

Metric specification and improvement of intersection models

Qilun Xu^{1,a}, Yuan Liu^{2,b}

¹*School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China*

²*Mathematics Teaching and Research Group, Leiyang No. 1 Middle School, Leiyang, 421899, China*
^axqilun615@163.com, ^bmaozai98617@163.com

Abstract: Egenhofer's intersection model for describing topological relations has a normative problem in metrics, and the inner, outer and boundary boundaries of different dimensional online objects do not conform to the definition of inner, outer and boundary based on metric-neighborhood language pairs in point set topology. Therefore, by establishing a mapping from the dimension to the lower dimension, the boundary, exterior and interior of the line object are inversely mapped in one dimension, so that the line object in the two-dimensional space still conforms to the three-part division of the metric specification. This process allows the new node degree to be used, and the resulting topological formal model is more distinguishable than other models.

Keywords: GIS, Topological Relations, 9-Intersection Model, Metric Specification

1. Introduction

Point-set topology is a mathematical language based on open sets and neighborhoods, which provides definitions for the interior, boundary, and exterior of a set. In a specific topological space, to conceptually determine the vicinity of a point ^{[1][2]}, i.e., the mathematical form of a neighborhood, it is necessary to specify a metric form. Therefore, according to different metric forms, the range of the interior, boundary, and exterior of a certain geometric object in the corresponding metric topology will be different. Based on point-set topology, an intersection model is proposed, which decomposes a geometric object into two subset parts: the interior and the boundary, while considering the exterior of the object. The operation type between subsets is defined as intersection, and a metric assumption is made ^[2]: "In a practical application where a certain topology of interest needs to be judged, the topology may not only have a universally applied metric." This treatment avoids the discussion of a unified metric and provides a preset definition of the interior, exterior, and boundary of points, lines, and surfaces. However, research by Deng Min ^[3] and others indicates that this leads to a contradiction between the preset definition of the boundary and interior in the intersection model in different dimensional Euclidean spaces and point-set topology.

All subsequent research improvements based on the intersection model have followed this assumption. For example, the DE-9IM model ^[4] that introduces dimension expression, the intersection model ^[5] that introduces internal segmentation, the 25-intersection model ^[6], and so on. These improved studies basically change or refine the partition rules (such as dividing into boundary, interior, hole area, etc.), increasing the total number of divided parts; or they add description rules for each intersection part (such as Euler number, dimension, node degree, etc.), thereby enhancing the model's ability to distinguish different topological relationships. However, they have not changed the metric assumption of the Egenhofer model, so in specific scenarios, they still have some places that can easily cause conceptual confusion. In practical applications of GIS, a certain distance between two objects in geospatial space, such as the Euclidean distance or Hausdorff distance in three-dimensional space, is established based on the corresponding metric. Based on the metric, it can be determined whether geographical objects intersect. Similarly, in the topology induced by a certain metric, the relationship between complex geometric objects composed of lines and surfaces (such as geometric objects composed of several simple lines tangent to the boundary of a surface) can be judged, or a fuzzy relationship that needs to be defined by a certain metric (for example, proximity, it needs to be judged how close two objects are to be considered close, which must introduce a unified metric). In these cases, Egenhofer's assumptions are inconvenient, so in these application scenarios, all current intersection models can be further mathematically standardized to ensure the rigor of discussing topological relationships and facilitate the

subdivision of topological relationships.

This study is based on the metric problem, further discusses the mathematical formalization norms of various intersection models, and improves the subdivision method of line-surface topological relationships based on the work of Zhou [5] and others. For the topological relationship between possible complex objects, the research gives a formalized description model of fixed metrics to better improve the expression of existing topological relationships in GIS applications.

2. Topology Relationship Model

In Geographic Information Systems, geometric objects can often be viewed as a set of points. The topological relationships between geometric objects can be defined and distinguished by the results of certain operations (such as intersection, union, difference, etc.) on subsets, once the rules for partitioning subsets are determined. These results can be used to determine whether they intersect or are separate. In addition to familiar relationships such as intersecting, separating, and adjacent, more complex topological relationships can be formed by combining the results of multiple operations on two sets of subsets, such as spanning, ring surrounding, crossing, etc. Furthermore, we can describe and distinguish the topological relationships between different types of spatial objects based on the relationships between these subsets, according to the needs of practical applications.

