

일반논문 (Regular Paper)

방송공학회논문지 제22권 제4호, 2017년 7월 (JBE Vol. 22, No. 4, July 2017)

<https://doi.org/10.5909/JBE.2017.22.4.438>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

한국 사행산업 관련 뉴스의 빅데이터 분석을 통한 인식 연구

문혜정^{a)*}, 김성경^{b)}

Study on Perceptions through Big data Analysis on Gambling related News in Korea

HyeJung Moon^{a)*} and SungKyung Kim^{b)}

요약

이 연구는 사행산업의 분야인 복권, 체육진흥투표권, 경마, 카지노에 대해 언론에서는 어떻게 다루어지고 있는지를 1990년부터 2015년까지의 뉴스데이터를 의미연결망 분석을 통해 밝혀보고자 하는 연구이다. 기사의 빈도와 연결성을 프레이밍과 시민관심 정도로 재조명 하여 기사에 대한 언론보도자의 의도와 시민의 인식차이를 밝히고, 이를 통해 정책적 특성과 개혁과제를 탐색하였다. 분석결과 복권의 경우 당첨번호, 당첨금, 조작의혹 등 당첨에 대한 부분이 주제인 ‘사회문제’ 형태였으며, 체육진흥투표권의 경우에는 사업입찰, 불법사이트, 발매대상 등 주로 사업추진과 불법사이트에 대한 ‘의무정보’ 종류였고, 경마의 경우 사업장, 홍보, 기사 등으로 사업홍보나 광고 관련 뉴스이었고, 마지막으로 카지노의 경우에는 불법, 도박장, 외국인 등 ‘주요정보’에 해당하는 논문이었다. 시대에 따라 1990년대에는 카지노, 2000년대에는 복권, 2010년대에는 경마에 대한 기사보도가 많아졌으며, 이에 대한 시민의 반응도 사업비리, 당첨, 시민운동 등의 차이가 있었다. 마지막으로 기사의 빈도와 연결성이 나타내는 프레이밍 정도와 시민의 관심은 ‘1. 홍보광고(경마), 2. 의무정보(스포츠베팅), 3. 사회이슈(복권), 4. 주요정보(카지노)’ 네 가지로 구분되었으며 이 중 사고, 비리 등 주요기사로 구분되는 사회문제가 주요 공공의제로 형성되는 것을 확인할 수 있었다.

Abstract

The purpose of this study is to understand the recognition of gambling industry through the semantic analysis of news data on lottery, sports betting, horse racing and casino that was reported between 1990 to 2015 in South Korea. This paper revealed the difference between journalists' intention and public's perception about news by analyzing the frequency and connectivity of news with framing and public's interest through semantic network analysis and explored the policy characteristics and innovation task. The result of analysis, news on lottery game mainly has been reported social issue related with win such as 'winning number', 'prize money', 'suspicion of manipulation' and etc. News on sports betting has been reported mandatory information related with business project and illegal site such as 'bidding', 'illegal site', 'sales target' and etc. News about horse racing has been reported the information about the business advertisement such as 'online race track' and 'promotion'. Lastly, casino related news has been reported 'major information' such as illegality, 'gambling place' and 'foreigner'. As a result of times series analysis, news about casino in the 1990s, news about lottery in the 2000s and news about horse racing in 2010s have been increased. Public's interest also has been moved to 'business scandal', 'winning game', 'citizens' campaign' and etc. Gambling related news has been classified by four types, 1. advertising publicity(horse racing), 2. mandatory information(sports betting), 3. social issue(public agenda, lottery), 4. major information(casino). We could get the insight that news can be formed a public agenda, when news is reported as a social issue with high frequency and public's interest like lottery related news.

Keyword : gambling, lottery, news, public policy, big data analysis, frame theory

Copyright © 2017 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I . Introduction

According to the National Gambling Control Commission, the sales volume of gambling industry in South Korea has been almost doubled to 20.5 billion USD(When we calculate that the exchange rate is 1 US\$ equals 1,000 KRW.) in 2016 compare to 12.5 billion USD in 2006^[1]. It can be understand that Koreans interest on gambling and lottery has been increased last decade. In the case of Lotto 6/45, the sales volume has been increased 10 times since the first issuance on December, 2002. The news coverage of major media on Lotto 6/45, sports betting, casino and horse racing in gambling industry has been increased during same period. It could be understood that the interest of public and major media have been grown up during these period. News is a representative object for understanding public perception^{[2][3]}. Understanding of popular culture is especially fundamental to analysis historical policy^[4].

