Development of a Web-based Trend Analysis System of Earthquake Disaster Researches Presented at the Past World Conferences on Earthquake Engineering

Shinya Kondo
Project Researcher, International Center for Urban Safety Engineering, Institute of Industrial Science, The University of Tokyo, Japan

Kimiro Meguro
Professor, Director of International Center for Urban Safety Engineering, Institute of Industrial Science, The University of Tokyo, Japan

SUMMARY:
In this study, the authors propose a web-based system aiming to analyse trends of past disaster researches. In particular, the authors plan to build "research paper database" and a "collection of keywords." The research paper database will consist of contents of research papers published in the Proceedings of the past World Conferences on Earthquake Engineering (WCEEs) with information regarding the authors. Each keyword in the collection of keywords will describe the research field on disaster and has set indexes such as "type of disaster," "type of countermeasures," and "type of object of damage." The disaster list will consist of disaster happened in the past and in the future. Furthermore, the proposed system has three functions and a full-text search of a research paper database: "trend analysis of disaster research," "impact analysis of disaster," and "search for related research."

Keywords: trend analysis, web-based system, the World Conference on Earthquake Engineering

1. INTRODUCTION

In our world, many disasters (earthquake, volcanic disaster, flood, etc.) occurred and cause extensive damage. Till date, many researchers from various academic fields have studied disaster and disaster countermeasures, aiming to decrease negative impact due to disasters.

Disaster research has wide fields from pure to practical sciences, and structural and non-structural measures. Researchers tend to make presentations on their research results at the conferences and meetings that their primary academic societies hold. Therefore, even if there are many useful research outcomes for disaster management, it is very difficult to overlook them from bird-eye as the number of societies is many and they are divided into many small fields. There is a need to establish an environment to overlook all related researches from bird-eye and share them.

In recent years, the service for research result searching and sharing has been developed in various fields and it becomes possible to check and get necessary information from libraries by remote. Google has opened a service of Google Scholar and it provides a simple way to broadly search for academic literature. From the Google Scholar, users can search across many sources in different fields: articles, papers and reports, books, abstracts and opinions, from academic publishers, professional societies, online repositories, universities and other web sites. Thomson Reuters is offering Web of Science. This service, which provides users with a search function with high speed citations and can easily identify author information can be easily identified, has been utilized researchers, administrators, university faculty, and students. And this service can be subject to important publications and influential journals more than 11,600 around the world dating back to 1900, to analyze the citation pattern to perform a search across the fields.

As above mentioned, general environment that can share the results of individual studies has been established. However, in order to use efficiently disaster research results for disaster management

around the world, we should grasp an overall picture and characteristic of disaster related researches.

In this study, as the first step of the research, the authors propose a web-based system aiming to understand an overall picture of earthquake disaster researches. This system has a function to analyse trends of past researches from the various viewpoints of disaster management. Furthermore, users can grasp the impact on the study of disasters, and it becomes possible to search related researches. In this paper, the authors try to build a prototype of proposed system using the research papers published in the proceedings of all the past World Conferences on Earthquake Engineering (WCEEs), and some results of analysis are introduced.

2. BUILDING DATABASE FOR PROPOSED SYSTEM

2.1. Research Paper Database

In order to make full-text search and mining of research papers and reports possible, we have established a research paper database using research papers published in the proceedings of all the past WCEEs. These papers are obtained from the website which National Information Centre of Earthquake Engineering (2009) has opened. Table 2.1 shows the number of research papers published in the proceedings of each conference. Each research paper is saved as a PDF file, and text data were extracted using “Adobe Acrobat 9.0” as OCR software.

