A Study on the Deduction of Relevant Keywords and Establishment of Marketing Strategy by SNS Analysis: Focused in Local Festivals

Hansol Park, Hyeonjin Song, and Kyung-Hee Lee Business and Data Convergence, Chungbuk National University, Cheongju, South Korea Email: skzk129@hanmail.net, com4217@naver.com, lee.kyunghee@gmail.com

Jihye Kim and Wan-Sup. Cho

BK21 Plus Big Data Service Model Optimization Team, Department of Management Information Systems, Chungbuk National University, Cheongju, South Korea Email: ttculture@naver.com,Wscho@chungbuk.ac.kr

> Jinhyuk. Kim Bigdata, Cheongju, Chungbuk National University, South Korea Email: hyuk0628@gmail.com

Abstract—This study suggests method to establish of marketing strategy by using relevant keywords derived from SNS big data analysis. For this study, 3 keywords (Organic, Expo, Eco-friendly) were used to collect data related to the organic-themed local festival for a year from Naver blog since Nov.01.2013. As a result, the local festival was related to 'Children', 'Baby', 'Mom', 'Coffee', 'Experience', 'Event', 'Booth', 'Detergent' and 'Washing'. The results suggest marketing strategies to promote local festival by using relevant keywords. There is meaning that using unstructured data such as SNS big data to establish marketing strategy.

Index Terms—SNS analysis, relevant keyword, marketing strategy

I. INTRODUCTION

SNS (Social Network Service) is a web-based interactive service in which individuals fill in personal profile online in an open or a close manner and share it with others to build relationship [1]. As Internet service has currently developed and more smart devices are used, a tremendous amount of data is pouring out of SNS and companies are active to find meaningful information in them. Namely, many companies expect that they can generate values and maximize profits through big data analysis in the field of marketing, CRM and SCM etc.

The present study intends to deduct relevant keywords through SNS big data analysis and use them to establish the marketing strategy that can activate local festivals. With such attempt, this study aims to establish consumeroriented marketing strategy, which is reliable and credible, by understanding people's minds, patterns and trends precisely. This paper is comprised of as follow: in chapter 2, the studies related to the text data mining and establishment of strategy by SNS analysis are investigated; in chapter 3, the method to establish the marketing strategy is proposed; in chapter 4, experiment results and analysis results are presented; in chapter 5, this study present the conclusion, limitations and future study.

II. RELATED RESEARCH

According to the results of big data analysis and reports from Samsung Economics Research Institute (SERI), companies can understand the structure of social network that is the passage of information between customers and find potential consumer segments through big data analysis. In addition SERI indicates that SNS big data analysis can allow companies to understand what roles and communication method and channel they have to play and use in online communities and secure underlying information on which effective marketing program can be designed and implemented [2]. Lately, we have often seen the cases that consumers' opinions that a company and an organization can't sense are understood and made available through SNS analysis.

SNS data are atypical. Atypical data can be analyzed with such big data analysis technique of text mining and Opinion Analysis. Of the, text mining is fundamental. This analytic method finds out meaningful information from atypical data in form of text through Natural Language Processing [3]. Using text mining has merits of enabling text-to-text relevance analysis from atypical data and understanding hidden context [4].

The present study examined the way to develop an effective marketing strategy by using the text mining technique to deduct 'relevant keywords'.

Manuscript received December 20, 2014; revised May 13, 2015.





Figure 1. Total process

A. Data Collecting

Web-based SNS data were collected by the web crawler (designed for this study) in Web Data Collecting JAVA language. In doing so, certain keywords were designated to extract data relevant to marketing objective and this data were stored in the database.

B. Text Mining

The data saved in database were processed in natural language with Kind Korean Morpheme Analyzer (KKMA) developed by the Intelligent Data Systems of Seoul National University. And then, only nouns were extracted from the analyzed data and stored in the database. These extracted words are relevant keywords.

C. Relevant Keyword Analysis

The stored relevant keywords in the database were computed for word of frequency. The frequency means the amount of SNS buzz and they were arranged in the order of buzz amount (from high to low frequency) and cleaned to eliminate words with no meaning.

D. Marketing Strategy

A marketing strategy was developed using the relevant keywords with high buzz amount. The strategy can be expected to be very reliable because it was developed with core relevant keywords.

IV. EXPERIMENT AND RESULT

The algorithm designed for the marketing strategy above was applied to a local festival with theme of organic foods and a strategy was sought that can promote the local festival effectively.

A. Collected Data

The present study searched NAVER blogs for the posted text that include 3 keywords (organic, Expo, eco-friendly) related to local festivals. The scope of the texts search is a year (2013.11.01~2014.10.31) and a total of

11,850 texts were collected. (Organic – 3,990, Expo – 3,986, Eco-friendly – 3,874).

B. Text Mining & Relevant Keyword Analysis

The collected texts went through text mining and relevant keyword analysis. And then KKMA extracted only nouns from the collected data and they were saved in the database. These relevant keywords were computed for buzz amount. The results were saved in the database as in Table I.

TABLE I. BUZZ AMOUNT STORED IN THE DB

Rank	ID	Word	Counter	Tag
1	21128	Environment	22853	NNG
2	8543	Use	10173	NNG
3	9735	Detergent	9725	NNG
4	16157	Product	7610	NNG
5	11411	Children	4804	NNG
6	21	Furniture	3406	NNG
7	11294	Baby	3213	NNG
8	9749	Washing	3071	NNG
9	9138	Think	3064	NNG
10	17558	Pure	2641	NNG

C. Analysis Result

Table II shows top 20 relevant keywords by 3 keyword (organic, Expo, eco-friendly) in order of frequency.