Schneider and Ingram suggested that policies could change depending on how the social perceptions of public interest on specific social issues are shaped^{[5][6][7][8]}. In particular, Codd, et al., and May said that agenda setting for policy decisions shows different type according to initiators of debate and public involvement on social issues^{[9][10]}. Types of the agenda setting on Lottery policy also has been

shown similar situation in South Korea^[11]. Based on these theoretical background, the purpose of this paper is studying public recognition and policy making on gambling industry using news data of Mediagaon¹⁾ from 1990 to 2015 through the semantic network analysis. What is the difference of media's expression on various different types of gambling? How has public view on gambling industry changed based on the increase in media's interest? For solving these questions, we discussed on the perception on gambling industry through semantic network analysis using the frequency of news reporting and the degree of public's reaction(public's interest).

II . Previous Studies and Theoretical Background

Study on the recognition or tendency through semantic network analysis has been increased in the field of journalism. In case study using semantic network analysis on social network service, there was a study done on tourist's using Social Network Service data^[12] to see how people view on the topic of reunification of North and South Korea^{[13][14]} and a study on content analysis about Literary therapy using text mining and opinion mining^[15]. There was a study on the frequency and text network analysis of the papers in library information science during 7 years since 2004^[16]. The study on perception through news analysis was a study on the changing perceptions of tourism in Korea through newspaper analysis [17]. There were also three papers on differences of contents and tones of arguments among newspapers using text mining analysis^[16], frame transition of personal information leakage from 1984 to 2014 using social network analysis approach^[18], and a semantic network analysis of the news-

a) 서울과학기술대학교(Seoul University of Science & Technology), 복권학회(Cooperative Lottery Institute), 아이엘피(Institute of Lottery Policy), 윌비솔루션(Will-Be Solution)

b) 크로센트(Crossent), 서울과학기술대학교 IT정책전문대학원 스포츠문화 복권정책학과(Department of Sport, Culture & Lottery Policy, Graduate School of IT Policy Seoul University of Science & Technology)

✉ Corresponding Author : HyeJung Moon

E-mail: hyejung.moon@gmail.com

Tel: +82-10-3262-9936

ORCID: <http://orcid.org/0000-0001-8265-3256>

※ 이 논문 또는 저서는 2015년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-과제번호)(NRF-2015S1A5B4A01036743).

· Manuscript received March 31, 2017; Revised May 8, 2017; Accepted May 8, 2017.

1) The Mediagaon(www.mediagaon.or.kr) was the company which has the news database of 61 media companies such as national daily, weekly newspaper, internet news, Television news, etc under Korea Press Foundation. The name of Mediagaon was changed to bigkinds(www.bigkinds.or.kr) at April 2016.

paper articles on big data^[19].

As we can see the review of precedent papers, semantic network analysis has been conducted in various areas. Semantic network analysis using newspaper articles frequently has been used to understand the framing of mass media on special subjects. Semantic network analysis is the method for applying social network analysis to the message on communication^[20]. Social network analysis is the method to unearth the relation between actors who has the force to derive social situation^[21]. By the same principle, we would like to understand the implications of the documents to be analyzed by understanding the structural relationship between the words that are the components of the message through semantic analysis. We also try to find cognition about gambling industry using these network analysis and semantic analysis. We also try to understand cognition on gambling industry using these network analysis and semantic analysis.

III. Research Design

Major research questions to be confirmed through big data analysis of news related to gambling are as follows.

- How is the main content of the news related to gambling industry differentiated by the fields of gambling industry?
- What are the characteristics of framing(frequency, intention) and network(connectivity, centrality) according to the topic of the news on gambling industry?
- What are the characteristics of framing and network of news when the contents of news related to gambling industry has been formed as public agenda?