Table 2.1. The Number of Research Papers Published in the Proceedings of the Past WCEEs

Conference	*Number of papers*
1st, United States, 1956	42
2nd, Japan, 1960	134
3rd, NewZealand, 1965	165
4th, Chile, 1969	167
5th, Italy, 1974	452
6th, India, 1977	669
7th, Turkey, 1980	743
8th, United States, 1984	845
9th, Japan, 1988	1004
10th, Spain, 1992	1191
11th, Mexico, 1996	1437
12th, New Zealand, 2000	1519
13th, Canada, 2004	2341
14th, China, 2008	3008

2.2. Collection of Keywords

In this section, the collection of appropriate keywords from disaster management viewpoints set in advance is described. Keywords are selected based on disaster prevention glossary in Japan and added from the proceedings of the disaster-related academic societies in Japan. Furthermore, the opinions of researchers in disaster fields are reflected to the collection of keywords. Keyword that is difficult to recognize the meaning in terms of disaster management in a single word, such as "rock", "water", "stress", etc. are removed. And keyword such as "disaster management" that has a broad concept and is difficult to classify research fields are excluded. Finally, the total number of keywords is 1,392.

To each keyword, indexes, such as "type of disaster," "type of countermeasures," and "type of object of damage" are set as shown in Table 2.2. "Type of disaster" means a kind of hazards, such as earthquake, volcano, flood, etc. that may become cause of damage and disaster (Table 2.3). "Type of measures" is a classification based on the concept of disaster life cycle (Meguro and Murao, 2008) as shown in Table 2.4. "Type of object of damage" means the target of disaster countermeasures and object of damage (Table 2.5).

Table 2.2. Example of indexes setting

Keywords	*Disaster*	*Countermeasures*	*Disaster effects*
Tsunami warning	Earthquake (Tsunami)	Disaster prediction and warning	
Disaster prevention learning		Preparation	
Dangerous area of slope failure		Mitigation	Natural slope (Landslide)

Table 2.3. The number of keywords set with the indexes "Type of Disaster"

Index name	*Number of keywords*
Volcanic disasters	58
Storm and flood	167
Earthquake	307
Large scale fire	126
Man-made disaster	10
International disaster relief	10
Others	8

Table 2.4. The number of keywords set with the indexes "Type of Countermeasures"

Index name	*Number of Keywords*
Hazard mechanism	288
Damage mechanism	313
Mitigation	153
Preparedness	105
Prediction and early warning	27
Damage assessment	49
Emergency disaster response	72
Recovery	36
Reconstruction	48
Information and communication	26

Table 2.5. The number of keywords set with the indexes "Type of object of damage"

Index name	*Number of keywords*
Natural slope (Landslide)	33
Earth structure（Embankment）	16
Ground deformation (Liquefaction)	10
Debris flow	13
Structure damage (Civil infrastructure)	55
Structure damage (Building)	53
Structure damage (Others)	3
Fire spread	12
Lifeline and system	21
Transportation (road, railway, harbor)	12
Human loss	13
Affected people (refugees)	24
Industrial damage (Business interruption)	11

3. DISTRIBUTION OF FIELD OF RESEARCH PAPERS

In this chapter, the distribution of research papers is calculated. The full-text search of the research paper database extracts keywords which consist of the collections of keywords. In this study, MeCab (Kyoto University and NTT communications) is used for full-text search. MeCab is the software for morphological analysis of Japanese text.

One research paper has been allocated to the research field based on the number of appearance of the keyword. In this study, this value is defined as the hitting ratio (Eqn. 3.1). Research fields are composed by items in two indexes. Two indexes (index “a” and index “b”) are selected from “type of disaster,” “type of countermeasures,” and “type of object of damage.”

$$\sum_i \sum_j m_{i,j,l} = 1 \quad (3.1)$$

($m_{i,j,l}$: Hitting ratio in item "i" (index "a") and item " j" (index "b") of paper "l")

Hitting ratio of the research paper is expressed as a product of the value of item "i" of index "a" and item "j" of index "b". This value is the number of appearance of the keyword divided by the total number of keywords (Eqn. 3.2).

$$m_{i,j,l} = \frac{x_{i,l}}{\sum x_{i,l}} \times \frac{y_{j,l}}{\sum y_{j,l}} \quad (3.2)$$

$$x_i = \frac{\text{The number of appearance of keywords (index "a", item "i")}}{\text{The total number of keywords (index "a", item "i")}}$$

$$y_j = \frac{\text{The number of occurrences of keywords (index "b", item " j")}}{\text{The total number of keywords (index "b", item " j")}}$$