V. DEVELOPING MARKETING STRATEGY USING RELEVANT KEYWORDS

The Table III shows a marketing strategy that was developed with the meaningful keywords of the total relevant keywords in Table II.

TABLE II.	CORE RELEVANT KEYWORD TOP 20	

Rank	Organic	Expo	Eco-friendly
1	Product	Photo	Detergent
2	Skin	Coffee	Product
3	Children	Experience	Children
4	Health	Children	Furniture
5	Ingredient	Yeosu	Baby
6	Baby	Booth	Washing
7	Oil	Event	Natural
8	Cosmetic	Baby	Diaper
9	Pure	Variety	Natural
10	Cream	Product	Health
11	Confectionery	Daejeon	Skin
12	Certification	Park	Paper Cup
13	Sugar	Exposition	Ingredient
14	Berry	Bio	Certification
15	Mom	Event(.English)	Soda
16	Cheese	Registration	Photo
17	Environment	International	Safety
18	Food	Beforehand	Apple
19	Strawberries	Industry	Design
20	Natural	Progression	Infant

TABLE III. STRATEGY ABOUT EACH CORE RELEVANT KEYWORD

Core Relevant Keyword	Strategy	
Children, Baby, Mom	Targeting females: pregnant women, housewives in their 30s to 40s	
Experience, Event, Booth	Experiential programs: 'Making Eco- Friendly Cosmetics', 'Picking Organic Strawberry'	
Coffee	Run a cafe that sells organic coffee in the promotion booth or offer experiential activity such as letting visitors roast organic coffee by themselves	
Detergent, Skin, Washing, Pure	Offer giveaways such as eco-friendly soup or hand cream	

VI. CONCLUSION

The present study proposed a way to develop a marketing strategy by using relevant keywords through SNS data analyses. A total of 11,850 text data including the keywords related to local festivals (Organic, Expo, Eco-Friendly) were collected from the data of a year in NAVER blogs. The collected data were processed with natural language (KKMA) to extract core relevant keywords. The extracted core relevant keywords were used to develop an effective marketing strategy. The results are as follows.

First, 'Children', 'Baby' and 'Mom' were the most relevant keywords. Accordingly, a marketing strategy that targets females, especially, pregnant women and housewives in their 30s to 40s was suggested for local festivals.

Second, 'Experience', 'Event' and 'Booth' turned out to be the most relevant keywords. Therefore, a marketing strategy that provides experiential programs in which visitors can participate was suggested for local festivals.

Third, 'Coffee' was the most relevant keyword. Based on it, a marketing strategy that runs a cafe dealing with organic coffee was suggested for local festivals.

Fourth, 'Detergent', 'Skin Washing', and "Pure' were the most relevant keywords. As a result, a marketing strategy that offers giveaways such as eco-friendly soup or hand cream was suggested for local festivals.

The study has some limitations. First, the findings may not be applied to all ages due to the characteristics of NAVER blogs - mainly used by certain age groups, from which relevant keywords were extracted. Second, the results of data analysis heavily rely on the representability of search words designated in collecting data. Therefore, a standard to secure the reliability of search keywords is necessary.

Developing a marketing strategy based on relevant keywords can be possible not only with SNS data but also any kinds of data on Web. Therefore, it is expected that this study will be followed by studies with more diverse types of data such as news and Goggle webpage as well as blog data.

ACKNOWLEDGMENT

This research was supported by the MSIP (The Ministry of Science, ICT and Future Planning), Korea, under the "SW master's course of a hiring contract" support program (NIPA-2014-HB301-14-1011)

REFERENCES

- D. M. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of Compute-Mediated Communication*, vol. 13, no. 1, pp. 210-230, 2008.
- [2] ITworld, "IDG tech report 빅데이터의 이해," Journal of International Data Group, pp. 5-6, 2012.
- [3] Y. J. Lee and J. H. Yoon, "A study on utilizing SNS big data in the tourism studies: Based on an analysis of key words for tourism information search," *International Journal of Tourism and Hospitality Research*, vol. 28, no. 3, pp. 5-14, 2014.
- [4] J. S. Kim, "Consumer preferences analysis based on multidimensional SNS big data analysis," ChungBuk National University, 2012.



Hansol Park received her B.A. degrees at the Department of Management Information Systems from Chungbuk National University, Korea, in 2010 and 2014, respectively. She is a master at the Business and Data Convergence, Chungbuk National University in Korea. Her research interests include database, bigdata and business intelligence.



Hyeon Jin Song received her B.A. degrees at the Department of Management Information Systems from Chungbuk National University, Korea, in 2010 and 2014, respectively. She is a master at the Business and Data Convergence, Chungbuk National University in Korea. Her research interests include database, bigdata and business intelligence.





Jinhyuk Kim received his B.A. degrees at the Department of Computer Science from Soongsil University, Korea, in 2004 and 2015, respectively. He is a master at the Big Data, Chungbuk National University in Korea. His research interests include big data, distributed computing and a lot of computer algorithms.

Wan-Sup Cho received his M.S. and Ph.D. degrees at the Department of Computer Science from KAIST, Korea, in 1987 and 1996, respectively. He is a professor at the Department of Management Information System, Chungbuk National University in Korea. His research interests include database, bigdata, business intelligence and ERP.



Kyung-Hee Lee received the Ph. D. degree in computer science from Chungbuk National University in 2004. From 2008 to now, she serverd as a research engineer in the Business Data Covergence of Chungbuk National Univeristy. Her research interests include database systems, bigdata analysis algorithm and technology, data mining and distributed computing platform.



Jihye Kim received her B.A. degrees at the Department of Social Work from Korean Bible University Korea, in 1999 and 2014, respectively. She is a master at the Department of Management Information Systems, Chungbuk National University in Korea. Her research interests include database, bigdata and business intelligence.