2.1. Intersection Model

The intersection model is based on point-set topology and expresses geometric objects as boundary, interior, exterior, above three parts, as: $A^\circ, \partial A, \neg A$. In the topological space where we want to handle topological relationships, for the topological relationship between spatial targets A and B , it is expressed in terms of an intersection matrix [2]:

$$R_{9IM}(A, B) = \begin{bmatrix} A^\circ \cap B^\circ & A^\circ \cap \partial B & A^\circ \cap \neg B \\ \partial A \cap B^\circ & \partial A \cap \partial B & \partial A \cap \neg B \\ \neg A \cap B^\circ & \neg A \cap \partial B & \neg A \cap \neg B \end{bmatrix} \quad (1)$$

Where each matrix element represents the intersection between a certain part of A and B . However, when actually storing the relationship matrix, it often requires parameterization of f , as functions in $9-IM$ or $Dim(x,y)$ in $DE-9IM$ model. And with this, it can express at least 2^9 types of spatial relationships in theory. To describe the interior, exterior, and boundary of sets A and B , we first need to specify the mathematical category of the discussion sets A and B . This means indicating the topology of A and B on the universal set U , as well as the metric d induced by τ .

2.2. Problems with the Intersection Model in Metrics

This section must be in one column. The intersection model discusses the internal, external, and boundary forms of points, lines, and planes under different metric topologies. This treatment may cause the 'interior' and 'boundary' of different dimensional geometries on the same metric topology to not necessarily strictly conform to the definition of point set topology. As pointed out in the research by Deng Min [7] and others, in the one-dimensional situation, the neighborhood of a point on the line is a certain section along the line, but when discussing the line and plane targets on the same Euclidean plane, the points on the Euclidean plane R^2 , generally have: $d_2(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$, then the base open neighborhood form of the metric space (X, d_2) has become a two-dimensional circle that does not include the boundary.

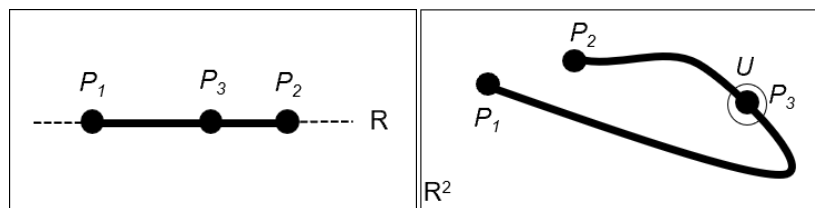


Figure 1: Simple line with interior point.

This is not a problem for a certain surface set $S_0 \subset R^2$, but for the line set on this metric space, the neighborhood obtained by the definition of metric d_2 , such as point p_3 on Figure 1, its neighborhood

$U(p_3, \varepsilon)$ should be an open plane circle around p_3 and is impossible to find an open set $U(p_3, \varepsilon)$ that does not contain the external points of this line to make $x \in (p_3, \varepsilon) \subset A$, so in this metric space, p_3 cannot meet the definition of internal points, but it can meet the definition of boundary points of this curve, that is, when $\forall \varepsilon > 0$ you can find points on and off the line in $U(p_3, \varepsilon)$ at the same time, which makes any simple line on the ordinary metric topology of R the boundary of this line, and the interior of this line is an empty set; this obviously contradicts the default definition of the intersection model that ‘the interior of the line target is the points on the line except the endpoints’.

For example, a certain surface in a three-dimensional Euclidean space, divided into the interior and boundary according to the three-dimensional Euclidean measurement, does not have an interior, and the boundary is itself, which is obviously contradictory to the ‘interior of the surface’ and ‘boundary of the surface’ preset by the intersection model. But this contradiction is not irreconcilable, it can be reconciled based on a certain mapping. Therefore, in the application scenarios where it is necessary to establish a common measure, further mathematical norms can be made. Generalized to R^n , for example, for different dimensional entities on the same dimensional Euclidean space, their boundary and internal division will have the above problems, where n-dimensional entities [8] [9] refer to:

$$\text{On the } \tau_d \text{ reduced by } d, \text{ } n - \text{dimensional volume is not zero } (n \geq 2) \quad (2)$$

$$\text{Boundary is homeomorphic to } S_n = \{x \in R^n \mid ||x|| = 1\} \quad (3)$$