The sum of hitting ratio of the research paper in a proceedings is the number of research papers considering hitting ratio in a proceedings (Eqn. 3.3).

$$M_{i,j} = \sum_{l=1}^{N_{ab}} m_{i,j,l} \quad (3.3)$$

($M_{i,j}$: The number of research papers considering hitting ratio in a proceedings in item "i" (index "a") and item " j" (index "b"))

(N_{ab} : The number of research papers having keywords (index "a" and index "b"))

The number of research papers considering hitting ratio in a proceedings divided by the number of research papers with keywords (index "a" and index "b") is the hitting ratio of a proceedings (Eqn. 3.4).

$$P_{i,j} = M_{i,j} / N_{ab} \quad (3.4)$$

($P_{i,j}$: Hitting ratio of a proceedings in item "i" (index "a") and item " j" (index "b"))

4. OUTLINE OF FUNCTIONS

The proposed system has three functions to analyse trends of past researches from the viewpoint of disaster management: “trend analysis of disaster research,” “impact analysis of disaster,” and “search for related research.”

4.1. Trend Analysis of Disaster Research

Using the function “trend analysis of disaster research,” users can see the hitting ratio of researches published in a proceedings. Figure 4.1 shows a distribution diagram of the hitting ratio of a proceedings. X-axis is set index "a", and Y-axis is set index "b". In this figure, the hitting ratio (%) of a proceedings in item "i" (index "a") and item "j" (index "b") is described at the respective area composed of item "i" and item "j". Area of the circle in respective area is proportional to the hitting ratio. In this study, three indexes are set. Therefore, users can analyse the trend of disaster research from the following three viewpoints.

- Analysis A: “type of disaster” and “type of countermeasures”
- Analysis B: “type of countermeasures” and “type of object of damage”
- Analysis C: “type of disaster” and “type of object of damage”

