

## 전투지휘훈련간 전투협조회의 분석방법론 연구1)

### A Study on the Methodology for Analyzing Battlefield Coordination Meeting during BCTP training

조성일<sup>2)</sup> · 신용<sup>3)</sup> · 문호석<sup>4)</sup>

Sungeel Jo · Yong Shin · Hoseok Moon

#### ABSTRACT

The data produced at command posts during BCTP training holds significant importance. However, it is often observed that this data is deleted post-training without thorough analysis. Addressing this issue, the present study structurizes the voice command information (text data) generated at Command Control Center(CCC) during such training, using text mining techniques. By extracting key features, the study reveals implications for the advancement of Battle Command.

The findings of this study offer valuable insights to commanders and staff in three key areas. Firstly, within the context of Battlefield Coordination meetings, this research demonstrates the feasibility of objectively and scientifically extracting summaries and key concepts from meeting discussions in a concise manner, facilitated by Topic Modeling. Secondly, the provision of analytical data concerning individuals, including commanders, holds potential for significant improvements in communication strategies. Thirdly, a chronological analysis of commanders' directives yields insightful information on the relationship between specific decisions and their outcomes in varying situations. This analysis is pivotal in identifying essential traits that influence battle outcomes, thereby augmenting the leadership capabilities of commanders and aiding in the development of a comprehensive, data-driven decision support model

Key Words : Battle Command Information, Text-mining, Network Analysis, Topic Modeling, t-SNE

---

논문접수일 : 2023년 11월 24일, 심사일 : 2023년 11월 27일, 게재확정일 : 2023년 12월 3일

1) 본 논문은 육군합성전장훈련체계의 텍스트마이닝 기술연구(제2023MNS03-04호)의 지원으로 작성되었습니다.

2) 국방대학교 국방과학학과 석사과정

3) 한국소프트스페이스(주) 연구소 부장

4) 국방대학교 국방과학학과 교수 / 교신저자(corresponding author)

# 1. 서론

갈수록 고도화되고 복잡도가 증가하는 미래 전장에서는 지휘관과 참모의 개인 전문성에 의존하는 지휘결심체계에 한계가 있을 수밖에 없다. 지휘통제 복잡도의 증가는 2차세계대전 이후 2차원의 영역에서 현재 JAD(Joint All - Domain Command and Control) 고차원영역으로 확장되었으며 대용량의 정보 속에서 최적의 의사결정을 내리기 위한 지휘결심지원 시스템이 필수적이다.[1]

지휘관의 결심지원을 위한 대표적인 전투지휘데이터는 육군 BCTP<sup>5)</sup>(Battle Command Training Program)단 훈련 간 CCC<sup>6)</sup>에서 전투참모단이 훈련하는 내용이다. 훈련부대는 구성(Constructive) 모의훈련기로 가상으로 진행되며 CCC활동은 실제 지휘관과 참모가 참여하여 진행한다.[2] 훈련간 다양한 형태의 데이터를 수집할 수 있는데 비정형데이터로는 전투협조회의 내용(음성), 작전명령(텍스트) 등이다.

이런 유용한 훈련데이터들을 활용하여 지휘관의 지휘능력을 점검하고 전투협조회의를 올바르게 이끌어가고 있는지 분석하여 지휘관의 지휘능력 향상 제고가 필요하지만, 지휘관의 결심을 지원하기 위한 시스템 개발은 미흡한 실상이다. 특히 전투지휘훈련간 지휘소에서 생산되는 데이터는 매우 중요한 정보이나 훈련 종료 후 분석 없이 삭제되고 있다. 일부 저장은 하지만 보안상의 이유로 열람이 불가하여 데이터 저장 시스템 또한 구축되어있지 않다. 따라서 데이터 과학자에 의한 분석을 통해 전

투결과에 영향을 미치는 중요한 특성이 무엇인지 추출함으로써 지휘관의 지휘능력 향상에 기여할 수 있는 노력이 필요하다.

본 연구의 범위는 텍스트데이터로부터 전투지휘에 관한 단어들을 정의하고 생산하여 특성을 추출하는 것이며 실제 전투지휘정보가 보안상의 이유로 축적되지 않아 대신 유사한 형태의 토론 텍스트데이터를 사용하였다. 토론 텍스트데이터는 국방 TV, 국방포커스의 ‘2023년 안보정세 전망 미.중 갈등 그리고 아시아’에 대한 내용으로 총 47분가량의 토론데이터를 네이버 클로바노트를 활용하여 <그림 1-1>과 같이 음성데이터를 텍스트데이터로 변환해서 수집한 다음 분석간 활용하였다.