Obviously, for a geometric entity A that satisfies the $k < n$ condition, its measure on R^n is less than or equal to 1, and any element of A at most constitutes its own boundary point. Now for the need of distinction, define $I_{Egen}(A), B_{Egen}(A)$ as the ‘interior’ and ‘boundary’ preset for a certain geometric object in the Egnhofer intersection model. $I_{Nor}(A), B_{Nor}(A)$ is the interior and boundary of a certain geometric object under general measurement. They are given their limited range according to different dimensions, where ℓ is a simple line, S is a simple surface, and A is a n-dimensional entity, In order to represent spatial entities of different dimensions, the distinctions in the division of their interior, exterior, and boundary in the intersection model and strict set topology are shown in Table 1.

Table 1: Comparison of the Egenhofer model with the point-set topology specification.

		R^1	R^2	R^3	R^4	...	R^n
<i>Egenhofer 9-IM Model</i>	$I_{Egen}(\ell)$	l°	without endpoints	without endpoints	without endpoints	...	without endpoints
	$B_{Egen}(\ell)$	∂l	endpoints	endpoints	endpoints	...	endpoints
	$I_{Egen}(S)$	-	S°	without linear feature	without linear feature	...	without linear feature
	$B_{Egen}(S)$	-	∂S	linear feature edge	linear feature edge	...	linear feature edge
	$I_{Egen}(A)$	-	-	-	-	...	A°
	$B_{Egen}(A)$	-	-	-	-	...	∂A
<i>Point-set Topology</i>	$I_{Nor}(\ell)$	l°	\emptyset	\emptyset	\emptyset	...	\emptyset
	$B_{Nor}(\ell)$	∂l	l	l	l	...	l
	$I_{Nor}(S)$	-	S°	\emptyset	\emptyset	...	\emptyset
	$B_{Nor}(S)$	-	∂S	S	S	...	S
	$I_{Nor}(A)$	-	-	-	-	...	A°
	$B_{Nor}(A)$	-	-	-	-	...	∂A

For example, considering the river in *Figure2* as a whole as an object considering topological relationships, its confluence points of each tributary still belong to the internal part in the 9IM model, and the collection of endpoints at the end of each tributary is regarded as the boundary. But for other objects, such as the confluence point that a certain object passes through, it is difficult to reflect this internal topological structure when actually analyzing the topological relationship with the 9IM model. But if you break down each tributary into independent objects according to a simple line, then use a model similar to 9IM, the resource consumption of calculating topological relationships will increase a lot, and each tributary object has to calculate the result of set operations with the three parts of other objects in the form of three parts inside, outside, and boundary.

The measure can flexibly define the concept of distance on the limited topology, and then change the division rules of part of the geometric object to actually calculate and judge the topological relationship,

especially for similar bifurcated lines and other objects with complex internal topological structures. If you want to further improve the rules for dividing objects, you have to discuss the issue of measurement. For objects with special geometric structures, such as self-intersecting lines under the same topology, the union of geometric objects with different topologies, and complex objects that are difficult to discuss boundaries, if you want to give the distinction rules of geometric objects based on the Egenhofer model, there is no ability to distinguish them.

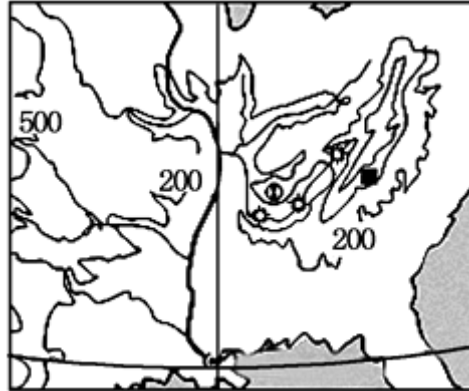


Figure 2: Example diagram of river basin.

Therefore, all the topological relationship models based on the Egenhofer intersection model, such as: 9IM, DE-9IM, or the V9I model based on Voronoi diagram^[10], the VW model using the target as a whole and its Voronoi area^[11], the E-WID model based on cross-set operations^[12], the RCC model based on the logic of the whole space^[13], these models are difficult to describe the relationship between these objects.

