Case-based Reasoning in Avian Influenza Risk Early Warning

Jie Zhang Weiqing Yin Narisorn Pongpanich Jie Lu and Guangquan Zhang

Faculty of Engineering and Information Technology, University of Technology Sydney
PO Box 123, Broadway NSW 2007, Australia
{Jiezhang; Jielu; Zhangg}@it.uts.edu.au, {Weiqiang.yin; Narisorn.Pongpanich}@student.uts.edu.au

Abstract

Techniques to predict epidemic diseases and generate early warnings are widely studied in current complex environments. Case-based reasoning (CBR) has advantages of using past experience to generate warnings for new situations. This paper presents a method by using CBR to map risks of avian influenza transmission of obscure reasons. This method applies fuzzy logic into CBR and combines CBR with expert rules to improve capacities of risk predictions. An illustrative system shows how to conduct risk prediction based on previous influenza cases, poultry trade history, and migratory birds' movement information.

Keywords: Avian influenza, early warning systems, risk analysis, case-based reasoning

1. Introduction

An early warning system (EWS) aims to examine all the accessible information and knowledge, verify abnormal situations, identify impending threats and finally generate and communicate warnings to decision makers [1]. It is often seen as a surveillance system which can gather information associated with the crisis [2]. EWSs have gradually turned out to be necessary equipments against the increasing tendency of various threats such as epidemic diseases, natural disaster and artificial hazards. Many advanced techniques facilitate the practice of launching an effective EWS including support vector machine, fuzzy logic, decision tree and artificial neural network [3]. CBR is a paradigm of applying previous cases as knowledge and adapting similar cases' solutions to a new case. It has the cutting edge of knowledge acquisition and the flexibility of building an intelligent EWS [4]. Epidemic disease EWS can take huge advantages of CBR methodology because of its capabilities of automatic knowledge accumulations and knowledge evolutions. In this paper, we focus on avian influenza EWS because the urgent need of mapping the risk of highly pathogenic avian influenza dispersal which have lead to 248 people death and billions dollars economic poultry trade lost by the end of 2008 [5].

The rest of the paper is organized as follows: Section 2 discusses related work. Section 3 presents our CBR avian influenza EWS architecture and describes the major procedures of CBR and fuzzy logic methods. Section 4 describes a prototype of avian influenza EWS. Finally conclusions are discussed in Section 6.

© Atlantis Press, 2009

2. Related Work

Epidemic EWS researchers have applied various intelligent technologies to improve predicting capabilities of impending breakout events. CBR and fuzzy logic (FL) can profoundly enrich the research by their knowledge processing potentials. We describe these in details in the following sections.

2.1. Epidemic EWS

Natural epidemic diseases and biological attacks can be a great threat to a nation. Many researchers focus on the epidemic EWS or epidemic models. ProMED system is an epidemic disease EWS which uses case-based reasoning, data mining technique to monitor health risks and analyze risk outcomes [6]. In USA, France and Japan, influenza surveillance systems for human and animal are well developed [6].

However, avian influenza virus transmission is extremely complicated to predict and the main reasons lie in: (1) The virus is special and hard to extinct [7]. The virus is varied by the hosts and its origin and has been changing continuously. Also the virus can cross different species, for example, from birds to mammals, but it is lucky to believe that it can not spread among mammals [7]. (2) Poultry farm, bird trade and wild birds migratory make it spread through all the possible channels. [7] [8] (3) Some other undiscovered reasons also contribute to these complex situations [8]. For example, the unreported and illegal trade of both wild birds and domestic poultry shields the truth. These reasons have sparked a timely need for EWS which can monitor and track new, unknown diseases, produce signals to detect possible crises at an early stage [6]. The question is how to take advantage of existing cases, experiences and knowledge to improve the prediction and convey the reliable early warning to people.

2.2. CBR method in EWS

Case-based reasoning is supposed to use past experiences which are stored in a case base as cases to support the current decision under the similar situations. This mechanism makes it more suitable for the medical decision making [9]. The CBR cycle consists of four steps [10]: 1) retrieving former similar cases, 2) reusing the similar solutions as references, 3) revising the solutions, and 4) retaining new learned cases. Two main challenge tasks are the retrieval and the adaption [10]. CBR techniques have also developed into both temporal and spatial application such as episode-based reasoning [11], spatial CBR applied to fire modeling to stop the fire [12]. The temporal and spatial CBR are both imperative and promising to the epidemic EWS.

