DEALING WITH SPATIAL RELATIONS FOR SPATIAL OBJECTS WITH RANDOMNESS IN GIS

DENG Min  
LI Chengming  
LIU Wenbao

KEY WORDS  spatial object; randomness; topologic relations

ABSTRACT  To design retrieval algorithm of spatial relations for spatial objects with randomness in GIS, this paper builds up the membership functions based on set theory idea, used for determination of topological spatial relations between random objects, such as between point and point, point and line or polygon, which provides theoretical basis for retrieving spatial relations between certain and random objects. Finally, this paper interprets detailed methods and steps of realizing them by means of some simple examples under the GIS’s environment.

1 Introduction

Spatial relations query is one of basic functions in GIS’s application. Most of current commercial GISs can only query spatial relations for spatial objects without any error or uncertainty, for example, to use computation geometry algorithm to determine whether a point falls inside a polygon (Worboys, 1995). However, because of inevitable errors of spatial data used to represent point, line and polygon in GIS, it is necessary to build up new algorithms to determine spatial relations for uncertain objects in GIS (Goodchild, 1989; Guptill, et al., 1995). Obviously, it is a fundamental problem on how to represent spatial objects with errors or uncertainties.

In general, uncertainty will be divided into fuzziness and random error. Therefore, all objects stored in GIS are classified as the following corresponding kinds: certain object, fuzzy object, random object and fuzzily random object. Here we take random object as an example, that is, only random error will be considered.

Some statistical models of representing random point, line and polygon have been constructed before (Liu, 1995; Shi & Liu, 1998; Liu, et al., 1998). On the basis of those work, this paper further discusses the problem of determination of spatial relations between random objects by means of set theory idea, that is, to determine quantitatively topological relations through constructing suitable membership functions.

2 Determination of topological relations between two points

In GIS, it is necessary to determine whether two points are the same. Below we presume that positional error models of all random points are circle normal distribution (Goodchild, 1991).

2.1 The case for two certain points

In general, we neglect positioning error for data edition and cartographic generalization, and to determine whether two points are the same point is
realized through comparison of the deviation $d$ between the points in a map with given critical value $d_0$. Its logical model is expressed as: if $d > d_0$ then $Z_1$ and $Z_2$ are not the same point, else $Z_1$ and $Z_2$ are the same point. According to the set theory, its membership is as follows:

$$A(d) = \begin{cases} 0 & d \notin A \\ 1 & d \in A \end{cases}$$  \hspace{1cm} (1)$$

where $A = \{x \mid x \leq d_0\}$ denotes a subset including all points with its deviation smaller than $d_0$, $A(d)$ denoting membership grade of element $d$ in subset $A$, if $A(d)$ equals to 1, $Z_1$ and $Z_2$ are the same points, that is to say, their different position in map is only caused by errors, such as cartographic error, or else, if $A(d)$ equals to zero, they denote two different point features.

2.2 The case for a certain point and a random point

Let $Z_1$ be a point without any error, for example, a point of laneway digging in design map, while $Z_2$ denotes a point with random error. Therefore, possible topological relationships between $Z_1$ and $Z_2$ are shown in Fig. 1. From the above assumption we know, error model of watering point $Z_2$ obeys circle normal distribution, and its planar projection is an equal density error circle (Goodchild, 1991). Moreover, along with the deviation away from the center increasing, its probability density becomes less. Therefore, it is more suitable to define the membership function used to determine topological relationship between $Z_1$ and $Z_2$ as:

$$A(d) = e^{-\frac{d^2}{2\sigma^2}}, 0 \leq d < +\infty$$  \hspace{1cm} (2)$$

where $\sigma$ denotes the standard error of $Z_2$, $d$ the deviation between $Z_1$ and $Z_2$, which can be attained by measure or computation. Therefore, the possibility degree of determining whether they are the same point will be computed by Eq. (2). For example, $A(d) = 0.8$ means that $Z_1$ and $Z_2$ are the same point with its probability of 0.8. And $A(d)$ will reduce with the decrease of $\sigma$ when $d$ is invariant. This index will be very useful when we make node snapping with a small deviation.

