty
16	Enlargement	Chairman	Army	Operation
17	USA	Member	Southern	National Defense
18	Role	General Assembly	Responsibility	Agency
19	Specialty	Independent	Picture	Practice
20	Change	Enter	Family	Work
Topic	Defense Diplomacy	Ceasefire Monitoring Group	Pro-Korean Activities	Specialization in Troop Deployment

도출된 토픽에 대하여 각각 살펴보면, Topic 1의 경우 ‘파병’, ‘고위급’, ‘국회’, ‘정부’, ‘요청’ 등이 주요 단어로 ‘이라크’와 ‘소말리아’ 같은 PKO 작전지역이 함께 도출된 것을 통해 파병 전반을 아우르는 ‘파병일반(General Deployment)’과 관련된 주제임을 확인할 수 있다. Topic 2는 ‘협외’, ‘한반도’, ‘교류’, ‘일본’, ‘중국’, ‘정상’ 등이 주요 단어로, 대북제재를 포함한 글로벌 이슈에 대한 동북아 주요 3국의 국방 정상 간 교류 등이 포함된 ‘동북아 국방협력(Northeast Asia Defense Cooperation)’과 관련된 주제임을 확인할 수 있다. Topic 3은 ‘일본’, ‘자위대’, ‘해군’, ‘공군’, ‘함정’, ‘해상’, ‘해병대’와 같은 단어들도 도출되는 것으로 보아 ‘다국적군 활동(Multinational Force Activities)’과 관련된 주제임을 확인할 수 있다. Topic 4는 ‘의료’, ‘재건’, ‘공병’, ‘진료’, ‘건설’의 단어들과 주요 PKO 작전지역들이 도출되는 것으로 보아 ‘민사작전(Civil Operation)’과 관련된 주제임을 확인할 수 있다. Topic 5는 ‘해외’, ‘외교’, ‘북한’, ‘강화’, ‘발전’, ‘관계’와 같은 단어들도 도출되는 것으로 보아 ‘국방외교(Defense Diplomacy)’와 관련된 주제임을 확인할 수 있다. Topic 6은 ‘분쟁’, ‘감시단’, ‘정전’과 같은 단어들과 ‘인도’, ‘파키스탄’ 같은 분쟁지역이 도출되는 것을 보아 ‘정전감시단(Ceasefire Monitoring Group)’과 관련된 주제임을 확인할 수 있다. Topic 7은 ‘레바논’, ‘동티모르’와 같은 PKO 작전지역과 ‘주민’, ‘문화’, ‘태권’과 같은 단어들도 도출되는 것으로 보아 ‘친한화 활동(Pro-Korean Activities)’과 관련된 주제임을 알 수 있다. Topic 8의 경우 ‘교육훈련’, ‘특수전’, ‘능력’, ‘전문’, ‘기관’ 같은 단어들도 도출되는 것으로 보아 파병 전문화 기관 및 교육훈련을 포괄하는 ‘파병전문화(Specialization in Troop Deployment)’와 관련된 주제임을 알 수 있다. 종합하자면 도출된 전체 토픽 8개 중 6개(동북아 국방협력, 다국적군 활동, 민사작전, 국방외교, 정전감시단, 친한화 활동)는 파병의 주요활동과

관련된 토픽이었으며, 2개(파병 일반, 파병 전문화)는 파병 관련사항 토픽이었다. 도출된 전체 토픽별 비율을 살펴보면 <Figure 3>과 같다.



<Figure 3> Distribution by Topic

비중 분석 간 이해를 돕기 위하여 기사 본문을 추가적으로 검토하였으며 최종적으로 도출된 분석결과를 통해 각 토픽의 비중을 살펴보면, Topic 7(친한화 활동) 21%, Topic 4(민사작전) 14%, Topic 2(동북아 국방협력) 14%, Topic 5(국방외교) 13%, Topic 8(파병 전문화) 12%, Topic 1(파병일반) 11%, Topic 6(정전감시단) 9%, Topic 3(다국적군 활동) 8% 순서로 높은 비중을 차지하는 것으로 나타났다. 가장 높은 비중을 차지하고 있는 토픽이 친한화 활동과 민사작전인 것으로 볼 때, 파병지의 주민들이 한국에 긍정적인 입장을 갖게 하여 우호적인 작전환경을 조성하기 위함이라 판단된다.

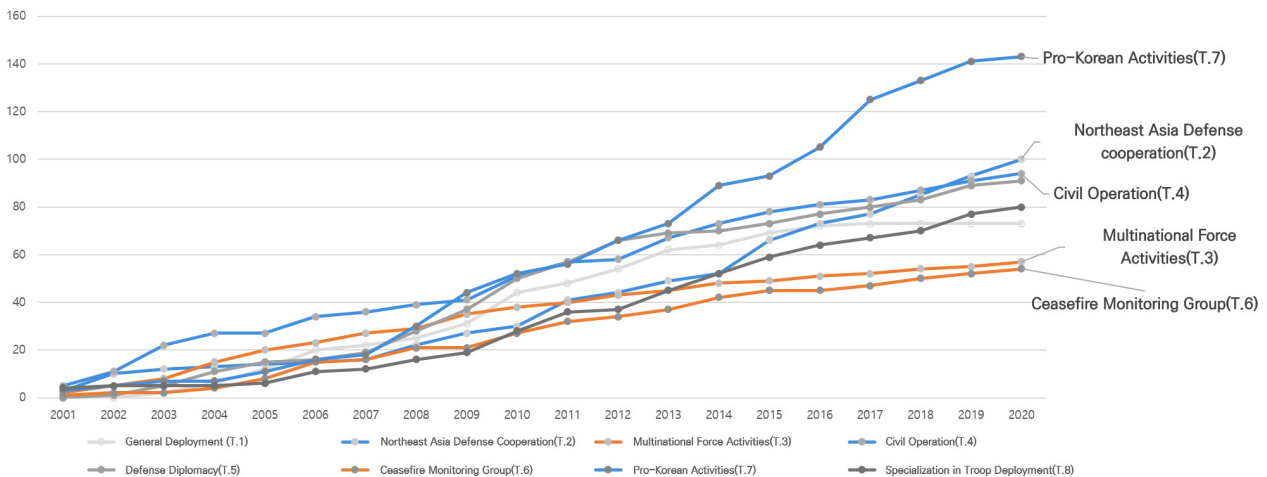
5.1.2 연도별 활동 추세 분석

연도별 활동 추세를 분석하기 위해 먼저 <Figure 4>에서 살펴본 Topic 별 비중을 연도별 누적 분포로 살펴보았다. <Figure 5>는 이를 나타낸 것으로 Topic 7(친한화 작전)은 분석 초기부터 지속적으로 증가하여 전반적으로 가장 많이 다룬 주제에 해당 된다. Topic 2(동북아 국방협력), Topic 4(민사작전), Topic 5(국방외교)는 점진적으로 증가하여 각 주제가 각각 비슷하게 다루어졌음을 확인할 수 있다. Topic 8(파병전문화), Topic 5(파병일반)는 앞서 다른 토픽들에 비하여 완만한 경사를 나타내고 있으며, Topic 3(다국적군), Topic 6(정전감시단)은 그 증가추세가 가장 낮게 나타남을 확인하였다.