<그림 1-1> 클로바 노트로 수집한 데이터

이번 연구의 목적은 육군 전투지휘훈련단(BCTP)의 전투지휘훈련간 지휘소(CCC) 음성지휘 정보(텍스트데이터)를 텍스트마이닝 기술을 이용하여 정량화 시키는 것이다. 정량화된 데이터를 이용하여 특성을 추출하고 전투지휘 발전을 위한 함의를 도출하여 향후 발전적인

5) 전투지휘훈련(Battle Command Training Program) : 부대의 지휘통제능력 향상을 위해 지휘통제분부를 대상으로 훈련용 가상모의 체계를 활용하여 실시하는 훈련. 전투지휘훈련단 주관으로 훈련통제와 사후검토를 실시하며 훈련에 사용되는 가상모의체계는 창조 21모델, 화랑21모델, 전투근무지원 모델 등이 있음. (군사용어, 2018. 12. 28., 육군본부)

6) CCC(Command Control Center) : 지휘관과 참모 등 지휘통제 요원이 예하 부대를 효과적으로 지휘 통제할 수 있도록 지휘소 일부로 편성된 기구. 일반적으로 본부, 정보, 작전, 화력, 작전지속지원, 지휘통신 등의 기능실로 편성됨. (군사용어, 2018. 12. 28., 육군본부)

분석을 위한 제언을 포함한다. 활용되지 못하고 있는 비정형데이터를 정형화시킴으로써 데이터 기반의 의사결정 지원모델 구축에 기여할 수 있을 것이다.

## 2. 관련 연구

텍스트데이터는 대표적인 비정형데이터로 신문기사, 보도, SNS, 회의록 등등 우리 주변에서 쉽게 접할 수 있다. 이런 텍스트데이터를 활용하기 위해 정형화시키는 연구가 다양한 분야별로 진행되어왔으며 대부분 연구자의 주관적 견해와 전문성이 포함되어 분석되어왔다.

관련된 연구를 살펴보면 감미아(2012)는 신문사별로 논조의 차이를 비교·분석하기 위해 단순 빈도분석, 군집화, 분류 등의 기법을 활용하였다. 단순 빈도분석을 통해 해당 신문사가 주요하게 다루고 있는 주제, 단어를 비교하고 키워드 군집화로 유사한 단어들끼리 묶어 네트워크 분석을 하였다. 분류에서는 Naive Bayes 분류기를 사용하여 앞선 군집화 결과와 논조의 차이가 어느 정도인지 분석하여 시사점을 제출하였으나 논조 분석간 긍정, 부정 두 분류로만 나누어 중립의 입장이 포함되지 않았기 때문에 결과가 다소 간결하였고 단어 선정간 TF-IDF 방법을 적용하지 않고 단순 빈도수만을 활용하였다는 한계가 있다.

이창용(2016)은 국방안보 분야에 대한 텍스트마이닝 연구가 활발하지 않음을 지적하면서 정부에서 발간하는 북한월간동향 문서를 텍스트마이닝과 군집분석을 활용하여 핵심단어사전 구축, K-means 핵심단어 군집화, 연관성분석을 하였다. 핵심단어의 대상으로 월별 변화량의 Elbow Point에 있는 단어수를 선정하였으며 군집분석간 Hartigan's Rule을 이용하여 k 값을 선정하였다. 단어 군집분석과 북한 도발과의 연관성을 분석하는 방법으로 비정형데이

터를 이용한 분석 방법론을 제시했다는 점에 의의가 있다.

이창용(2016)의 '텍스트마이닝을 이용한 북한 보도동향과 북한 도발과의 연관성 분석'에서는 북한 보도내용을 텍스트마이닝과 시계열 분석을 통해 북한 도발과의 연관성을 분석하였다. 단어빈도수와 감성분석, 긍정부정 표현 등을 핵심단어 사전으로 구축하였으며 구축된 단어들을 시계열적으로 표현하여 시점별 북한의 실제 도발사례와 비교하여 연관성을 분석하였다. 특히 시계열분석을 통해 통계량의 변화를 시간에 따라 비교하는 방안을 제시하는 등 과학적인 기법을 최대한으로 사용하여 주관적인 판단을 줄이려는 노력을 하였으나 분석간 연구자의 주관이 들어갈 수밖에 없음을 한계로 보았다.

방한술(2019)은 토픽 모델링을 활용하여 토픽을 시계열적으로 표현하는 방법론에 대한 연구를 하였다. 네이버 뉴스기사에서 '북핵'과 관련된 기사를 바탕으로 핵심단어를 선정하고 토픽모델링과 주요 토픽을 선정한 다음 토픽을 시계열적으로 표현하고 군집화 하였다. 핵심단어 선정간에 TF-IDF를 활용하여 기존 연구자의 정성적인 판단에만 의존했던 대표단어 선정 방법에 객관적인 방법을 제시하였다. 특히 토픽모델링 결과에 대한 시계열적인 분석은 시기별 토픽의 관심도 변화를 정량적 분석에 기여하였다. 한계점으로 적정 토픽수 선정간 2가지의 방법(Deveaud, Cao)만 적용하였으며 큰 사회적 흐름을 뉴스 기사만으로 분석하려 하였기에 제한사항이 있었다.

