A Design of Semantics Filter for GIS

Moeko Nerome[1], Tomohide Yabiku[1,2], Yoshitaka Matsuda[1,2],
Dongshik Kang[1,2], Hayao Miyagi[1,2] and Kenji Onaga[1]
[1]Okinawa Research Center, Telecommunications Advancement Organization
The Okinawa Prefecture South General Building 12th Floor,
1 Asahi-machi, Naha, Okinawa, 900-0029, Japan
Tel/Fax. +81-98-862-3986, E-mail: moe@eva.ie.u-ryukyu.ac.jp
[2]Department of Information Engineering, University of the Ryukyus
1 Senbaru, Nishihara, Okinawa, 903-0213, Japan

Abstract: We propose a semantics filter aimed at filtering unnecessary data for a user. In the Geographic Information Systems (GIS), transfer of spatial data for each user is one of the important problems. We design a system that filters spatial data by using user's information adaptively. Our system derives the degree of geographical knowledge and the priority of buildings by using fuzzy reasoning. Furthermore, this system computes the priority of roads by deriving the contiguity relation between a building and a road. This paper describes a method for filtering of spatial data.

1. Introduction

In recent years, the field of GIS has attracted attention as computer throughput is improved. GIS is defined as a computer system that stores and links geographically referenced data with graphic map features to allow a wide range of information processing, display operations, map production, analysis, and modeling[1].

Since spatial data is accompanied not only by the positional information and geometric information on object, but by the attribute information on itself as well, the amount of spatial data becomes huge. Therefore, reduction of transfer cost and improvement in transfer response are required. Especially, it is difficult to transfer the spatial data frequently to the user of mobile GIS, with the constraints in communication environment and memory capacity. Thus, in the map information service currently provided on the internet, the spatial data which is filtered beforehand is prepared by the information provider. The current research under study is also aimed at looking into the possibility of a provider setting a value on each spatial data, and transferring the selected data such that the value is maximum[2]. The spatial data transferred in such a way will be determined by the subjectivity of the information provider. Therefore, it is difficult to provide information in consideration of the feature, demand, conditions, etc. which differ from one individual to another.

This paper proposes a semantics filter aimed at extracting only the most needed spatial data autonomously according to each user's conditions, demand, and so on. By filtering redundant data using the semantics filter, the amount of data is reduced and an intelligible map is offered to a user.

Since the judgment criteria for reasoning the data which a user desires is ambiguous, our semantics filter uses fuzzy theory as the filtering technique. Furthermore, an object-oriented GIS[3], which can treat each data as an object, is used. This paper describes the filtering method of the data by fuzzy reasoning using the user's information, and how to determine the priority of roads from the priority of buildings.

2. Spatial data transfer system using semantics filter

In this research, a semantics filter is defined as the filter which extracts only the data which is meaningful for a user. In order to interpret the spatial data which a user needs and to show the suitable data, it is important to make a decision on the criteria necessary for the filtering process. The criteria of filtering are roughly divided into the following 3 parts:

1. Purpose of using spatial data
2. User's geographical preliminary knowledge
3. Priority of spatial data

According to the first item above, it is possible to extract data which suits the purpose to some extent beforehand by the provider. For the last two parts, since many patterns are assumed and formulation of the relevant equations is difficult, then the judgement standard is also ambiguous. Therefore, in this research, we design the "Semantics Filter" which determines the data shown to a user by transferring only the data suitable for the individual. Fuzzy reasoning is employed to construct the semantics filter, due to the fact that it is not necessary to prepare many rules, and it excels in dealing with ambiguity of human being. Since the semantics filter derives the degree of geographical knowledge and the priority of each object from general user's information such as age, sex, residence, etc. dynamically, only the suitable spatial data is extracted. It is possible to filter not only building objects but road objects by using the object-oriented GIS as development environment.

In addition to the semantics filter which we proposed before[4], road objects also are filtered in this research. The flow of the semantics filter proposed in this paper is shown in Figure 1.