다음으로 도출된 토픽을 바탕으로 전체기간 대상의 시계열적 변화를 분석하였다. 분석에 앞서 2021년의 경우, 코로나19 감염병의 영향으로 파병요원의 감염 위험으로 인한 활동 자제와 회원국의 재정지원 감소 등 PKO활동 수행에 있어 어려운 상황임을 고려했을 때, 분석에 있어 유의미한 결과를 도출하는 것이 어렵다고 판단하였다. 따라서 2001년부터 2020년까지를 분석구간으로 설정하고, 이를 중심으로 연도별 활동 추세를 살펴보았다. <Figure 5>는 도출된 토픽 8개에 대한 연도별 PKO활동 추세를 그래프로 나타냈으며, 용이한 분석을 위해 전체기간 중 5년을 주기로 하여 일정 수치 이상의 유의미한 토픽들을 가장 높은 수치와 가장 낮은 수치로 도출된 토픽 위주로 분석하였다.

전체기간 중 일정 수치 이상의 유의미한 토픽들은 다음과 같다. 분석 초기에는 Topic 2(동북아 국방협력), Topic 4(민사작전)가 도출되었으며, 중기에는 Topic 7(친한화 활동), 후기에 갈수록 Topic 7(친한화 활동), Topic 2(동북아 국방협력)가 연달아 높은 수준으로 도출되었다.

전체기간 중 5년을 주기로 가장 높고, 가장 낮게 나타난



<Figure 4> Cumulative Distribution by Year

토픽들은 다음과 같다. 2001년에는 Topic 4(민사작전)가 가장 높은 수치로 도출되었으며, Topic 1(파병일반), Topic 5(국방외교)가 가장 낮은 수치로 도출되었다. 2006년은 Topic 1(파병일반)이 가장 높은 수치로 도출되었으며, Topic 2(동북아 국방협력)와 Topic 5(국방외교)가 가장 낮은 수치로 도출되었다. 2011년은 Topic 2(동북아 국방협력)가 가장 높은 수치로 도출되었으며, Topic 3(다국적군)이 가장 낮은 수치로 도출되었다. 마지막으로 2016년에는 Topic 7(친한화 활동)이 가장 높은 수치로 도출되었으며, Topic 6(정전감시단)이 가장 낮은 수치로 도출되었다.

도출된 결과를 종합적으로 살펴보면, Topic 7(친한화 활동)은 지속적으로 높은 수준이 도출되었는데 Topic 4(민사작전)도 일정 부분 높은 수준으로 도출되고 있음을 확인할 수 있다. 이를 통해 파병부대의 민사작전과 친한화 활동이 어느 유형의 파병에서든 보편적으로 나타나고 있음을 확인할 수 있다.

또한, Topic 2(동북아 국방협력)를 통해 동북아 3국(한·중·일)의 PKO활동 참가 및 확대가 2000년대 들어 지속적으로 증가되었음을 확인하였다.

Topic 1(파병일반)의 경우, 아프가니스탄·이라크 전쟁이 발발하여 다국적군이 파병된 2000년대 초반과 동명부대, 청해부대, 아크부대, 아라우부대가 파병된 2007년부터 2011년까지 유의미한 증가를 보여준다.

Topic 3(다국적군)의 경우, 중동지역에 파병이 진행된 구간에 증가한 후 이후 파병이 없자 감소하는 추세를 보인다. Topic 5(국방외교)는 국방협력 목적으로 아크부대 파병이 논의 및 실행된 구간 상승하였으며, Topic 8(파병 전문화)은 전문 파병부대 필요성에 대한 논의가 활발히 진행된 구간, 즉 국방대학교 내 PKO센터 설립 구간과 특전사 내 파병전문부대 ‘국제평화지원단’이 창단된 구간에 상승

세를 보인다.

5.2 연결망 구조 분석

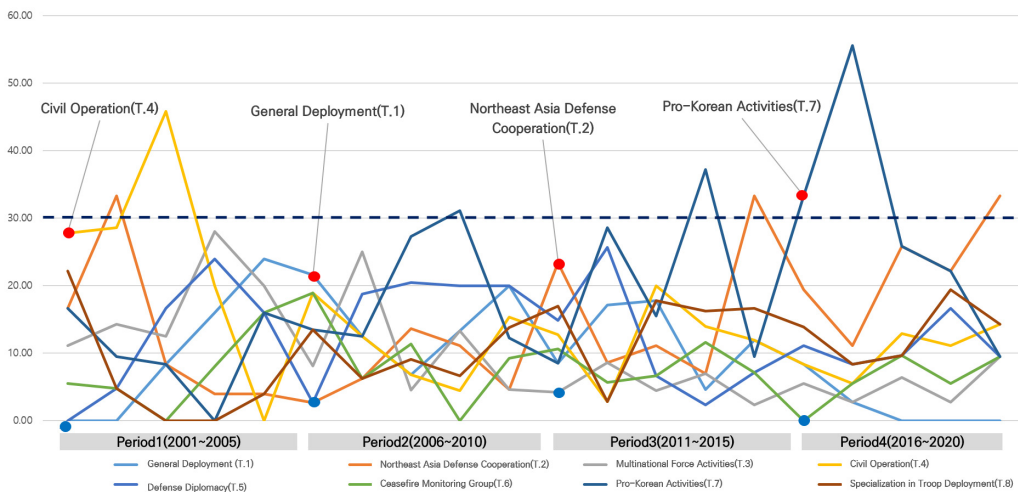
5.2.1 키워드 네트워크 분석 및 시각화

PKO활동의 시계열적 변화를 파악하기 위해 구간별 보도 자료를 분류하였고, 각 구간별 네트워크를 구성하였다. 구간은 국방일보 홈페이지 기사의 시작인 2001년부터 2020년까지를 5년 단위로 1구간(2001~2005), 2구간(2006~2010), 3구간(2011~2015), 4구간(2016~2020)로 구분하였다. 노트에 해당하는 키워드는 토픽 8개에 대한 각 토픽별 상위 키워드 20개를 추출하여 총 112개로 구성하였으며, 이에 대한 구성은 <Table 6>과 같다.

