Case Database Construction and Temporal-spatial Analysis of Mass Incidents in China Based on GIS

Lida Huang, Jianguo Chen, Zewei Zhang, Hongyong Yuan,
Qing Deng and Feng Sun

March 16, 2018

Case Database Construction and Temporal-spatial Analysis of Mass Incidents in China Based on GIS*

Lida Huang
Institute of Public Safety Research,
Department of Engineering Physics,
Tsinghua University
Beijing 100084, China
hld14@mails.tsinghua.edu.cn

Jianguo Chen
Institute of Public Safety Research,
Department of Engineering Physics,
Tsinghua University
Beijing 100084, China
chenjianguo@tsinghua.edu.cn

Zewei Zhang
Institute of Public Safety Research,
Department of Engineering Physics,
Tsinghua University
Beijing 100084, China
sdtazzw@126.com

Hongyong Yuan
Institute of Public Safety Research,
Department of Engineering Physics,
Tsinghua University
Beijing 100084, China
hy-yuan@tsinghua.edu.cn

Qing Deng
Institute of Public Safety Research,
Department of Engineering Physics,
Tsinghua University
Beijing 100084, China
dq13@mails.tsinghua.edu.cn

Feng Sun
College of Info Sciences & Tech
Penn State University
State College 16803, USA
feng@psu.edu

ABSTRACT

As one· of the important research areas of public safety and sociology, the study of mass incidents is inseparable from the support of a large number of real cases data. However, there is a lack of available and reliable mass incidents databases in academic circles in China. Using the news retrieval approach, this paper builds a case database of mass incidents in China from 1998 to 2014 based on the GIS system. Then the spatial-temporal distribution rules of these mass incidents data are analyzed from the perspective of geography. Our research effectively integrates the temporal attributes, spatial attributes, organizational attributes and violence attributes of mass incidents. The results show the hot areas and temporal distribution of mass incidents and mass violence, and reveals the main triggers of these events.

CCS CONCEPTS

• GIS → Information systems → Database → Statistics

KEYWORDS

Mass incidents; Case Database; Spatial-temporal distribution; Triggers; Mass Violence.

1 INTRODUCTION

During the period of social transition, China has been faced with a lot of mass incidents caused by different social contradictions. These events not only have a negative impact on social stability, but also put greater demands on government administration. In recent years, more and more researchers have begun to analyze mass incidents based on case study. For examples, Cai [1] and Su [2] discussed the labor protests and pension protests followed the SOE reforms; Guo [3] and Ong [4] discussed the mass incidents caused by land disputes; Tong [5] and Deng [6] analyzed the environmental protests over serious pollution; Ma [7] analyzed the mass incidents caused by the city management officials (known as ChengGuan in Chinese). All these researches have advanced our understanding of mass incidents in China. However, as most works were based on case studies or small-scale opinion surveys, the existing literature has left ample space for further inquiries. We remain unable to get a big picture of the mass incidents in China. What is the spatial-temporal distribution and of mass incidents? What is the proportion of each type of the incidents on a national scale? Is there any rule of mass violence? Answering these questions must rely on a comprehensive case database of mass incidents.

Up to now, neither in political circles nor academic circles, there is no one who has published any database on mass incidents in China. By consulting and reading literatures, this study summarizes the existing databases related to mass incidents. The existing databases can be divided into two categories: statistical databases which only count the number of daily, monthly or annual mass incidents, and case databases which record the occurrence time, location, size, trigger factor, violence level and other attributes of each incident. For statistical database, "The Social Blue Book" [8], published by the Chinese Academy of Social Sciences, estimated the annual number of mass incidents in China from 1993-2009; "Report on China's Legal Development 2014" [9] counted 871 mass incidents from 2000 to 2014; Hu [10] counted over 11800 mass incidents from 2005 to 2015; Leetaru [11] counted over 100,000 protests from 1979 to now. Since the statistical methods are different, these databases are significantly

different. Hu and Leetaru only counted the number of the news and didn't check the repetitiveness, so their data is quite larger than others. As these databases only record number and few attributes like location and time, it is hard to do deep analysis. For case databases, Xiao [12] arranged 1189 mass incidents from 1975 to 2009; Shao [13] sorted out 5708 mass incidents. However, these databases are not open to public. Scholars in Harvard University published 1008 strikes from 2004.03 to 2012.12 in China. This database has high research value, but only the strike data is not comprehensive enough for mass incident study. Therefore, it is necessary to build an available and scientific case database of mass incidents in China.

In this paper, we develop a case database of mass incidents from 1998 to 2014. Then, based on the GIS system, we reveal the spatial-temporal distribution rules of mass incidents in China. We also analyze the triggers and mass violence, which can help to prevent and control mass incidents. The remainder of this paper is organized as follows. The next section describes the construction methods of our database and shows the database framework. In Section 3, the geo-statistical results of the database are presented and the incident triggers and causes of mass violence are discussed. Then, Section 4 concludes.

