Analysis on trends and future signs of smart grids

Chankook Park^a, Seunghyun Cho^{b*}

^a Korea Energy Economics Institute, 405-11 Jongga-ro, Jung-gu, Ulsan, 44543, South Korea ^b NH Future Management Research Institute, 16 Saemunan-ro, Jung-gu, Seoul, 04516, South Korea

Abstract

In such a wide range, understanding the current status and future prospects of smart grid issues requires much effort and time. Recently, with the spread of big data in the form of text and development of computing technology, efforts to automate trends review and future prospects are continuing. In particular, efforts are being made to derive future signs through text mining using the concept of weak signals. However, since the related research is still in the early stage, there are many improvement tasks. Currently, the output derived from the future signal derivation methodology using text mining is a keyword type. Since the keyword can have various meanings, there is a difficulty in practically interpreting the keyword. In this study, we examined trends of smart grid industry, explore future signs, and try to improve the interpretation of the future sign keywords through text mining. As a result, through the recent three years of smart grid news articles, we reviewed various keyword-type signals in relation to smart grid technologies, smart grid application scopes, smart grid regions and countries. Also, we suggested an alternative to interpret keywords in a more easy way and have examined some points to consider for developing more advanced methods in the future.

Keywords: Smart grids, weak signal, future sign, text mining

1. Introduction

Smart grids are transforming the energy sector regarding business models and industry structure [1-4]. The smart grids are recognized as the evolving intelligent electric power network. The U.S. Department of Energy (DOE) defined the smart grids as "an intelligent electricity grid-one that uses digital communications technology, information systems, and automation to detect and react to local changes in usage, improve system operating efficiency, and, in turn, reduce operating costs while maintaining high system reliability." [5] International Energy Agency (IEA) emphasizes that improved monitoring, control and automation technologies of smart grids contribute to enabling new business models while unlocking system-wide benefits including reduced outages, improved response times, deferral of investment in the grids themselves and the integration of distributed energy resources [6].

Overall, investment in smart grid technologies has risen to 12% in 2016 compared to 2014 [6]. However, investment in smart distribution networks is growing relatively slowly, and smart meter penetration rates are uneven across countries. In addition, regulatory changes and new business models are required for the smart grids to play a more central role in energy transition [6,7]. In order for the smart grids to play a leading role in the transition to cleaner energy, it is necessary to solve challenges in various issues. First, in the technical sector, there are issues such as cyber security, storage technology innovation, data management, efficient communication networks, and stability maintenance. Second, in a socio-economic sector, there are issues such as high capital investment costs, stakeholder engagement, consumer acceptance, privacy protection, and the possibility of electricity tariff hikes. Other issues include regulations, policies, and the cultivation of manpower [7].

In the face of such diverse issues, policymakers and stakeholders need to quickly identify and respond

^{*} Manuscript received January 25, 2019; revised February 10, 2020.

Corresponding author. Tel.: +82-2-2080-6909; E-mail address: commodity.cho@gmail.com.

doi: 10.12720/sgce.9.3.533-543

to changes in technology, market, policy, and regulatory issues in a timely manner to effectively implement smart grid projects.

However, in such a wide range, understanding the current status and future prospects of smart grid issues requires much effort and time. We tend to rely on experts to identify current trends or future prospects. However, if we rely on a small number of experts, we may encounter limitations due to the narrowness of their knowledge. Even if we ask for opinions from various experts, there is a costly problem [8].

Recently, with the spread of big data in the form of text and development of computing technology, efforts to automate trends review and future prospects are continuing [8,9,10]. In particular, efforts are being made to derive future signs through text mining using the concept of weak signals. However, since the related research is still in the early stage, there are many improvement tasks. Currently, the output derived from the future signal derivation methodology using text mining is a keyword type. Since the keyword can have various meanings, there is a difficulty in practically interpreting the keyword [8,10]. In this study, we examine trends of smart grid industry, explore future signs, and try to improve the interpretation of the future sign keywords through text mining.