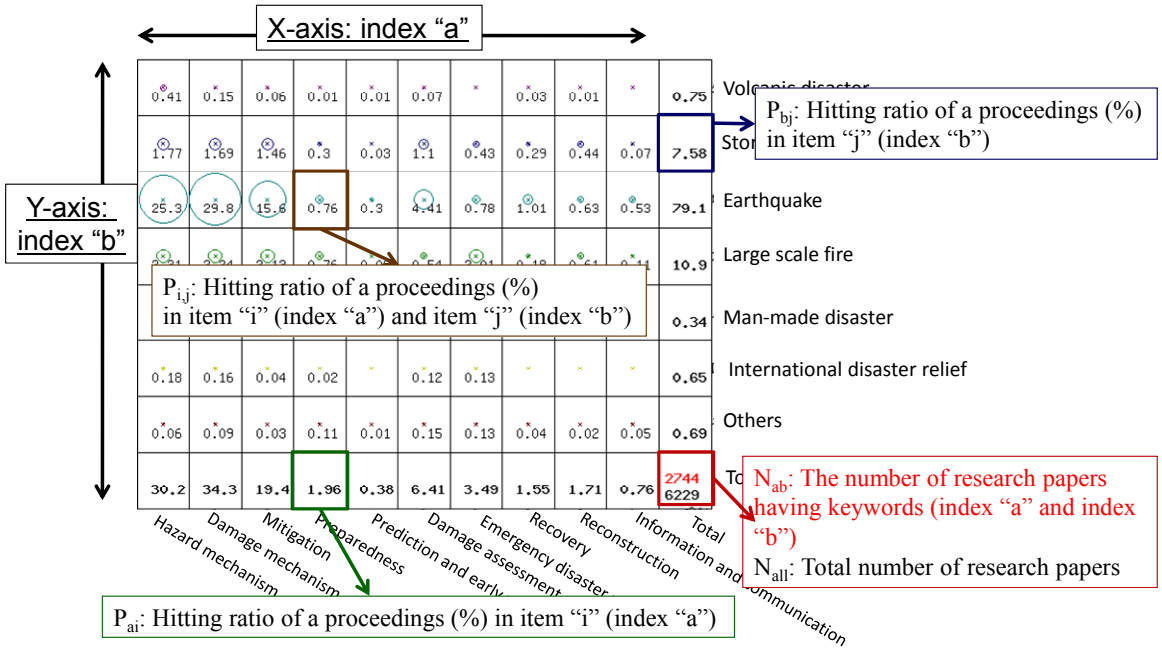

Figure 4.1. A distribution diagram of the hitting ratio of a proceedings

4.2. Impact Analysis of Disaster

In order to prepare for future disasters, it is necessary to organize problems and research results about past disasters. “Impact analysis of disaster” is intended to analyse how long one disaster keeps impact on disaster research. Using this function, users can see the time series variation of the number of research papers related to past disasters. Specifically, a "disasters list" which is the summary of the name of past disasters is built, and research papers with the name of disaster are extracted by the full-text search of research paper database. In this paper, earthquake disasters occurred since 1900 killing more than a thousand people (United States Geological Survey: USGS) are used to a "disasters list". The number of disasters extracted is 125.

4.3. Search for Related Research

In order to take advantage of a research paper on disaster countermeasures, it is necessary to extract also research papers related to its contents. The function “search for related research” is intended to search related research papers which have keywords or the name of past disasters. When users view the page of a certain research paper, they can see the list of research papers related to its contents.

5. APPLICATION (WCEE)

In this paper, two results are described. The first is the result of trend analysis of disaster research papers published in the proceedings of the all past WCEEs using the function "trend analysis of disaster research". The second is the extraction result of research papers related to the past earthquake disaster using the function “impact analysis of disaster.”

5.1. Trend Analysis of Disaster Research

Figure 5.1 shows distribution diagrams of the hitting ratio from the 1st to the 14th WCEE proceedings. The X-axis shows "type of countermeasures" and the Y-axis shows "type of disaster effect." Contents observed from these distribution figures are as below.

- In all conferences, the main research fields are composed of "Structure damage (building)" and four types of countermeasures such as "Hazard mechanism", "Damage mechanism", "Mitigation", and "Damage assessment."
- Research field of WCEE is getting wider comparing to before.
- The most popular research field consists of "Structure damage (building)" and "Damage mechanism" until 2000, since 2004, it has become "Structure damage (building)" and "Mitigation."