2 CASE DATABASE

2.1 Data Collection

Refering to reference [13], the data collection process includes three steps, keywords generation, news retrieval and data arrangement. First, we manually select 48 news reports of mass incidents from 1998 to 2014 and extract keywords from them. Keywords generation consists of words segmentation and frequency count. The keywords used in this paper includes nouns (eg. event, police, police stations, Public Security Bureau, government agencies, arrears of wages, criminals, rumors, disputes, mass, unclear truth, anti-law) and verbs (eg. hit, incite, beat, onlookers, gather, conflict, disturbance, trouble, siege, calm down). Then, based on these keywords, we use WiseSearch, a news search platform which can obtain over 4000 kinds of news reports including web news and newspapers, to grab related news. The last step is data arrangement. The initial search for mass incidents is not all in line with the requirements. Therefore, we sort out the database in detail, deleting irrelevant cases and duplicate cases, and finally obtain a case database containing 3417 mass incidents.

2.2 Database Framework

After gathering the news data, we need to encode them. The first step is building the database framework. The framework includes four aspects, background factors, organization factors, police disposal factors and results factors. In this paper, we focus mainly on triggers and mass violence. Triggers mean the causes of mass incidents. There are 16 triggers classified in our study: law enforcement disputes (except for disputes caused by city management officials, which are listed individually), accidents (like traffic accidents, medical accidents, industrial accidents and accidental deaths), labor disputes, disputes caused by city management officials, disputes caused by programmes or resources, land disputes, property disputes, grass-roots governance, industrial actions (like taxi strikes), business management disputes, environmental issues, flight delays, sport events (like fans fights), lootings, abreacting mood after sever disasters and ethnic or religious issues. Mass violence refers to the extent of violence of group behaviors, which can be divided into two classes. One is non-violent actions like marches, demonstrations, strikes, making banners, shouting slogans, blocking roads and so on. The other is violent actions like beating, smashing, looting and burning.

Based on the database framework, we make an encode table, which introduces the rule of judging each attribute in detail. Then we recruit 20 volunteers independently and train them to use the encode table. For cross validation, each piece of data is coded by two volunteers. After two weeks of work, all of the data have been coded.
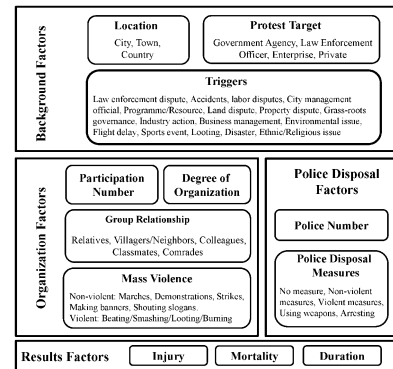

Figure 1: Database framework.

3 GEO-STATISTICAL ANALYSIS

This section analyzes the spatial-temporal distribution rules, the violence distribution and triggers of mass incidents. The software we used is ArcGis10.2.2.

3.1 Spatial Distribution

As shown in Figure 2, mass incidents in China show an obvious regional feature. South China (Guangdong, Guangxi and Hainan) is the hottest region of mass incidents, with a total of 1249 cases, accounting for 36.5% of that happened all over the country. In East China (Shandong, Jiangsu, Shanghai, Anhui, Zhejiang, Jiangxi and Fujian), there are 763 mass incidents (22.3%), ranking second. There are 410 mass incidents (12.0%) in Southwest China (Yunnan, Guizhou, Sichuan and Chongqing), ranking third. The followings are North China (Hebei, Beijing, Tianjin, Shanxi and Inner Mongolia, 9.8%), Central China (Henan, Hubei and Hunan, 8.2%) and Northwest China (8.0%). The last is Northeast China (2.1%). In addition, it is worth noting that there are 1033 mass incidents in Guangdong, accounting for 30.2%, ranking first in the country. The followings are Yunnan and Beijing, accounting for 5.9% and 5.5% respectively.