2.3. Improvement of the Metric Mathematical Norm of the Intersection Model

In the intersection model, the two planar objects A and B are defined in a certain topological space. The interior of the simple line object (non-self-intersecting, non-branching) needs to be defined according to the usual Euclidean metric d on (R^2, d) . That is, we need to give a continuous mapping $f : L \rightarrow I$ for a simple line object L on a metric topology τ on R^2 , and an interval $I = [a, b]$, $a > 0$. Each point on I is used to label the corresponding points on L , thus obtaining $d = \rho(x, y) = |x - y| \geq 0$. The distance between two points on the interval is used to measure the position of a point within the simple line (distance from the endpoint, i.e., the boundary point). After defining the metric d , we give the definition of the neighborhood we want on the metric topology τ , where U is on the topology τ_d induced by d :

$$U(P, \delta) = \{P \in \tau | d(x, y) < \delta\} \quad (4)$$

$U(P, \delta)$ is the preimage of U_f , where:

$$U_f(x_0, \varepsilon) = \{x \in R | |x - x_0| < \varepsilon\} \quad (5)$$

In the case of a simple line $\ell \subset 2R^2$, it is clear that the interior point P has $x_0 = f(P) \neq a, b$, x_0 is not an endpoint. In the case where ℓ does not intersect itself, the continuous mapping f can be a bijection in the practical sense (the condition for homeomorphism is stronger than this condition, homeomorphism also stipulates that f^{-1} is also continuous, but it is not necessarily continuous in our argument). According to the definition of continuous mapping, due to the connectedness of R^2 itself, it can define the interior points on the line under the one-dimensional space that conforms to $f(\ell)$, and at the same time, it has:

$$\varepsilon < \frac{b+a}{2} - x_0, \quad x_0 \in U(x_0, \varepsilon) = \{x \in f(L) | |x - x_0| < \varepsilon\} \subset f(L) \quad (6)$$

Therefore, we can use a certain continuous mapping f to make the original intersection model satisfy the definition of the neighborhood condition of the line set's inner point under the general Euclidean measure of R^2 in $U(P, \delta)$. So in the research of topological relations based on the Egenhofer intersection model, what is actually discussed is that each A and B are divided into interior and boundary $(\partial A, A^\circ)_{\tau_1}, (\partial B, B^\circ)_{\tau_2}$ on two metric topologies τ_1, τ_2 , and then their respective $f^{-1}[\partial A], f^{-1}[A^\circ], g^{-1}[\partial B], g^{-1}[B^\circ]$ corresponding images, these images are on a certain metric topology τ on R^2 , this topology is a topological space that needs to perform other operations, such as

relying on a certain measure to define the proximity relationship, define the fuzzy relationship, the intersection of the subset of the set obtained from these images, rather than the strict point set topology meaning of $\partial A, A^\circ, \partial B, B^\circ$ on R^2 , that is:

$$I(A) = f^{-1}[A^\circ_{\tau_1}], I(A) \subset \tau; B(A) = f^{-1}[\partial A_{\tau_1}], B(A) \subset \tau; \quad (7)$$

$$I(B) = f^{-1}[B^\circ_{\tau_1}], I(B) \subset \tau; B(B) = f^{-1}[\partial B_{\tau_1}], B(B) \subset \tau; \quad (8)$$

At this time, after establishing the point set mapping, the elements in the intersection matrix formula (1), such as $I(A), B(A), E(A), I(B), B(B), E(B)$, are all planar point sets on R^2 , and they satisfy the distinction rules for the interior, exterior, and boundary of objects preset by Egenhofer in its intersection model on the plane. When a line target A is in the two-dimensional vector space IR^2 that is, the codimension is 1, the boundary of the line target A is itself, and the interior is an empty set, which can be better explained by the concept of neighborhood. In IR^2 , the neighborhood of any point P_i on the line target can be defined as the circular domain c with P_i as the center and an infinitesimal positive number ε as the radius. Obviously, it is impossible for $C\varepsilon \in A$ to hold, that is to say, there is no neighborhood on the line target, that is, the interior of the line target is an empty set, and the boundary is the line target itself. In this case, the description and distinction of the topological relationship between line targets by the 4-intersection model and the nine-intersection model will not make much sense, but the 33 situations distinguished by the nine-intersection model are indeed different topological relationships. For the same two spatial targets, the possible topological relationship situations in the low-dimensional embedding space still hold in the high-dimensional space, and the types of topological situations in the high-dimensional embedding space may increase or remain unchanged, but will not decrease.