2.3 The case for two random points

As above-mentioned, the designed position of spatial point is without errors. However, for spatial data stored in GIS databases, for example, well canister, telegraph pole, hydrant, if they are captured by means of measurement, digitizing and so on, then random errors are included in these data. Their possible topological relations are illustrated graphically as Fig. 2. Intuitively, to determine whether two random points are the same point is dependent on overlapping areas of their density distribution (Liu, 1995). Therefore, it is more suitable to construct following membership function:

$$A(d) = k \int_{m}^{n} e^{-\frac{x^2}{2\sigma^2}} \left( \frac{d-x}{2\sigma} \right)^2 dx$$  \hspace{1cm} (3)$$

where $k = \left( \int_{-n}^{n} e^{-\frac{x^2}{2\sigma^2}} \left( \frac{d-x}{2\sigma} \right)^2 dx \right)^{-1}$ is a normalized factor; $n$ is often taken as three times standard error, $m = -n + d$, $d$ is the distance between $Z_1$ and $Z_2$, which becomes known by computation, taking value as $[0, +\infty)$; while $\sigma_1$, $\sigma_2$ are standard error of $Z_1$, $Z_2$ respectively; $x$ is an integral variable, taking value in the interval $[m, n]$. Therefore, if $\sigma_1 = \sigma_2 = \sigma$, Eq. (3) is simplified as follows:

$$A(d) = k \int_{m}^{n} e^{-\frac{1}{2\sigma^2}(2x^2 - 2xd + d^2)} dx$$  \hspace{1cm} (4)$$

This index will be very useful when we make node snapping with a small deviation.
3 Determination of topological relations between point and line segment

In GIS, it is necessary to determine topological relations between a point feature and a line feature. For example, to determine whether two lines are intersected or crossed is realized through determining whether the intersection point of two lines is inside these two lines respectively. Apparently, the results are classified as two kinds: point belonging to line, noted as \( P_i \in l_i \), or not belonging to line, i.e. \( P_i \notin l_i \) (Worboys, 1995).

3.1 The case for a certain point and a random line segment

The determination of topological relations between certain point and random line segment is realized by computation of vertical deviation distance from point to random line segment, as illustrated in Fig. 3. Here, \( P_1 \) and \( P_2 \) are two certain points, \( Z_1Z_2 \) is a random line with random error. We can see from Fig. 3, \( P_1 \) and \( P_2 \) are closer to \( Z_1Z_2 \). Obviously, it is probable that \( P_1 \) or \( P_2 \) are on the line \( Z_1Z_2 \). Therefore, its membership function of topological relations is similar to Eq. (2). It may be also regarded as the determination of topological relations between a certain point and its pedal point in random line segment. Here \( d \) denotes vertical deviation distance between \( P_i \) and \( Z_1Z_2 \); \( \sigma_i \) positional standard error of two base points in random line segment respectively.

\[
\sigma = \sqrt{1 - 2t + 2t^2} \sigma_i, i = 1, 2 \tag{5}
\]

where \( t = \frac{Z_1Z_i}{Z_1Z_2}, 0 \leq t \leq 1; \sigma_i \) are positional standard error of two base points in random line segment respectively.

![Fig. 3](image_url)

Fig. 3 Graphic illustration of topological relations between certain points and random line segment

3.2 The case for a random point and a certain line segment

As for determination of topological relations between random point and certain line, its quantitative determination method is similar to the mentionned in Section 2.2. Here, pedal point \( Z_i \) is a certain point.

3.3 The case for a random point and a random line segment

For the determination of topological relations for random point and random line segment, its membership function for quantitative determination of this kind of topological relation is the same to Eq. (3). Here, pedal point is a random point.

4 The determination of topological relations between point and polygon

In GIS, there is often the query on whether the points are in the same polygon. For example, to query whether a fire point is within some district, and to query whether a point of laneway digging is in an area threatened by water. When a point and a polygon are both certain, there will be three possible results, namely, 1) point falling inside the polygon, 2) point falling outside the polygon, 3) point falling on the boundary of the polygon (Worboys, 1995; Goodchild, 1991).