홍세은(2020)은 치안분야에 텍스트마이닝 및 클러스터링 기법을 적용하였다. 연구는 비정형 데이터인 112 신고내용을 각각의 사건 유형별로 백터화한 다음 유사도에 따라 군집화하여 경찰 초기대응의 범위를 설정하는데 기여하였다. 군집화간에 사용한 데이터는 TF-IDF값이며 t-SNE를 활용하여 저차원으로 원데이터의

거리정보를 최대한으로 유지한 채 타점한 다음 군집화하였다. 이를 통해 클러스터링 결과 군집 그래프는 사건 유형들의 위치로 사건유형간 얼마나 유사한지 탐색적으로 확인할 수 있었다. 한계점으로 데이터의 질에 대해 논할 수 있는데 112 신고내용의 특성상 사람이 작업하기 때문에 오타자와 전문적인 은어 등은 분석하는데 제한사항으로 작용하여 데이터 축적뿐만 아니라 양질의 데이터 축적을 위한 필요성에 대해 지적하였다.

강동구(2022)는 코로나19 시기 자영업자 관련 신문기사를 텍스트마이닝 하여 사회적 인식 변화에 대해 연구하였다. 빅카인즈에서 신문 기사를 수집하여 네트워크 분석, 중심성, 상관관계, 빈도수, 워드클라우드, 감성분석 등의 방법을 적용하였다. 이런 연구를 통하여 분석결과와 자영업자 관련 정책을 비교함으로써 이해하

는데 활용할 수 있고 네트워크 분석으로 중심성이 높은 단어들이 어떤 특정 사회적 현상과 관계가 있는지 분석할 수 있었다. 한계점으로 자료의 수집에서부터 분석까지 텍스트롬(Textom)<sup>7)</sup>, UCINET-6<sup>8)</sup> 일관된 환경에서 작업하지 못했다는 점과 감성분석간 텍스트롬에서 어떤 사전을 사용하는지 공개하지 않아 언급하지 못했다는 점이 있다.

앞선 관련연구들의 사례에서처럼 신문기사, 뉴스, SNS 등 다양한 비정형 텍스트데이터를 분석하기 위한 연구가 활발히 진행되고 있었으나 우리 군에서 중요한 데이터인 전투지휘데이터를 어떻게 정형화시킬 것인가에 대한 연구는 없었다. 전투지휘데이터야말로 전투결과와 결부지어 우리 군의 교육훈련분야 발전에 도움이 될 수 있다. 따라서 본 연구는 전투지휘훈련간 전투협조회의 데이터를 어떻게 분석해서 활용

<표 3-1> 연구대상 데이터

Name	Time	Contents
윤지원	0:16	중국의 부상과 미중간 전략 경쟁의 심화로 동북아 안보 정세가 급변 ...
태상호	0:23	미중 경쟁과 함께 안보 이슈의 중심에 선 아시아 ...
황영진	0:27	인도 태평양 지역에서 미국과 중국을 중심으로 한 블록 형성 가속화 ...
윤지원	0:33	2023년 미중 갈등과 동북아 안보 정세 변화 전망해 보겠습니다.
이재현	1:12	네 안녕하세요, 아산정책연구원 이재현입니다.
이강규	1:15	안녕하세요. 한국국방연구원에서 국방전략연구실장을 맡고 있는 이 ...
정구연	1:21	네 안녕하세요, 강원대학교 정책학과 정구연입니다 ...
윤지원	1:25	앞으로 10년 미국인들은 결정적 10년의 시대를 살고 있다. 이 말이...
정구연	1:45	지금 미국은 여러 전략 보고서를 통해서 중국을 유일한 경쟁자라고 ...
윤지원	2:26	전방위적으로 이제 미중 간의 갈등은 심화될 수밖에 없다. 멈출 수 ...
태상호	2:34	미중 전략 경쟁이 심화가 되면서 더 이상 안보가 국가 차원 혹은 ...
...	...	...
이재현	45:05	제가 아까 말씀드린 것처럼 강대국 경쟁 사이에서 다변화라는 게 ...
윤지원	45:50	예 러시아 우크라이나 전쟁으로 인해서 글로벌 공급망 대란은 더욱 ...

