



Topics Analysis of Media Coverage Related to the Safety Sector of Namhae-gun

Seong-Bhin Joo[#]

Department of Police Administration, Dong-Eui University, 176 Eomgwangno, Busanjin-gu, Busan, Korea

ABSTRACT

This study is to draw objectively draw the functions for each topic through safety related new articles of Namhae-gun and present policies to create safe Namhae-gun. We examined the status and trends in the safety sector of Namhae-gun from the view point of our society and analyzed how the central topic has changed over the time. Safety among major study scopes was analyzed by dividing "crime", a major task of police organization and 'fire', a major task of fire station. Topic modeling that can analyze safety related contents in Namhae-gun reported as article during the study period was used as the study method and we identified the drawn topics to discuss the relevance of each period and policy issues. Based on the analysis results, the implications related to the safety sector in Namhae-gun are as follows. First, we should find active countermeasures for improving disordered environment such as understanding characteristics of crime occurrence occurred in the local community and having perception about the fact that the risk increases due to the hollowing out of the local community. Second, it is necessary to seek appropriate countermeasures that are suitable for the local environment such as understanding on location of fires considering the regional characteristics and managing vulnerable spaces in the public and private sectors.

Key words: topics analysis, media, safety, crime, fire, Namhae-gun

Introduction

This study aims at each topics analysis about safety related new articles of Namhae-gun and present policies to create safe Namhae-gun. We examined the status and trends in the safety sector of Namhae-gun from the view point of our society and analyzed how the central topic has changed over the time. Safety among major study scopes was analyzed by dividing "crime", a major task of police organization and "fire", a major task of fire station. Topic modeling that can analyze safety related contents

in Namhae-gun reported as article during the study period was used as the study method and we identified the drawn topics to discuss the relevance of each period and policy issues.

Recently, the study using big data on "crime" has been increasing. However, the scope of study is focused on the analysis of literature and empirical study is relatively insufficient. In addition, the study topic are mainly about personal information protection due to the introduction of big data policy, issue of basis right infringement and its policy suggestions. So the analysis using actual big data is not sufficient. Study with big data on "fire" is more insufficient. Most studies are focused on system building and development

⁺ The 1st and Corresponding Author: Seong-Bhin Joo, E-mail. tjdqlsw@deu.ac.kr

Received: Apr. 15, 2020 / Revised: Apr. 29, 2020 / Accepted: Apr. 29, 2020

© 2020 Crisis and Emergency Management: Theory and Praxis. All rights reserved.

<Table 1> Crime occurrence status

type	2009	2010	2011	2012	2013	2014	2015	2016	2017
homicide	1	2	1	1	2	0	3	1	1
robber	2	5	3	2	0	0	3	0	1
rape	15	10	12	6	18	12	9	14	13
theft	149	233	153	174	155	119	107	77	100
assault	210	214	197	169	140	174	150	134	116
Total	377	464	366	352	315	0	0	0	0

* Source: Nanhae Police Station(2020).

and evaluation of artificial neural network technology and its stability in terms of engineering. As such, there are not enough studies on how to prepare strategies for safety sector by using big data in the social science sector, and even there are some studies on this, those only introduced the usability of big data in preparing countermeasures.

Therefore, this study is different from existing studies in that it examines the direction of safety-related policies through the analysis of central topic related to safety sector, especially Namhae-gun, contained in the news articles.

Theoretical Background

Recently, the crimes and fires which regional characteristics are reflected has been frequently occurred in Namhae-gun.

Crime Status of Namhae-gun

Crime Occurrence Status

The status of crimes occurred in Namhae-gun between 2009 and 2017 is as follows.

Specifically, we can find that the number of occurrence has been increased or decreased continuously. That is, the number of crime occurrence has been decreased in 2012 and 2013 and part of crime occurrence has been increased, but maintained in 2014 and 2015.

Fire Occurrence Status

The status of fires occurred in Namhae-gun between 2009 and 2017 is as follows. Specifically, we can find that the number of fire occurrence has been increased or decreased continuously. That is, the number of fire occurrence has been decreased before and

<Table 2> Fire occurrence status

(unit: case, 1,000won)

Year	Occur	Damaged property	Number of casualties
2010	78	215,008	3
2011	103	471,921	3
2012	88	590,115	1
2013	102	421,475	4
2014	103	436,594	4
2015	116	664,167	3
2016	98	339,667	1

* Source: KOSIS(2020).

after 2011 and it has been continuously decreased since 2015.