The content that is reported on the mass media, especially news, depends very much on the framing that the media creates. However news is re-reported, deepened and spread according to public reaction to news. In other words, news is initially framed by a small number of major media companies, but public reaction makes news solid and spread. This has a difference in centrality and connectivity in the social network theory. With this premise, we tried to identify public perception through news related to gambling industry.

The subject of analysis is 33,749,890 news items from 1990 to 2015 searched by keyword related to gambling industry from database of Mediagaon that accumulated articles of 61 media companies for 35 years. Once searching and selecting based on major game type in Korean gambling industry, the title of news retrieved by lottery(18,963), sports betting(2,277), horse racing(10,720) and casino(11,189) were analyzed after data cleansing except news unrelated to gambling industry. Looking at the number of articles searched by the type of gambling industry by period, lotto and sports betting had few or no article, and even in the case of horse racing and casino, there were not much number of articles in the 1990s. So the news before 1990s was excluded from the analysis. The analysis procedure is as follows:

1. Morpheme analysis of the collected news using our own Java²⁾ program.
2. Obtaining the co-occurrence matrix of words for each article to visualize the refined words (workbook³⁾ making for network analysis) using our own PHP(Hypertext Preprocessor)⁴⁾ program based on web server system.
3. Network analysis to derive between centrality from

2) Java is a general-purpose computer programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible.

3) workbook is the matrix of all the combinations of keywords for semantic analysis.

4) PHP is a server-side scripting language designed primarily for web development but also used as a general-purpose programming language

connectivity using NodeXL⁵⁾.

4. Visualizing and performing semantic network analysis to identify cluster of words using NodeXL.

We excluded keywords which have the frequency of occurrence less than two in whole newspaper articles for visualization of more refined data. News were initially reported by a small number of major media companies. However public's reaction has affected how news became solid and spread. In other words, the frequency with which news is reported on a particular topic is largely determined by media's intention[22]. And the centrality that indicates how much news has been linked to other news indicates the degree to which the news has been re-quoted and spread by public since it was first reported. This means that existing news is spread and solid is related to the degree of public reaction to the news. In this study, we designed the analysis framework by reinterpreting the number of times being reported and the degree of connectivity of the news as the intended framing and naturally reacted public perception. The analysis framework of this study is shown in <Table 1>.

The following four categories can be classified by the correlation between degree of news coverage and spread of news. 1. Advertising publicity, 2. mandatory informa-

tion, 3. social issue, 4. major information. 1. Advertising publicity is likely to be the intended framing information that media company has frequently distributed news for the purpose of spreading the information but public's interest is low. 2. mandatory information that has to be compulsorily reported to a small number of people because media has no will to spread and public is less interested. 3. Social issue is likely to be popular information that news is continually spread because media is willing to spread information and public reaction is high. Lastly, 4. Major information is likely to be an important issue in society. This is because although media company has low intention to report the news, rather, public is more interested in it and even if it is reported less at the beginning, it is spread and cited in the end. In particular, the last 4. major information is likely to be a major reform subjects requiring change in society. Based on this premise, we will identify public perception and find major policy issue about the gambling industry through semantic analysis.

IV. Result of News Analysis

As described in the research design, the frequency and centrality for each subject indicate the degree of news re-

표 1. 분석틀

Table 1. Analysis Framework

<p style="text-align: center;">high ↑ media spread intention, framing characteristic (number of report, intention) ↓ low</p>	<p style="text-align: center;">spread intention ↓ low public's interest ↓ 1. advertising publicity</p>	<p style="text-align: center;">spread intention ↓ high public's interest ↓ 3. social issue</p>
	<p style="text-align: center;">calm intention ↓ low public's interest ↓ 2. mandatory information</p>	<p style="text-align: center;">calm intention ↓ high public's interest ↓ 4. major information</p>
<p style="text-align: center;">public's reaction/interest low ← news network (connectivity/centrality) → high</p>		

5) NodeXL is one of famous Open Source Software which provide network analysis include analysis on social network service.

port and public reaction. In other words, the high frequency of key word means there are many cases reported by the media company. First, we performed the semantic network analysis for each gambling industry. Second, we expressed the group in network graph using intention of the media company. Third, we expressed the link and the connectivity using the relation how news has been spread among other news.