7) 네덜란드 암스테르담대학의 로엣 리셀로드프(Loet Leydesdorff)박사가 빅데이터 처리를 위해 개발한 소프트웨어의 변형된 프로그램, 강동구(2022)

8) Windows 운영체제용 소셜네트워크 데이터 분석을 위한 소프트웨어 패키지, 강동구(2022)



<그림 3-1> 연구절차

할 수 있을까에 대해 연구하였으며 과학적인 기법을 통해 단시간내 분석정보를 제공할 수 있는 방안을 제시하였다.

### 3. 연구 방법

#### 3.1 연구 대상

실제 전투지휘정보는 보안상의 문제로 축적되지 않고 있다. 따라서 그와 유사하다고 판단한 국방 TV의 ‘안보정세 전망 미·중 갈등 그리고 아시아’<sup>9)</sup> 토론 자료를 활용하였다. 수집방법은 네이버클로바 노트를 사용하였다. 네이버클로바노트는 회의록 관리서비스로 음성문자 변환서비스(STT)를 제공하며 음성을 텍스트로 정리해준다. 클로바노트는 2020년 11월 처음 서비스를 제공한 하이퍼클로바 기반 음성인식 기술로 회의자료나 대화 음성데이터를 텍스트 데이터로 변환시키는데 효율적으로 활용할 수 있다.<sup>10)</sup> 이후 CSV파일로 정리한 결과 <표 3-1>과 같다.

t-SNE분석간 군집분석에 사용할 데이터는 '00년도 사단장 취임사데이터를 활용하였다. 총 33개사단의 취임사 데이터를 활용하였으며 취임사에는 지휘관의 강조사항을 포함한 지휘중점, 의도 등이 있어 음성데이터보다는 부족하지만 성향을 비교해볼 수 있는 이점이 있다.

#### 3.2 연구 절차와 방법

전투지휘훈련간 전투협조회의 텍스트데이터 분석 연구절차는 <그림 3-2>와 같다.

자료수집 단계에서는 음성데이터는 네이버클로바에서 STT(Sound To Text)를 적용하여 텍스트화해서 수집하였고 취임사데이터는 공개되어있는 자료를 정리하였다.

전처리 단계에서는 Stemming(어간 추출)과 Lemmatization(표제어 추출), 불용어단어 제거(Stopwords) 등의 과정을 거쳐서 분석 가능한 형태로 가공하였다.

핵심단어 추출 단계에서는 형태소 분리, 빈도수, TF-IDF<sup>11)</sup>기반 단어사전을 구축하였다. TF-IDF를 활용한 이유는 단순 빈도수만을 고려

9) 유튜브 국방TV 채널, “[2023 국방포커스] 2023년 안보정세 전망 미·중갈등 그리고 아시아” ([https://youtu.be/RCinvTEHCw?si=8\\_U9Ve1HcjmmrLj](https://youtu.be/RCinvTEHCw?si=8_U9Ve1HcjmmrLj)) (검색일 : 2023.10.01.)

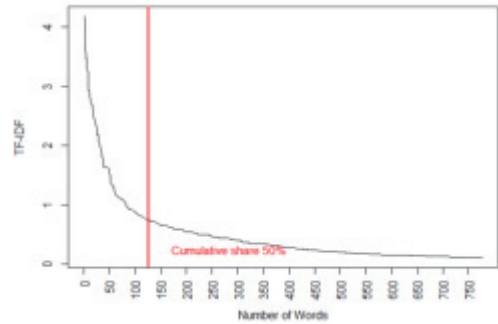
10) 전자신문, “네이버, 클로바노트 출시 3년 만에 누적 노트 3000만개 넘어” (<https://www.etnews.com/20230922000166>) (검색일 : 2023.10.01.)

11) TF-IDF(Term Frequency -Inverse Document Frequency) : 정보 검색과 텍스트 마이닝에서 이용하는 가중치로, 여러 문서로 이루어진 문서군이 있을 때 어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타내는 통계적 수치 (<https://ko.wikipedia.org/wiki/Tf-idf>) (검색일 : 2023.10.01.)

해서 핵심단어를 선정할 때보다 소수문서에 집중적으로 출현하는 단어도 중요한 단어로 선정해주는 이점이 있기 때문에 좀 더 핵심적인 단어를 선정할 수 있도록 해준다. 즉, Term - Document Matrix(이하 TDM)에서 TF-IDF누적빈도 50%를 이용하여 핵심단어를 선정하였다.

이후 시사점 도출단계에서 핵심단어분석과 네트워크 분석, 토픽모델링, t-SNE를 활용하여 군집분석을 실시하였다. 워드클라우드를 토른 핵심단어 시각화, 단어분포를 확인하고 네트워크 분석에서는 단어 간에 숨겨진 패턴, 관계를 확인하였다. 토픽 분석간에는 작전명령 내에 있는 주요 토픽을 탐색하였다.

126개 단어가 전체 단어 발생비율 중 약 56.2%를 차지하였다. 발생빈도가 높은 126개 단어의중 상위 30개 단어는 <표 4-1>과 같다.