<Table 6> Number of Articles and Word Count Per Period

Type	Entire (2001 ~2020)	Period-1 (2001 ~2005)	Period-2 (2006 ~2010)	Period-3 (2011 ~2015)	Period-4 (2016 ~2020)
No. of Articles	692	113	207	212	160
Word Count	112	109	110	111	112

다음 단계로 각 기간별 키워드 연관 네트워크 분석을 진행하기 위해 동시출현 네트워크를 형성하였으며, 이를 위한 분석으로 연결중심성(Degree Centrality), 위세중심성(Eigenvector Centrality)을 통해 분석하였다. 연결중심성 값이 클수록 기간 내 기사에 있어서 중심의 대상일 가능성이 크며, 해당 키워드가 배제되는 경우 내용의 구성이 어렵게 된다. 더불어 연결중심성이 낮더라도 위세중심성이 높은 경우는 네트워크 내에서 높은 영향력을 가졌다는 의미를



<Figure 5> PKO Activity Trends by Year

나타낸다[11]. 분석 결과는 연결중심성 순위를 기준으로 상위 30개의 키워드를 위세중심성의 순위와 함께 정렬하였다.

동시출현 네트워크는 PFnet을 활용하여 분석에 불필요한 네트워크를 제거하고 시각화에 유리하게 정제하였으며 이를 Fruchterman & Reingold 알고리즘을 활용한 Spring Map을 통해 시각화하였다. 노드의 크기는 해당 키워드가 등장한 기사의 수를 의미하며, 링크의 두께는 연결된 키워드의 동시출현 네트워크 값을 의미한다.

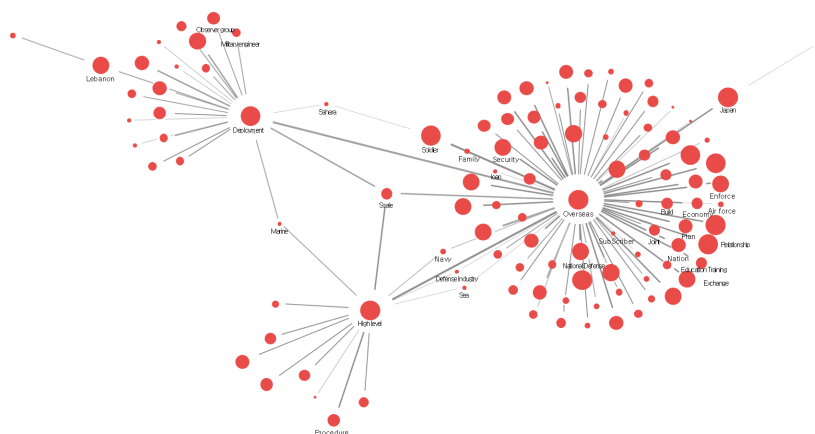
5.2.2 전체기간 네트워크 분석

전체기간의 분석 결과를 살펴보면, 상위 3개 ‘해외’, ‘파병’, ‘고위급’ 키워드의 연결중심성의 값이 매우 높은 것으로 나타났으며, 그 이하는 값이 급격하게 낮아지는 것을 볼 수 있다. 이러한 분석 결과는 위세중심성에서도 같은 경향을 갖고 있다. 이러한 경향은 ‘해외’, ‘파병’, ‘고위급’ 키워드가 사용자 사전 처리 과정에서 광범위하게 이루어지며 높은 중심성을 갖게 되었다고 볼 수 있으나, 전체기간에 걸쳐 매우 중요한 키워드이며 본 연구에서는 이들 키워드가 배제되는 경우 PKO동향에 대한 분석에 어려움이 있다는 것을 의미한다. <Table 7>은 전체기간 분석 결과를 나타낸 값이다.

이러한 분석 결과는 <Figure 6>과 같이 시각화한 네트워크에서도 나타난다. 높은 중심성을 가진 상위 3개의 키워드 ‘해외’, ‘파병’, ‘고위급’을 중심으로 네트워크 지도를 형성하고 있으며, 이는 3개의 상위 키워드가 PKO동향을 분석함에 있어 중요한 키워드임을 의미한다. 즉, 상위 3개의 키워드는 중요한 키워드임과 동시에 높은 대표성을 나타내는 키워드로도 해석할 수 있으며, 각 키워드가 연결된 노드와 네트워크 등을 미루어 보아 각각의 상위 키워드는 파병의 경우 실질적인 PKO활동의 수행, 고위급은 대내외적인 교류 및 의사결정, 해외는 대내외에서 발생하는 PKO활동의 전반적인 사건에 대한 대표성을 나타낸다고 해석할 수 있다.

<Table 7> Entire Period Degree Centrality, Eigenvector Centrality Analysis

Entire Period(2001~2020)			
NO	Keyword	Degree	Eigenvector
1	Overseas	0.738739	0.269153(1)
2	Deployment	0.189189	0.255402(2)
3	High Level	0.126126	0.231308(3)
4	Soldier	0.027027	0.191652(4)
4	Scale	0.027027	0.082065(48)
5	Family	0.018018	0.02677(104)
5	Lebanon	0.018018	0.085687(43)
5	Defense Industry	0.018018	0.029611(100)
5	Sahara	0.018018	0.030261(99)
5	Security	0.018018	0.146924(11)
5	Japan	0.018018	0.148238(10)
5	loan	0.018018	0.019371(109)
5	Navy	0.018018	0.04189(91)
5	Marine	0.018018	0.02328(106)
5	Sea	0.018018	0.031428(98)
16	Subscriber	0.009009	0.032852(97)
16	Observer Group	0.009009	0.083974(47)
16	Enforce	0.009009	0.124115(16)
16	Build	0.009009	0.078758(51)
16	Economy	0.009009	0.097003(33)
16	Plan	0.009009	0.098305(31)
16	Air Force	0.009009	0.05067(81)
16	Joint	0.009009	0.073523(56)
16	Military Engineer	0.009009	0.056985(74)
16	Procedure	0.009009	0.092538(37)
16	Relationship	0.009009	0.154427(7)
16	Exchange	0.009009	0.090648(39)
16	Education Training	0.009009	0.183236(5)
16	Nation	0.009009	0.102635(29)
16	National Defense	0.009009	0.144324(12)