The structure of this study is as follows. We discuss the existing literature of weak signals and future sign briefly in Chapter 2. We explain the methodology of this study in Chapter 3. In Chapter 4, we show research results. In Chapter 5, we discuss the way to improve the future sign detection method through text mining, and we draw conclusions in Chapter 6.

2. Literature Review

Future sign detections through text mining are mainly based on a weak signal concept. A weak signal can be understood as a signal that suggests a possible change in the future. The term weak signal was proposed by Ansoff [11]. He defined "a weak signal as a system of possible changes in the future and asserted that strategic planning was a rational when a historical trend developed gradually, but not effective in dealing with unexpected events." [8]

Coffman [12] presented the concept of a weak signal in a more concrete way. He emphasized that the weak signal has seven characteristics. "1) An idea or trend that will affect how we do business, what business we do, and the environment in which we work, 2) New and surprising from the signal receiver's vantage point (although others may already perceive it), 3) sometimes difficult to track down amid other noise and signals, 4) a threat or opportunity to your organization, 5) often scoffed at by people who "know", 6) usually has a substantial lag time before it will mature and become mainstream, 7) Therefore, an opportunity to learn, grow and evolve." [12].

Hiltunen [13] used the concept of a future sign to better organize the weak signal. Hiltunen [13] described the future sign in three dimensions: 'signal', 'issue', and 'interpretation.' The 'signal' referred to by Hiltunen [13] means the visibility or number of future signs and the 'issue' is the number of events that can indicate how far future signs are spreading. The 'interpretation' means how many people understand the meaning of future signs. Hiltunen [13] noted that if signals, issues, and interpretations are all at a low level, they are classified as weak signals. If the values of all three dimensions are high, it is converted to a strong signal [8].

Later, beyond the conceptual studies of a weak signal, studies were conducted to utilize text mining to explore future signs more efficiently. Yoon [9] suggested a method to detect a weak signal utilizing text mining. Yoon classified a keyword, which has a low term frequency (TF) and low document frequency (DF) but a high growth rate, as a weak signal based on Hiltunen's definition of the weak signal. In contrast, a keyword with a high increase rate as well as a large number of TF and DF was classified as a strong signal [8]. Park and Kim [10] defined keywords with high frequency and low rate of growth as not-strong-but-well-known signals, and keywords with low frequency and low growth rate of both terms and documents are defined as latent signals.

Park and Cho [8] applied the future sign detection method using text mining to smart grid industry analysis. They collected 14, 325 news articles from 2014 to 2016 for a total of three years to explore

future signs in terms of smart grid components, smart grid stakeholders, and smart grid applications scope. However, when they collected news articles, only 'smart grid' is used as a search keyword. It is necessary to add synonyms such as' smart power grid ',' smart electric grid ', and intelligent grid'. In addition, authors interpreted the future sign-related keywords in an arbitrary way, so it is necessary to reinforce the objectivity of study.

3. Methodology

3.1. Text data collection and refining

We crawled the last three years of news articles about the smart grids using Google search engine. When we crawled the news articles on Google, we crawled them by setting a date range of one day to collect as many news articles as possible. The keywords entered into the Google search are "smart grid," "smart power grid," "smart electric grid," and "intelligent grid." The content of a news article we crawled consists of approximately 200-300 characters. While there is a limit to the amount of the text data, there are little major problems finding trends, as the documents have key sentences related to the search keywords. After collecting the news text data, duplicate news articles were removed by double checking the text. Finally, the total number of news articles used was 16,413: 3,000 in 2016, 4,756 in 2017, and 8,657 in 2018.

The collected text data was refined using R 'tm' package [14]. All terms are converted to lower case, and symbols, numbers, stopwords, etc. were removed. In analyzing keywords, it is ineffective to see all the terms contained in the collected documents. Therefore, we selected the top 1,000 terms and focused on keywords related to smart grid technologies, application scopes, and regions and countries.

