

FUTURE DISASTER SCENARIO USING BIG DATA: A CASE STUDY OF EXTREME COLD WAVE

S.J. PARK, D.W. KIM, J.H. KIM, J.H. CHUNG & J.S. LEE
Safety Division, National Disaster Management Research Institute, South Korea.

ABSTRACT

The ability to predict the future was considered a very important factor for humanity since long ago. Future prediction, which was non-scientific, took on significant developments with the advancement of science and technology. Nevertheless, predicting the future is still not an easy task. Therefore, it is more essential to develop diverse future scenarios for establishing policies with a clear vision on the personal, or even national levels rather than attempting to precisely predict a specific future event.

Particularly, future research plays a crucial role in the field of disaster management to prevent national crises. Future disasters could also result in an unimaginable scale of damages due to the complex network development of our society. Thus, it is necessary to develop scenarios in advance from the perspective of potential damages caused by disasters.

Future scenario development largely comprises quantitative and qualitative methods, which are applied identically in the field of disaster management. Quantitative method is developed using various statistical methods based on numerical data, while qualitative method is developed based on the intellect of a group of experts. In this study, the latter method is used because of the unpredictable nature of disasters. Furthermore, in order to provide a solution for the biased opinions that may occur from the group of experts, big data is used to propose a method for developing future disaster scenarios. The results from this method are preferred to efficiently develop future disaster scenarios, because the opinions of the group of experts are mostly biased.

Keywords: big data, cold wave, future disaster, future scenario.

1 INTRODUCTION

The future is filled with uncertainties, and it is nearly impossible to predict the future. Nevertheless, the human race has always had a desire to foresee the future since long ago because knowing the future was regarded as very important to ensure better lives for the people. In the past, predicting the future by means of astrology or palm reading was considered a non-scientific practice. However, due to the recent advancement of scientific technology, prediction of the future has become an area of academic study since not too long ago. Nevertheless, predicting the future is not an easy task. Many people are devoting their lives to studying the future to ensure better lives tomorrow. The future is uncertain. However, the process of establishing a vague outline of the future could assist in establishing policies with a clear vision on the personal and even national levels [1]. With respect to studying the future, it is necessary to devise a method to come up with diverse future scenarios rather than attempting to predict an uncertain future because scenario development is far more effective compared to proposing an absolute prediction of an uncertain future.

Such future scenario development is very necessary in the area of disaster management. Disasters could lead to major social chaos and national crisis. Thus, if we could devise strategies and policies to counter disastrous situations by developing future disaster scenarios, the rate of damage caused by disasters could be minimized. The rapid advancement of society is leading to more complex



This paper is part of the Proceedings of the International Conference on Big Data
(Big Data 2016)
www.witconferences.com

social situations and the scope of potential disasters is also becoming greater. Particularly, climate change, aging society, deterioration of facilities, and other such environmental/social trends are adding to the complexity of social changes. Such aspects could cause situations where traditionally experienced disasters lead to other unexpected disasters [2–4]. Take for example, the major earthquake in Eastern Japan in 2011, which caused a Tsunami and led to the Fukushima nuclear accident. This incident clearly shows us that a disastrous event in our society causes a chain reaction to increase the rate of the severity of damage. In other words, the disasters that may take place in the future will lead to consecutive damages due to the major, complex, and intricate network established in our society. Therefore, the task of developing future disaster scenarios requires the general understanding of disasters itself, as well as of the diverse mechanisms of our society.

There are many ways to develop future scenarios, but it could be largely divided into a quantitative method and a qualitative method [5]. Quantitative method refers to a method used to predict the future using various statistical methods based on numerical data. On the other hand, qualitative future scenario development method requires a group of experts from various fields to use their intelligence to develop future scenarios.

There are cases of future disaster scenarios that have been developed through qualitative and quantitative methods. First, the climate change scenario [6, 7] developed by IPCC is a chief example of a future scenario developed using statistical methods based on numerical data. Climate change scenario is a scientific future scenario developed based on the prospect of socio-economic advancement and greenhouse gas emissions. In addition, England developed a future flood scenario for as near as 2030 to as far as 2100 based on the greenhouse gas emissions scenario [8]. The advantage of such scenarios developed based on past data is that it is scientific and quantitative. However, it does have its limitations for studying future disasters to identify and foresee unexpected situations in the future.