4.1 The case for a certain point and a random polygon

As is shown in Fig. 4, both \( P_1 \) and \( P_2 \) are certain points and \( O_1 \) is a random polygon with random error. Statistical error model of random polygon \( O_1 \) is shown in Liu, et al. (1998). Considering that its possibility of ‘point falling inside polygon’ is 0.5 when a point feature is on the boundary of polygon, thus it is suitable to extend Eq. (2) as follows:

\[
A(d) = \begin{cases} 
\frac{1}{2} e^{-\frac{d^2}{2\sigma^2}}, & d \leq 0 \\
1 - \frac{1}{2} e^{-\frac{d^2}{2\sigma^2}}, & d > 0 
\end{cases} \tag{6}
\]

where \( d \leq 0 \) and \( d > 0 \) denote respectively point lying outside or inside the polygon \( O_1 \), and the definition of \( \sigma \) is similar to Eq. (5).
4.2 The case for a random point and a certain polygon

Possible topological relations between random point $P_1$ and certain polygon $O_1$ are shown in Fig. 5. The methods for determining these topological relations are similar to that mentioned in Section 3.1, and the only distinction is that the pedal point on polygon’s boundary is a certain point.

$$A(d) = \begin{cases} \int_{-3\sigma_1}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx, & d \leq -3\sigma_1 \\ \int_{-3\sigma_1}^{-d} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx + \int_{d}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx, & -3\sigma_1 < d \leq 0 \\ \int_{-3\sigma_1}^{-d} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d+x)^2}{2\sigma_1^2}}\right) dx + \int_{d}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d+x)^2}{2\sigma_1^2}}\right) dx, & 0 < d \leq 3\sigma_1 \\ \int_{-3\sigma_1}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d+x)^2}{2\sigma_1^2}}\right) dx, & d > 3\sigma_1 \end{cases}$$

where $k = \left[\int_{-3\sigma_1}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} dx\right]^{-1}$ is a normalized factor, $\sigma_1$ denotes positional standard error of random point $P_1$, and $\sigma_2$ denotes positional standard error of pedal point $Z_t$ in random boundary line segment $Z_1Z_2$, which is a random point.

Above we only discuss the relationships between a point and a polygon by means of one of its boundary line segments. When one point lies in the neighborhoods of two line segments, it is suitable to take their maximum values. Here it is noted that we only compute its probability of all possible relationships between a point and a polygon on a map by means of boundary line segments after determining their basic relations, such as “point falling inside polygon”, while these basic relationships will be directly attained by means of computational geometry methods.

4.3 The case for a random point and a random polygon

When positional data of point and polygon features are all with random errors, relatively positional relations between them have five possible cases. Hence, it is necessary to extend the membership function in order to determine their topological relations, i.e.

$$\begin{align*}
A(d) &= \begin{cases} 
\int_{-3\sigma_1}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx, & d \leq -3\sigma_1 \\
\int_{-3\sigma_1}^{-d} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx + \int_{d}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx, & -3\sigma_1 < d \leq 0 \\
\int_{-3\sigma_1}^{-d} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d+x)^2}{2\sigma_1^2}}\right) dx + \int_{d}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d+x)^2}{2\sigma_1^2}}\right) dx, & 0 < d \leq 3\sigma_1 \\
\int_{-3\sigma_1}^{3\sigma_1} e^{-\frac{x^2}{2\sigma_1^2}} \left(1 - \frac{1}{2} e^{-\frac{(d-x)^2}{2\sigma_1^2}}\right) dx, & d > 3\sigma_1
\end{cases}
\end{align*}$$

5 Examples

Firstly, we should determine those spatial features necessary to query their spatial relations, and draw reliability circle of point’s positional error, reliability band of line’s positional error and reliability donut of area’s positional error respectively based on three times standard errors (Liu, 1995). Then, we compute membership function value of their topological relations with the above-mentioned formula, and to determine resulting description of topological relations through comparison of these computation values with 0.5.

5.1 The determination of spatial relations between two points

The steps of the determination of spatial relations between two points are as follows.

1) Compute distance between points $(x_1, y_1)$ and $(x_2, y_2)$ with Euclidean formula:
\[ d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{\Delta x_{12}^2 + \Delta y_{12}^2} \] (8)

2) Evaluate standard error by means of the method in Liu (1995);
3) Compute membership grade \( A(d) \) by Eqs. (1)-(3). The results are shown in Table 1. Here, if computed value of \( A(d) \) is larger than 0.5, then we determine these two points \((x_1, y_1)\) and \((x_2, y_2)\) to be the same points; or else, they are not.