2.3. Fuzzy logic and CBR

Fuzzy set theory [13] was introduced to manipulate ambiguous, uncertain and imprecise value in real life. Fuzzy logic which applies linguistic variables eases the knowledge elicitation process in intelligent system [14]. Both FL and CBR have their own advantages and drawbacks and the integration of two methods can gain better performance than employing only one paradigm. FL and CBR combination systems originate in the early 1990s by using fuzzy features and fuzzy pattern matching algorithms [15]. From then on fuzzy attribute, fuzzy similarity measures are gradually combined into FL-CBR hybrid intelligent systems [14]. FL and CBR can be perfect integrated in medical intelligent system lie in the fact that both two methods have flexibilities of manipulating knowledge [14]. FL's capabilities of processing linguistic terms expand the scope of the data the system can reach, while CBR's case retaining provides the possibilities of accumulating knowledge. Both two methods can improve knowledge processing abilities and can aid to implement a system more close to the real world. Avian influenza early warning system exactly need these capabilities to overcome the difficulties of information insufficiency and imprecision. In the integration system, FL components are often embedded into CBR structure to compensate each other [14].

3. CBR-based Avian Influenza EWS Structure and Procedures

We combine both CBR and FL in this system to improve the capabilities of the EWS. The system functions are on the basis of a case base and a knowledge base. Here we describe the structure of the system and its function procedures.

3.1. CBR-based avian influenza EWS structure

The CBR-based EWS has two separate subsystems which share the case base as showed in Fig 1. One subsystem is Windows application which includes major CBR components, the other subsystem is a WEB presenting and data management platform which can feed the case base with the real events and present the system results to users.

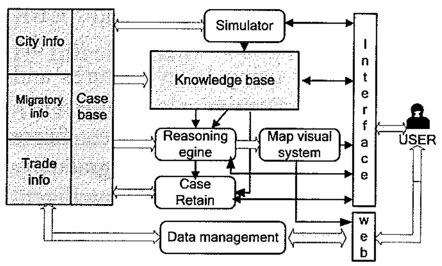

Fig. 1: CBR-based avian influenza EWS structure

The major subsystem contains a case base, a knowledge base, a reasoning engine, a case retaining component, a simulator and a map visual component. We describe each component as follows:

- Case base comprises city information, poultry trade information and wild birds migratory information. City information covers indicators which indicate facts of the city itself associated with the epidemic outbreak such as population density, poultry density, wild bird density, farm information, date, temperature, humidity and risk level. Migratory information includes the indicators of poultry trade start city, destination, poultry type and amount. Wild bird migratory information contains indicators of date, wild bird type, migratory from city and to city.
- Knowledge base involves reasoning knowledge, solution adaption knowledge and case retaining knowledge.
- Simulator can generate the risk level of a specific city by the SIR model. We will discuss it later.
- Reasoning engine and retaining component fulfills the major CBR functions.
- Map visual components present the results by visual map interface.

The above are the components of the system. The mechanism of the system and the method is discussed in next Section.

3.2. CBR risk evaluation mechanism

CBR system has the previous cases as experiences. In avian influenza EWS, we select three main fact groups as cases. Beside the information of city itself, the poultry trade and wild bird migratory also applied to obtain the final result. The event happed can be recorded into these three case bases. While only happened events are not enough for describe risk levels of every day, we also use SIR model and fuzzy theory to simulate and generate the risk level between the two

sequential events. A SIR model is presented as follows.

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \lambda I \qquad (1)$$
$$\frac{dR}{dt} = \lambda I$$

Where S denotes the number of healthy individuals who are susceptible; I denotes the number of already infected individuals who can transmit the disease to the healthy ones; and R the individuals who are removed from the infection cycle cured, or simply through their demise. The size of the whole population *(S+I+R)* is constant. β is infectivity parameters and λ is recovery parameters. If we only consider poultry in one city, the assumptions of the model can be accomplished. When an event happens, we can know whole affected poultry size, the number of birds dead at that time. Then we can simulate and generate the infected poultry size *(I)* which can be used to estimate the risk level of the city day by day (Fig. 2). If there are two sequential events happen, the maximum of two estimation value is applied for the risk level at a particular time. Risk value can be calculated by (2) and infectious rate can be calculated by (3).

$$Risk(t) = \begin{cases} I(t)/10 & I < 10 \\ \log_{10}(I(t)) & I \geq 10 \end{cases} \qquad (2)$$

$$Rate(t) = \frac{I(t)}{S(t) + I(t) + R(t)} \qquad (3)$$

Fuzzy linguistic values can be generated by fuzzy membership functions. We apply five risk levels *{normal(N)-0, low risk(LR)-1, medium(M)-2, high risk(HR)-3, very high risk(VH)-4}* showed in Fig 3.