We represent an outline of the semantics filter, which includes the determination of the priority of roads, as follows:

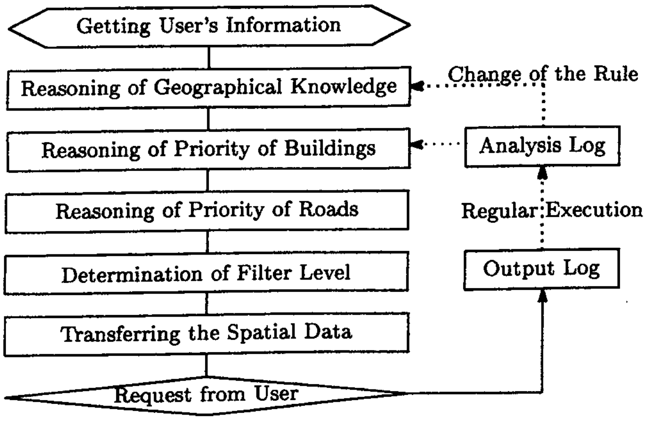

Figure 1. Filtering flow.

Step1: Getting user's information
The user's information such as age, sex, occupation, and place of residence region, is inputted beforehand.

Step2: Reasoning of geographical knowledge
The degree of geographical knowledge is guessed by fuzzy reasoning.

Step3: Reasoning of the priority of buildings
The priority of building object is determined by user's information and the type of building.

Step4: Determination of the priority of roads
The priority of a road object is computed by deriving the contiguity relation between a building and a road.

Step5: Determination of a filter level
The filter level is determined by a user's degree of geographical knowledge and by a user's terminal environment.

Step6: Transferring the spatial data
According to a determined filter level, spatial data objects are filtered.

Step7: Request from user
The displayed result is changed by user's request.

Step8: Output log
Access logs are stored in database by user.

Step9: Analysis log
All the collected access logs is periodically analyzed. The results of analysis are used to renew the rules.

Here, the degree of geographical knowledge shows the degree by which a user knows about the region. And the priority of buildings and roads show the degree of importance for a user. We explain the details of Step2 to Step6 in Section 3.

3. Design of semantics filter

This section describes reasoning of the degree of geographical knowledge and the priority of the buildings of the user by fuzzy reasoning using the user's information. Moreover, the method of determining the priority of roads from the priority of buildings is proposed. Finally, our method of the data based on the priority of buildings and the priority of roads is explained.

Table 1. Rules for reasoning of geographical knowledge.

Rule1:	If the distance is **close** and the user is **young**, then the user may be **knowing** the region **well**.
Rule2:	If the distance is **just a little bit close** and the user is **young**, then the user may be **knowing** the region.

3.1 Reasoning of the geographical knowledge

In order to reason the degree of knowledge, a reasoning rule is generated from the "user's age" and the "distance between the user's residence and the region desired". An example of the reasoning rules is shown in Table 1.

In Table 1, the user's degree of geographical knowledge is reasoned from "distance between the user's residence and the region desired" and "user's age". Each fuzzy variable is defined by a triangular membership function as shown in Figure 2.

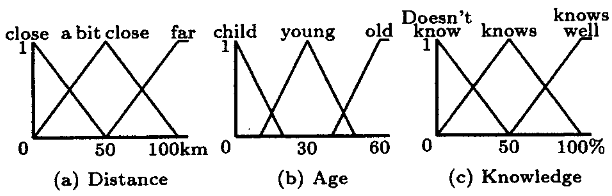

Figure 2. Triangular membership function.

We consider the case where the spatial data of X-city is offered to a 20-year-old user who lives in Y-town using these rules and membership functions. Here, the distance between the central point of X city and the central point of Y town is assumed to be 20km.

For rule 1, as shown in Figure 3, grade W_1 is derived by calculating the compatibility of the input "20km" and the membership function "close", and grade W_2 is derived by calculating the compatibility of the input "20-year-old" and the membership function "young". Since the operator of the two conditions in rule 1 is AND, W_2 is obtained ($W_2 < W_1$). And P is obtained by computing the minimum operation of W_2 and the membership function "knows well".

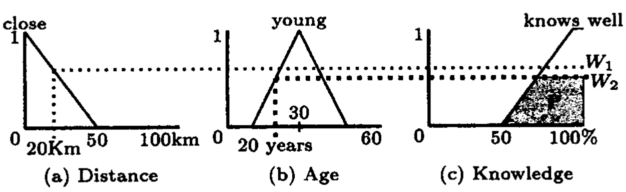

Figure 3. Calculation of rule 1.

Similarly, Q is derived from the analysis of rule 2, as shown in Figure 4. The final conclusion is obtained by computing the logical sum (OR) of the result P and Q. The obtained result is a fuzzy set. In order to derive a crisp value, defuzzification must be done, as shown in Figure 5.