There is a lot of news in the lottery field, and the spread of news is also high afterwards<Fig. 1>. The subjects

(number of reports) of news that have been much reported are related to the winning result such as win(5,039), number(3,851) and bonus(1,365). The news of highly spread subject(connection centrality) was mostly hopeful contents such as jackpot(28,512), great spot(6,871) and luck(3,637) and it was the field with the greatest number of reports. This implies that the intention of the media company to spread relevant information was high and the public has a great interest in the news. This 'social issue' news was formed as a 'public agenda' as follows. Korean govern-

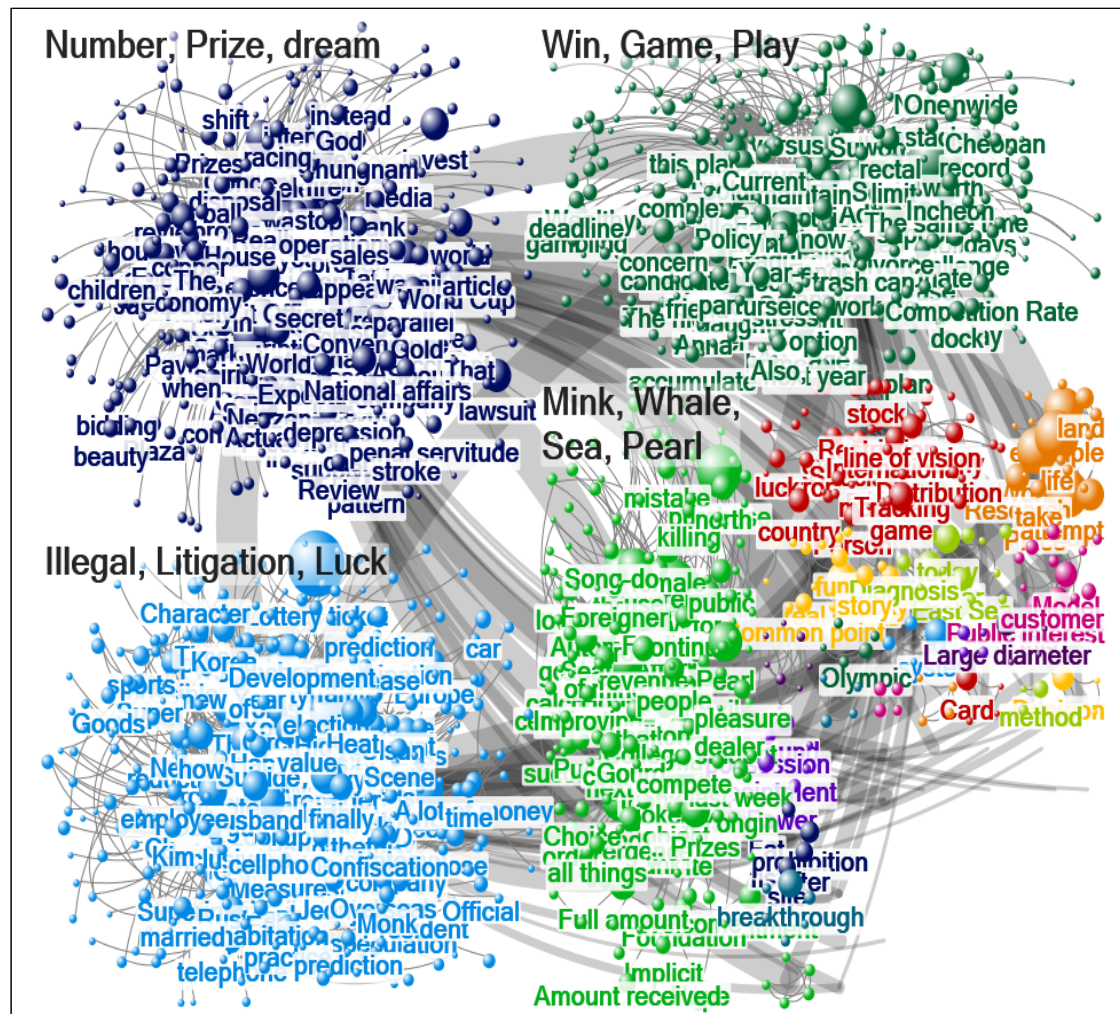
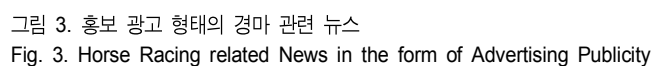