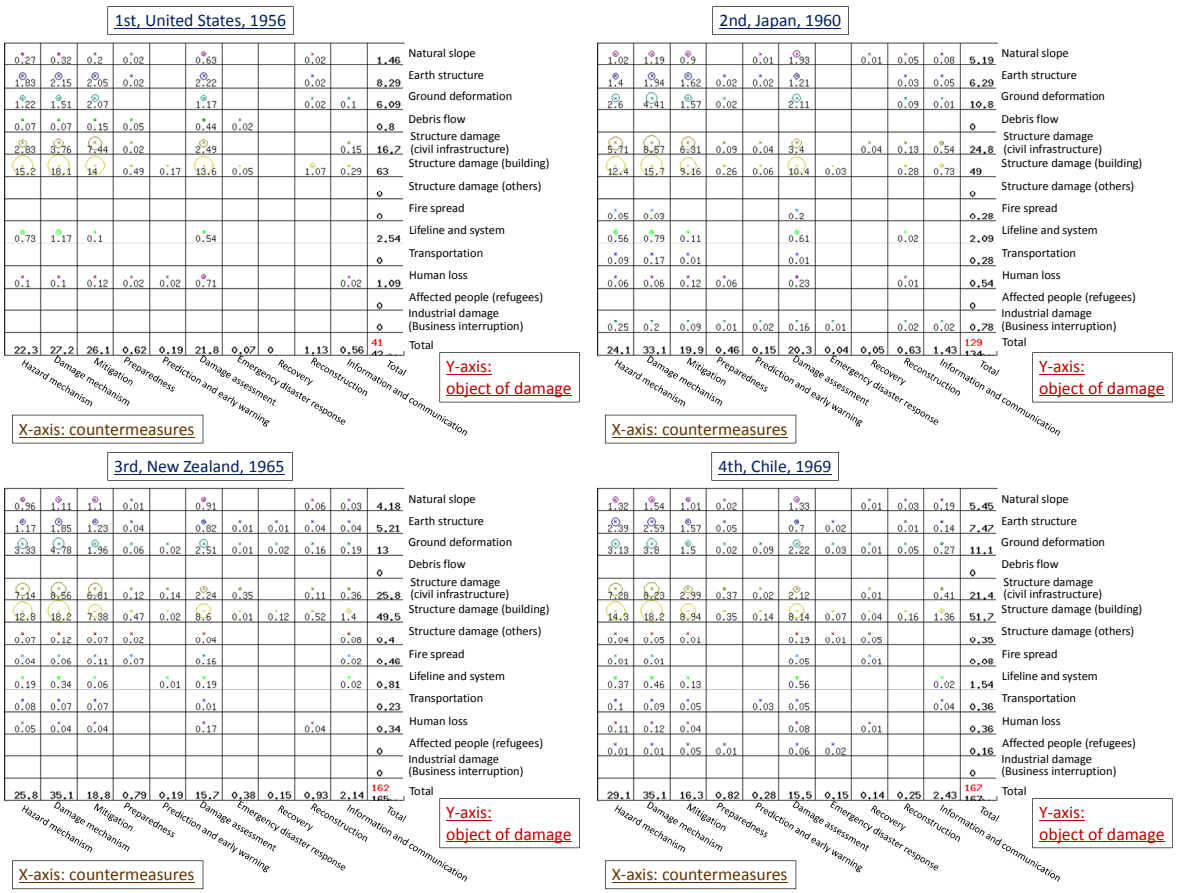

Figure 5.1 (a). Distribution diagrams of the hitting ratio of all the past WCEE proceedings