On the other hand, the qualitative future disaster scenario has a risk register in the risk management program specified in ISO 31000, which develops future scenarios based on the intellect of expert groups. The major advantage of the risk register method is that it puts unexpected events in developing future disaster scenarios, and this method has been used in the United States, England, and Canada [9–11]. However, this method is solely dependent on human opinions, which means that there could be biased opinions depending on the group of experts. Nevertheless, the development of future disaster scenarios through a qualitative method based on the intellect of a group of experts is considered appropriate due to the nature adopted for predicting future disasters.

Then, how can we overcome the limitations present in developing future disaster scenarios using qualitative methods? We decided to incorporate the use of big data. That is, prior to depending solely upon the intellect of experts, rough scenarios were created using all data with regard to past disasters, which are available, for developing a final future disaster scenario. By doing so, we attempted to overcome the limitations present in the qualitative future disaster scenario development method.

2 PROCEDURE OF FUTURE DISASTER SCENARIO USING BIG DATA

This is an introduction of the procedure using big data, for future disaster scenario development. First, check for data that could be used for future disaster scenario development. Due to the nature of future disaster studies, non-definite data such as news contents are more useful compared to definite forms of data. The reason for using news data is because it clearly indicates the cause, damages, and cause-and-effect relationship, which are not indicated in various reports with respect to disaster events, and it also indicates the time and place of the disaster. Therefore, texts are retrieved from the accumulated news data of 29 domestic press agencies based on keywords related to the disaster scenario development.

News data categorized by science, economy, society, etc. are processed through a stage known as text parsing. The parsing process divides the news text sentences into words, and the words are divided into parts of speech. It is then saved as data. The divided text data is put through a filtering process to extract the desired information. Filtering excludes unnecessary terms and parts of speech from the words divided into parts of speech for analysis and generates a statistical value of the words from the news data. These two processes create and save a table that illustrates the role and frequency of words, and the information of the relevant document (Fig. 1). In addition, this process allows for topic modeling, which illustrates the main topic of the numerous texts, as well as network analysis, which observes the relationship between words. SVD algorithm is used in the Topic modeling stage. SVD stands for Singular Value Decomposition, which is created based on the eigenvector value attributable to words categorized according to specific topics in the term-document matrix saved in a table format, as shown in Fig. 1 (Fig. 2). Through this method, topic modeling and network analysis are possible. This technique of acquiring desired information through a series of text-based processes – parsing, filtering, topic modeling, and network analysis – is called Text mining.

We primarily created a causal loop draft, as shown in Fig. 3, using the data acquired through text mining based on news text data accumulated over the past 10 years [12]. Although it requires redundant work of having to review the news contents once again, analyzing the text filtered through text mining is not very time consuming. While creating a causal loop draft based on the text mining results, we organized a group of experts from areas of related keywords based on the generated draft. It is preferable to have experts from diverse fields with extensive expertise. A part of the group of experts must include individuals who can understand the general overview of the causal loop to

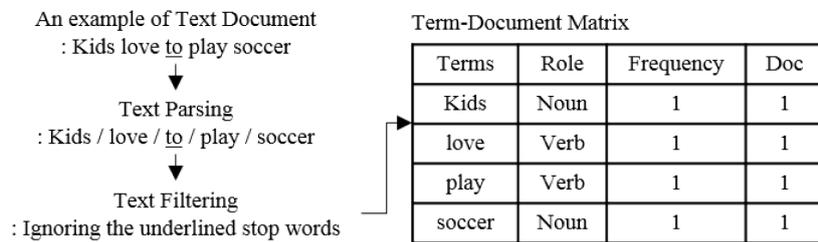


Figure 1: An example of text parsing and text filtering.