<그림 4-1> TF-IDF 누적 50% 단어 선정 결과

## 4. 연구 결과

### 4.1 핵심단어 분석 결과

총 776개의 전체 단어중에 TF-IDF 누적비율이 50%를 차지하는 단어 126개를 <그림 4-1>과 같이 추출하였으며 발생빈도가 높은

핵심단어는 다음 적용할 텍스트마이닝 기법들에 중요한 기준이 되며 여기서 선정된 126개의 단어를 토대로 분석이 이루어진다. 핵심단어를 Word cloud한 결과 <그림 4-2>과 같으며 ‘미국’, ‘중국’, ‘지역’, ‘생각’ 등의 단어가 부각되는 것을 볼 수 있는데 빈도수 만큼 크게 표현되기 때문에 직관적으로 이해가 가능하다. 이를 통해 복잡한 분석 결과를 누구나 쉽게 이

<표 4-1> 핵심단어 Top 30 Words

순위	핵심단어	빈도	순위	핵심단어	빈도
1	미국	93	16	측면	24
2	중국	86	17	문제	21
3	지역	57	18	차원	16
4	생각	46	19	한미일	16
5	미중	46	20	행정부	15
6	국가	43	21	세계	14
7	경쟁	41	22	동맹	14
8	안보	34	23	북한	13
9	전략	33	24	강화	12
10	경제	33	25	교수	11
11	인도	29	26	정세	11
12	상황	29	27	제재	11
13	태평양	28	28	질서	11
14	입장	28	29	심화	10
15	협력	24	30	군사력	10

해할 수 있는 형태로 단순화시켜주어 커뮤니케이션 뿐만 아니라 토론 중요 개념과 대상이 무엇인지 이해할 수 있는 용이한 Tool로써 활용 가능하다.



<그림 4-2> 핵심단어 Word cloud 결과

<표 4-2> 상관계수 분석 결과

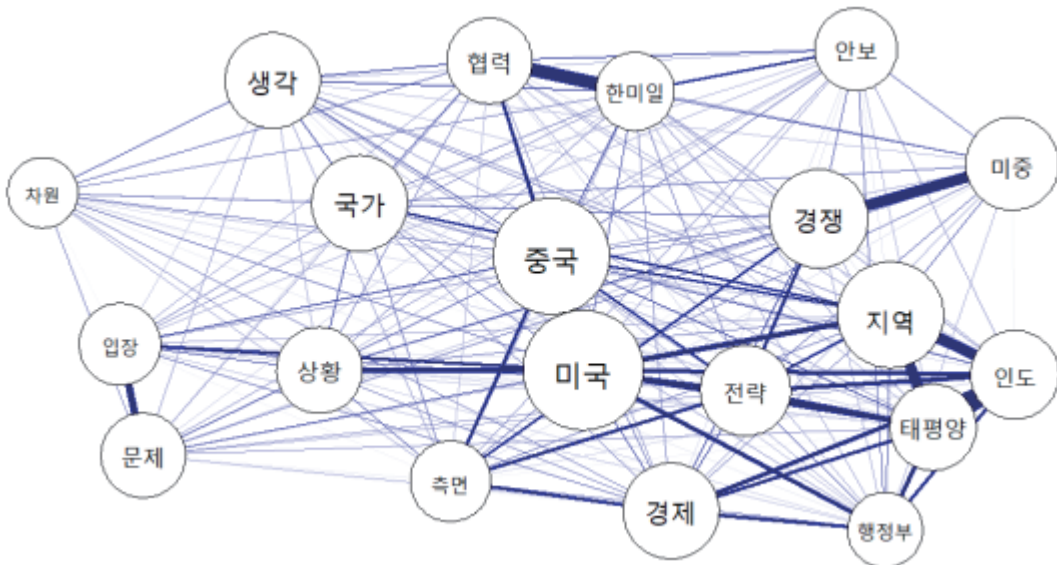
단어1	단어2	상관계수
인도	태평양	0.9976
협력	한미일	0.9158
지역	태평양	0.9114
지역	인도	0.9055
미중	경쟁	0.8816
미국	상황	0.7892
미국	중국	0.7777
미국	지역	0.7538
입장	문제	0.7506
미국	행정부	0.7423

<표 4-2>를 보면 ‘인도-태평양’의 관계가 가장 밀접하게 함께 언급되었음을 알 수 있고 ‘협력-한미일’ 단어도 상관계수가 0.9158로 밀접한 것으로 보아 한미일간의 협력적인 관계가 중요함을 알 수 있었다. 또한 ‘미중-경쟁’의 상관계수가 0.8816로 높았는데 이는 현재 경제적, 안보적 대립관계에 놓여있는 미중갈등 상황을 보여주는 것이라 유추해볼 수 있다.