There may be two problems in the existing formalized description of topological relationships: on the one hand, it is the topological distinction of a single spatial target, and on the other hand, it may be the model used for the formalized description of topological relationships. Therefore, how to identify and distinguish the topological differences of the constituent elements of the line target in IR^2 is very important, which is also the basis for establishing the topological relationship distinction model. The norm of the metric is the premise of using some topological invariants, such as the connectivity and Euler number on the topology τ induced by the metric d ^[14]. Compared with the simple straight and curved line targets in the IR^2 space, for the line targets with more complex topological structures such as self-intersecting lines and bifurcated lines, these topological invariants can be used to distinguish points with different topological characteristics on the line target.

Taking connectivity as an example, the establishment of connectivity depends on the topology τ induced by the metric d . After the aforementioned norm metric d is established, we can start discussing the topological relationship between complex geometric objects on τ_d . If we discuss in the two-dimensional vector space IR^2 , then the connectivity of any point P_i on the line set l can be determined as follows, let $Card(A)$ represent the number of elements in set A , then we have:

$$Dn(P_i) = Card(\partial U(P, \delta) \cap l) \quad (9)$$

Here, D represents the connectivity, and U is on the topology τ induced by the metric d . In IR^2 , the method for determining the connectivity of any point on the line target is similar to that in IR^1 . However, in IR^3 , the difference is to calculate the number of intersection points between the boundary of the base open neighborhood U of the point and the line target. By analogy, in IR^n , the method for determining the connectivity of any point on the line target is to calculate the number of intersection points between the neighborhood on τ induced by a certain metric d_n corresponding to n dimensions and the line target^[15]. In this case, the expression of the topological relationship between the line target and other geometric objects will be more distinguishable.

Under the expression of the same relationship matrix in the original nine-intersection model, by improving the expression of the metric topology, introducing topological invariants such as connectivity dependent on the metric can reflect a topological structure of the line target itself. Referring to the expression of the two-dimensional situation, if you want to use the nine-intersection model to represent the topological relationship between three-dimensional spatial entities, for the geometric body in three-dimensional space, its boundary often appears in the form of a surface. In mathematics (topology), a surface is a two-dimensional manifold, that is, it is a spatial object that is locally similar to a plane^[16]. Examples in three-dimensional space include the boundaries of three-dimensional solid objects, such as spheres, cylindrical surfaces, conical surfaces, etc.

However, not all surfaces can be used as the boundary of a geometric body in three-dimensional space. For example, the Möbius strip and the Klein bottle are some non-orientable surfaces, that is, they do not

have two different surfaces or directions. Such surfaces cannot be used as the boundary of a geometric body in three-dimensional space because they do not distinguish between inside and outside. But in actual geographical space, self-intersection or self-cutting situations are rarely considered.

If the surface in three-dimensional space is a single-leaf surface, that is, it does not have self-intersection or self-cutting parts, then it can be represented by a parametric equation, that is: $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, where u, v are two parameters in plane space. In this way, any point (x, y, z) on the surface can be established with a one-to-one continuous mapping with a point (u, v) in plane space. Another method is that if the surface in three-dimensional space is an orientable surface, that is, it has a continuous and non-zero normal vector field, then it can be represented by a scalar function, that is, $z = f(x, y)$ where x and y are two parameters in plane space^[17]. In this way, any point (x, y, z) on the surface can be established with a one-to-one continuous mapping with a point (x, y) in plane space.

3. Conclusions

In those cases, the expression of the topological relationship between the line target and other geometric objects will be more distinguishable. Under the expression of the same relationship matrix in the original nine-intersection model, by improving the expression of the metric topology, introducing topological invariants such as connectivity dependent on the metric can reflect a topological structure of the line target itself.

In the *9-IM*, *DE-9IM* intersection model, the matrix elements of the aforementioned intersection matrix are not as subsets of X in the metric space (X, \mathbf{d}) , but a discrete logical function that judges whether the intersection is empty and how many dimensions the intersection has. Because the composition of the intersection itself contains information about the topological relationship formed by two geometric objects, for the classification of these components, such as considering the subdivision of feature nodes, additional information is introduced, such as feature nodes, which are also obtained on the boundary of geometric objects as additional information, and the establishment of discrete mapping, all of which will cause the loss of component information in the subset. In addition to the true or false values T or F that judge whether the intersection is empty, the matrix elements of 9IM can also be discrete mapping function values of dimensions -1, 0, 1, 2, etc. When its matrix elements are established according to the actual needs and the results obtained by some operators on the intersection space, the description ability of the entire intersection model for spatial topological relationships will be greatly improved.

