Semantic Symbols Extraction Model for Emergency Hazard Map

Lijian Sun[1], Jie Zhao[2], Lihong Shi[1], Zheng Gong[2], Yi Zhu[1], Agen Qiu[1]

1. Chinese Academy of Surveying and Mapping
E-Government GIS Center
Beijing China
e-mail:sunlj@casm.ac.cn

2. National Administration for Code Allocation to Organization
Department of Data Processing
Beijing, China
e-mail:zjvsfd@sina.com

***Abstract*— Emergency event information released in map is necessary for emergency management and disaster reduction. A new method for emergency map symbols extraction from symbol collection based semantic analysis is presented in this paper. A novel map symbol semantic matrix is introduced to measure the degree of the semantic representation between the symbol meaning and the emergency event conception. The necessary content factors of emergency released in map are conduced by analyzing the emergency content construction and statistic result of the emergency content published by national disaster reduction center. The fuzzy comprehensive evaluation is proposed to extract the map symbol class. The sub function $f_i(k_i)$ is designed to confirm the risk of factor-ki . The simulation results verify the feasibility of emergency semantic model and its extraction model.**

Keywords-Symbol; Semantic Construction; Emergency; Fuzzy comprehensive evaluation.

I. INTRODUCTION

Natural disasters or emergency hazards occur frequently in the world [1], such as earthquake, typhoon, mud-rock flow, land-slip of the mountains, falling and the downthrow of the earth, etc. These are significant and realistic threats to people's wealth and life. Most of the natural disasters and emergency hazards are intimately related with spatial. To marked those hazard information including characters, impact range and the response processing by symbol on a map, it is of important practical and immediate significance to the disaster prevention and reduction. [2]. At present, researchers have done little work on this area [3]. The dissemination of emergency information processing system is mostly manual intervention, in a low automation level [2], so it is necessary to do some work in this area. The sematic of cartographic symbols includes spatical character and reference characters[4].

In this paper, a cartographic symbol extraction model on emergency hazard map n based semantic analysis is presented. A novel map symbol semantic matrix is introduced to measure the degree of the semantic representation between the symbol meaning and the emergency event conception. Further the fuzzy comprehensive evaluation and sub function are proposed to extract emergency map symbol from symbol library, by evaluating the relationship between the symbol semantic representation and the emergency event construction.

II. SEMANTICS RELATED TO EMERGENCY SYMBOL

Map symbols, composed of some graphics with different shape, size and color, are atlas language and could express geo-information effectively. Some researchers have done some work on map symbols linguistics, while most of them focused on the organization mode of graphic elements and symbol design. In this session, research on semantic representation of map symbol will be carried out by analyzing relationship between the entire semantic representation of the map symbol and the emergency content construction.

The map symbol for emergency release is thematic symbol. Spatial and representation are the basic attributes of those symbols.

The spatial character is the position on the map. According to the distribution mode of the object, the pattern of the map symbol spatial distribution includes dot, line, and poly. The dot distribution refers to the emergency whose spatial distribution mode is a point, for example, "January 2, 2008, 20 pm, a great fire occurred in U-City, D-Square", the D-Square is just a point in small scale map. The line distribution refers to the emergency whose spatial distribution mode is the line, for example, "January 6, 2008, Ice-run appeared in the Yellow River in Ningxia section, the whole length is about 234 Km". The poly distribution refers to the emergency whose spatial distribution mode is poly, for example, "from January 12,2008, a great snow storm fell in Anqing, Chizhou, Tongling district". The Anqing, Chizhou, and Tongling districts, which suffered from a great snow storm, are an area on a large scale map.

Symbol representation is to assess the relationship between the symbol and the emergency event, which is the ideographic expression. In this paper, the symbol representation measure is studied, not the symbol design. The symbol representation includes two aspects: symbol credibility and symbol (press) degree.

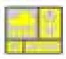
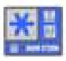
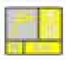
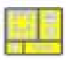
Definition: Symbol credibility: The real event is material, while the symbol is abstract, so there is gap between them. Symbol credibility (φ_{ij}) is used to measure this relationship.