Big Data and R

Concept

The term “big data” refers to data that is so large, fast or complex that it’s difficult or impossible to process using traditional methods. The act of accessing and storing large amounts of information for analytics has been around a long time. But the concept of big data gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three V’s(SAS Institute Inc., 2020. 04. 29.).¹⁾

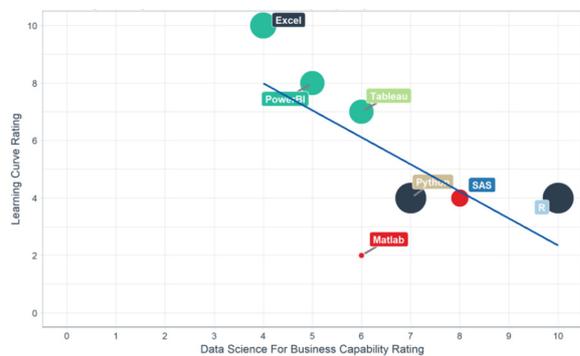
R is a programming language developed by Ross Ihaka and

- 1) ① Volume: Organizations collect data from a variety of sources, including business transactions, smart (IoT) devices, industrial equipment, videos, social media and more. In the past, storing it would have been a problem. but cheaper storage on platforms like data lakes and Hadoop have eased the burden.
- ② Velocity: With the growth in the Internet of Things, data streams in to businesses at an unprecedented speed and must be handled in a timely manner. RFID tags, sensors and smart meters are driving the need to deal with these torrents of data in near-real time.
- ③ Variety: Data comes in all types of formats - from structured, numeric data in traditional databases to unstructured text documents, emails, videos, audios, stock ticker data and financial transactions

<Table 3> Concepts of Big Data

	References	Definition
Foreign	IDC (2011)	A new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis.
	Mckinsey (2013)	Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.
	Gartner (2012)	High-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.
Domestic	KRIHS (2013)	A new technology to store, manage, and process exponentially increasing information caused from various information production and development of information sharing technology.
	KIPA (2013)	All kinds of structured and unstructured datasets for forecasting, risk management, and productivity enhancement.

* Source: Lee, et. al.(2014: 15).



<Figure 1> DS4B tools: capability vs learning curve

Robert Gentleman in 1993. R possesses an extensive catalog of statistical and graphical methods. It includes machine learning algorithms, linear regression, time series, statistical inference to name a few. Most of the R libraries are written in R, but for heavy computational tasks, C, C++ and Fortran codes are preferred. R is not only entrusted by academic, but many large companies also use R programming language, including Uber, Google, Airbnb, Facebook and so on. R has a longer learning curve but had a massive business capability rating (Guru99, 2020. 04. 29.).

Use of Big Data in the Safety Field

Recently, the study using big data on 'crime' has been increasing. However, the scope of study is focused on the analysis of literature and empirical study is relatively insufficient. In addition, the study topic are mainly about personal information protection due to the introduction of big data policy, issue of basis right infringement and its policy suggestions. So the analysis using actual big data is not sufficient. Study with big data on 'fire' is more insufficient. Most studies are focused on system building and development and evaluation of artificial neural network technology and its stability in terms of engineering. As such, there are not enough studies on how to prepare strategies for safety sector by using big data

in the social science sector, and even there are some studies on this, those only introduced the usability of big data in preparing countermeasures.

Therefore, this study is different from existing studies in that it examines the direction of safety-related policies through the analysis of central topic related to safety sector, especially Namhae-gun, containe in the news articles.