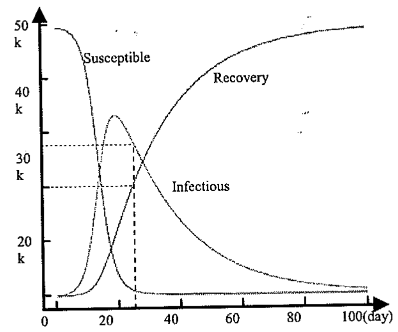

Fig.2 If the affected whole poultry size is 50k and on the spot time 20k poultry died. After simulation, we can identify 28k poultry infected at that time and the other value can be calculated day by day.

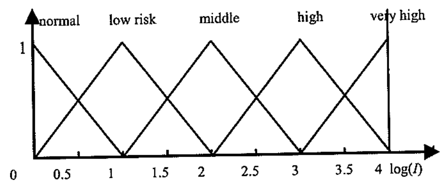

Fig.3 Risk level fuzzy function

Poultry trade amount plus city infected poultry value can be calculated by

$$I^c(t) = I(t) + \sum_{i=1}^{k} N_i(t).Rate_i(t) \qquad (4)$$

Suppose we have known the trade plan of poultry size N_i from city i at time t from k cities, then the city risk combining with poultry trade risk can be calculated by (2) with the parameter $I^c(t)$.

Wild bird migratory risk level can be estimated by experts because wild bird migratory risk is hard to determine by calculation or simulation. But wild bird migratory has common seasonal laws related to the changing weather, therefore these factors can only be considered in the migratory seasons.

3.3. CBR operating mechanism

CBR system has the famous 4R cycle as retrieve, reuse, revise and retain. We will describe each in details. Firstly, retrieving and reusing process is easy to implement.

The distance between two cases can be calculated by

$$Dist = \frac{\sum_{i=1}^{n} w_i .norm(| f_i^n - f_i^o |)}{\sum_{i=1}^{n} w_i} \quad (5)$$

Suppose we have n indicators, w_i is the weight of the ith indicator, f_i^o, f_i^n are the value of the ith indicator of old and new cases respectively and $norm(| f_i^n - f_i^o |)$ is normalized distance between two indicators. We combine three most similar cases' weighted *Risk(t)* as the final risk value.

Revising process is implemented by involving the trade and migratory data. We first combining trade data with the city data by formula (4), then the combined risk $Risk^c(t)$ can be obtained. The final combined risk of $Risk^c(t)$ and $Risk^m(t)$ is obtained by using fuzzy logic method. The fuzzy rules are listed in Table 1. The defuzzified method is centroid.

Table 1 Fuzzy combined rules

		$Risk^c(t)$				
		N	L	M	H	VH
$Risk^m(t)$	N	N	L	M	H	VH
	L	N	L	M	H	VH
	M	N	L	M	H	VH
	H	N	M	M	H	VH
	VH	L	H	H	H	VH

The Retain process will proceed while new facts are provided. These new evidence can be applied by the simulator then the most recent risk value will be change by the simulator if it against the real happened events.

4. System Prototype Implementation

The CBR-based avian influenza EWS prototype has been implemented. We only use five cities in China to show the prediction processes of this system. The early warning code can be executed automatically and the system can finally generated the risk value of each city daily. The final risk levels are presented by different color (Fig. 4). The red line between each city denotes the wild bird migratory and blue line between two cities denotes poultry trade directions. Some details of the CBR process can be examined by the user (Fig 5).