그림 1. 사회이슈 형태의 복권 관련 뉴스

Fig. 1. Lottery related News in the form of Social Issue

The number of news reported is high but the degree of spread is relatively low in the horse racing filed<Fig. 3>. The major subjects are related to the racetrack facilities such as online-racetrack(2,087), work(1,198), racing(1,035) and Busan(1,009, the provincial city which established new race track). The news of highly spread subject(connection



centrality) was related to specialty such as jockey(23,233), horse(9,744) and trainer(4,975). This is the intention of media to provide specific information such as publicity or advertisement to the public. However public reaction was not so high for such publicity information. So the news about horse racing stayed at the publicity level. This phenomenon means that horse racing is a special industry targeting a specific customer.

The number of news reported is relatively low but the spread of specific news is high about casino<Fig. 4>.

The number of reports is high in the order of gambling place (1,094), Kangwon Land(700, casino company) and foreigners(638). The news about Macau(28,935), local access(41,116) and local permission(15,075) was that connection centrality is high. In other words, media company wants to dismiss the news but became a social issue because public has high interest in the specific issue. This can be understood that a company operating a casino can control the media well but cannot control public's interest.

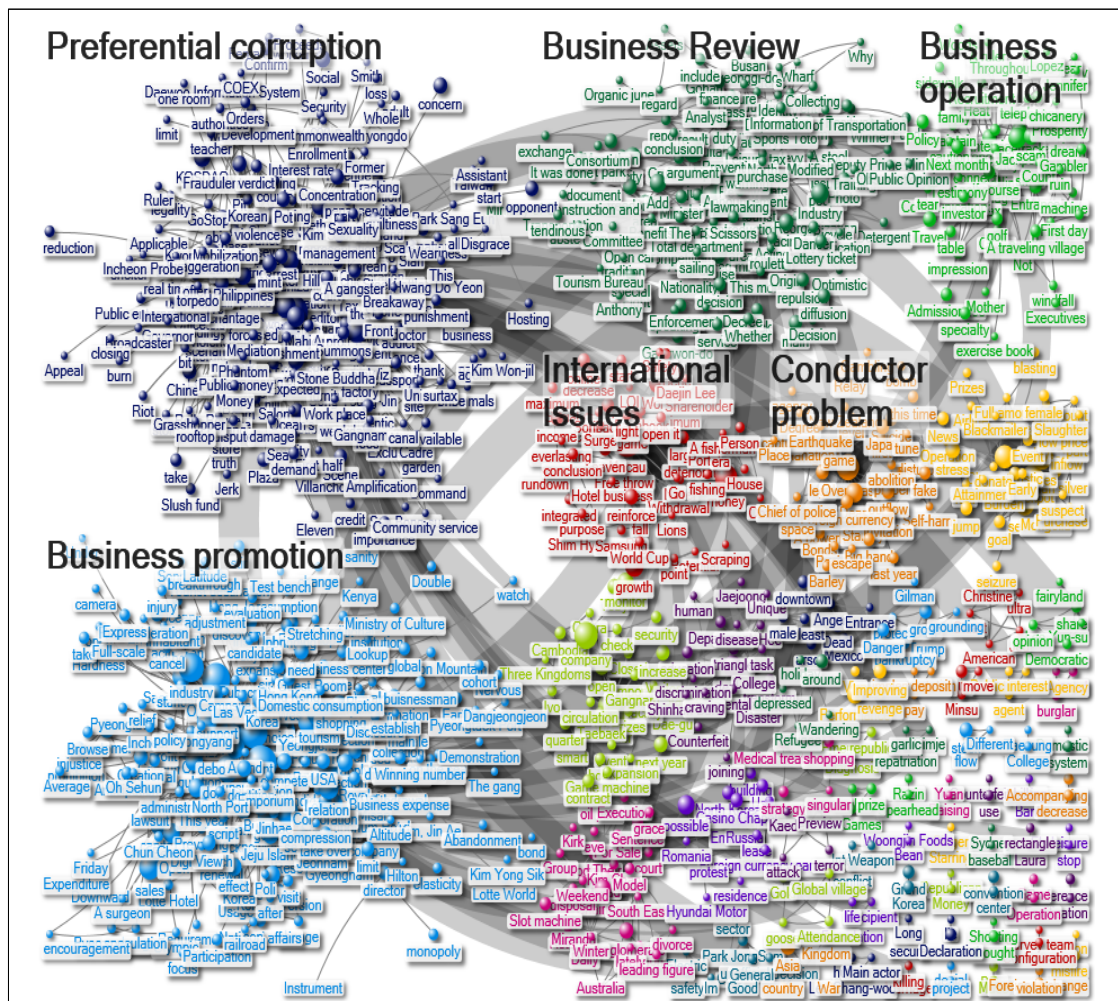


그림 4. 주요정보 형태의 카지노 관련 뉴스
Fig. 4. Casino related News in the form of Major Information

표 2. 사행산업별 뉴스 프레임링과 연결성

Table 2. News Framing and Connectivity by Gambling Industry

spread intention ↑ framing characteristic ↓ dismissal	1. advertising publicity i.e.) horse racing : racing, park	3. social issue → public agenda i.e.) lottery : win, jackpot
	2. mandatory information i.e.) sports betting : manipulation, Orion	4. major information i.e.) casino : gambling, illegality
low ← news spread (public's interest) → high		

V. Conclusion and Implication

<Table 2> shows the result of news analysis about the gambling by the framing characteristics based on the number of reports and the intention of media and the degree of public's interest based on network connectivity.