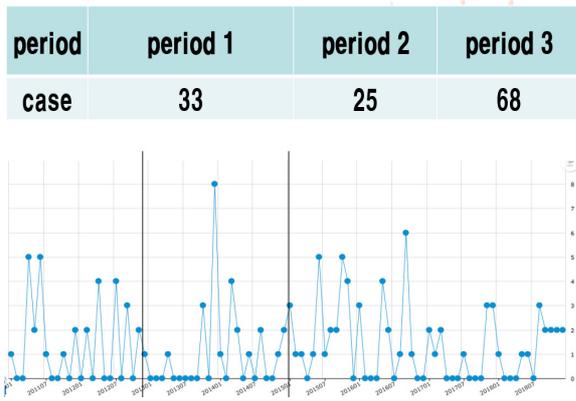
Research Design

Collecting Data

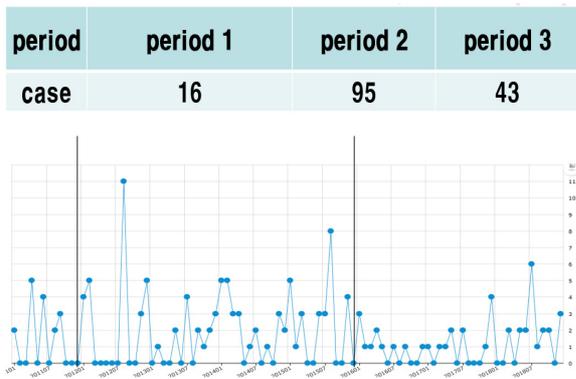
This study collected news articles related to safety in Namhae-gun through package R. As mentioned earlier, the collection period of the study subject was mainly divided into 'crime' and 'fire'. Specifically, the 'crime' was divided into period 1 with 2011-2012, period 2 with 2013-2014 and period 3 with 2015-2018 when the number of crime occurrence was increased or decreased based on the crime statistics in Namhae-gun.

'fire' was divided into period 1 with 2011, period 2 with 2012-2015 and period 3 with 2016-2018 when then number of fire occurrence was increased or decreased.

The data was collected by partly crawling the articles in NAVER, a portal site and the most of articles was collected through 'BIG KINDS.' This was to supplement the limitations of many articles which are extracted by providing only by some media since with limited access. In terms of collection unit, the articles contained keywords such as 'Namhae-gun', 'crime' and 'Namhae-gun', 'fire' at the same time in the text and title for the mentioned periods were extracted.



<Figure 2> Crime period



<Figure 3> Fire period

<Table 4> Collecting data Results

Type	Media
Metropolitan newspaper	Kyunghyang Shinmun, KukminIlbo, Donga Ilbo, Munhwa Ilbo, Seoul Shinmun, Swgye Ilbo, Chosun Ilbo, Hankyoreh Shinmun, HankookIlbo
Economic Newspaper	Maeil Business Newspaper, Seoul Economic Daily, Money Today, Asia Business Daily, AjuNews, Financial News, Korea Economic Daily, Herald Economy
Broadcasting station	KBS, MBC, SBS, YTN, JTBC
Special paper	Journalist association of Korea, Dongascience, Womennews, Digitaltimes, Electronic Newspaper, Korea Herald

Research Design

In addition, the refining process for words were went through. First, the list of words was organized in order of frequency and then incorrect and unnecessary words which are not related to study topic were deleted to arrange the words to a level suitable for the final analysis. This was also a way to increase the validity of the study results. For example, the postposition or bound nouns (ex. to, for, is, are, etc.) were processed as stopword.

```
setwd("C:/Users/Administrator/Desktop/R")
f <- file("ji.txt")
fl <- readLines(f)
close(f)

head(fl)
usesejongpic()

nouns <- sapply(fl, extractNoun, USE.NAMES=F)
nouns <- nouns[nchar(nouns)>=2]
head(nouns)

K<-10
set.seed(1004)
num.iterations <- 1000

corpus <- lexicalize(nouns, lower=TRUE)
corpus$vocab<-gsub("\\(", "", corpus$vocab)
corpus$vocab<-gsub("\\)", "", corpus$vocab)
corpus$vocab<-gsub("\\.", "", corpus$vocab)
corpus$vocab<-gsub("\\-", "", corpus$vocab)
corpus$vocab<-gsub("\\_", "", corpus$vocab)

result <- lda.collapsed.gibbs.sampler(corpus$documents, K, corpus$vocab,
num.iterations, 0.1, 0.1, compute.log.likelihood = TRUE)

top.words <- top.topic.words(result$topics, 500, by.score = TRUE)
print(top.words)

write.csv(top.words,file="LDA.csv")
```

<Figure 4> R codes sample

In addition, the computed words were examined to unify two or more words with the same meaning or words that are expressed in different ways (ex. crime rate, crime) by considering spacing and abbreviation. Finally, this study was analyzed for a total of 20,057 morphemes (nouns, adjectives, adverbs and verbs).