<Figure 6> Spring Map(Entire Period)

5.2.3 기간별 네트워크 분석

<Table 8>, <Table 9>에서 나타난 바와 같이 구간 중심성 분석 결과를 살펴보면 전체기간 상위 3개의 키워드 이외에도 ‘군인’, ‘동티모르’, ‘미국’ 등 30개의 모든 키워드에 걸쳐 고른 중심성 값을 갖고 있다. 그 중 1구간에서는 ‘실시’, ‘동티모르’, ‘협동’, ‘공군’ 등 작전과 작전 수행과 관련된 키워드의 연결중심성과 위세중심성 모두 높은 것을 볼 수 있다.

2구간 중심성 분석 결과는 1구간과 전반적으로 유사하나 ‘고위급’, ‘파병’에 대한 중심성이 가장 높은 것으로 나타났다. ‘역할’, ‘관계’, ‘발전’, ‘시작’, ‘발전’ 등 이들을 뒷받침하는 키워드 역시 높은 중심성을 갖고 있다.

<Table 8> Period 1 Degree Centrality, Eigenvector Centrality Analysis

Period 1(2001~2005)			
NO	Keyword	Degree	Eigenvector
1	Overseas	0.694444	0.366678(1)
2	High Level	0.657407	0.34966(2)
3	Deployment	0.601852	0.313271(3)
4	Soldier	0.453704	0.24029(4)
5	Relationship	0.407407	0.205745(6)
6	Japan	0.37037	0.205975(5)
7	Government	0.361111	0.176877(8)
8	Discussion	0.351852	0.183244(7)
9	Security	0.314815	0.162849(10)
10	Education Training	0.296296	0.171237(9)
11	War	0.287037	0.16103(11)
12	National Defense	0.277778	0.151987(12)
13	Practice	0.25	0.131207(16)
13	Policy	0.25	0.136954(14)
15	East Timor	0.231481	0.144735(13)
16	Enforce	0.222222	0.117032(20)
16	USA	0.222222	0.129657(17)
18	Role	0.212963	0.133209(15)
19	Specialty	0.203704	0.108432(22)
19	Collaboration	0.203704	0.118691(18)
21	Development	0.194444	0.11245(21)
22	Plan	0.166667	0.0947(25)
22	Evaluation	0.166667	0.102455(23)
24	Army	0.157407	0.117581(19)
25	Society	0.148148	0.095014(24)
25	Operation	0.148148	0.083446(30)
27	Capability	0.138889	0.086132(28)
28	Air Force	0.12037	0.078957(32)
28	Way	0.12037	0.087354(27)
28	Iraq	0.12037	0.085983(29)

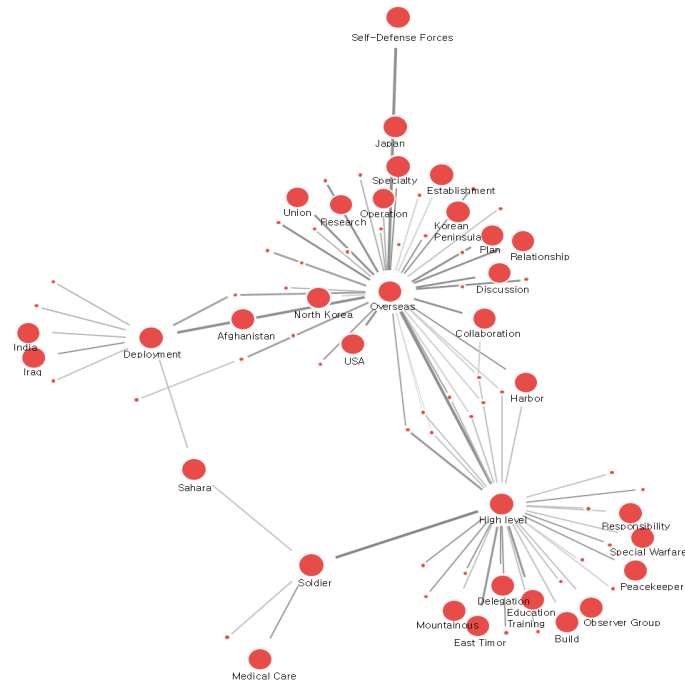
<Table 9> Period 2 Degree Centrality, Eigenvector Centrality Analysis

Period 2(2006~2010)			
NO	Keyword	Degree	Eigenvector
1	High Level	0.93578	0.26716(1)
2	Deployment	0.926606	0.238392(3)
2	Overseas	0.926606	0.262515(2)
4	Soldier	0.880734	0.189014(4)
5	Education Training	0.825688	0.169014(7)
5	Japan	0.825688	0.164709(8)
5	Discussion	0.825688	0.178288(6)
8	Government	0.798165	0.181811(5)
9	Security	0.788991	0.162797(9)
10	Operation	0.770642	0.131193(17)
11	Society	0.752294	0.140096(13)
11	Role	0.752294	0.14338(12)
13	Enforce	0.743119	0.132007(16)
14	Relationship	0.724771	0.15096(10)
14	Development	0.724771	0.146008(11)
14	Union	0.724771	0.116261(25)
17	National Defense	0.715596	0.133725(14)
17	Start	0.715596	0.115542(26)
19	Capability	0.706422	0.114371(27)
19	USA	0.706422	0.128175(18)
19	War	0.706422	0.117436(23)
22	Economy	0.697248	0.125308(20)
23	Plan	0.688073	0.120373(22)
23	Procedure	0.688073	0.111506(30)
23	Diplomacy	0.688073	0.122991(21)
26	Scale	0.678899	0.113272(29)
26	Policy	0.678899	0.12753(19)
26	Collaboration	0.678899	0.133674(15)
29	Evaluation	0.66055	0.111135(31)
30	Strategy	0.651376	0.114233(28)