이처럼 네트워크 분석을 통해 단어간의 상관관계로 토론에서 주요하게 논의되었던 내용이

## 4.2 네트워크 분석

단어와 단어간 상관관계를 분석하는 네트워크 분석 결과는 <그림 4-3>과 같다. 네트워크는 단어와 단어의 연결을 통해 의미 있는 문구를 구축하고 단어와 단어의 연결정도는 상관계수를 통해서 결정한다.



<그림 4-3> 네트워크 분석결과 그래프

무엇인지, 왜 그런 말이 언급되었는지를 비교 해서 중요한 시사점을 얻을 수 있었다.

토픽수를 최적의 토픽수로 선정하였다.

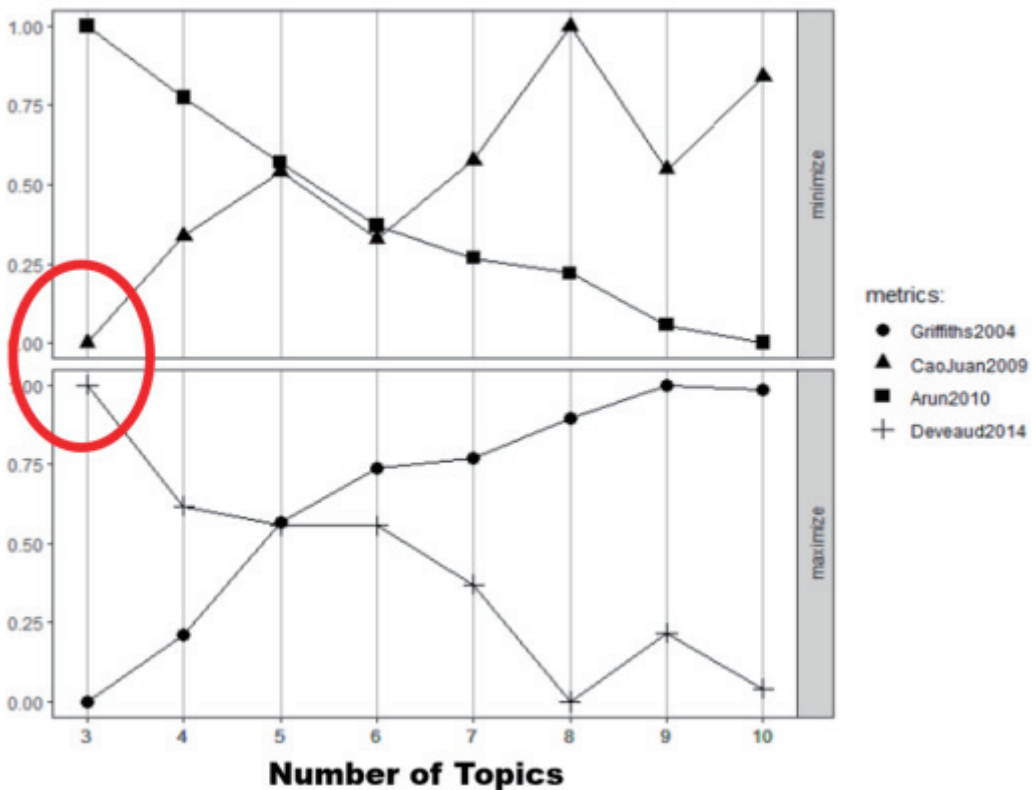
### 4.3 토픽 모델링 결과

LDA를 활용하여 회의간 토픽(주제)를 도출함으로써 회의에서 중요하게 다루어졌던 내용이 무엇인지 파악할 수 있다. 최적의 토픽수를 구하기 위해 4개의 방법을 적용하였다. Griffiths2004와 Deveaud2014는 클수록 좋고, CaoJuan2009와 Arun2010은 작을수록 좋은 성능 결과를 보여주며 이중 다수결로 최적의 토픽수를 선정하고 동일한 빈도수일 경우 가장 작은 값을 최적의 토픽 수로 선정하였다.[3] 이번 연구간 <그림 4-4> 그래프를 보면 토픽 3개가 2표, 토픽 9개가 1표, 토픽 10개가 1표를 받은 것을 알 수 있는데 가장 많은 득표를 한 3개의

<표 4-3> LDA 토픽모델링 결과

토픽	상위 10개 단어
토픽1 (태평양 동맹간 전략적 상황)	미국, 중국, 전략, 지역, 인도, 태평양, 동맹, 행정부, 생각, 상황
토픽2 (미-중 경제 갈등)	중국, 생각, 국가, 미중, 미국, 경쟁, 문제, 경제, 인도, 입장
토픽3 (지역적 안보 경쟁)	지역, 미국, 안보, 경쟁, 측면, 상황, 협력, 미중, 경제, 차원