3.2. Future sign-keyword extraction^c

According to Yoon [9], the keyword portfolio is made up of a keyword emergence map (KEM) and a keyword issue map (KIM) as shown in Fig. 1. To make these maps, separate statistics are used as follows. The degree of visibility (DoV) is reflected to the KEM as a proxy of visibility; the degree of diffusion (DoD) to the KIM as a proxy of diffusion. Each index is calculated respectively by the following equations [8].

$$DoV_{ij} = \left(\frac{TF_{ij}}{NN_j}\right) \times \{1 - tw \times (n - j)\}$$
(1)

$$DoD_{ij} = \left(\frac{DF_{ij}}{NN_j}\right) \times \{1 - tw \times (n - j)\}$$
⁽²⁾

Where, NN is the total number of documents, tw is time weight, n is an entire time period, j is a point of time. The above DoV and DoD are the variables that are obtained respectively with the term frequency and the document frequency divided by the total number of documents to calculate the rate at a certain point of time. Yoon suggested tw to determine the size of the time weight and applied 0.05 to his own research on solar cells. If a time period of three years is analyzed annually, the most recent year has zero for (n - j), resulting in 1 for $\{1 - tw \times (n - j)\}$ and the previous year is 0.95. In this study, the time weight was set to 0.05 as in Yoon. However, it may not matter in the analysis not to apply the time weight. A value other than 0.05 can be applied as a time weight, and the results can be separately interpreted [8].

In KEM, the x-coordinate represents the average term frequency and the y-coordinate the average (geometric mean) growth rate of DoV. When the quadrants are divided by the medians of the respective values, the area corresponding to the first quadrant has a high average term frequency and a high average growth rate of DoV, indicating a strong signal, and the area corresponding to the second quadrant has a low average term frequency and a high average growth rate of DoV, meaning a weak signal. The keyword

^c In this section, the methodology of Park and Cho [8] is cited.

located in the third quadrant can be identified as a latent signal that is not yet significantly noticeable. The keyword in the fourth quadrant, which has a high term frequency but a low growth rate of DoV, can be classified as a not-strong-but-well-known signal because it is already familiar to people but its increase rate of DoV is not high [8-10].

The KIM shows the average document frequency on the x-coordinate and the average growth rate of DoD on the y-coordinate, mapping the selected keyword in the same way as the KEM. If a keyword has a low average document frequency but a high average growth rate of DoD, it is a weak signal; if it has a high average document frequency and a high average growth rate of DoD, it is a strong signal [9]. Likewise, if the word is positioned in the third quadrant, it is a latent signal; if it is located in the fourth quadrant, it is a keyword that is well known but has a low average growth rate of DoD. Finally, based on the future signs identified in the KEM and the KIM, a weak signal, a strong signal, and a not-strong-but-well-known signal are to be examined [8,10].



Fig. 1. Process identifying the future sign [8,10].

3.3. Future sign-keyword interpretation

This study attempted to derive sentences or documents most similar to future sign-related keywords in order to understand more practically the meaning of the keywords derived as future signals. We used Global Vectors for Word Representation (GloVe) to extract sentences or documents similar to keywords. GloVe is a word embedding methodology developed by Stanford University in 2014. GloVe is known to improve the shortcomings of Word2Vec developed by Google in 2013. Word2Vec has a better performance than the existing Latent Semantic Analysis (LSA) in measuring the similarity between word vectors embedded in a low dimensional vector space. However, since learning and analysis are performed only within the window (indicating surrounding words) specified by the user, it is difficult to reflect the co-occurrence information of the whole corpus.

Pennington et al. [15] criticized LSA and Word2Vec as follows. "While methods like LSA-efficient leverage statistical information, they are relatively poorly on the word analogy task, indicating a suboptimal vector space structure. Methods like skip-gram may do better on the analogy task, but they poorly utilize the statistics of the corpus since they train on separate local context windows instead of global cooccurrence counts." According to Pennington et al. [15], the model "efficiently leverages statistical information by training only on the nonzero elements in a word-word co-occurrence matrix, rather than on the entire sparse matrix or on individual context windows in a large corpus. The model produces a vector space with meaningful substructure. It also outperforms related models on similarity tasks and named entity recognition".