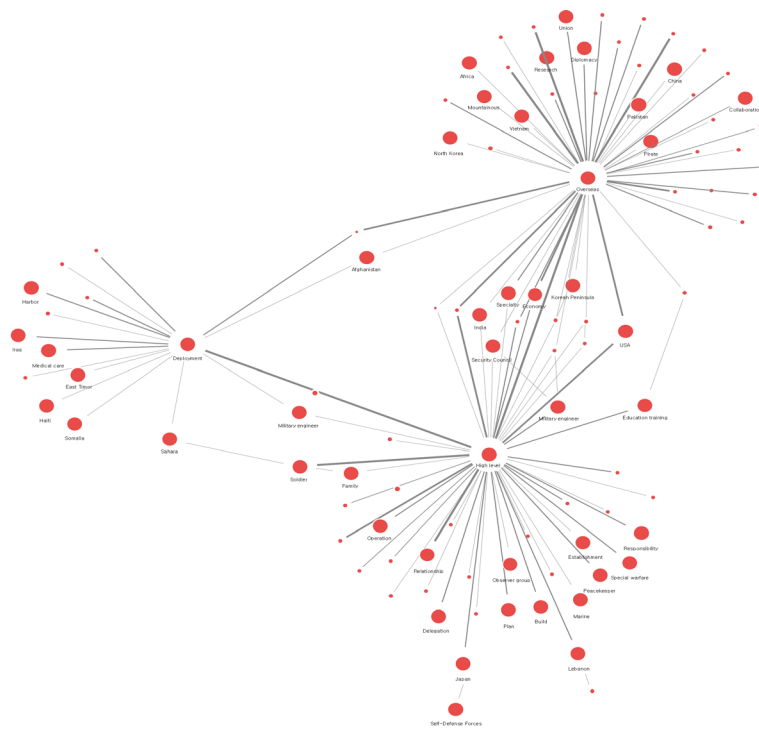
<Figure 7>과 같이 1구간의 네트워크 지도를 전반적으로 살펴보면, ‘해외’와 ‘고위급’ 사이에 위치한 키워드인 ‘교류’, ‘방안’, ‘확대’ 등은 그 키워드와 네트워크 지도에서 위치를 통해 국방협력 또는 국방외교의 의미를 갖는 것으로 해석 가능하다. 1구간 네트워크 지도에서 링크 값이 높은 대표적인 네트워크를 살펴보면 ‘해외’, ‘파병’, ‘이라크’로 이어지는 네트워크나 ‘해외’, ‘고위급’, ‘군인’, ‘의료’로 이어지는 네트워크이다. 이들 모두 2000년 이전부터 진행되었던 파병부대의 민사작전 등을 대표하는 네트워크로 볼 수 있으며, 이러한 점은 앞선 토픽모델링의 연도별 동향 분석과 마찬가지로 해당 구간의 PKO 활동의 동향이 민사작전의 수행이 중점임을 알 수 있다.

다음으로 <Figure 8>의 2구간 네트워크 지도에서 눈에 띄는 점은 ‘해외’, ‘고위급’ 간 관계가 높아졌으며, 그 사이 위치한 키워드 중 특히 ‘미국’, ‘안전보장이사회’ 등과의

교류가 많았음을 알 수 있으며, ‘고위급’과 ‘파병’의 네트워크 역시 높은 관계를 갖는 것을 통해 PKO활동 전반에 걸친 대내외적 교류와 더불어 실질적인 PKO활동의 수행



<Figure 7> Spring Map(Period 1)



<Figure 8> Spring Map(Period 2)

까지 이루어져 있음을 알 수 있다. 특히 2구간의 대표적인 네트워크가 ‘해외’, ‘재건’, ‘파병’, ‘의료’와 ‘협외’, ‘고위급’, ‘미국’, ‘해외’라는 점은 이러한 분석을 뒷받침한다. 2009년 청해부대 파병, 2010년 아이티 지진으로 인한 동명부대 파견 등이 실제로 이루어졌다는 점과 당시 정부의 PKO활동에 대한 전반적인 강화 등이 이루어졌다는 점은 이러한 분석에 영향을 미친 것으로 보인다.

3구간과 4구간의 중심성 분석 결과는 2구간의 중심성 분석 결과의 상위 30개의 키워드 대부분과 중복된다는 점에서 상당 부분 유사하다. 다만 3구간 중심성 분석 결과의 특징으로는 ‘확대’, ‘발전’, ‘위상’ 등이 더욱 높은 중심성을 가지며, ‘대표단’이 상위에 있다는 점을 통해 국방외교 및 대내외 협력에 대한 부분이 강조되었다고 볼 수 있다. 또한

<Table 10> Period 3 Degree Centrality, Eigenvector Centrality Analysis

Period 3(2011~2015)			
NO	Keyword	Degree	Eigenvector
1	Overseas	0.963964	0.282357(1)
2	High Level	0.945946	0.255539(2)
3	Deployment	0.927928	0.245184(3)
4	Soldier	0.900901	0.209187(4)
5	Education Training	0.882883	0.20334(5)
6	Discussion	0.828829	0.176506(6)
7	Collaboration	0.81982	0.173772(7)
8	Relationship	0.810811	0.169193(8)
9	Japan	0.801802	0.148751(10)
10	National Defense	0.792793	0.152392(9)
11	Role	0.792793	0.146135(13)
12	Government	0.783784	0.14692(12)
13	Union	0.774775	0.133704(17)
14	Enlargement	0.765766	0.130114(18)
15	Development	0.756757	0.148368(11)
16	Society	0.756757	0.134468(16)
17	Delegation	0.747748	0.139025(14)
18	Operation	0.72973	0.121543(21)
19	Status	0.72973	0.118531(23)
20	War	0.720721	0.110661(28)
21	USA	0.711712	0.118853(22)
22	Security	0.711712	0.13868(15)
23	Capability	0.702703	0.122958(20)
24	Reconstruction	0.693694	0.11106(27)
25	Dispute	0.675676	0.105369(29)
26	Start	0.666667	0.111308(26)
27	Enforce	0.657658	0.115148(25)
28	Nation	0.657658	0.1168(24)
29	Evaluation	0.657658	0.123604(19)
30	Battle	0.63964	0.099883(35)

‘평가’의 위세중심성 값이 연결중심성에 비해 상당히 높다는 점은 네트워크에서 높은 영향력을 가지고 있으며, 구간을 미루어 보아 지금까지의 PKO활동에 대한 평가 또는 보도가 전면적으로 이루어진 것으로 해석해 할 수 있다.