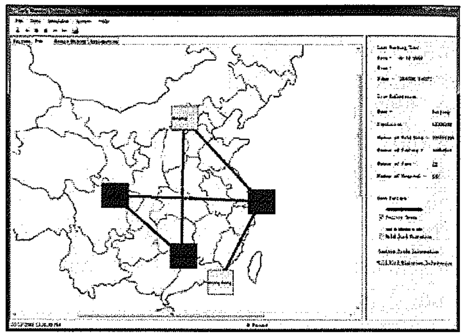

Fig. 4 The map result of city risk level

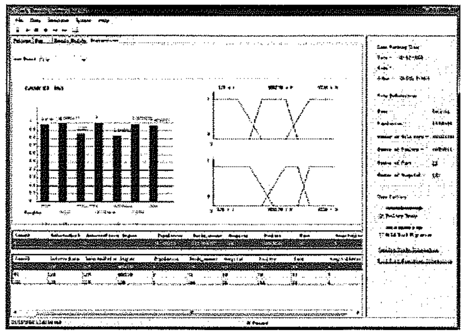

Fig. 5 The result of CBR processes

5. Conclusions and Further Study

Generating avian influenza early warning is a challenging task. An avian influenza EWS prototype has been implemented based on CBR method, FL theory and simulation technique. One important facet is applying combining methods can facilitate addressing this issue. The system architecture and processes are based on primarily CBR paradigm, while we also introduce fuzzy logic and simulation to finally improve the resolution. The other facet is CBR and FL can address the knowledge issues under the obscure mechanism of avian influenza background. In the future we will improve the research by selecting and refining factors applied in the system when providing the early warning.

Acknowledgement

The work presented in this paper was supported by Australian Research Council (ARC) under Discovery Project DP0880739.

References

[1] G. Zhang, X. Yang, and J. Lu, "A Support Vector machine (SVM) Based Classification Approach For Early Warning System," *in the 8th International FLINS Conference*, Madrid, Spain, 2008, pp. 549-554.

[2] WHO, "Early warning systems," 2009.http://www.who.int/csr/labepidemiology/projects/earlywarnsystem/en

[3] T. Y. Kim, K. J. Oh, I. Sohn, and C. Hwang, "Usefulness of artificial neural networks for early warning system of economic crisis," *Expert Systems with Applications*, vol. 26, pp. 583-590, 2004.

[4] S. K. Pal and S. C. K. Shiu, Foundations of Soft Case-Based Reasoning: John Wiley & Sons, Inc., 2004.

[5] WHO, "Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1) Reported to WHO," 2009.http://www.who.int/csr/disease/avian_influenza/country/cases_table_2009_01_19/en/index.html

[6] L. Gierl and R. Schmidt, "Geomedical warning system against epidemics," *International Journal of Hygiene and Environmental Health*, vol. 208, pp. 287-297, 2005.

[7] M. Gauthier-Clerc, C. Lebarbenchon, and F. Thomas, "Recent expansion of highly pathogenic avian influenza H5N1: a critical review," *Ibis*, vol. 149, pp. 202-214, 2007.

[8] A. M. Kilpatrick, A. A. Chmura, D. W. Gibbons, R. C. Fleischer, P. P. Marra, and P. Daszak, "Predicting the global spread of H5N1 avian influenza," *Proceedings of the National Academy of Sciences (PANS)*, vol. 103, pp. 19368-19373, 12/19 2006.

[9] R. Schmidt, S. Montani, R. Bellazzi, L. Portinale, and L. Gierl, "Cased-Based Reasoning for medical knowledge-based systems," *International Journal of Medical Informatics*, vol. 64, pp. 355-367, 2001

[10] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *AI Communications*, vol. 7, pp. 39-59, 1994.

[11] M. Sánchez-Marré, U. Cortés, M. Martínez, J. Comas, and I. Rodríguez-Roda, "An Approach for Temporal Case-Based Reasoning: Episode-Based Reasoning," *in Case-Based Reasoning Research and Development*, 2005, pp. 465-476.

[12] B. Knight, S. Taylor, M. Petridis, J. Ewer, and E. R. Galea, "A knowledge-based system to represent spatial reasoning for fire modelling," *Engineering Applications of Artificial Intelligence*, vol. 12, pp. 213-219, 1999.

[13] L. A. Zadeh, "Fuzzy sets," *information and control*, vol. 8, pp. 338-353, 1965.

[14] A. Khanum, M. Mufti, M. Y. Javed, and M. Z. Shafiq, "Fuzzy case-based reasoning for facial expression recognition," *Fuzzy Sets and Systems*, vol. 160, pp. 231-250, 2009.

[15] P. Bonissone and S. Ayub, "Similarity measures for case-based reasoning systems," *in IPMU '92—Advanced Methods in Artificial Intelligence*, pp. 159-172, 1993.