There are several approaches in algorithm that implements the topic modeling of big data, but this study applied 'Latent Dirichlet Allocation (hereinafter referred to as 'LDA')', a statistical model of language.

Result

Top Advent Words in related Articles

Crime (period 1, period 2, period 3)

The computed words were examined to unify two or more words with the same meaning or words that are expressed in different ways (ex. crime rate, crime, etc.) by considering spacing and abbreviation. The results by period are as follows.

Period 1 appear frequently that 'Changwon Prosecutor's Office Jinju Branch Office', 'Indict without detention', 'Bribery', 'Farming union delegate', 'Sexual assault' etc. Period 2 appear frequently that 'Tongyeong', 'Gangnam-dong, Jinju city', 'Reformatory', 'Chairman', 'CCTV Integrated Control Center' etc. Period 3 appear frequently that 'Natural disasters', 'Infectious diseases', 'Domestic', 'Gimhae City', 'Traffic Accident' etc.

<Table 5> Results by period

	Period 1	Period 2	Period 3
1	Changwon Prosecutor's Office Jinju Branch Office	Tongyeong	Natural disasters
2	Indict without detention	Gangnam-dong, Jinju city	Infectious diseases
3	Bribery	Reformatory	Domestic
4	Farming union delegate	Chairman	Gimhae City
5	Sexual assault	CCTV Integrated Control Center	Traffic Accident
6	Judicial action	Handicapped person	Life safety
7	Garlic harvest	injury and robber	Yangsan City
8	Subsidy 600 million	juvenile intermediate treatments	DNA
9	Subsidiary business	Police office	Official
10	County headman lady	Criminal punishment	Namhae Police Station

<Table 6> Results by period

	Period 1	Period 2	Period 3
1	Namhae Fire Station	Property damage	Natural disasters
2	Jo Jong Je	Fire administration	Life safety
3	Business agreement	Namhae Fire Station	Infectious diseases
4	Union president	Maritime pension	Fire extinguisher
5	Gwangyang Port	Casualties	Gyeongsangnam-do
6	Firefighters	Hadong-gun	Fire Administration
7	Agricultural accident insurance	Gyeongnam	Hadong-gun
8	Seolbong ferry	Hadong Hwagae Market	Property damage
9	Tourist bus	Pig barn	Traffic Accident
10	Supporting Organization	volunteer fire fighting team	Lowest grade

<Table 7> Topics analysis of related articles by period 1

	Political crime	Crime response	Fraud
W1	United States-Korea Free Trade Agreement	Police	Farming
W2	Gyeongnam Headquarters	Report Crime	Namhae
W3	Ye Sang Kyu	citizen	Incumbent official
W4	Aide	Commendation	Law
W5	Dismissal	Sexual Assault	Subsidy
W6	Hadong	Card loan	Prejudice
W7	Absentee voting	Loan	Misuse
W8	Intervention	Voice Phishing	Town hall

Fire (period 1, period 2, period 3)

The computed words were examined to unify two or more words with the same meaning or words that are expressed in different ways (ex. fire fire station, fires etc.) by considering spacing and abbreviation. Period 1 appear frequently that 'Namhae Fire Station', 'Jo Jong Je', 'Business agreement', 'Union president' etc. Period 2 appear frequently that 'Property damage', 'Fire administration', 'Namhae Fire Station', 'Maritime pension' etc. Period 3 appear frequently that 'Natural disasters', 'Life safety', 'Infectious diseases', 'Fire extinguisher' etc. The results by period are as follows.

Topics Analysis of related Articles by Period

Crime

Period 1 (2011-2012)

<Table 7> is a table listing the 4 topics and the top 8 words that correspond to individual topics extracted through LDA analysis. The last three selected topics were 'Political crime', 'Crime response', 'Fraud'. The results by period 1 are as follows.