φ_{ij} is the credibility for symbol i representing the case j. φ_{ij} is continuous not binary, $0 \le \varphi_{ij} \le 1$, i∈ symbol, j∈ case, φ_{ij} is dimensionless. The greater φ_{ij} is, the more little the gap is, that is the symbol is more similar to the event.

φ_{ij} is often set by expert's subjective judgment, it is a definition value.

Definition: Symbol (press) degree: Generally speaking, a different emergency is often represented with different degree's symbols, according to the emergency press. The class of emergency symbol could be set by different shape and graphics construction and the press degree of emergency map symbol could be set by different color. For instance, the level of different meteorological disaster symbols is distinguished by four color-blue, yellow, orange, and red (according to the regulation of meteorological disaster symbols for prediction, published by China Meteorological Administration in 2004) [6] .

TABLE I. METEROLOGICAL DISAER SYMBOLS COLLECTION (PART)

Meteorological disaster	Risk Level			
	IV Slight	III Serious	II Magnitude	I Destroy
Typhoon				
Rain Storm				
Snow Storm				
Cold Wave				
Gale				
Sand Storm	------			
Drought	------	------		
Frost				-----

In Table 1, the meteorological disaster level *IV*(Slight) is marked with blue, the level *III*(Serious) is marked with yellow, the level *II* (Magnitude) is marked with orange, and the level *I* (destroy) is marked with red.

III. EMERGENCY SEMANTIC CONSTRUCTION MODEL

Emergency is an out-of-order incident, which people have little knowledge and information about [7]. A range of issues arising from the emergency are of ill-structured or unstructured problems [12]. If the emergency could not be responded effectively, it will lead to crisis. Emergency is often regarded as pre-crisis. Compared with general incidents, emergency incidents have the following three characteristics: emergent, severe ,and urgent.

Some linguists have done useful exploration and research on the analysis of the emergency content, such as Zeng Qingqing[5], who defined two information chains from the perspective of researchers—the main information chain and the secondary information chain. Yang Erhong analyzed the emergency content by setting the key words. The key elements in emergency content construction will be deduced by analyzing the relationship between the emergency meaning and the semantic representation of map symbols.

Event in context refers to language description for the special matter which people are concerned about, it belongs to the union meaning description [4]. Event is made up of event words and event parameters. In other words, the behavior generally described by verbs also includes event word, location, participants and so on.

Event word is to flag the property of the event, which is the key difference compared with other events. For example, "January 16, 2008 16:50 pm, one person was missing in a land-slip of the mountains, in Guxiang village, Xinshui town, Daning country". In this context, the event word is "land-slip of the mountains", which is the key word to distinguish the emergency. Generally speaking, a content collection may be confirmed by the type of the emergency when the event word appears.

Definition: Event word collect A:

$$A = \{a_1, a_2, , , a_n\} \tag{1}$$

where $a1, a2, ..., an$ are the elements of A, a_i is the event word. For example, the coalmine accident is defined as: Coalmine Accidents ={gas accident, collapse, colliery flooding, ...}.

The emergency hazard type should be certain from the context. If the type is not indexed in event word library, it could be pushed in the event word library, thus the event word library is open for extend.

Sometimes, the event word is not in context, for example, the context—"the houses have been razed to the ground, most of people was homeless", we knew the state of the house and the victims, but we could not detect the disaster type from the content, it may be earthquake, volcano or flood. This is called the event word missing, and the symbol could not be extracted from the emergency content when event word missing.

A. *Event parameters*

An integrated emergency should have a completed event word centric expression model that is given by:

$$a_i = f(p_{i1}, p_{i2}, , , p_{in}) \tag{2}$$

where p1, p2,,, pn are the event parameters to describe the emergency.

The components of the emergency event are always the mapping results from the part parameters in the whole event parameters. For example, the event parameters of the emergency- land-slip of the mountains includes :

" p_1 -Location",

" p_2 -Impact range",

" p_3 -The time of occurrence",

" p_4 -The duration",

" p_5 -The dead and hurt people",

" p_6 -The economic losing", etc.

For example, the context-"January 16, 2008 16:50 pm, one people was missing in a land-slip of the mountains, in Guxiang village, Xinshui town, Daning country", the event parameters in above include location, the time of occurrence and the number of missing.