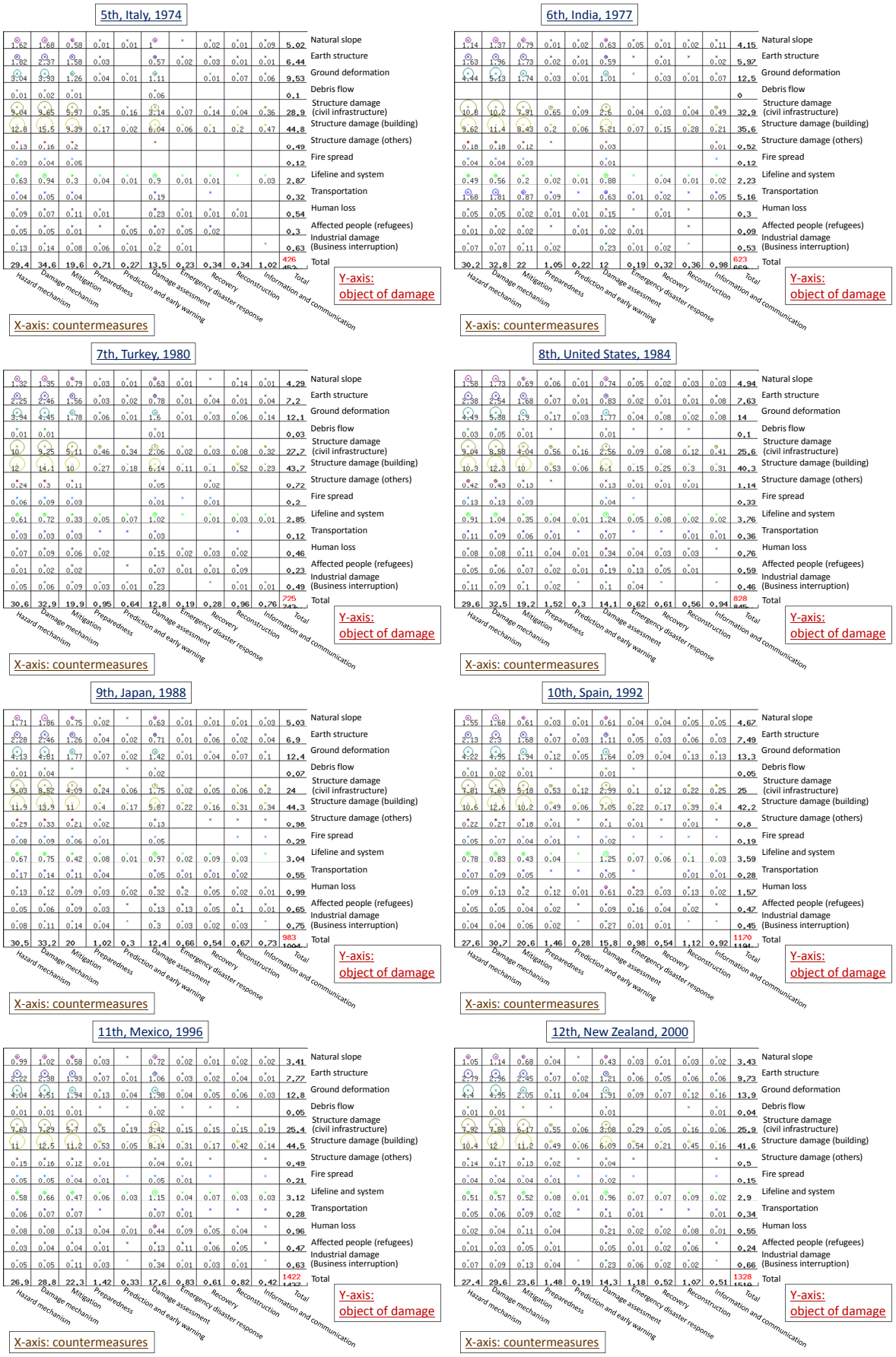

Figure 5.1 (b). Distribution diagrams of the hitting ratio of all the past WCEE proceedings

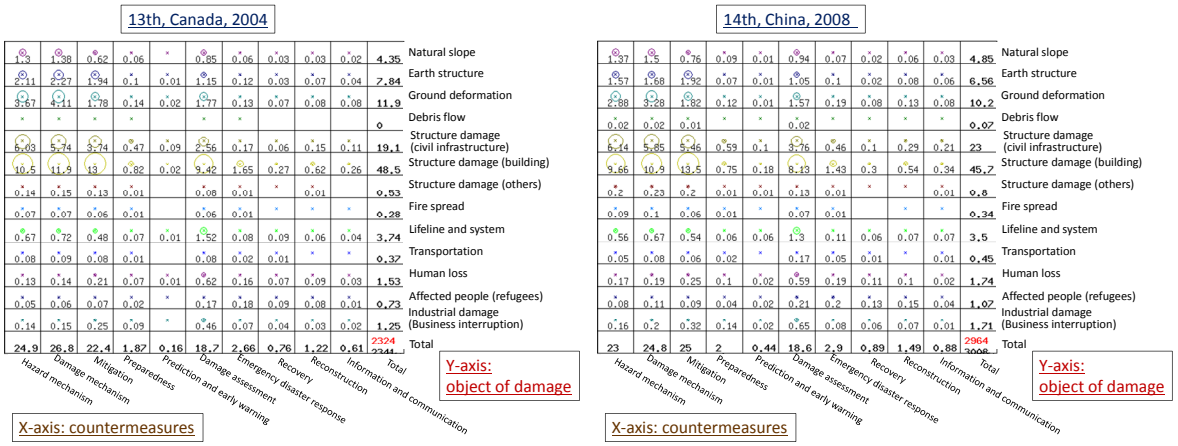

Figure 5.1 (c). Distribution diagrams of the hitting ratio of all the past WCEE proceedings

5.2. Impact Analysis of Disaster

Figure 5.2 shows the number of research papers related to the past earthquake disaster in descending order using the function "impact analysis of disaster." This is intended for disaster with 5 papers or more. The number of research papers related to the 1995 Kobe Earthquake in Japan is the largest. And other earthquakes in Japan such as the 1923 Kanto Earthquake and the 1948 Fukui Earthquake are well-researched.

Figure 5.3 shows the process of the number of research papers which are presented at each conference. The target earthquake disasters have related research papers more than 20. Study of the 1995 Kobe Earthquake has been an increase in the number of research papers from 1996. Research about earthquake that occurred before the 1st conference tends to be continuously until 14th conference. Many researchers presented research papers related to the 1976 Tangshan Earthquake at the 14th China conference.