New Perspectives on Risk Analysis and Crisis Response

Proceedings of the 2nd International Conference on Risk Analysis and Crisis Response

October 19–21, 2009
Beijing, China

Edited by

Chongfu Huang
Jonathan B. Wiener
Jinren Ni

ATLANTIS PRESS
AMSTERDAM – PARIS

ISBN: 978-90-78677-34-5

This book is part of the series ***Advances in Intelligent Systems Research*** (ISSN 1951-6851) published by Atlantis Press.

Aims and scope of the series

During the past decade computer science research in understanding and reproducing human intelligence has expanded from the more traditional approaches like psychology, logics and artificial intelligence into multiple other areas, including neuroscience research.

Moreover, new results in biology, chemistry, (surface) physics and gene technology, but also in network technology are greatly affecting current research in computer science, including the development of intelligent systems. At the same time, computer science's new results are increasingly being applied in these fields allowing for important cross-fertilisations.

This series aims at publishing proceedings from all disciplines dealing with and affecting the issue of understanding and reproducing intelligence in artificial systems.

Also, the series is open for publications concerning the application of intelligence in networked or any other environment and the extraction of meaningful data from large data sets.

Research fields covered by the series include: * Fuzzy sets * Machine learning * Autonomous agents * Evolutionary systems * Robotics and autonomous systems * Semantic web, incl. web services, ontologies and grid computing * Biological systems * Artificial Intelligence, incl. knowledge representation, logics * Neural networks * Constraint satisfaction * Computational biology * Information sciences * Computer vision, pattern recognition * Computational neuroscience * Datamining, knowledge discovery and modelling for e.g. life sciences.

© ATLANTIS PRESS, 2009
www.atlantis-press.com

ISBN 978-90-78677-34-5

This book is published by Atlantis Press, scientific publishing, Paris, France.

All rights reserved. No part of this book may be reproduced, translated, stored or transmitted in any form or by any means, including electronic, mechanical, photocopying, recording or otherwise, without prior permission from the publisher.

The 2nd International Conference on Risk Analysis and Crisis Response

October 19–21, 2009, Beijing, China

THEME: Theories and Applications of Risk Analysis and Crisis Response

ORGANIZED BY	:	Risk Analysis Council of China Association for Disaster Prevention
HOSTED BY	:	Peking University
SPONSORED BY	:	Society for Risk Analysis Society for Risk Analysis-Europe Society for Risk Analysis-Japan
SUPPORTED BY	:	National Natural Science Foundation of China

Organization

General Chair: — **Chongfu Huang**, Beijing Normal University, China
Co-Chairs: — **Jonathan B. Wiener**, Duke University, USA
— **Roberto Bubbico**, "La Sapienza" Univeristy of Rome, Italy
— **Akihiro Tokai**, Osaka University, Japan

Programme Committee

Chairman: — **Jinren Ni**, Peking University, China
Co-Chairs: — **Olivier Salvi**, European Virtual Institute for Integrated Risk Management, Germany
— **Haibo Wang**, Texas A&M International University, USA
Members: — **Alistair Borthwick**, University of Oxford, UK
— **Luis Cifuentes**, P. Catholic University of Chile, Chile
— **Yongsheng Ding**, Donghua University, China
— **Jiali Feng**, Shanghai Maritime University, China
— **Christopher Frey**, North Carolina State University, USA
— **David Hassenzahl**, University of Nevada at Las Vegas, USA
— **Gordon Huang**, University of Regina, Canada
— **Cengiz Kahraman**, Istanbul Technical University, Turkey
— **Daniela Leonte**, University of New South Wales, Australia
— **Xilin Liu**, Sun Yat-sen University, China
— **Siuming Lo**, City University of Hong Kong, China
— **E Lopez-Vazquez**, Autonomous University of the State of Morelos, Mexico
— **Andre Maisseu**, World Council of Nuclear Workers, France
— **Marc Poumadere**, Institut Symlog, France
— **Guosheng Qu**, National Earthquake Response Support Service, China
— **Ortwin Renn**, University of Stuttgart, Germany
— **Sviatoslav Timashev**, Russian Academy of Sciences, Russia
— **Guangqian Wang**, Tsinghua University, China
— **Hao Wang**, China Institute of Water Resources & Hydropower Research, China
— **Jin Wang**, Liverpool John Moores University, UK
— **Dash Wu**, University of Toronto, Canada

	—	**Zongzhi Wu**, China Academy of Safety Science and Technology, China
	—	**Bin Zhang**, Beijing Municipal Institute of Labour Protection, China
	—	**Zongfang Zhou**, University of Electronic Science & Technology of China, China
Secretariat:	—	**Xiaoling Mao**, Peking University, China
	—	**Meiping Tong**, Peking University, China
	—	**Zheng Huang**, Peking University, China

Foreword

This volume collects the papers presented at RACR-2009, the Second International Conference on Risk Analysis and Crisis Response, held in Beijing, China, on 19-21 October 2009. We organized this conference in order to promote advances in methodology and practice of risk analysis, and to provide an international forum for sharing and debating the theories and applications of risk analysis and crisis response.