B. *Event attributes*

Event attributes are used to describe the state of the event. The event modality, press degree, frequency and the state are focused on. The event modality is the possibility of emergency occurrence. The event modality mode is determined when the emergency has occurred, while for the predicted emergency, it should be marked by title or other form of annotation. For example, the context from the China Central Meteorological Station, "in the next 2-3 days, the 8th typhoon will generate in the northern of the South China Sea, the wind in typhoon center will be expected to reach 10-12 class. " is a predictive emergency. The event press degree is used to measure the hazardous extent of the emergency. It is in the content of the emergency sometimes, for example, "January 8, 2008, a serious traffic accident occurred in the Hexu expressway in Anhui province section". (The press degree in this emergency is serious). While sometimes it is not direct in the emergency content, for example," January 7, 2008, two were hurt in the traffic accident in 312 State Road in Yongshou country section". Sometimes the event press degree could be detected from the content if there is enough information in. The event frequency is the number of the emergency occurred. For example, "January, 20, 2008, two blasts occurred in a chemical factory, fortunately nobody was hurt", The event frequency of this blast is two times. The event state is used to describe the current state of the emergency incident, for example, "January, 10, 2008, a coalmine flooding accident occurred in K-Country, D-city. Now the rescue work is still on."

The statistic result of the emergency context for release with map symbols is as follows: (From national disaster reduction center, public)

TABLE II. THE SUMMARY STATICTIS RESULTS OF THE NUMBER OF EMERGENCY CONTEXT FOR RELEASE FROM JANUARY1,2008 TO JANUARY 27,2008,CHINA

Emergency	Event Word	Emergency Space		Emergency Attribute			
				Event extent		Time information	
		Location	Impact range	Numbere Infor	Extent Infor	Occurrence frequency	Duratio Infor
Earthquak	11	11	2	11	0	11	0
Snowstorr	58	58	0	0	20	58	1
Fire	1	1	0	0	1	1	1
Over-wate Accident	2	2	0	2	0	2	1
Traffic accident	4	4	0	4	1	4	0
Drought	4	4	0	4	0	4	0
Ice-run	1	1	1	1	0	1	0
Coalmine accident	1	1	0	1	0	1	0
Explosion	1	1	0	1	0	1	0
land-slip o mountains	1	1	0	1	0	1	0
Crash	1	1	0	1	0	1	0
Cold Wave	3	3	0	3	0	3	0
Total	92	92	3	29	22	92	3
Scale	100%	100%	3.26%	31.52%	23.91%	100%	3.26%

In Table 2, the event word, event space, event occurrence time are all emerging in the emergency context, combination of the research result of the map symbol, a conclusion is drawn that the emergency content for release in map should at least include the event word, space and occurrence time. The event word and the event space have to be marked on the map. The event occurrence time is sometimes marked out of the map, usually as title, annotation. The event extent is necessary for the level emergent.

IV. SOFTWARE MODEL DESIGN FOR SYMBOL EXTRACTION

The basic principle of the map symbols extraction algorithm is scattering the event according to the attributes of the elements. Then they can be converted to quantitative data according to the discrete elements and can be mapped to the symbol library through the mapping model . This paper will further extend this mapping as an evaluation.

The result of scattering event is the formation of an event object , the event object including (attribute 1 attribute 2, ... attribute n).Each symbol in the symbol library is associated with a semantic routing model .Semantic routing model is a table, which saves semantic categories and similarity of each symbol. The attribute of the event object can be into the evaluation index, the sign which gets a high score in this indicator is the symbol to be extracted. Then setting attributes (color, size, etc.) of the extracted symbol according to other semantic events in the object (degree of harm).The model is shown in Figure 2:

The Procedure Design Language (PDL) for map symbol extraction is as follows:

(1) Analyzing the emergency content, getting the elements collection $A = \{a_1, a_2,,, a_n\}$.

if (Event word $\notin A$ | Event space $\notin A$ | event occurrence time $\notin A$), A could not be released, then finish.

(2) Map symbol extraction algorithm:

if($\exists$ Event extent) $f1 = f$ (event word, event extent, symbol collection)

Else $f1 = f$ (event word, default event extent, symbol collection)

(3) if($f1 \neq$ null)， Symbol space position ： $d = f($ Event space) , then $f1$ is marked at d .