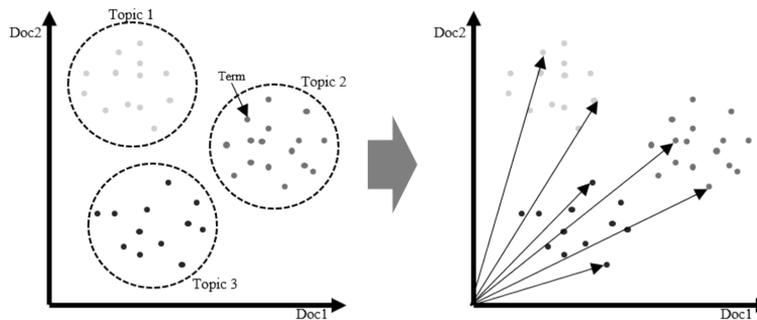


Figure 2: Description of SVD algorithm for Topic Modeling.

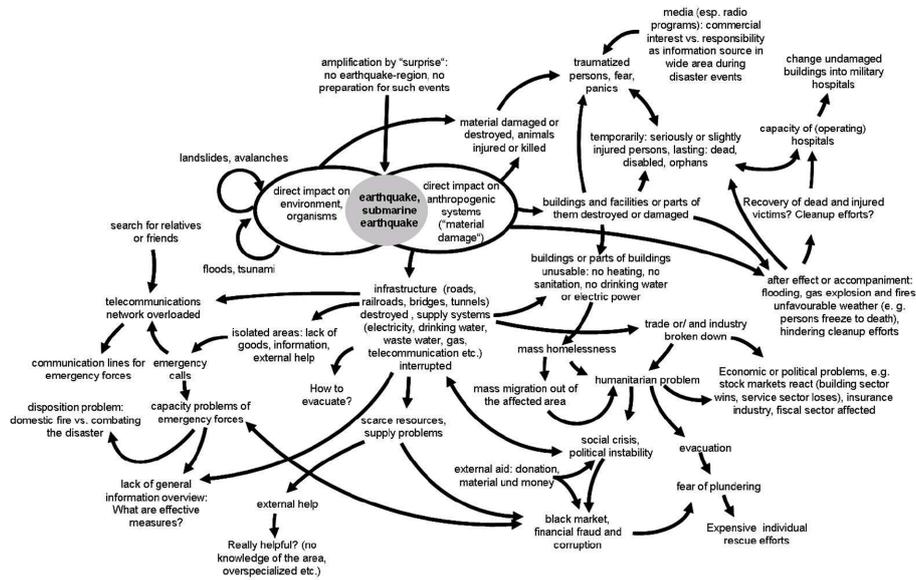


Figure 3: Causality loop of earthquakes (Helbing et al. [12]).

propose a direction for the scenario. The intellect of the organized group of experts is used to produce a final causal loop through several updates of the causal loop draft, based on which a future disaster scenario is proposed.

3 FUTURE DISASTER SCENARIO ON EXTREME COLD WAVE

This study introduces an example of winter cold wave related future disaster scenario development. As described previously, first and foremost, all data required to develop a cold wave related future disaster scenario were reviewed. With regard to the cold wave related big data, texts from news data concerning annual records of disaster reports and the resulting damage expenses and casualties accumulated from 29 press agencies from 2004 to 2015 were collected. Disaster reports were used for the purpose of understanding the situation at the time, and news data were generally used to develop the future disaster scenario. Thus, approximately 48,000 news articles were extracted from 10 years of news data based on the keywords ‘cold wave’ and ‘heavy snowfall’. Then, the extracted texts were put through processes of text parsing, filtering, topic modeling, and network analysis. We created a cold wave related causal loop draft based on the information acquired through the text mining process (Fig. 4).

We categorized the elements affected by cold waves based on this draft. Based on the categorized areas, experts from academic circles and research organizations who were involved in cold wave related researches were hired. A group comprising 19 experts were organized from diverse areas (e.g. transportation, agriculture, health, climate, administration, etc.) (Table 1). The group included individuals who had an understanding of the general overview of the cold wave related causal loop.