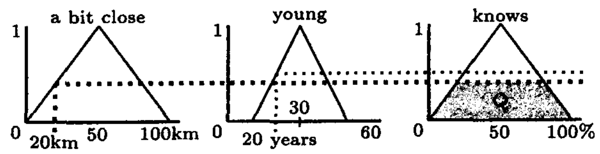

Figure 4. Calculation of rule 2.

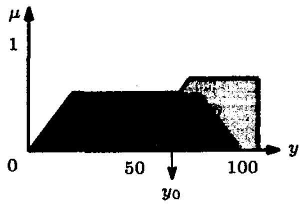

Figure 5. Defuzzification

The centroid method of defuzzification is represented by the following equation[5][6]:

$$y_0 = \frac{\int y \cdot \mu(y)dy}{\int \mu(y)dy} \quad (1)$$

The value derived by equation (1) shall be the degree of geographical knowledge.

3.2 Reasoning of the priority of Buildings

Methods for the reasoning of priority of the buildings are roughly divided into the following two ways:

- Reasoning from the geometrical information or topological information, which the building has.
- Reasoning from user's age, occupation, etc. and attribute information, which the building has.

In this research, we generate reasoning rules based on the latter method using user's age, sex and the type of building. The example of the reasoning rules is shown in Table 2.

By using these rules, the priority of each building is dynamically reasoned by the same method as reasoning of geographical knowledge. In order to reduce the amount of calculation concerning a reasoning, the reasoning using topological information without relation to a user's characteristics is performed beforehand.

Table 2. Rules for reasoning priority of buildings

Rule1:	If a building is near a crossing, then the priority is high.
Rule2:	If a building is a little near a road, then the priority is a little high.
Rule3:	If the area of a building is large, then the priority is a little high.
Rule4:	If a user is old and a building is an amusement facility, then the priority is low.
Rule5:	If user is a woman and a building is a beauty parlor, then the priority is high.

3.3 Determination of the Priority of Roads

In this paper, to decrease in the amount of data, we propose a method for determination of the priority of a road object, which leads to filtering of the unnecessary roads. It can filter unnecessary roads by proposal method. We consider that the road near the building having high priority is important. Therefore, the priority of a road is determined in accordance with the priority of a building. We have to take account of following.

- Derivation of the contiguity relation between a building and a road.
- Grouping some road objects, since a road many consist of several road objects.

We show the determination algorithm of the priority of a road as follows:

Step4-1: Grouping some road objects
The connection relationship between objects is derived by using the information of starting point and ending point of a road.

Step4-2: Saving the topology information
Each object saves the connected road objects.

Step4-3: Derivation of the contiguity relation between a building and a road
Distance between a building and a road is defined as altitude from center of the building to the road. Here, altitude is l. Contiguity definition is shown as follows:

If $l < L$ then "the building borders on the road"

where, L is threshold value.

Step4-4: Determination the priority of a road
The value of the priority of a road selects the maximum value of buildings bordering on the road .

In the Step4-3, if L is a constant value, distance from center of the large building to the road is far. In this case, it may be determined that the building doesn't border on the road even if the front of the building almost touches the road. In this paper, the L is computed by object. So, we compute the L using the following equation:

$$L_{OB(i)} = \alpha \times S_{OB(i)} \quad (2)$$

where, $S_{OB(i)}$ is a superficial measure of the building object ($OB(i)$) and α is constant value.

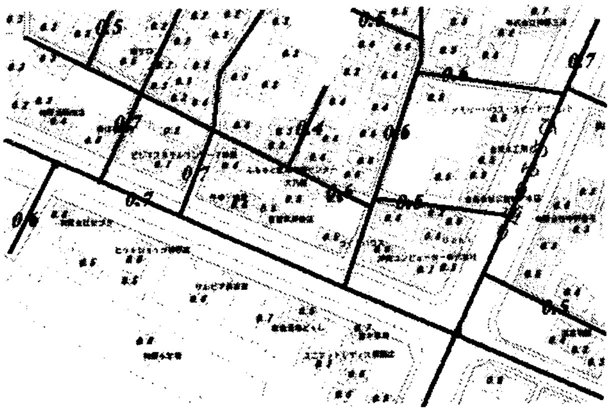

Figure 6. The result of the determination of priority value.