‘안보정세 전망 미-중 갈등 그리고 아시아’ 토론에서 중요한 토픽들을 정리해보면 <표 4-3>와 같이 3개로, 토픽1은 미-중-인도-태



<그림 4-4> 최적의 토픽수 선정 결과

평양 동맹간 전략적 상황을 고려하여 ‘태평양 동맹간 전략적 상황’, 토픽2는 토픽2는 중국-미국 경제 경제적 경쟁 문제에 대한 내용으로 보아 ‘미-중 경제 갈등’, 토픽3은 토픽3은 지역적 안보 경쟁적인 상황에 대한 내용으로 보아 ‘주변국간의 안보 경쟁’이 논의되었음을 알 수 있었다. 실제로 토론간 해당 주제에 대한 논의가 이루어졌으며 주관적인 사람의 개입 없이도 과학적인 기법으로 주요주제를 객관적으로 도출해낼 수 있었다.

이처럼 토픽의 분석이나 개별 토픽을 단시간 내 분석하여 주제에 대한 분석이 가능하였다. 추후 많은 데이터가 학습 된다면 참모AI에 의해 특정 상황별 토픽 추천으로 의사결정을 지원할 것이고 작전반응시간에 대한 단축이 가능할 것이다.

#### 4.4 지휘관 취임사 자료 분석 결과

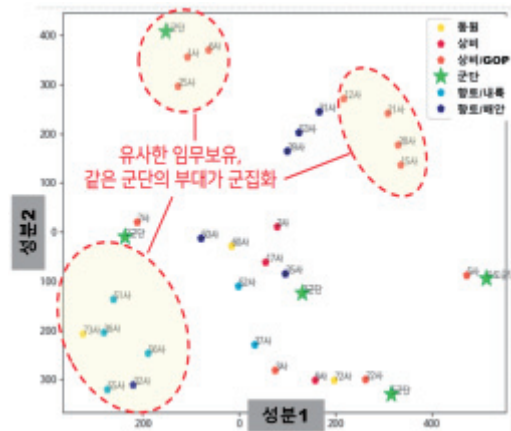
t-SNE는 차원이 높은 공간에서의 데이터간의 거리정보를 최대한으로 보존하면서 차원을 축소시켜 저차원에 타점시키는 비선형적인 방법이다.[4] 군단장, 사단장 취임사 핵심단어 사전 구축 결과 Word cloud는 <그림 4-5>와 같다. 사단장 취임사가 주로 다뤄졌기 때문에 ‘사단’이 시각적으로 가장 부각되었고 Word cloud를 통해 지휘관 취임간 주요 관심사항으



<그림 4-5> 취임사 데이터 Word cloud 결과

로 ‘장병’, ‘전우’, ‘임무완수’, ‘훈련’등이 언급되었음을 알 수 있었다.

취임사데이터에서는 감정이 대화에 비해 상대적으로 적은 문서 데이터로 제한적 해석이지만 군집분석 결과 같은 그룹내에 위치한다면 유사한 성향으로 분류할 수 있다. <그림 4-6>의 결과를 보면 유사한 임무를 보유하거나 같은 군단의 부대가 군집화 된 결과를 볼 수 있는데 이러한 군집분석 결과를 활용한다면 유사한 성향의 지휘관은 어느 부대인지 분석해서 정보를 제공할 수 있다. 뿐만 아니라 전투협조회의 간 작전명령지와 회의내용간의 유사도를 분석함으로써 명령하달의 일관성 측면에서 작전명령지가 전투협조회의 내용을 충분히 반영하고 있는지 비교분석할 수 있을 것이다.



<그림 4-6> 취임사 데이터 군집분석 결과

### 5. 결론

본 연구에서는 우리 군에서 중요하게 다뤄야 할 전투지휘정보 중 전투협조회의내용을 정형화하고 분석하는 방법론에 대해 제시하였다.

본 연구의 결과는 지휘관과 참모에게 유용한 정보를 다음 3가지 측면에서 기여할 수 있다. 첫째, 전투협조회의에서는 주제 도출을 통해

회의결과 요약 및 주요 개념 도출을 과학적인 기법으로 객관적으로 단시간내 도출해낼 수 있었다. 둘째, 지휘관을 포함한 개인에 대한 분석 정보를 제공함으로써 커뮤니케이션 향상에 활용할 수 있을 것이다. 셋째, 지휘관 지시사항에서 시간대별 지시사항 목록을 작성하여 시계열적으로 분석하고 상황별 지시사항을 설명변수, 전투결과를 반응변수로 상황별 어떤 의사결정이 특정 결과를 주는지에 대한 분석정보를 제공할 것이다.