### Table 1 Determination of topologic relations between points

| \( Z_1 \) | \( Z_2 \) | \( A(d) \) |
|---|---|---|
| \( x_1/m \) | \( y_1/m \) | \( \sigma_{x_1}/cm \) | \( \sigma_{y_1}/cm \) | \( x_2/m \) | \( y_2/m \) | \( \sigma_{x_2}/cm \) | \( \sigma_{y_2}/cm \) |
| CP/RP | 1000.00 | 800.00 | 0.00 | 0.00 | 990.00 | 80.00 | 1.70 | 1.70 | 0.332 |
| RP/RP | 960.00 | 800.00 | 1.70 | 1.70 | 88.00 | 80.00 | 1.70 | 1.70 | 0.773 |

Note: \( CP \) denotes certain point; \( RP \) denotes random point.

5.2 The determination of spatial relations between point and line segment

The detailed steps of the determination of spatial relations between point and line segment are as follows:
1) Compute vertical distance \( A(d) \) from point \( P_i(x_i, y_i) \) to line segment \( Z_1Z_2 \) with the following formula:
   \[ x_i = x_2 + \frac{m}{n} \Delta x_{21}, y_i = y_2 + \frac{n}{m} \Delta y_{21} \] (9)
   where 
   \[ m = \Delta x_{21} + \Delta y_{21} \]
   \[ n = \Delta x_{21} \Delta x_{11} + \Delta y_{21} \Delta y_{11} \]
2) Evaluate standard error \( \sigma \) by Eq. (5);
3) Compute membership grade \( A(d) \) by Eqs. (2) and (3). The resulting data are listed in Table 2. Here, if computation value of \( A(d) \) is larger than 0.5, then we determine that point \( P_i(x_i, y_i) \) falls in the line segment \( Z_1Z_2 \), denoted as \( P_i \in Z_1Z_2 \); or else, \( P_i \notin Z_1Z_2 \).

### Table 2 Determination of topologic relations between point and line segment

| \( P_i \) | \( Z_1 \) | \( P_i \) | \( Z_1 \) |
|---|---|---|---|
| \( x_1/m \) | \( y_1/m \) | \( \sigma_{x_1}/cm \) | \( \sigma_{y_1}/cm \) | \( x_1/m \) | \( y_1/m \) | \( \sigma_{x_1}/cm \) | \( \sigma_{y_1}/cm \) | \( x_2/m \) | \( y_2/m \) | \( \sigma_{x_2}/cm \) | \( \sigma_{y_2}/cm \) | \( A(d) \) |
| CP/RL | 1000.00 | 800.00 | 0.00 | 0.00 | 1500.00 | 1200.00 | 3.96 | 3.96 | 500.00 | 400.00 | 3.96 | 3.96 | 0.212 |
| RP/CL | 1022.10 | 827.60 | 0.00 | 0.00 | 1500.00 | 1200.00 | 3.96 | 3.96 | 500.00 | 400.00 | 3.96 | 3.96 | 0.490 |
| 1240.00 | 1010.00 | 0.00 | 0.00 | 1500.00 | 1200.00 | 3.96 | 3.96 | 500.00 | 400.00 | 3.96 | 3.96 | 0.741 |
| RP/RL | 1022.10 | 827.60 | 1.70 | 1.70 | 1500.00 | 1200.00 | 0.00 | 0.00 | 500.00 | 400.00 | 0.00 | 0.00 | 0.258 |
| 1253.70 | 1044.60 | 2.50 | 2.50 | 1500.00 | 1200.00 | 0.00 | 0.00 | 500.00 | 400.00 | 0.00 | 0.00 | 0.443 |
| 1200.00 | 960.00 | 1.70 | 1.70 | 1500.00 | 1200.00 | 4.95 | 4.95 | 500.00 | 400.00 | 4.95 | 4.95 | 0.352 |
| RP/RL | 1231.20 | 921.00 | 1.70 | 1.70 | 1500.00 | 1200.00 | 4.95 | 4.95 | 500.00 | 400.00 | 4.95 | 4.95 | 0.553 |
| 1262.50 | 811.90 | 2.50 | 2.50 | 1500.00 | 1200.00 | 4.95 | 4.95 | 500.00 | 400.00 | 4.95 | 4.95 | 0.892 |

Note: \( CL \) denotes certain line; \( RL \) denotes random line.