<Table 8> Topics analysis of related articles by period 2

	Crime area	Juvenile crime	Crime prevention	Violent crime	Crime victim	Crime response
W1	Provincial	Unlicensed	Prevention policy	Imitation	Jailbird	Area
W2	Gyeongnam	Middle school student	Integrated Control Center	adolescent	Extortion	Automobile tax
W3	Area	Dismiss with a caution	CCTV	Jinju	Proportional representation	Order
W4	Namhae	Blank ammunition	Enforcement	Adult	Case	Midnight
W5	Human rights	Amerson pacific	Construction	criminal syndicate	Family	Perfect crime
W6	Keeper	Provincial	Car theft	Violence	Child	Neighbor
W7	event	Namhae	Violent crimes	Charge	Violence criminal	Kindliness
W8	Municipality	CCTV	Close	Police	Gang	Tongyong City
W9	Alected	Gallop	Guard	International	Tongyong City	Namhae
W10	Audit	Group	Public safety	Public safety	Crime rate	CCTV

<Table 9> Topics analysis of related articles by period 3

	Sexual slavery	Crime prevention	Crime victim	Spatial crime I	Spatial crime II	Response status
W1	Wednesday rally	Joint Enforcement	Project	Crime	Around forty	Crime
W2	Namhae	Namhae	Safety	Made in China	Vacant house	Natural disaster
W3	Renovation	Illegality	Children	Trepang	Elderly	Traffic accident
W4	Gyeongnam	Daepo car	Woman	Ancestors' Shrine	Dementia	Safety Index
W5	Dawn	County headman	Evaluation	Namhae	Disappearance	Disaster
W6	Convenience store	CCTV	Old woman	Municipality	Kyungnam	Kimhae
W7	Robber	Concentration crackdown	Sancheong	Park jong yeoul	Sex offense	Infectious diseases
W8	Yun Mi hyang	Amount in arrears	Elderly	Police	DNA	Life safety
W9	Human rights	Levy	Weapon	Embezzlement	Report	'Good' level
W10	municipality	Agencies&Organizations	Victim	Kyungnam bank	Countermeasure	Community

Period 2 (2013-2014)

<Table 8> is a table listing the 6 4 topics and the top 10 words that correspond to individual topics extracted through LDA analysis. The results by period 2 are as follows.

The last six selected topics were 'Crime area', 'Juvenile crime', 'Crime prevention', 'Violent crime', 'Crime victim', 'Crime response'.

Period 3 (2015-2018)

<Table 9> is a table listing the 7 4 topics and the top 10 words that correspond to individual topics extracted through LDA analysis. The last six selected topics were 'Sexual slavery', 'Crime prevention', 'Crime victim', 'Spatial crime I', 'Spatial crime II', 'Response status'. The results by period 3 are as follows.

Fire

Period 1 (2011)

<Table 10> is a table listing the 7 4 topics and the top 10

words that correspond to individual topics extracted through LDA analysis. The last six selected topics were 'Period', 'Countermeasure', 'Self - control safety & management', 'Country fire I', 'Country fire II', 'Country fire III'. The results by period 1 are as follows.

Period 2 (2012-2015)

<Table 11> is a table listing the 5 topics and the top 10 words that correspond to individual topics extracted through LDA analysis. The last five selected topics were 'Prevention activity', 'Victim psychology', 'Damage target', 'Crime hot spot', 'Countermeasure'. The results by period 2 are as follows.

Period 3 (2016-2018)

<Table 12> is a table listing the 6 topics and the top 10 words that correspond to individual topics extracted through LDA analysis. The last six selected topics were 'Community correspondence', 'Municipality correspondence', 'Correspondence condition', 'Linked accident', 'Self - control safety & management', 'Crime hot spot'. The results by period 3 are as follows.

<Table 10> Topics analysis of related articles by period 1

	Period	Countermeasure	Self - control safety & management	Country fire I	Country fire II	Country fire III
W1	House	Namhae Fire Station	Fire	nonghyup	Fire	Namhae
W2	Elderly	Fire Damage	Namhae	Bean	Yeosu	House
W3	Household	Business agreement	Woamn volunteer fire fighting team	Crop	Seolbong ferry	Pig
W4	Hope	Chuseok	Safety education	Accident insurance	Gwang yang port	Fire
W5	Chuseok	Holidays	Idong-myeon	Chuseok	man-made disaster	Chuseok
W6	Holidays	Fire	Pigsty	Traffic accident	Barn	Holiday
W7	Traffic accident	Yeosu	Kang Chan Kyu	Fire	Chuseok	Traffic accident
W8	Fire	Gwang yang port	Inauguration	Yeosu	Traffic accident	Occur
W9	Farming	Seolbong ferry	Yeosu	Holiday	Namhae	Yeosu
W10	occurrence	Network	Blind spot	Occur	Pigsty	Offing