The intellect of the organized group of experts has been used to develop a final causal loop through numerous updates of causal loop draft (Fig. 5). During the process of taking advantage of the intellect of the group, we conducted interviews with industry experts considering the risk of unprecedented

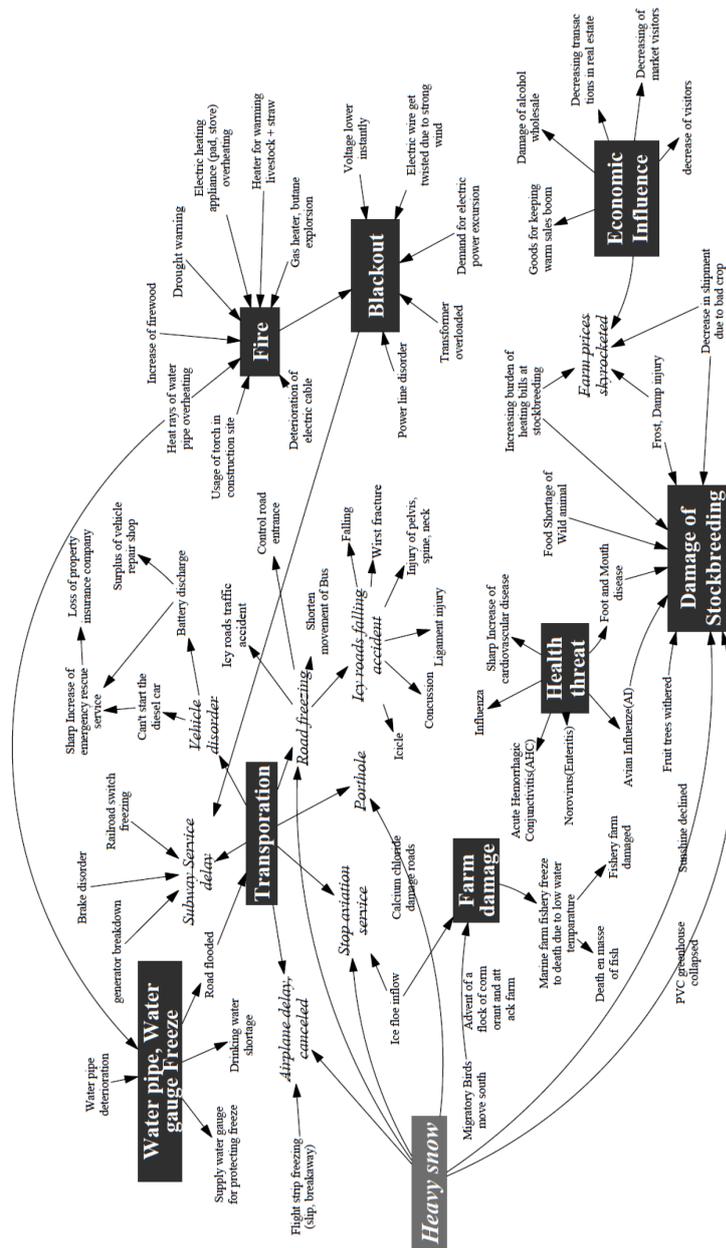


Figure 4: Causal loop draft related with extreme cold wave.

Table 2: Proposal of notable future disaster scenarios based on causal loop.

Field	Disaster scenario
Transportation	<ul style="list-style-type: none"> • Road standstill / Reduction of snow removal resources / Degradation of restoration equipment performance → Problems in supplying resources → Increased potential of isolations on roadways • Icy road → Difficulties in snow removal and restoration → extensive time consumption for restoring → Additional damages and losses → Large-scale traffic chaos
Transportation & Industry	<ul style="list-style-type: none"> • Strong cold wave and heavy snow → Logistics transportation system errors → Reduced sales • Strong cold wave and heavy snow → Difficulty in commuting → Work hour reduction → Reducing productivity
Agriculture & Fishery	<ul style="list-style-type: none"> • Temperate fruit tree growing region and warm current fish species are moving and expanding to the north → Severe unexpected cold waves → Large scale mortality due to relative vulnerability to low temperatures
Health & Environment	<ul style="list-style-type: none"> • Human body adapting to warmer climates → Unexpected cold waves → Significant shock • Increase in the number of the senior / Rapid increase of air pollutant emission from China during the winter season → Northwesterly winds blowing → Increase the rate of damage → Increased number of patients → Medical facilities overloaded → Secondary damages
Facilities	<ul style="list-style-type: none"> • Industrial facilities and infrastructures developed during the 1970s → Deteriorating over time → Heavy snowfall → Collapse accidents • After the winter season → Frozen facilities begin to melt and thaw → Collapsing buildings and facilities
Facilities & Forestry	<ul style="list-style-type: none"> • Cold waves → Freezing reservoirs in mountain / Rupturing water mains → Potential shortage of drinking water • Accompanied by heavy snow, cold wave → Difficulty in entering the mountain of large fire trucks → Difficulty of emergency water supplies
Forestry	<ul style="list-style-type: none"> • Frequent droughts due to climate change → Decreasing of the rate of precipitation in winter and spring → Reduced amount of water in catchments → Difficulty in procuring water to suppress forest fires • Cold wave → Freezing the water catchments → Making the situation even worse • Accompanied by gusty winds, recent cold waves → Raising further concerns for major forest fires.
Energy	<ul style="list-style-type: none"> • Sever winter drought → Hydroelectric power suspension → Winter energy crisis • Severe unexpected cold wave → Increased demands of gas energy → Winter energy crisis