이러한 결과는 4구간에서도 이어진다. 상위 3개의 키워드와 순위가 같으며, 다만 ‘주민’, ‘교류’ 등의 키워드는 앞선 1구간에서 3구간의 상위 키워드에서는 나타나지 않으며, 이는 대민 또는 민간 차원의 교류가 1구간에서 3구간과 달리 많아졌다는 의미로 볼 수 있다.

<Table 11> Period 4 Degree Centrality, Eigenvector Centrality Analysis

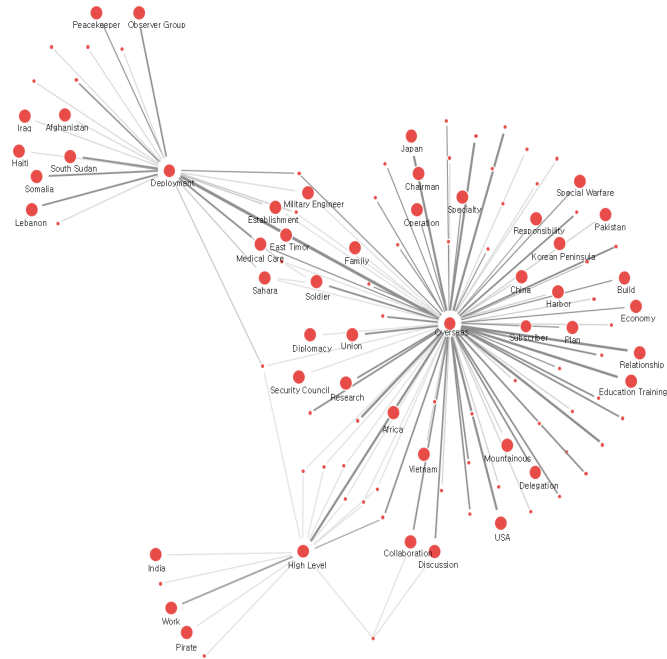
Period 4(2016~2020)			
NO	Keyword	Degree	Eigenvector
1	Overseas	0.81982	0.320435(1)
2	High Level	0.765766	0.292821(2)
3	Deployment	0.720721	0.246096(3)
4	Education Training	0.666667	0.241419(4)
5	Soldier	0.603604	0.202101(5)
6	Role	0.585586	0.164783(11)
7	Society	0.576577	0.165662(9)
8	National Defense	0.567568	0.184532(8)
9	Delegation	0.567568	0.16535(10)
10	Collaboration	0.558559	0.187477(6)
11	Discussion	0.54955	0.186292(7)
12	Relationship	0.522523	0.15186(14)
13	Government	0.522523	0.151618(15)
14	Evaluation	0.522523	0.14773(16)
15	Security	0.495495	0.162662(12)
16	Japan	0.495495	0.126084(20)
17	Development	0.486486	0.159032(13)
18	Enforce	0.45045	0.145903(17)
19	Specialty	0.441441	0.128678(18)
20	Lebanon	0.396396	0.11534(23)
21	South Sudan	0.387387	0.116014(21)
22	War	0.378378	0.098914(29)
23	Citizen	0.378378	0.100469(27)
24	Enlargement	0.378378	0.127885(19)
25	Exchange	0.369369	0.115473(22)
26	Status	0.36036	0.091234(34)
27	Union	0.351351	0.093825(31)
28	Peacekeeper	0.351351	0.101266(26)
29	USA	0.342342	0.092133(32)
30	Policy	0.342342	0.108927(25)

<Figure 9>의 3구간 네트워크 지도에서의 특징은 ‘해외’, ‘파병’ 간 또는 ‘해외’, ‘고위급’ 간 네트워크 값이 상당히 높다. 특히, ‘해외’ ‘고위급’ 간 네트워크 값이 높다는

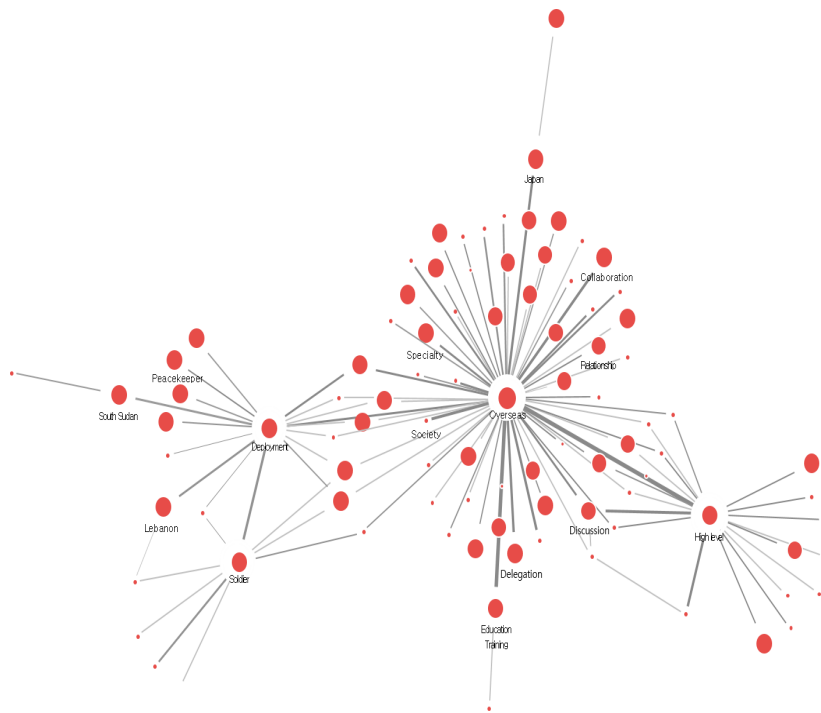
것을 미루어 보아 대외적인 교류 및 의사결정이 많이 이루어진 것으로 볼 수 있다. 반대로 ‘파병’, ‘고위급’ 간 네트워크는 2구간과 달리 많이 약해진 것으로 나타난다. 이는 앞선 2구간에서 전개되었던 PKO활동이 진행됨에 따라 파병활동에 대한 의사결정보다는 파병활동 진행 자체가 중점적으로 다루어진 것으로 해석된다. 실제 2011년부터

2015년까지 PKO활동보다 동북아 국방협력 및 외교 등 대외 활동이 중점적으로 이루어진 점을 통해 이러한 분석 결과가 도출된 것으로 보인다.