(4) Mark the event occurrence time in map as label or title

or other illustration.

The algorithm and technical of position on map is quite mature, the event occurrence time is often marked as title or annotation. The map symbol extraction in step (2) is studied in this paper.

Define symbol as a bivariate function, that is,

$$S = f(B,V) \tag{3}$$

where B is the symbol ' s class and V is its extent.

Define Symbol semantic matrix, $B(i) = (\varphi_{ij})_{m\times n}$, φ_{ij} is the degree of credibility for symbol i substituting the case j .

Definition: Symbol attributes matrix, $L(i)_{1\times m} = (l_{kj})_{1\times m}$, that shows the relationship between attribute k and the class i. l_{kj} =1, when k consistent with j in class i, otherwise l_{kj} =0. For example, class i is {flood, snow storm, earthquake, volcano, explosion, ice-run,,,}, then the $L(earthquake)_{1\times m} = \{0,0,1,0,0,0,,\}$.

A. Symbol's class extraction

The fuzzy comprehensive evaluation, which is a method to classify the examples according by some indicators, is introduced to extract the map symbol class.

Suppose the count of event word is n, its attribute matrix is $L(i)$, (i=1,2,…n), then the extracted analysis matrix $X = |\eta_1, \eta_2, ..., \eta_r|^T$, $\eta_i = (L(i)A(i))^T$, where i=1,2,,,r. r is the count of symbol in symbol collection.

The processing steps are as follows:

1. Confirm the fuzzy relationship matrix Ri (including the membership function and the result), i=1,…,n.

2. Confirm the weighed distribution vector A, A= (a1,a2,,,,an).

3. Get the evaluation result Bi by blue processing, Bi = A ° Ri, Bi= (bi1,bi2,,,,bin).

where the samples belong to the class k*, when b_{ik}^* = $\max_k |b_{ik}|$.

B. Confirm the risk of symbol

The event degree is to explain the hazardous extent of the emergency, including the directly or indirectly information, indirect information, such as "7 people injured in a traffic accident, in Badong country, January 3, 2008", the directly information is that, "January 8, 2008, a serious traffic accident occurred in the Hexu expressway in Anhui province section".

(1) If the event press degree is not in the emergency context, define the event extent is the lowest.

(2) If the press degree l is in the context, the v is extracted directly. The higher the event extent level is, the more hazardous the emergency is.

(3) If the extent level is not direct in the context, while the number information about extent is in. Supposing that the information number k about extent is an indicators collection, that is, $K = \{k_1, k_2,,, k_n\}$ n is the count of the elements in k. Then, the sub-event extent to k_i could be confirmed according by the industry standard.

For example, the rainfall extent defined by China Meteorological Administrator is that: 1th grade is from 4.17mm/h to 8.33mm/h, 2 th is from 8.33mm/h to 16.67mm/h, 3 th is from 16.67mm/h to 33.33mm/h, 4 th is from 33.33mm/h to larger. The function is like that:

$$f_i(k_i) = \begin{cases} 1 & 4.17 < k_i \le 8.33 \\ 2 & 8.33 < k_i \le 16.67 \\ 3 & 16.67 < k_i \le 33.33 \\ 4 & k_i > 33.33 \end{cases} \tag{4}$$

The function (4) is shown in Figure 1.

The event extent with indicators collection k is that,

$$V = \max\{ f_1(k_1), f_2(k_2),,, f_n(k_n)\} \tag{5}$$

V. EXPERIMENTS AND ANALYSESE

Since the event word and press degree are necessary for emergency release. Suppose there are event words t1, t2, t3, t4 and press degree indicators k1, k2, k3, k4 in the emergency context. The simulation processing $f1 = f$ (event word, event extent, symbol collection) is as follows:

A. Symbol's class extraction

Six map symbols in collection as examples will be evaluated. After analysis from the symbol mapping table, the relationship result between the symbol semantic and the event word is as follows:

TABLE III. THE RELATIONSHIP BETWEEN SYMBOLS AND EVENT WORD

symbol \ event word	t1	t2	t3	t4
1	0.9	0.0	0.5	0.2
2	0.1	0.9	0.4	0.3
3	0.0	0.1	0.3	0.7
4	0.1	0.9	0.0	0.6
5	0.2	0.4	0.1	0.0
6	0.0	0.2	0.1	0.1