We obtained the word-to-word distance matrix based on the word vector matrix derived from GloVe. The smaller the value in the distance matrix, the shorter the distance between words. In other words, the similarity between words is high. Based on the distance matrix, a weight matrix is constructed to represent the similarity between words, and the weight matrix is multiplied by the term-document matrix to determine the similarity between the document and the word. This work was carried out using R

programming.

4. Results

Through the recent three years of smart grid news articles, this study examined not-strong-but-wellknown signals, strong signals, and weak signals in relation to smart grid technologies, smart grid application scopes, smart grid regions and countries. Moreover, in order to facilitate the interpretation of each keyword, the most similar sentences were derived.

4.1. Smart grid technologies

Table 1 shows the signal classifications in KEM and KIM and the common signals in terms of smart grid technologies. The keyword classified as a strong signal is 'cyber security', 'data', "battery", 'analytics', 'sensors', 'remote terminal unit (RTU)', and 'submetering', which have high term frequency and document frequency and recently shows great DoV and DoD growth rates as well. Table 2 shows sentences that best represent the meaning of each strong signal.

The keyword categorized as weak signals is 'self-healing', 'blockchain', 'machine learning.' For self-healing, reports on market prospects were the main focus, and blockchain was emphasized as a rapid spread in peer-to-peer (P2P) electricity trading. Machine learning has been emphasized in various aspects such as energy efficiency, cyber security. Internet of things (IoT) and big data were strong signals a few years ago, but they are now classified as not-strong-but-well-known issues [8].

Table 1. Signals in relation to smart grid technologies

Category	Strong signal	Weak signal	Not-strong-but-well-known signal
KEM	cyber security, data, battery,	self-healing, computing,	IoT, storage, solution, meter,
	analytics, sensors, RTU,	blockchain, machine	communications, cloud, big data,
	submetering	learning	automation, control
KIM	cyber security, data, sensors,	self-healing, hardware,	IoT, solution, storage, cloud, meter,
	battery, analytics, RTU,	blockchain, machine	big data, communications,
	submetering	learning, automatic	software, control
Common signal	cyber security, data, battery, analytics, sensors, RTU, submetering	self-healing, blockchain, machine learning	IoT, storage, solution, meter, communications, cloud, big data, control

Table 2. Representative sentences in relation to smart grid technologies

Strong signals (technologies)	Representative sentences
cyber security	Cyber security of the smart grid has been an area of concern for the power utility sector, due to frequent exchange of sensitive information that takes place via communication networks such as the internet, intranets, extranets, and corporate networks.
	However, smart grids are designed to mitigate the risk of an attack happening and to lessen the damage that can be caused should hackers penetrate a smart grid's network security.
data	The Smart Grid platform began gathering data about natural gas supply agreements at the beginning of 2017.
	WePower will be leveraging Estonia's existing advanced smart grid infrastructure, particularly the 100% smart meter coverage with hourly reading and a central data hub.