The description of the subset has been quite detailed, and the object that actually has a topological relationship, the situation of the intersection has been detailed enough in the case of completing the subset components, so compared to other models, one is mathematically normative and unified with point set topology, there is no ambiguity in the metric, and the second is that it has certain advantages in terms of distinction.

References

- [1] Egenhofer M, Franzosa R .*Point Set Topological Spatial Relations*[J]. 1991.DOI: 10.1080/2693799108927841dx.doi.org.
- [2] David M. Mark & Max J. Egenhofer (1994) *Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing*, *Cartography and Geographic Information Systems*, 21:4, 195-212, DOI: 10.1559/152304094782540637.
- [3] Deng Min, Feng Xue-zhi, Liu Wen-bao. *Fundamental Problems and Research Progress in Formal Description of Topological Relations*[J]. *Computer Engineering and Applications*, 2004, 40(1):5. DOI:10.3321/j.issn:1002-8331.2004.01.003.
- [4] Deng M ,Cheng T ,Chen X ,et al.*Multi-level Topological Relations Between Spatial Regions Based Upon Topological Invariants*[J].*Geoinformatica*, 2007, 11(2):239-267.DOI:10.1007/s10707-006-0004-x.
- [5] Zheng Heng, Zhou Xiaoguang. *Conflict Detection of OSM Data Based on E-WID Line/Line Subdivision Topological Relationship*[J]. *Geoinformation World*, 2020, 27(1):6. DOI:CNKI:SUN:CHRK.0.2020-01-004.
- [6] Cheng Jishu. *Point-set Topology* [M]. Science Press, 2008.
- [7] Deng Min, Feng Xuezhao, Chen Xiaoyong. *Hierarchical Model for Formalized Description of Topological Relationships between Face Targets* [J]. *Journal of Surveying and Mapping*, 2005, 34(2):6. DOI:10.3321/j.issn:1001-1595.2005.02.010.

- [8] Deng Min, Li Zhilin, Li Yongli, etc. 4-Cross Model for Describing Topological Relationships between GIS Line Targets [J]. *Journal of Wuhan University (Information Science Edition)*, 2006. DOI:CNKI: SUN: WHCH.0.2006-11-001.
- [9] Li, Jian. *Research on the Topological Relationship Model Between Multiple Spatial Regions* [D]. Jilin University, 2013.
- [10] Chen Jun , Li Chengming , Li Zhilin & Gold CM (2000) Improving 9-intersection model by replacing the complement with voronoi region, *Geo-spatial Information Science*, 3:1, 1-10, DOI: 10.1007/BF02826800.
- [11] Liu W , Chen J , Yan C ,et al. Model for calculating topology stable area of the plane discrete points[J].*Acta Geodaetica et Cartographica Sinica*[2023-12-08].DOI:10.1007/s11783-011-0280-z.
- [12] Zhou Xiaoguang, Chen Fei, Chen Jun. Line/Area Topology Relationship Subdivision Method and Application Introducing Node Degree[J]. *Journal of Surveying and Mapping*, 2015, 44(04):445-452. , 2020, 37(10):57-63.
- [13] Ivo, Düntsch, Gunther,et al.A Necessary Relation Algebra for Mereotopology[J].*Studia Logica*, 2001.DOI:10.1023/A:1013892110192.
- [14] Clementini, E., Felice, P.D., & Oosterom, P.V. (1993). *A Small Set of Formal Topological Relationships Suitable for End-User Interaction*. *International Multi-Conference on Systems, Signals & Devices*.
- [15] Peipei Li. *Research on fitting, parameterization and shape optimization problems in curve modeling* [D]. Shandong University, 2012. DOI: CNKI: CDMD: 1.1012.461424.
- [16] Jingwei Shen. *Research on Description, Calculation, and Inference of Three-Dimensional Topological Relationships* [D]. Nanjing Normal University, 2011. DOI: 10.7666/d.y1924372.
- [17] Min Deng. *Extended model of topological relationship in vector GIS: theory and method* [J]. *Journal of Surveying and Mapping*, 2004, 33(002):188-188. DOI:10.3321/j.issn:1001-1595.2004.02.019.