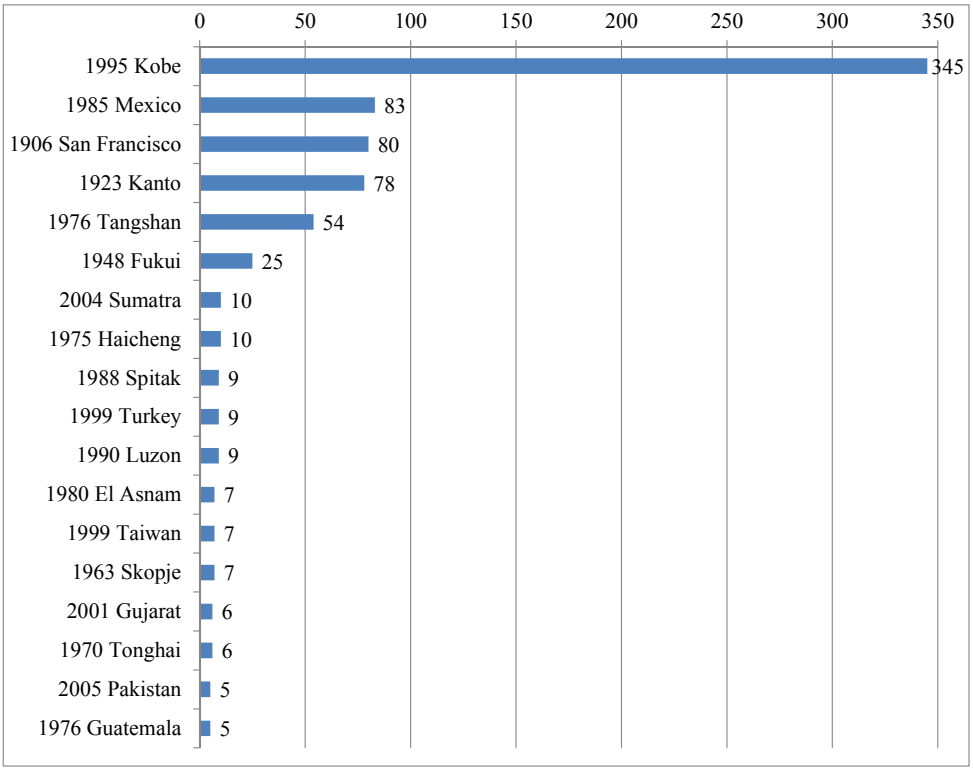

Figure 5.2. The number of research papers related past earthquake disaster

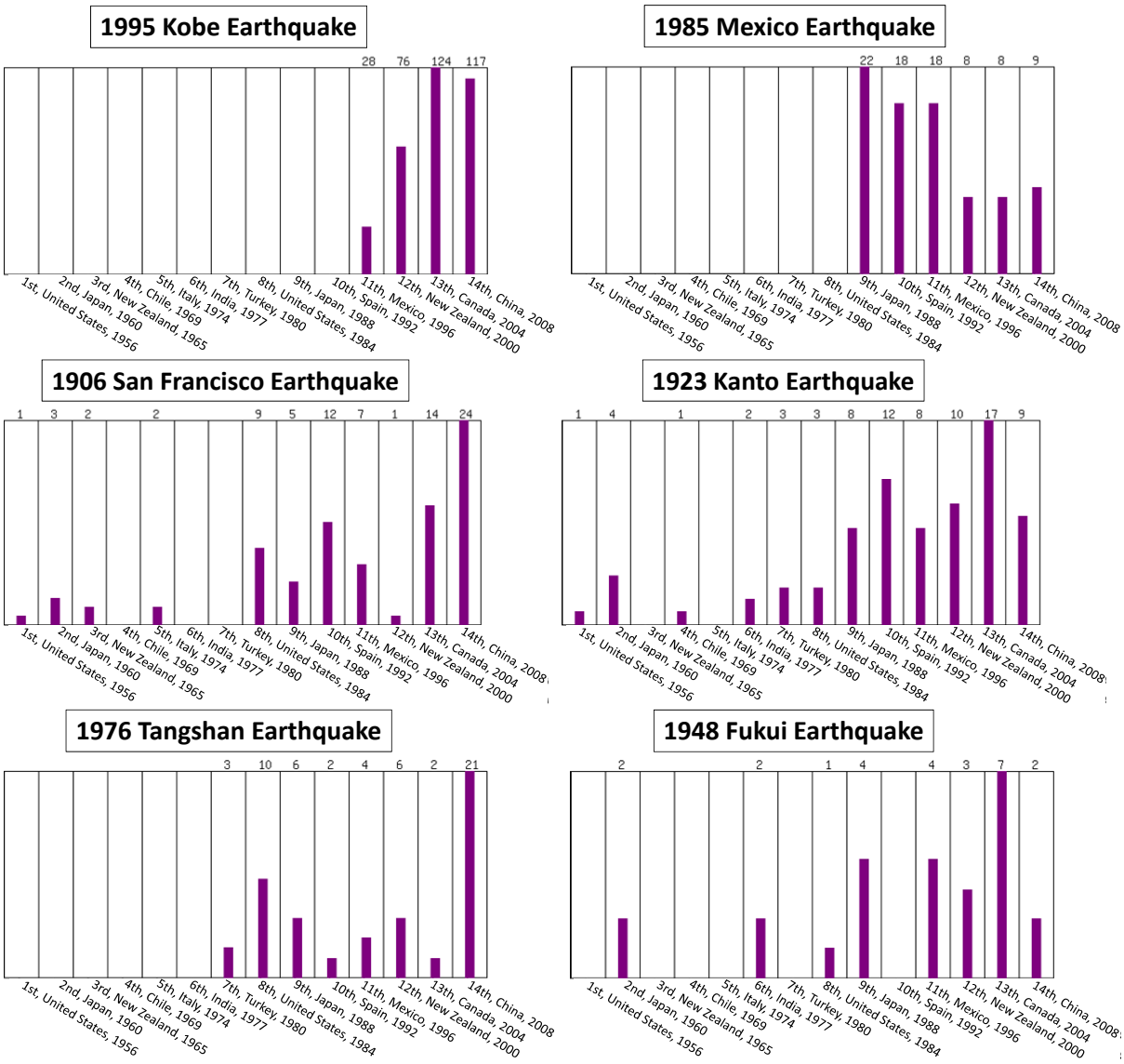

Figure 5.3. The process of the number of research papers which are published in each proceedings

6. CONCLUSIONS

In this study, the authors proposed a web-based system aiming to understand and grasp overall picture of disaster researches from bird-eye. Using this system, users can analyse the trend of disaster research and grasp the changes of the number of research papers related to past earthquakes. In this paper, the authors tried to apply the proposed system to analyse all research papers published in the proceedings of all past WCEEs. As a result, the trend of disaster research field and research on past disasters from the 1st to the 14th conferences become clear. In the future, the proposed system will be improved to assist implementation of proper disaster countermeasures using disaster research database.

REFERENCES

Google (2004). Google Scholar, http://scholar.google.co.jp/schhp?hl=en

Thomson Reuters. Web of Science, http://apps.webofknowledge.com/WOS_GeneralSearch_input.do?locale=en_US&last_prod=WOS&SID=Y1p%40DG%40EoaPf863f8gk&product=WOS&highlighted_tab=WOS&search_mode=GeneralSearch

National Information Centre of Earthquake Engineering (2009). WCEE ONLINE PROCEEDINGS, http://www.nicee.org/wcee/

Kimiro Meguro and Osamu Murao (2008). Urban city and disaster prevention, The Open Univeristy of Japan, Japan.
Kyoto University and NTT Communications. MeCab: Yet Another Part-of-Speech and Morphological Analyzer, http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html
United States Geological Survey (2012). Earthquakes with 1,000 or More Deaths since 1900, http://earthquake.usgs.gov/earthquakes/world/world_deaths.php