cold wave disasters that may be caused by global megatrends such as climate changes, aging societies, urbanization, etc. Notable future disasters scenario based on the final causal loop is proposed in Table 2 shown below.

4 DISCUSSION & CONCLUSION

Based on the results of the future disaster scenario using big data, we were able to conclude that a single influential factor could cause chain reactions in various unexpected areas. This research also demonstrated that text-based big data could be used for creating future disaster scenarios. This aspect played a crucial role supplementing the limitations of a future scenario through qualitative methods based on the intellect of an expert group. In other words, it gave consideration to the biased opinions found in qualitative methods resulting from solely depending on the intellect of the expert group. Also, a scenario based on big data was efficient in the fact that it minimized the time and scope of organizing the expert group. Based on in-depth interviews with the expert group, they pointed out that the draft scenario created based on news data played a decisive role in completing the future disaster scenario.

Nevertheless, the results of the future disaster scenario proposed in this study are still insufficient to be used as effective policy support tools to persuade policy-deciding authorities. Support tools developed using quantitative methods based on non-definite and definite data are required to support such shortcomings. In future studies, numerical statistics method will be used in addition to develop future disaster scenarios to be used as effective tools for decision making.

REFERENCES

- [1] Choi, H.S., Theory and methodology of futures study. *Korean Sociological Association*, **13**(1), pp. 197–230, 2012. (In Korean).
- [2] Hur, J.Y. & Lee, J.H., Study on building a disaster management system responding to future disaster -using expert survey method. *Crisis and Emergency Management*, **10**(10), pp. 173–195, 2014. (In Korean).
- [3] Boin, A., Ekengren, M. & Rhinard, M., Protecting the union: analysing an emerging policy space. *Journal of European Integration*, **28**(5), pp. 405–421, 2006.
<http://dx.doi.org/10.1080/07036330600979573>
- [4] Quarantelli, E.L., Lagadec, P. & Boin, A., A heuristic approach to future disasters and crises: new, old, and in-between types. *Handbooks of Disaster Research*, Springer: New York, pp. 16–41, 2007.
- [5] Kosow, H. & Gaßner, R., Methods of future and scenario analysis; overview, assessment, and selection criteria. *German Development Institute*, **39**, pp. 1–133, 2008.
- [6] IPCC, IPCC Fourth Assessment Report: Climate Change 2007 (AR4), 2007.
- [7] IPCC, IPCC Fifth Assessment Report, 2014.
- [8] Government Office for Science, Future Flooding: executive summary, 2004.
- [9] London Resilience Team of Greater London Authority, London Community Risk Register, 2011.
- [10] Public Safety Canada, All Hazards Risk Assessment Methodology Guidelines, 2012.
- [11] Homeland Security, Threat and Hazard Identification and Risk Assessment Guide – Comprehensive Preparedness Guide (CPG) 201, 2nd edn, 2013.
- [12] Helbing, D., Ammoser, H. & Kühnert, C., Disasters as extreme events and the importance of network interactions for disaster response management. *Extreme Events in Nature and Society*, pp. 319–348, 2006.
http://dx.doi.org/10.1007/3-540-28611-x_15