battery	Where wireless sensor terminals are deployed over wide areas and in remote locations, as in agricultural or smart grid applications, concerns over battery life and maintenance costs quickly arise.
battery	Battery Energy Storage Systems for Smart Grid Market Report cover detailed competitive outlook including the market share and company profiles of the key participants operating in the global market.
analytics	Besides this, increasing adoption of smart grids is also fueling the growth of global smart grid analytics market.
anarytics	Moreover, rising advancements in Internet-of-Things and Big Data analytics plays a vital role in the elevation of the global market.
	Smart Grid has sensors on poles that detect outages, according to Witter. "That automatically reroutes power to a different line to customers.
sensors	Smart grid sensors are compact electrical devices used for detecting and monitoring power system equipment. They track the power line temperatures
RTU	The competitive market research study on Global Remote Terminal Unit (RTU) in Smart Grid Market contains descriptions of the leading players including key financial metrics and analysis within the market.
	Latest research study from HTF MI with title Asia-Pacific Remote Terminal Unit (RTU) in Smart Grid by Manufacturers, Regions, Type and Application, Forecast to 2023.
submetering	Global Electricity Submetering For Smart Grid Market Research Report 2018 presents an in- depth assessment of Electricity Submetering For The report Electricity Submetering For Smart Grid Market by Manufacturers highlights the essential market dynamics of Electricity Submetering
	essential market dynamics of Electricity Submetering
Weak signals (technologies)	Representative sentences
self-healing blockchain	Detailed market study on the "Global Self Healing Smart Grid Market" Research Report 2018- 2025 By QY Research Store.
	Global Self Healing Smart Grid Market report keep a basic overview of the industry including definitions, classifications, applications and
	WePower will integrate its blockchain and smart contract powered green energy trading platform into Elering's Estfeed
	Grids such as this are being built around the world to combat a number of energy dilemmas, as well as easing the introduction of renewable energy sources, and are usually deployed together with smart grid technologies such as blockchain.
machine learning	Artificial intelligence and its cousin, machine learning, enable energy providers to make smarter decisions about
	The smart grid is already transforming the industry, but with machine learning and reality models exploiting the IIoT, the smart grid has themarket.

4.2. Smart grid application scopes

Looking at the major areas where smart grids are applied, it seems that 'microgrid' is receiving much attention now as a strong signal. Through similar sentence extraction, we confirm that the microgrid market is developing and various projects are going on. In Park and Cho [8], terms such as 'city', 'distribution', and 'renewable' were classified as strong signals. The use of intelligent technologies in the power distribution sectors has been emphasized continuously, and the city is receiving the spotlight in the smart city dimension [16]. In the renewables sector as well, the smart grids are being recognized as an important means of adjusting intermittent power and maintaining a balance of power supply and demand [17]. However, the terms have been converted to not-strong-but-well-known signal in the current situation.

In the weak signal, 'blackouts' has been derived. Recently, smart grid projects have been actively promoted in developing countries, and the benefit of smart grid for blackouts reduction has been

emphasized. Also, projects in countries such as the Philippines have been mentioned.

Table 3. Signals in relation to smart grid application scopes

Category	Strong signal	Weak signal	Not-strong-but-well-known signal
KEM	substation, microgrid	voltage, blackouts	city, distribution, transmission, renewable, generation, building, solar, vehicles, charging, transportation, consumption
KIM	charging, microgrid, voltage	natural gas, blackouts	city, distribution, transmission, renewable, generation, vehicle, solar, building, consumption
Common signal	microgrid	blackouts	city, distribution, transmission, renewable, generation, building, solar, vehicle, consumption

Table 4. Representative sentences in relation to smart grid application scopes

Strong signals (application scopes)	Representative sentences
microgrid	In battery storage, the top funded company in the third quarter was Advanced Microgrid Solutions, which raised \$34 million from Energy Impact Partners, Southern Company, The Japanese firm had previously installed a similar solar power microgrid at Fujisawa Sustainable Smart Town, 30 miles west of Tokyo. There, Panasonic hooked up 1,000 new residences to the smart grid, which monitors energy usage in real time.
Weak signals (application scopes)	Representative sentences
blackouts	A Japanese company Tokyo Electric Power Company and its affiliate firm Takaoka TokoCo., Ltd (TKTK) partnered with the National Electrification Administration (NEA) for a pilot project to roll out a smart grid technology to help the Philippines improve its ability in preventing power interruptions or blackouts. The adoption of smart grids is expected to increase substantially over the forecast period owing to the several benefits they offer including the reduction in blackouts, higher efficiency during distribution, and reduction in electricity theft.