<Table 11> Topics analysis of related articles by period 2

	Prevention activity	Victim psychology	Damage target	Crime hot spot	Countermeasure
W1	Area	Fire engine	Namhae	Incident	Namhae
W2	Affection	Disappearance	Fire	Marine Pension	Gyeong nam
W3	Culture project	Hapcheon	Safety	Busan	Fire
W4	Apple tree	Infectious disease	Single seniors	One-bedroom	Holiday
W5	Talent show	Lowest	Inspection	Crash	Hamyang
W6	Children's Day	Public facilities	Arson	Sseason	Gyeongsangnam-do Fire Department
W7	Daldongnae	Cooking	Woman	Vacation	119 Hope of House
W8	Residential environment	Drinking	Children	Barn	Return home
W9	Improvement	Anxiety	Transportation	Tearoom	Emergency service
W10	Tree planting	Fear	Lifesaving	Accident	Uiryeong

<Table 12> Topics analysis of related articles by period 3

	Community correspondence	Municipality correspondence	Correspondence condition	Linked accident	Self - control safety & management	Crime hot spot
W1	Fire	Province governor	Safety	Crash	Namhae	Safety
W2	Volunteer fire fighting team	The acting	Namhae	Truck	Fire extinguishing	Tong yeong
W3	Company	Local election	Natural disaster	Casualties	Welfare	House
W4	Product	Samdong-myeon	Facilities	Ladder truck	Fire	Storage
W5	Physiognomy	Community	Sancheong	Official	Voluntary	Vacant house
W6	Neighbor campaign.	Single seniors	Safety Grade	Evacuation	Community	Farm
W7	Donation	Mayor	Lowest	Overthrow	Consultative body	Market
W8	Casualties	Inauguration	Serious	Poisonous substance	Volunteer fire fighting team	Car
W9	Service	Spot inspection	Vulnerable groups	Leak	Happiness	Facilities
W10	moonsutemple	Recantation	Gimhae	Beach	Resident	Gyeongnam

Conclusion

Based on the analysis results, the implications related to the safety sector in Namhae-gun are as follows.

Main subject of this study are mainly two. The first was about major keywords of media reports related to safety of Namhae-gun to examine the major keywords of media reports related to crimes in Namhae-gun and major keywords of media reports related to fires in Namhae-gun. The second was about what changes were

made for major topics of each period in media reports related to safety of Namhae-gun. In terms of major topic of each period of media reports related to crimes in Namhae-gun, we should understand the characteristics of crime occurrence occurred in local community and seek active countermeasures for improving disordered environment by perceiving the increase of risk due to the hollowing out of local community.

Then, we should understand the location of fire should be discussed for the major topic of each period of media reports related to fires in Namhae-gun. We also seek appropriate countermeasures suitable for the regional environment such as strengthening the management and supervision of vulnerable space in public and private sector, improving environmental design and expanding organizations and manpower.

References

- Gartner. 2012. Gartner IT Glossary. Accessed Nov 15. <http://www.gartner.com/it-glossary/big-data>.
- Guru99. <https://www.guru99.com/>
- Korea Institute of Public Administration. 2013. Utilization of Big Data for Realization of Gov. 3.0. Seoul.
- Korea Research Institute for Human Settlement. 2013. Big Data Utilization for Monitoring Territorial Policy Responses and Predicting Policy Demand Anayang.
- KOSIS. <http://kosis.kr/>
- Nanhae Police Station. <https://www.gnpolice.go.kr/nh/>
- SAS Institute Inc. <https://www.sas.com/>
- Yu, Seon Cheol, Won Wook Choi, Dong Bin Shin, and Jong Wook Ahn. 2014. A Study on Concept and Services Framework of Geo-spatial Big Data. *Journal of Korea Spatial Information Society*. 22(6): 13-21.

Seong Bhin Joo (tjdqlswn@deu.ac.kr)

He received his Ph.D. from Dongguk University, Korea in 2015. He is a Professor of the Department of Police Administration at Dong-Eui University, in which he has taught since 2015. His interesting subject and area of research and education is public safety network, active policing, national security crime.