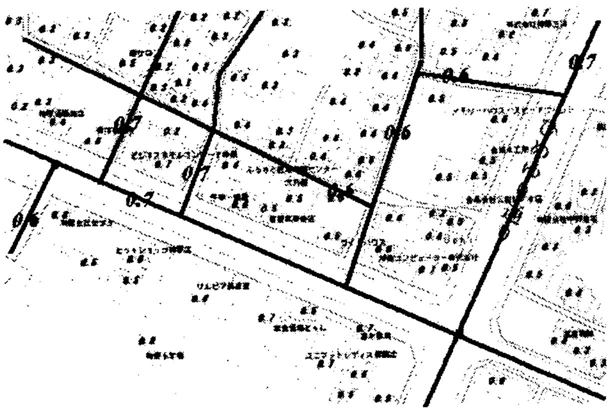

Figure 7. The result of filtering.

In the equation (2), the value of $L_{OB(i)}$ is determined by the superficial measure of the building object. Therefore, a problem of contiguity relation depending on the distance from the center of building can be solved by using equation (2).

We apply the semantics filter to an object-oriented GIS to investigate the performance of our filter. The simulation results involving a determination of the priority of a road are shown in Figure 6 and Figure 7. Figure 6 presents the result of the determination of the priority value. Figure 7 is the result of filtering, for the case, filter level = 0.5. Comparing the two figures, the roads of low priority are filtered. Therefore, the roads of low priority are assumed unimportant objects. From the simulation results, our filter can lead to a decrease in amount of data by filtering the unnecessary roads. Furthermore, we consider that our filter can create the simple map by displaying only important objects for user.

3.4 Determination of a filter level

The filter level is decided based on the degree of geographical knowledge derived by fuzzy reasoning. If the degree of user's geographical knowledge is high (if the user knows about the region well), the filter level is set to high and more objects are filtered. Conversely, if the degree of geographical knowledge is low, the filter level will be set to low and many objects will be displayed. The relation between the filter level and the degree of geographical knowledge is given by the following equation:

$$\gamma = \frac{\beta}{100} \tag{3}$$

Here, β expresses the degree of geographical knowledge and γ expresses the filter level.

3.5 Transferring the spatial data

The objects are filtered according to the filter level. Let a set of all the objects included in the region be $F = \{f_0, f_1, \cdots, f_n\}$, and the priority of each object be $v(f_i)$. The set $S(\subset F)$ of the object displayed by the filtering process is then derived as follows:

$$S = \{ f_i \mid v(f_i) \geq \gamma \}, \qquad 0 \leq \gamma \leq 1 \tag{4}$$

The equation (4) shows that the priority $v(f_i)$ given to the object can never be lower than the filter level γ.

4. Conclusion

In this research, we proposed the semantics filter which transfers only the most needed spatial data for a user. The proposal technique reasons a user's degree of geographical knowledge and the priority of buildings from clearly undistinguishable user's information using fuzzy reasoning, and computes the priority of roads from the contiguity relation between a building and a road. By setting the filter level corresponding to the degree of geographical knowledge, it is possible to show the spatial data suitable for the user. In future, we will examine the proposal technique in detail in order to improve the processing speed. Moreover, we will consider the generation method of the reasoning rule in log analysis.

References

[1] Yue-Hong Chou, *Exploring spatial analysis in geographic information systems*, Onword, 1997.

[2] Shohei Miyagi, Morikazu Nakamura, Kenji Onaga, "Optimal Spatial Object Collection for Mobile Geographical Information Systems", *Proceeding of ITC-CSCC '01*, Vol. II, pp. 1231–1234, 2001.

[3] Yoshinori Nie, Morikazu Nakamura, Hayao Miyagi, Kenji Onaga, "A Design and Implementation of Spatial Data Tranfer System with Object Oriented Technology -Ryutan", *Theory and Applications of GIS*, Vol. 8, No. 2, pp. 51–59, 2000.

[4] Yoshitaka Matsuda, Hayao Miyagi, Dongshik Kang, Junya Uema, Kenji Onaga, "Extraction of Spatial Data Using Meaning Filter", *Proceeding of FAN Symposium '00 in Tokyo*, pp. 417–420, 2000. (in Japanese)

[5] Stamatios V. Kartalopoulos, *Understanding neural networks and fuzzy logic*, IEEE Press, 1996.

[6] George J. Klir and Bo Yuan, *Fuzzy sets and fuzzy logic*, Prentice Hall PTR, 1995.