5.3 The determination of spatial relations between point and polygon

The concrete steps of the determination of spatial relations between point and polygon are as follows.
1) Compute vertical distance \( A(d) \) from point \( P_i(x_i, y_i) \) to polygon’s boundary lines segment \( Z_1Z_2 \) with Eq. (9);
2) Evaluate standard error \( \sigma \) by Eq. (5);
3) Compute membership grade \( A(d) \) by Eqs. (6) and (7). The resulting data are listed in Table 3. Here, if computation value of \( A(d) \) is larger than 0.5, then we determine that point
\( P_i(x_i, y_i) \) falls in the polygon; if equals 0.5, then \( P_i(x_i, y_i) \) falls on one of boundaries of polygon; or else, point \( P_i(x_i, y_i) \) falls outside the polygon.

| \( P_i \) | \( x_i/m \) | \( y_i/m \) | \( s_1/cm \) | \( s_2/cm \) | \( x_i/m \) | \( y_i/m \) | \( s_1/cm \) | \( s_2/cm \) | \( x_i/m \) | \( y_i/m \) | \( s_1/cm \) | \( s_2/cm \) | \( A(d) \) |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| CP/RA | 117.62      | 98.97       | 0.00        | 0.00        | 150.00      | 120.00      | 3.54        | 3.54        | 50.00       | 40.00       | 3.54        | 3.54        | 0.303       |
| RP/CA | 117.62      | 98.97       | 3.54        | 3.54        | 150.00      | 120.00      | 0.00        | 0.00        | 50.00       | 40.00       | 0.00        | 0.00        | 0.374       |
| RP/RA | 113.75      | 103.81      | 3.54        | 3.54        | 150.00      | 120.00      | 4.95        | 4.95        | 50.00       | 40.00       | 4.95        | 4.95        | 0.145       |

Note: CA denotes certain area; RA denotes random area.

6 Conclusion

1) The determination of topological relations between two point features is dependent on their distance as well as their precision. However, the determination of topological relations between point and line or polygon is related to their positional precision, vertical distance and relatively planar position between point and line or polygon.

2) If the distance between point and point, line, polygon is larger than double positional standard error, it is unsuitable to determine whether they are "the same points" and point falls on the line or in the polygon.

3) This paper is only limited to discuss the determination of topological relations for point/point, point/line and point/polygon. Future researches are to focus on the problems about the determination of topological relations between line/line, line/polygon and polygon/polygon.

References

1 Worboys M. GIS, a computing perspective. London: Taylor & Francis, 1995
2 Burrough P A, McDonell R A. Principles of geographical information systems. Oxford: Oxford University Press, 1998, 220-240
3 Blakemore M. Generalization and error in spatial databases. Cartographica, 1984 (21):131-139
4 Goodchild M F. Modelling error in object and fields. In: Goodchild, Gopal eds. Accuracy of Spatial Databases. London: Taylor & Francis, 1989, 107-113
5 Guptill S C, Morrison J L. Elements of spatial data quality. Oxford: Elsevier Scientific, 1995
6 Liu W B. Modeling uncertainty in GIS and remote sensing [Ph. D Thesis]. Wuhan: Wuhan Technical University of Surveying and Mapping, 1995 (in Chinese)
7 Shi W Z, Liu W B. A new stochastic model for handling positional uncertainties of line feature in GIS. Acta Geodaetica et Cartographica Sinica, 1998, 27(1):37-44 (in Chinese)
8 Liu W B, Dai H L. Analytical representation of positional uncertainties of line features in GIS. Acta Geodaetica et Cartographica Sinica, 1998, 27(3):231-237 (in Chinese)
9 Liu W B, Dai H L. A new method of positional error donut for planar polygon in GIS. Acta Geodaetica et Cartographica Sinica, 1998, 27(4):338-344 (in Chinese)
10 He X G. Theory and technology of handling fuzzy knowledge. Beijing: National Defence Press, 1998 (in Chinese)
11 Goodchild M F. Issues of quality and uncertainty. In: Muller J C. Advances in Cartography. Elsevier Science Publishers Ltd., 1991, 113-138