4.3. Smart grid regions and countries

We have reviewed which areas in the region or countries have recently emerged in relation to smart grids. India, China, the Americas, and Asia have had a lot of frequency, but the rate of increase has not been high in the last three years. On the other hand, Africa and Japan are classified as strong signals because both their frequencies and the increase rate are high.

France, Swiss and the Philippines were classified as weak signals. As Table 6 shows, France is focused on smart grid investments, and Switzerland has many news articles on technology development of collection of energy usage data. In the Philippines, cooperation with Japan and future smart grid investment plans are mentioned.

Category	Strong signal	Weak signal	Not-strong-but-well-known signal
KEM	Africa, Japan	France, Swiss, Philippines	India, China, America, Europe, Asia, Canada
KIM	Europe, Africa, Japan	France, Swiss, Philippines	India, China, America, Asia
Common signal	Africa, Japan	France, Swiss, Philippines	India, China, America, Asia

Table 5. Signals in relation to smart grid regions and countries

Strong signals (regions & countries)	Representative sentences
Africa	The USTDA is an agency involved in the overall U.S. Power Africa Initiative to connect U.S smart grid supplier with projects to modernize the electricity grid in cities in Sub Saharan Africa and to bring microgrid solutions to remote areas on the continent.
	The report deep dives into the dynamics of EMEA (Europe, Middle East and Africa) Smart Grid Optimization Solutions providing useful and unique
Japan	The project has been initiated under intensive relations between Japan and India and the Prime Minister's Make in India Programme.
	The Smart Grid Pilot Project, being set up with the assistance of Japan, would ensure uninterrupted power supply to domestic as well as industrial consumers of the city
Weak signals (regions & countries)	Representative sentences
France	New Delhi, Mar 1 Power discom TPDDL has joined hands with French company Electricite de France (EDF) and its wholly-owned subsidiary Enedis to explore the opportunities in Smart Grid technology in India.
	With more than 508 million euros invested in refurbishment and smart grid deployment activities in the region, France is recently drawing a
Swiss	Swiss scientists have developed a system for an all-encompassing smart building to collect data on people's energy usage, and then send the data to a smart electric grid to allocate resources optimally, according to Swiss Federal Institute of Technology in Lausanne (EPFL)
	Landis+Gyr (Swiss: LAND.SW) reported continued growth with public power utilities during 2017, signing 46 contracts for smart grid technology and services.
Philippines	A Japanese company Tokyo Electric Power Company and its affiliate firm Takaoka TokoCo., Ltd (TKTK) partnered with the National Electrification Administration (NEA) for a pilot project to roll out a smart grid technology to help the Philippines improve its ability in preventing power interruptions or blackouts.
	According to this report, by the end of 2025, Indonesia targets to accomplish around 90% electrification and few other countries like Malaysia and Philippines are also having similar plans.

Table 6. Representative sentences in relation to smart grid regions and countries

5. Discussion

5.1. Considering the change of signals

In this study, when deriving various signals in the form of a keyword, even if it reflected changes within a specific period, it did not deal with the change of the signals beyond that specific period. Park and Kim [10] and Park and Cho [8] emphasized the need to continue a dynamic study of how each signal will change in the future. In the Park and Cho [8], which analyzed smart grid industry trends based on news articles from 2014 to 2016, 'big data', 'IoT', and 'analytics' were captured as strong signals in smart grid technologies. However, two years later, in this new study, 'big data' and 'IoT' were converted to a not-strong-but-well-known signal and only 'analytics' remained a strong signal.

On the other hand, 'gateway' that was selected as a weak signal in Park and Cho [8] was selected as latent signal in this study. As such, over time, the strong current signal may continue to be a strong signal or turn into a not-strong-but-well-known signal. The weak signal may be changed to a strong signal or it may become a latent signal with its importance lowered. To understand the changes of the signals, it is necessary to acquire data for a longer period of time, analyze change patterns of the signals, and predict how signals will change according to the patterns.