The issue of our era is risk. This has only become clearer since RACR-2007, the First International Conference on Risk Analysis and Crisis Response held in Shanghai 2007. At that time, we noted the challenges posed by risks such as terrorism, SARS, earthquakes, tsunamis, and bird flu. Today, after the financial crisis of 2008, it is even more evident that globally interconnected risks – and, in part, flawed or neglected risk analyses – are key drivers of the major crises in credit markets, climate change, and world conflict. At the same time, the concerns and priorities of citizens and governments vary around the world, so there can simultaneously be agreement on risk as a major overarching issue and yet disagreement on which specific risks deserve the most urgent responses. Moreover, the different strategies used to assess and manage different risks in different countries can present a source of friction – and an opportunity for testing, learning and borrowing the best strategies from each other.

From the many submissions received, the RACR-2009 program committee conducted a careful review and selected 114 papers to publish in these proceedings. The authors come from more than 16 countries, including Australia, Belgium, Canada, China, France, Greece, Italy, Japan, Korea, Mexico, the Netherlands, Russia, Slovakia, Turkey, the United Kingdom, and the United States. The papers address a wide range of topics, including the assessment and management of risks to health, safety, environment, security, and finance. The high quality and quantity of these papers gives us optimism that through sound analysis, we can find solutions to risk and crisis issues.

Special thanks are due to all the contributors and referees for their kind cooperation in helping to prepare this book. We express particular gratitude to the Programme Committee and its co-chairs, Olivier Salvi (France) and Haibo Wang (USA); to Xiaoling Mao (China) and the conference secretariat; and to the general co-chairs of the conference, Roberto Bubbico (Italy) and Akihiro Tokai (Japan). Together, their leadership and hard work has made RACR-2009 a success.

Chongfu Huang (China)
Jonathan B. Wiener (USA)
Jinren Ni (China)

Contents

A Note on the Difference between Disaster Prediction and Risk Assessment in Natural Disasters*

Chongfu Huang[1,2,3]

[1]State Key Laboratory of Earth Surface Processes and Resources Ecology (Beijing Normal University), Beijing 100875, China
[2]Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education of China, Beijing Normal University, Beijing 100875, China
[3]Academy of Disaster Reduction and Emergency Management, Ministry of Civil Affairs & Ministry of Education, the Peoples' Republic of China, Beijing 100875, China

Abstract

Many people consider risk assessment as to predict disaster or calculate the probabilities of adverse incidents. In this paper, it demonstrates that a natural disaster prediction is to model some objects which would meet hazard, damage or loss. For risk assessment, it is to model some aspects of the scene in the future associated with some adverse incidents. The research results also show that, in terms of mathematics, a model for disaster prediction can be expressed in the form of an explicit function while a model for risk assessment takes the form of an implicit function.

Keywords: Disaster prediction, risk assessment, probability, scene

1. Introduction

Since the September 11 attacks[1], more and more people have been thinking of that today's society is a risk society[2], filled with challenges as well as opportunities.

Now, the term "risk assessment" has become a fashion label for natural disaster reduction. A motley variety of models are employed to assess risks. Then, the models for disaster prediction[3][4], GIS models for superimposing maps[5], the models for probability estimation[6] and the models for fuzzy comprehensive evaluation[7] appear frequently in the so-called risk assessment. In other words, within concept of many people, risk assessment is to predict disaster or calculate the probabilities of adverse incidents.

In fact, any risk assessment is to model some aspects of the scene in the future associated with some adverse incident, and different aspects for assessment leads to different scene.

In this paper, we show a note on the difference between disaster prediction and risk assessment in natural disasters: the former focuses on incidents, but the latter on scenes. Any disaster prediction is to model some objects which would

* This work was supported by National Natural Science Foundation of China (No.40771007), National Key Project of Scientific and Technical Supporting Programs Funded by Ministry of Science & Technology of China (No.2006BAD20B01-02) and The National High Technology Research and Development Program of China (No.2009AA12Z124).

© Atlantis Press, 2009