Then the matrix X :

$$x = \begin{bmatrix} 0.9, 0.0, 0.5, 0.2 \\ 0.1, 0.9, 0.4, 0.3 \\ 0.0, 0.1, 0.3, 0.7 \\ 0.1, 0.9, 0.0, 0.6 \\ 0.2, 0.4, 0.1, 0.0 \\ 0.0, 0.2, 0.1, 0.1 \end{bmatrix}$$

Suppose the evaluation factors collection is set as follows:

TABLE IV. THE EVALUATION FACTORS TABLE

level \ evaluation factor	t1	t2	t3	t4
I	0.3	0.3	0.3	0.3
II	0.6	0.6	0.6	0.6
III	0.9	0.9	0.9	0.9

Define that the bigger the value of the evaluation factor is, the better the expression effect is.

Confirm the membership and weigh:

Define a factor membership function for evaluation is like:

$$u_{ij} = \begin{cases} 1 & 0 \le d_i \le c_{ij} \\ \frac{c_{ij+1} - d_i}{c_{ij+1} - c_{ij}} & c_{ij} \le d_i \le c_{ij+1} \\ 0 & c_{ij} < d_i \end{cases} \qquad (6)$$

where u_{ij} is the factor i that belongs to the membership degree j. d_i is the value of the factor i. c_{ij} is the criterion value of i with degree j, i = 1,2,3,4，j=1,2,3.

So, the symbol K 's membership matrix with all the evaluating factors is defined as $R^{(k)}{}_{4\times3} = (\mu_{ij})_{4\times3}$,k=1,2,…,6. The weight parameter matrix Ak = (a1,a2,a3,a4), where a_i is like:

$$a_i = \frac{x_i / a_i}{\sum_{i=1}^{n} x_i / a_i} \qquad (7)$$

Fuzzy comprehensive evaluation is presented next.

To evaluate symbol K , the evaluation matrix is generated by synthesizing Ak and $R^{(k)}{}_{4\times3}$. In this paper, the average weigh model M(*,+) is adopted, that weighed all the factors. Then Ak is as follows:

$$B_{6\times3} = \begin{bmatrix} A_1 \circ R^{(1)}{}_{4\times3} \\ A_2 \circ R^{(2)}{}_{4\times3} \\ \cdots \\ A_6 \circ R^{(6)}{}_{4\times3} \end{bmatrix}$$

, put the data into, then

$$B_{6\times3} = \begin{bmatrix} 0.23, 0.21, 0.56 \\ 0.39, 0.08, 0.53 \\ 0.36, 0.42, 0.22 \\ 0.06, 0.38, 0.56 \\ 0.81, 0.19, 0.00 \\ 1.00, 0.00, 0.00 \end{bmatrix}.$$

It shows that symbols 0 and 3 are better than others, that is B = 0 or B = 3.(0,3 are the row order) , B(3,2) > B(0,2), so B = 3.

B. Confirm the risk of symbol

From Eq.(4), the conclusion is drawn that,

$$V = \max\{ f_1(k_1), f_2(k_2), f_3(k_3), f_4(k_4)\},$$

where $f_i(k_i)$ is a linear classified function, usually defined by industrial standard.

After confirming B and V, according to (3), the target symbol could be calculated. Generally speaking, the function $f(B,V)$ is a mapping function to the symbol, B is the row, V is the col.

The electronic map with this method is shown in Figure 3 (dot symbol and line symbol in map).

VI. CONCLUSIONS

In this paper, the context and semantic construction about the emergency for release was studied. The map symbol semantic matrix is introduced to measure the relationship between the symbol semantic representation and the emergency event. According to the research, the necessary conditions for the emergency release in map are deduced. The fuzzy comprehensive evaluation is proposed to extract the symbol's class and the classified function for the factor of the event extent is used to confirm the event press degree. In this experiment, the symbol semantic matrix values are set by subjective definition, the value will be set by the standard or the expert database after the enough resource has been collected in future. By analyzing the example, the effectiveness and practice of the present experimental method are proved. The extension of the normalization of performance value and efficiency compared with other method would be our future direction.

ACKNOWLEDGMENT

This paper is partially supported by the China National Key Technology R&D Program with no. 2012BAH24B02.