5.2. Interpretation of removed keywords during future sign detection

After the weak signals and the strong signals in KEM and KIM were respectively extracted, the

keywords giving the common signals in both maps were selected as the final future signs. As Park and Kim [10] and Park and Cho [8] noted, "the advantage of this method is that since the final keywords are located in the same quadrant in both maps, it is possible to more clearly see whether the keywords have the features of their quadrant. However, even if a certain keyword is a significant keyword that should be paid attention to if it is located in a different quadrant from another in the two maps, it is removed from future signs. For example, when a certain keyword is taken as a weak signal in KEM and as a strong signal in KIM, there is no method to analyze this keyword. In this case, it is necessary to examine the implied meaning of the keyword by naming it as a separate signal" [8].

When a keyword is not in agreement in both maps, it is not necessary to name all the combinations of the keywords as a new signal. When there are many types of signals, they can complicate the interpretation, increasing the confusion in analyzing future signs. If a keyword is located in the first quadrant as a strong signal in KEM, but located in the fourth quadrant as a not-strong but well-known signal in KIM, it can be recognized as a signal of the fourth quadrant. In other words, although this word cannot be said to have a high term frequency growth rate over time, it definitely has high frequency, and hence it can be interpreted as a fourth quadrant signal. Otherwise, we can develop a new analysis framework that integrates KEM and KIM to avoid the problem of missing some signals [8].

5.3. Establishment of classification frame for energy sector text

In order to meaningfully interpret any keywords obtained by text mining, it is necessary in advance to provide the environment to extract important keywords. The most basic task is to remove stop words in each language. A stop word is a word that does not have to be searched by a keyword, such as articles, prepositions, postpositions, and conjunctions. Text mining tools such as R and Python provide basic functions to perform this task. If there are words that are not removed by the default function of these tools, it is necessary to create a separate index for those words to be automatically removed [10].

It is important to ensure the categorization of the combination words that are often used in the energy sector and the extraction of synonyms as one word, as well as the removal of unnecessary words. For example, combination words such as 'energy efficiency', 'energy conservation', and 'energy storage' should be recognized as a single word rather than splitting them. And 'cyber security', 'cyber attack', 'cyber terror', 'cyber-terrorism', and 'hacking' can all be integrated as one combination word of 'cyber security.' Unless this word dictionary is applied in advance, a synonym or a combination word with a similar meaning can be calculated separately, and finally their importance may be recognized as low [10].

This study expresses idioms as a single word and replaces certain words. But further text classification should be made in order to derive more meaningful keywords in the smart energy sector. A framework for more abundant text classification will be established in future studies [10].

5.4. Additional considerations

The set-up of keyword groups for mutual comparison should be considered in order to increase the relevance of future sign searching results through text mining. In this study, the keyword groups were first divided into the smart grid technologies, application scope, and regions and countries. Park and Cho [8] focused on components, stakeholders, and applications scope of smart grids. Yoon [9], in examining the future signs of solar cells, classified the keywords into product/technological components, environmental factors, and business needs. This framework can vary in accordance with the purpose of study. It is important to perform a comparative analysis of keywords by setting up an appropriate classification framework in accordance with the purpose of the research.

Meanwhile, this study used the median as the standard of the classification in categorizing a future sign into a weak signal or a strong signal. However, verification of appropriateness is additionally necessary for choosing the median as the standard of division. In addition, more consideration is necessary on whether a different standard should be used for the x-coordinate of KEM and the x-coordinate of KIM, or for the y- coordinate is of KEM and the y-coordinate of KIM.

6. Conclusion

The purpose of this study is to search for the future signs of the smart grid industry using a technique using text mining and suggest what changes could be made in the smart grid sector. This trial might provide insights into smart grid stakeholders. There is also another significant research purpose of discussing the challenges of the future sign method through text mining. Thus, we suggested an alternative to interpret keywords in a more easy way and have examined some points to consider for developing more advanced methods in the future.