<Figure 10>의 4구간 네트워크 지도 또한 ‘파병’, ‘고위급’ 간 네트워크가 강하지 않다는 것이 특징이며, ‘해외’, ‘파병’, ‘군인’, ‘주민’으로 이어지는 네트워크의 값이 크다



<Figure 9> Spring Map(Period 3)



<Figure 10> Spring Map(Period 4)

는 점을 보았을 때, 앞선 중심성 분석 결과와 마찬가지로 민간 지원 작전 분야의 PKO 활동이 강화되었음을 알 수 있다. ‘해외’, ‘베트남’, ‘고위급’, ‘안보’ 네트워크는 강화되는 동북아 국방협력 중 특히 베트남과의 협력 강화에 대한 부분임을 알 수 있고, 2015년부터 2020년까지의 대한민국의 외교활동에 있어 남방외교 정책으로 인한 동남아, 특히 베트남과의 국방 교류가 많아졌다는 점을 보아, 이러한 부분이 네트워크 지도에 반영된 것으로 볼 수 있다.

6. 결 론

본 연구에서는 한국군 파병부대 동향분석을 통해 활동 초기부터 현재까지의 흐름을 정리하고 주요 동향을 분석하는 것을 목적으로 연구를 수행하였다. 이를 위하여 국방일보에서 한국군 파병부대 및 PKO 활동 관련 기사를 수집하여 토픽모델링을 수행하였고, 도출된 주요 키워드를 바탕으로 사회연결망 분석(SNA)을 통해 그 연관성과 인과관계를 분석하였다.

본 연구의 주요 결과는 다음과 같다. 2001년부터 2020년까지 20년간의 국방일보 PKO 활동·파병부대 관련 기사에 대한 토픽모델링 분석을 통하여 8개의 토픽을 도출하였으며, 각 토픽 명명은 국방일보 기사 본문과 국방부 자료를 참고하였다. 그 결과 Topic 1(파병일반), Topic 2(동북아 국방협력), Topic 3(다국적군 활동), Topic 4(민사작전), Topic 5(국방외교), Topic 6(정전감시단), Topic 7(친한화 활동), Topic 8(파병 전문화)로 6개의 주요활동과 2개의 관련사항 주제를 도출하였다. 도출된 8개 토픽 중에서 Topic 7(친한화 활동)과 Topic 4(민사작전)가 21%, 14%로 가장 높은 비중을 차지하고 있음을 확인하였으며, 다음으로 Topic 2(동북아 국방협력)와 Topic 5(국방외교)가 각각 14%, 13%를 높은 비중을 보였다.

이후 도출된 토픽들에 대하여 누적 분포와 연도별 분석을 실시하였는데, Topic 7(친한화 활동)과 Topic 4(민사작전)이 분석 구간 전반에 걸쳐 일정 기준 이상으로 상승하였으며, 이를 통해 대부분의 한국군 파병부대가 친한화 활동과 민사작전을 보편적으로 수행하고 있음을 알 수 있다. Topic 2(동북아 국방협력)의 높은 도출 빈도 또한 한·중·일 3국이 국제평화와 국가이익을 위해 협력하여 파병활동을 확대하고 있음을 증명한다. 이외에도 각각의 토픽들이 구간별 이슈에 따라 주목할 만한 상승세를 보였음을 확인하였다.

이와 같은 경향은 키워드 네트워크 분석 결과에서도 나타난다. 전 구간에 걸쳐 ‘해외’, ‘파병’, ‘고위급’의 중심성이 가장 높았으며, 이들을 중심으로 키워드 네트워크 지도가 형성되었다. 1구간에서는 ‘동티모르’, ‘의료’ 등의 키워드로 대표되는 민사작전, 2구간에서는 ‘협의’, ‘안전보장

이사회’, ‘아이티’, ‘재건’ 등으로 대표되는 파병 진행, 3구간에서는 ‘확대’, ‘발전’, ‘위상’, ‘대표단’ 등으로 대표되는 국방외교, 4구간에서는 ‘베트남’, ‘안보’, ‘교류’ 등으로 대표되는 동북아 협력 등이 분석 결과로 나타났다. 실제로 해당 구간에서의 정부의 기조와 사건 등을 미루어 보았을 때, 토픽모델링과 키워드 네트워크 분석 모두 실제 사건에 부합하였다.

본 연구는 지난 20년간 한국군 평화유지활동을 비정형화된 텍스트 문서인 국방일보를 기반으로 빅데이터 기법을 적용하여 주제를 살펴보았다는 점에서 의의가 있다. 그러나 국방일보 외에도 일반 언론과 다양한 방송 매체 등에 기고 및 보도된 내용을 포함하여 분석해야만 보다 폭 넓고 심도 있는 한국군 파병 활동을 전반적으로 파악할 수 있다는 점에 한계가 있다. 또한, 단순히 동향을 파악하는 수준에 머물지 말고 본 연구를 통해 이러한 연구가 가지는 학문적, 사회적, 군사적 측면에서의 의의 역시 명확하게 제시할 필요가 있다.

그럼에도 불구하고 본 연구에서 도출한 주제는 향후 우리나라 평화유지 활동의 발전 방향을 모색할 수 있는 토대로 활용할 수 있을 것으로 기대된다. 한국군의 평화유지활동은 글로벌 안보환경의 변화와 유엔의 국제평화활동의 패러다임 변화 그리고 우리나라가 국제사회에서 차지하고 있는 위상 등 여러 가지 요소에 의하여 영향을 받아왔다. 따라서 지금까지의 한국군의 평화유지활동에 대하여 살펴보는 것은, 향후 안보환경변화와 유엔의 정책 그리고 우리나라 국력에 맞추어 이에 적합한 파병정책 및 발전전략을 도출하는데 의미가 있기 때문이다.

Acknowledgement

This work was supported by Chungnam National University.

This paper is a supplement to the paper presented at the 11th National University Student Defense Policy Excellent Paper Presentation(Korea Institute for Crisis Management Analysis).

References

- [1] Araw Contingent Public Affairs Department, Dreams and hope for the wounded Philippines!, Korea National Defense University PKO Center, *PKO Journal*, 2014, Vol. 8, pp. 25-27.
- [2] Cho, Y.M., An Analysis of UN PKO activity and a suggested Pragmatic Direction of Korean PKO, *Korean Journal of International Relations*, 2010, Vol. 50, No.