We also analyze the distribution of the mass incidents involving over 100 people and over 500 people respectively. In our opinion, the former has certain seriousness and representativeness, and the latter refers to the large-scale event which threats social stability seriously. For the mass incidents involving more than 100 people, the overall distribution is similar to that of all mass incidents. South China (42.3%), East China (20.4%) and North China (10%) are the top three regions, followed by Central China (9.2%), Southwest China (8.8%) and Northwest China (7.7%), and the last is Northeast China (1.6%). There are only 336 large-scale mass incidents which involve more than 500 people. The provincial distribution characteristics are more obvious for these events. Most of them distribute in Guangdong, Zhejiang, Yunnan and Sichuan, among which Guangdong and Zhejiang rank top two, accounting for 32.7% and 6.8% respectively.
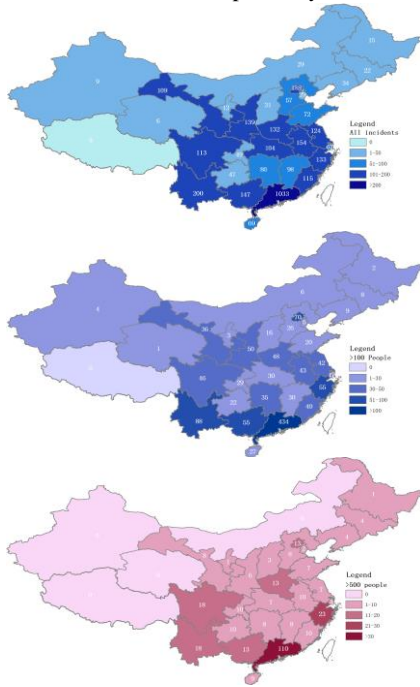

Figure 2: The distribution of all mass incidents in China from 1998-2014.

3.2 Triggers and Mass Violence

This section analyzes the triggers and the violence distribution of mass incidents. Figure 3 shows the trigger distribution of mass incidents. It can be seen that law enforcement disputes, accidents, labor disputes, disputes caused by city management officials, disputes caused by programmes or resources and land disputes are the main six triggers of mass incidents, which account for 87.4%. All of these six triggers belong to the right defending category. For the mass incidents involving more than 100 people, disputes caused by city management officials, land disputes and labor disputes rank the first three places. And for the large-scale mass incidents, disputes caused by city management officials and accidents rank the first two places, and land disputes ranks third.
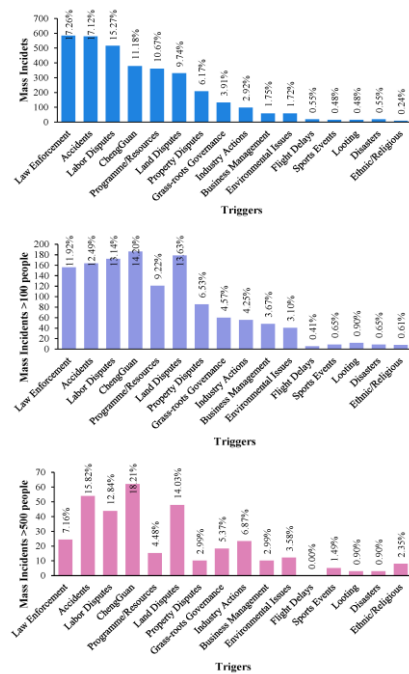

Figure 3: Triggers of mass incidents in China.

Figure 4 shows the distribution and triggers of violent mass incidents in China. As for the regional distribution, South China (35.1%) and East China (22.9%) are the top two hottest regions. As for the provincial distribution, Guangdong, Yunnan and Beijing rank top three, which accounts for 28.5%, 6.5% and 5.0% respectively. For the triggers, it can be seen that law enforcement disputes, accidents and disputes caused by city management officials are the three major triggers and they provoke more than 54.8% of mass incidents. In fact, disputes caused by city management officials is a kind of law enforcement disputes. That is to say, the law enforcement dispute is definitely the most significant factor for violent mass incidents.
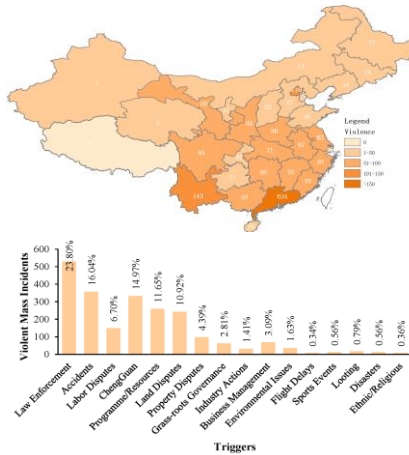
3.3   Temporal Distribution