본 연구의 한계점으로 실제 전투협조회의 자료는 보안상의 이유로 녹음되고 축적되지 않고 있어 다루지 못하였지만, 데이터가 없다고 해서 중요한 분야에 대해 연구하지 않을 수 없었기에 그와 유사한 패턴을 가지는 국방TV 토론 자료와 지휘관 취임사를 통해서 분석하였다.

추후 연구 방향으로 KCTC(Korea Combat Training Center) 결과와 같은 실제와 유사한 전투환경의 데이터를 반응변수로 연계하여 참모 AI 학습에 활용한다면 지휘관 의사결정 지원에 활용할 수 있을 것이다. 또한 전투지휘 데이터 저장체계 확립이 병행되어 전투지휘 음성데이터를 자동저장하는 ‘데이터 저장소’를 구축함으로써 활용되지 못하고 버려지는 데이터가 없도록 함으로써 축적된 데이터를 활용하여 많은 양의 정보를 단시간내 처리하고 적시 적절한 판단을 지원하는 결심지원 참모로써 기능할 수 있는 모델을 구축할 수 있을 것이다.

## 참 고 문 헌

- [1] Hoehn, John R. "Joint all-domain command and control: Background and issues for Congress." Congressional Research Service 46725 (2021): 18.
- [2] 육군본부. 『야전교범 1-1 군사용어사전』, (2019).
- [3] Griffiths, Thomas L., and Mark Steyvers. "Finding scientific topics." Proceedings of the National academy of Sciences 101.suppl\_1 (2004): 5228-5235.
- [4] 김기풍, 김충락, "t-SNE에 대한 요약", 2023, 응용통계연구
- [5] Tan, Ah-Hwee. "Text mining: The state of the art and the challenges." Proceedings of the pakdd 1999 workshop on knowledge discovery from advanced databases. Vol. 8. 1999.
- [6] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." Journal of machine Learning research 3.Jan (2003): 993-1022.
- [7] Arun, Rajkumar, et al. "On finding the natural number of topics with latent dirichlet allocation: Some observations." Advances in Knowledge Discovery and Data Mining: 14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010. Proceedings. Part I 14. Springer Berlin Heidelberg, 2010.
- [8] 감미아, and 송민. "텍스트마이닝을 활용한 신문사에 따른 내용 및 논조 차이점 분석." 지능정보연구 18.3 (2012): 53-77.
- [9] Le, Quoc, and Tomas Mikolov. "Distributed representations of sentences and documents." International conference on machine learning. PMLR, 2014.
- [10] 이창용, 문호석. "텍스트마이닝을 이용한 북한 보도동향과 북한 도발과의 연관성 분석." 국방연구 59.4 (2016): 103-124.

- [11] 이창용, 진효원, 문호석. "텍스트마이닝을 이용한 북한 동향 핵심단어 사전 구축 및 분석." 한국경영과학회 학술대회논문집 (2016): 4817-4822.
- [12] Choi, Ki-won, et al. "Hangeul Stem Extraction Algorithm for Text Mining Based on Natural Language Processing." Proceedings of the Korean Institute of Information and Commucation Sciences Conference. The Korea Institute of Information and Commucation Engineering, 2017.
- [13] 방한술, 문호석. "텍스트마이닝을 이용하여 텍스트의 주요 토픽을 시계열적으로 표현하는 방법론 연구." 한국데이터정보과학회지 30.6 (2019): 1259-1276.
- [14] 홍세은, et al. "112 신고 백터화를 통한 텍스트마이닝 기반의 군집분석과 이를 활용한 사건유형 간 유사성 탐색." 경찰학연구 20.3 (2020): 63-86.
- [15] 권연하, 박세진, 이현숙. "텍스트마이닝 기법을 활용한 코로나 19 발생 이후 교육격차의 쟁점 분석." 학습자중심교과교육연구 21.6 (2021): 625-644.

## 저자 소개



**조성일 (E-mail: enjoy\_all@naver.com)**

2016 육군3사관학교 전자공학 학사  
 현재 국방대학교 군사운영분석전공 석사과정  
 관심분야 : 국방 빅데이터 텍스트마이닝,  
 통계적 머신러닝, 다변량 자료분석



**신용 (E-mail: yshin@kssi.co.kr)**

2003 전북대학교 재료공학 학사  
 현재 한국소프트스페이스(주) 연구소 부장  
 관심분야 : 국방 빅데이터 텍스트마이닝,  
 통계적 머신러닝, 다변량 자료분석



**문호석 (E-mail: hsmoon0329@kndu.ac.kr)**

1994 육군사관학교 화학과 학사  
 2003 고려대학교 전자공학 석사  
 2006 고려대학교 산업공학 박사  
 2010 고려대학교 통계학 박사  
 현재 국방대학교 국방과학학과 교수  
 관심분야 : 데이터과학, 국방 AI, M&S