REFERENCES

[1] N. Sun and L. Sh. Li, "The Study of Hierarchy Assessment Theory about Meteorological Disasters". 2007 Annual Conference of the Chinese Meteorological Society, July.2007, pp. 85-93.

[2] S. Stein, R. Geller, and M. Liu, "Why earthquake hazard maps often fail and what to do about it, "Tectonophysics, July 2012, pp. 1-48.

[3] X. Li and Z. Li, "Preliminary Discussion on the Theoretical Framework of the Earthquake Symbology System Building," Technology for Earthquake Disaster Prevention, July.2012, pp. 37-41.

[4] A. Schaff, "Introduction to Semantic," Pergamon Press, Oxford, New York, 1962

[5] Q. Zeng and E. Yang, "Event Anonotation and Analysis about the Content in Sudden Events Discourse," Advances of computational Linguistics in China, July 2009, pp. 600-605.

[6] China Meterological Administration, Symbols for meteorological disaster warning signal. China Meteorological Administration, 2004,(4),http://zwgk.cma.gov.cn/web/showsendinfo.jsp?id=1011.

[7] D.S. Mileti, "Natural Hazards and Disasters – Disasters by Design A Reassessment of Natural Hazards in the United States," Joseph Henry Press, Washington DC, 1999.

[8] M.Me, et al.," Geological disaster analysis and risk evaluation by GIS," Journal of Geosciences Translation, March.1995, pp. 72-79.

[9] A. Bagga, "Analyzing the Complexity of a Domain With Respect To An Information Extraction Task," Proceedings of the MUC-7, 1998,(6) http://www.muc.saic.com.

[10] National Disaster Reduction, Disaster Information. http://www.jianzai.gov.cn/.

[11] Z. Hu and H.W. Yan, Analysis on Linguistics Mechanism for cartographic symbols and its application. Geography and Geo-Information Science, January.2008, pp. 17-20.

[12] L.J. Sun, Y. Zhu, and X.Diaodong Liu, Research on the Theory of E-Government Emergency Information Publication Rapidly Based on Hierarchical Workflow. Geomatics World, June.2009, pp. 16-21.

[13] ACE.ACE Chinese Annotation Guidelines for Event. http://www.ldc.upemn.edu/Projects/ACE/docs/Chinese-Events-uidelines_ V5.5.1.pdf.2005C.

[14] X.H. Tan, The Application of Fuzzy Comprehensive Evaluation and Grey Assembly Classifying in lithologic Classification . Anhui geologic, April.1996, pp.71-76.

[15] L.N. Dang, F. Wu and Xuedong Li, Representation pf map symbols based on description method. Journal of Geomatics Oct, July 2007, pp.16-18.

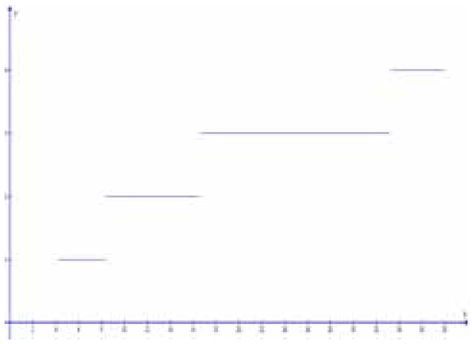


Figure 1. Example of Graphic for classified function

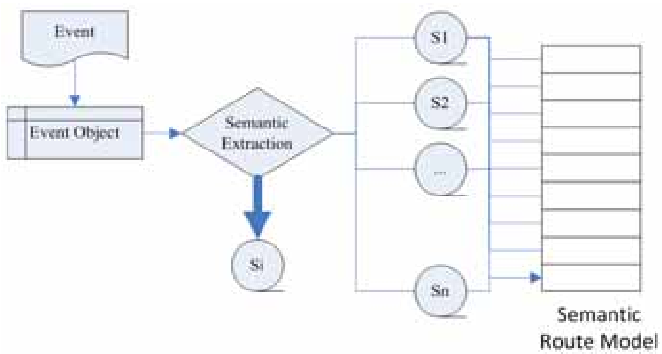


Figure 2. Software design model for symbols extraction

Figure 3. Emergency released in map by symbol