To study the temporal distribution rules of mass incidents, we compare the distribution of mass incidents in China from 2000-2004, 2005-2009 and 2010-2014 in Figure 5. In 2000-2004, there are only 663 mass incidents all over the country, in which Guangdong and Shanxi rank top two, accounting for 27.0% and 8.0% respectively. 2005-2009 is the peak period of mass incidents. The number of mass incidents reaches 1812 during this period, which is almost three times as many as that in 2000-2004. In this period, the numbers of mass incidents in all regions increased. Among them, the largest increases are in North China and South China, which is 289% and 263% respectively. Remarkably, mass incidents number in Beijing increases by 468% in this period, which is from 25 to 142. This may be caused by the 2008 Olympic Games. In 2009-2014, there are 942 mass incidents in the whole country. Except for Northeast China, mass incidents in all regions reduce obviously. Jilin was the only province where mass incidents increase, which may be due to the "TongGang Event" in 2009 and the economic breakdown in these years.
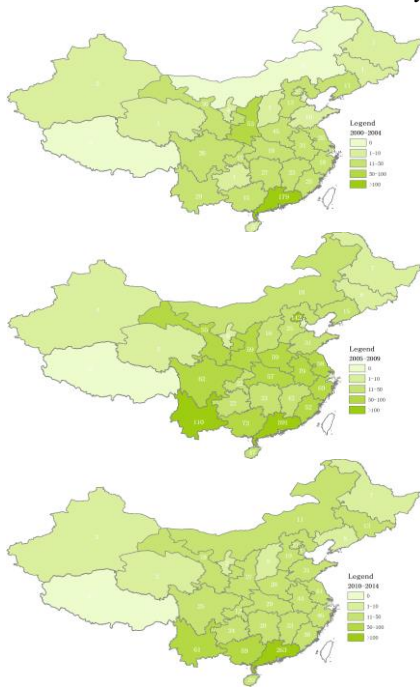

Figure 5: Temporal distribution of mass incidents in China.

4   CONCLUSIONS

Based on the GIS system, this paper builds a case database of mass incidents in China from 1998 to 2014. Then using the geo-statistical analysis methods, we study the spatial-temporal distribution rules of these mass incidents data. Several conclusions have been received in this paper.

First, mass incidents mainly concentrate in South China, especially in Guangdong Province. The second is East China. The third are North China, Central China, Southwest China and Northwest China. Northeast China ranks last. Large-scale mass incidents involving more than 500 people mainly distribute in Guangdong, and Zhejiang. Second, over 87.4% of the mass incidents belong to right defending events, among which law enforcement disputes, accidents, labor disputes, disputes caused by city management officials, disputes caused by programme or resources and land disputes are the major triggers. Besides, law enforcement dispute is also the biggest cause of mass violence. Third, 2005-2009 is the peak period of mass incidents. During this period, the number of mass incidents in Beijing increases by 468%, which may be caused by the 2008 Olympic Games. In 2009-2014, the number of mass incidents in almost all provinces reduce obviously. However, in Jilin, it increases. This may be due to the "TongGang Event" and the economic breakdown in Northeast.

The following research can discuss other attributes like the degree of organization and group relationships, and analyze the evolution of mass incidents based on the case database on GIS.

REFERENCES

[1] Cai Yong-shun. The Resistance of Chinese Laid-Off Workers in the Reform Period. The China Quarterly, 2002, 170: 327-344.

[2] Su Yang, He Xin. Street as Courtroom: State Accommodation of Labor Protest in South China. Law & Society Review, 2010, 44(1): 157-184.

[3] Guo Xiao-lin. Land Expropriation and Villagers' Complaints in Northeast Yunnan. China Quarterly, 2001, 166: 422-439.

[4] Ong L H. State-led Urbanization in China: Skyscrapers, Land Revenue and "Concentrated Villages". The China Quarterly, 2014, 217: 162-179.

[5] Tong Jing-rong, Zuo Lan-dong. Weibo Communication and Government Legitimacy in China: A Computer-Assisted Analysis of Weibo Messages on Two 'Mass Incidents'. Information, Communication & Society, 2014, 17(1): 66-85.

[6] Deng Yan-hua. Environmental Protest in Rural China: A Case Study of the Huazhen Incident. PhD thesis, The Chinese University of Hong Kong (CUH K), HongKong, China.

[7] Ma Ye-feng, Deng Qing, Wang Xin-zhi, et al. 2014. Keyword-Based Semantic Analysis of Microblog for Public Opinion Study in Online Collective Behaviors. In WAIM 2014 International Workshops: BigEM.: 44-55.

[8] Li Pei-lin, Chen Guang-jin. Chinese Academy of Social Sciences. The Social Blue Book. Social Sciences Literature Press, 2010.

[9] Li Lin, Tian He. Report on China's Legal Development 2014 . Social Sciences Literature Press, 2015.

[10] Hu Jun, Shu Xue-ming, Shen Shi-fei, et al. 2016. Research of Spatial Distribution Rules of Mass Incidents Based on GIS. In Proceedings of the Second ACM SIGSPATIAL International Workshop on the Use of GIS in Emergency Management. ACM: 10.

[11] Leetaru K, Schrodt P A. 2013. GDELT: Global Data on Events, Location, and Tone 1979–2012. In ISA Annual Convention. 2(4): 1-49.

[12] Xiao Tang-biao. What Have Caused Violence in Group Events - A Tentative Analysis of 1189 Group Event. Journal of Jiangsu Administration Institute, 2013(2): 46-55.

[13] Shao Dong-ke. A Study on Police Weiwen and Popular Protests Tactic Change in China. PhD thesis, Tsinghua University(THU), Beijing, China.