The analysis of news related to gambling industry is divided into following four categories. 1. advertising publicity, 2. mandatory information, 3. social issue(public agenda), 4. major information. Among them, social issue is the news that is set to public agenda. In other words, news could be formed as a public policy when media's intention is strong and public's reaction is high. It is similar case about consolidation process of agenda setting when Korea Lottery Association was formed in 2006 with strong government's intention and high publics' interest. In addition, we can understand that company about sports bets and casinos have controlled media and news well because they are an operator under the Ministry of Culture, Sport and Tourism, which is the government agency that supervises media companies. However no one could control public's interest on gambling and lottery industry. And there occurs the difference in public's interest on news by the degree of player's distribution. The public shows high interest in news related to lottery or casino which has wide player's distribution. On the other hand, the public shows lower interest in the news related to sports betting or horse racing because player's distribution is narrow.

In this paper, we analyzed the difference and change in the media reporting framing and public perception of the gambling industry based on news for the last 30 years. In addition, we could explore the policy characteristics and reform subjects by analyzing the frequency and connectivity of articles with framing and public's interest based on the results of semantic network analysis. There was a lot of research on the perception change through semantic network analysis in the existing news analysis. However, this paper is the first attempt to study the setting conditions of public agenda by reinterpreting the frequency and connectivity of network as the importance of information and public reaction. We can also provide to government sector new method to measure the public's cognition about gambling industry instead existing survey.

참 고 문 헌 (References)

- [1] National Gambling Control Commission, 2014 Gambling Statistics, National Gambling Control Commission, 11-1371045-000037-10, 2014.
- [2] Chunsob Kwak, Ilsoo Kyung, Hyunji Lee, "A Study on the Recognition of Information Acceptor about Civil Defence Alert Broadcasting", *Journal of Broadcast Engineering*, Vol.20, No.6, pp.827-836, 2015.
- [3] Eung Sook Kim, "Global Media Environments and Glocalism Contents as Alternatives for Cultural Diversity", *Journal of Broadcast Engineering*, Vol.12, No.5, pp. 480-490, 2007.
- [4] Moon Man-Ki, Kim Tae-Yong, "Development of Digital Games Based on Historical Material and its Design Components - With History Based Games of 5 Countries", *Journal of Broadcast Engineering*, Vol.12, No.5, pp.460-479, 2007.
- [5] Schneider, Anne & Ingram, Hellen., "Behavioral Assumptions of