- 1, pp. 165-190.
- [3] Cho, W.K., Social composition and health behaviors of medical knowledge: focusing on the abstract of breast cancer papers and topic modeling of patient forum posts, [dissertation], [Seoul, Korea], Seoul National University, 2017.
- [4] Choi, S.Y., and Ko, E.J., Analysis of <Korean Journal of Journalism & Communication Studies> from 1960 to 2018 using Metadata with Dynamic Topic Modeling, *Korean Journal of Journalism & Communication Studies*, 2019, Vol. 63, No. 4, pp. 7-42.
- [5] Gwak, K. Y., Analysis of Social Network, Chungarm, 2017.
- [6] Heo, N.S., Plan to expand the international role of the Korean military: Focusing on peacekeeping activities, National Assembly Defense Committee, 2002.
- [7] Scott, J., Analysis of Social Network, Communication Books, 2012.
- [8] Jeong, S.J. and Park, B.R., A Topic Modeling Approach to the Analysis Research Trend on Career Transition of Middle-aged Adults, *The Journal of Career Education Research*, 2020, Vol. 33, No. 4, pp. 101-124.
- [9] Kim, H.T., Meaning and strategic development plan of dispatching UAE ARK troops, Korea International Peace Operations Center, *Peacekeeping Operation Journal*, 2015, Vol. 10, pp. 22-25.
- [10] Kim, Y.H. and Kim, Y.J., Analysis of Social Network, Parkyoung Book, 2016.
- [11] Kwag, A.R., Analysis of Research Trends of South Korean Master's and Doctorate degree dissertations on Horticulture Therapy Using Text Mining and Semantic Network Analysis, [dissertation] [Seoul, Korea], Dankook University, 2020.
- [12] Lee, S.H., international conflict and the United Nations: focusing on the role of peacekeeping activities, *New Asia*, 2004, Vol. 11, No. 3.
- [13] Freeman, L.C., The development of social network analysis: A study in the sociology of science, Empirical Press, 2004.
- [14] Lee, S.H., A Study on the Dividing of Korean International Peace Activities, *Asia Research*, 2013, Vol. 56, No. 2, pp. 188-225.
- [15] Lee, K.H., Jung, H.J, and Song, M., Weighted topics-method network analysis of key journals in the field of literature informatics, *Journal of the Korean Society of Literature and Information*, 2015, Vol. 49, No. 3, pp. 457-488.
- [16] Lim, S.W., A design framework for Product-Service System using HOQ and patent data: A case study of augmented reality-HMD, [dissertation], [Seoul, Korea], Hanyang University, 2017.
- [17] Ministry of Defense, The Korean Armed Forces' overseas dispatch instructions, 2018.
- [18] Ministry of Defense, https://www.mnd.go.kr/mbshome/mbshome/mnd/subview.jsp?id=mnd_010901040000.
- [19] Newman, D. J. and Block, S., Probabilistic topic decomposition of an eighteenth century American newspaper, *Journal of the American Society for Information Science and Technology*, 2006, Vol. 57, No. 6, pp. 753-767.
- [20] Oh, J.S., A Study of Intangible Cultural Heritage Community through Social Network Analysis-Focused on the Item of Pansori, *The Journal of Cultural Policy*, 2017, Vol. 31, No. 1, pp. 158-183.
- [21] Oh, J.S., A Study on the Intangible Cultural Heritage Community Knowledge Network through Social Network Analysis: Focusing on Jeongseon Arirang, *Korean Journal of Cultural Heritage Studies*, 2019, Vol. 52, No. 3, pp. 172-187.
- [22] Oh, J.S., A Study on Research Trends and Knowledge System in Cultural Contents Studies Using Topic Modeling and Text Network Analysis Based on Big Data, *The Journal of Cultural Policy*, 2020, Vol. 34, No. 2, pp. 35-69.
- [23] Park, C.S., and Chung C.W., Text Network Analysis: An example of identifying shared meaning among policy stakeholders through socio-cognitive network analysis, *Journal of Governmental Studies*, 2013, Vol. 19, No. 2, pp. 73-108.
- [24] Park, J.S., Kim, C.S., Kwak, K.Y., Investigation of Research Trend in Hotel Domain Using Text Mining and Social Network Analysis, *Journal of Tourism and Leisure Research*, 2016, Vol. 28, No. 9, pp. 209-226.
- [25] Park, J.D., A Study on Issue Tracking on Multi-cultural Studies Using Topic Modeling, *Journal of the Korean Society for Library and Information Science*, 2019, Vol. 53, No. 3, pp. 273-289.
- [26] Park, D.S., An Evaluation of 50 Years of the ROK Militarys Overseas Dispatch and Future Policy Suggestions, *Korean Journal of Military Affairs*, 2020, Vol. 8, pp. 107-146.
- [27] Park, J.H., A Study toward Measures for Deploying the Reserve of the Korean Armed Forces to International

- Peacekeeping Operations, [dissertation], [Nonsan, Korea], Konyang University, 2021.
- [28] Song, S.J., Understanding of the UN peace maintenance activities, Yeon Kyeong Publishing, 2006.
- [29] Seol, H.J., A Study on the Research Trends on Military Studies: Focusing on the Analysis of Review of Korean Military Studies, *Review of Korean Military Studies*, 2019, Vol. 8, No. 2, pp. 213-232.
- [30] Seo, Y., Analysis of Defense R&D Technology Trends Using Patent Information, [dissertation], [Daejeon, Korea] Chungnam National University, 2021.
- [31] Seo, H.O., Defense Daily, https://kookbang.dema.mil.kr/newsWeb/20210823/3/BBSMSTR_000000100138/view.do.
- [32] Yoon, J.W., Watchmen of international peace, Korean troop dispatch activities and defense cooperation, *Defense & Technology*, 2016, pp. 124-131.
- [33] Yang, K.W., Research Trend Analysis of 'International Commerce and Information Review' Using SNA-based Keyword Network Analysis, *International Commerce and Information Review*, 2017, Vol. 19, No. 1, pp. 23-42.

ORCID

Donghyeon Jung | <http://orcid.org/0009-0002-7232-7515>
 Chansong Kim | <http://orcid.org/0009-0000-8266-7535>
 Kangmin Lee | <http://orcid.org/0009-0002-0614-4446>
 Soeun Bae | <http://orcid.org/0009-0000-2219-4107>
 Yeon Seo | <http://orcid.org/0009-0007-8407-8970>
 Hyeonju Seol | <http://orcid.org/0000-0001-6953-9766>