As a result, in the smart grid technologies, keywords such as 'cyber security', 'analytics', 'sensors', etc. appeared as strong signals. It suggests that technologies related to effective and safe data utilization are attracting much attention in the field of smart grids as well. On the other hand, the keyword 'self-healing', 'blockchain', 'machine learning' was represented as a weak signal. Therefore, we can identify that more intelligent technologies are getting more and more attention.

In the smart grid application scopes, while the keyword 'microgrid' was selected as a strong signal, the keyword 'blackouts' was derived as a weak signal. In the smart grid regions and countries, recently, Africa and Japan are attracting much attention, while France, Switzerland, and the Philippines are not attracted attention as much, but the interest of the countries is rapidly increasing.

This study also examined the challenges that the future sign detection method has. As the existing studies [8,10] emphasized, it is difficult to know in detail about which of the weak signals will be converted into strong signals in the future. Of course, because the future has many uncertainties, it may not be possible to fully predict which signals will be converted into strong signals. However, if we analyze a large amount of information, we can understand the pattern of the signals in the future and understand the transition from a weak signal to a strong signal.

Another limitation is that we have to overcome the problem of throwing out some keywords in future sign detection [10]. The methodology of this study selects the signals extracted commonly from KEM and KIM. However, some signals are removed in this process. Therefore, we should consider changing the way of interpreting keywords or developing a new future sign detection framework. Furthermore, we emphasized the establishment of a classification frame for energy sector text and suggest additional considerations.

This study assessed how the smart grid-related issues are changing through various signals and it developed the existing method to better understand the keyword-type signals. In the next study, we need to develop a framework for future sign detection that can overcome the problem of missing some keywords and provide a basis for determining which weak signals are converted to strong signals.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

A designed the research; B analyzed the data; A,B wrote the paper; all authors had approved the final version.

References

- [1] IEA (International Energy Agency). Smart Grids Technology Roadmap. Paris: IEA, 2011.
- [2] MIT. The Future of the Electric Grid. Cambridge: MIT, 2011.
- [3] Accenture. Global perspectives on Smart Grid Opportunities. Dublin: Accenture, 2013.
- [4] Bichler, M. Smart grids and the energy transformation. Berlin: SEFEP, 2012.
- [5] Campbell, RJ. The Smart Grid: Status and Outlook. Congressional Research Service. CRS Report, Federation of American Scientists, 2018.
- [6] IEA (January 2019). Smart grids: Tracking Clean Energy Progress. [Online]. Available:

https://www.iea.org/tcep/energyintegration/smartgrids/

- [7] Kappagantu R, Daniel SA. Challenges and issues of smart grid implementation: A case of Indian scenario. Journal of Electrical Systems and Information Technology, 2018; 5:455-467.
- [8] Park C, Cho S. Future sign detection in smart grids through text mining. Energy Procedia, 2017; 128:79-85.
- [9] Yoon J. Detecting weak signals for long-term business opportunities using text mining of Web news. *Journal Expert Systems* with Applications, 2012;39:12543-12550.
- [10] Park CK, Kim HJ. A Study on the development of the new energy industry through the Internet of things: searching future signs using text mining. Ulsan: KEEI, 2015.
- [11] Ansoff HI. Managing strategic surprise by response to weak signals. Californian Management Review, 1975;18:21-33.
- [12] Coffman B. Weak signal research. Part I. Introduction. Journal of Transition Management, 1997;2.
- [13] Hiltunen E. The future sign and its three dimensions. Futures, 2008;40:247-60.
- [14] Feinerer I, Hornik K, Meyer D. Text mining infrastructure in R. Journal of Statistical Software, 2008; 25(5):1–54. http://www.jstatsoft.org/v25/i05/.
- [15] Pennington J, Socher R, Manning CD. GloVe: Global vectors for word representation. In: Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014:1532-1543.
- [16] Geisler K. The Relationship Between Smart Grids and Smart Cities. IEEE, 2013.
- [17] IRENA (The International Renewable Energy Agency). Smart grid and renewables-A guide for effective deployment. Abu Dhabi: IRENA, 2013.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License (CC BY-NC-ND 4.0), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.