- Policy Tools". *Journal of Politics*, Vol.52, No.2, pp.510-529, 1990.
- [6] _____, "Social Construction of Target Populations: Implications for Politics and Policy". *American Political Science Review*, Vol.87, No.2, pp.334-347, 1993a.
- [7] _____, Social Constructions and Policy Design: Implications for Public Administration. In James L. Perry (eds.) *Research in Public Administration*. JAI Press, 1993b.
- [8] _____, Policy Design for Democracy. Lawrence, KS: The University of Kansas, 1997.
- [9] Cobb, Roger, Ross J. K. & Ross Macc H., "Agenda Building as a Comparative Political Process". *American Political Science Review*, Vol.70, pp.126-135, 1976.
- [10] May, P. J., "Reconsidering Policy Design: Policies and Publics", *Journal of Public Policy*, Vol.11, No.2, pp.187-206, 1991.
- [11] Moon, HyeJung, "Study on the Agenda Setting of Lottery Policy". *Korean Society and Public Administration*, Vol.23, No.1, pp.287-316, 2012.
- [12] Kam, Miah & Min Song, "A Study on Differences of Contents and Tones of Arguments among Newspapers Using Text Mining Analysis", *Bibliographic Info: J Intell Inform Syst*, Vol.18, No.3, pp.53-77, 2012.
- [13] Oh, Ick Keun, Tae Sook Lee & Chae Nam Chon, "A Study on Awareness of Korea Tourism through Big Data Analysis", *Journal of Tourism Sciences*, Vol.39, No.10, p.107, 2015.
- [14] Song, TaeMin, "Analysis of the trend of recognition of national unification using social big data", *Issue & Focus*, Vol.269, pp.1-8, 2015.
- [15] Choi, Kyoung-Ho & Jeong-Hye Park, "The Analysis of Public Awareness about Literary Therapy by Utilizing Big Data Analysis", *Journal of Digital Convergence*, Vol.13, No.4, pp.395-404, 2015.
- [16] Cho, Jane, "A Study for Research Area of Library and Information Science by Network Text Analysis", *Journal of the Korean Society for Information Management*, Vol.28, No.4, pp.65-83, 2011.
- [17] Min, Kyung Ick & Tae Sook Lee, "A Study on the Changing Perceptions of Tourism in Korea through Newspaper Analysis", *Journal of Tourism Sciences*, Vol.36, No.10, pp.81-104, 2012.
- [18] Jeong, Seo Hwa & Hyun Suk Cho, "A study on frame transition of personal information leakage", *Journal of Digital Convergence*, Vol.12, No.5, pp.57-68, 2014.
- [19] Choi, Yoon-jung & Kweon, Sang-Hee, "A Semantic Network Analysis of the Newspaper Articles on Big data", *Journal of Cybercommunication Academic Society*, Vol.31, No.1, pp.241-286, 2014.
- [20] Wang, W. & Rada, R., "Structured hypertext with domain semantics", *ACM Trans. Inform. Syst.*, Vol.16, pp.372-412, 1998.
- [21] Wasserman, S. & Faust, K., *Social network analysis: methods and applications*, Cambridge University Press, 1994.
- [22] Entman, R.M., "Framing: Toward clarification of a fractured paradigm", *Journal of Communication*, Vol.43, pp.51-58, 1993.

저 자 소 개

문 혜 정



- 현재 : 서울과학기술대학교 IT정책대학원 스포츠문화복권정책학과 겸임교수
- 현재 : 복권학회협동조합 이사장, 아이엘피 대표, 월비솔루션 이사
- ORCID : <http://orcid.org/0000-0001-8265-3256>
- 주관심분야 : 복권, 사행산업, 빅데이터, IT정책, 데이터 분석

김 성 경



- 현재 : 서울과학기술대학교 스포츠문화복권정책학과 석사졸업
- 현재 : 크로스넷 수석
- ORCID : <http://orcid.org/0000-0001-9543-8184>
- 주관심분야 : 빅데이터, 사행산업, 클라우드