EFFICIENT TOPOLOGICAL DATA MODELS FOR SPATIAL QUERIES IN 3D GIS

Seokho Lee [a,] *, Jiyeong Lee [a]

[a] Dept. of Geoinformatics, University of Seoul, Siripdae-gil 13, Dongdaemun-gu, Seoul 130-743, Korea – (shlee627, jlee)@uos.ac.kr

Commission IV, WG IV/8

KEY WORDS: GIS, Three-dimensional, Performance, Topological Data Model, Spatial Query, Database

ABSTRACT:

Currently, demands on 3D GIS increase, so spatial data analysis in 3D space also is required. Especially as large-scale and complex indoor space has been developed, it is more important to analyze human behaviors in complicated 3D indoor space to find accident spots or calculate evacuate routes. In order to do that, it is essential to represent topological relationships among the 3D entities in spatial data modeling. So far, topological relationships have been expressed by feature-based data model based on B-rep. However, this data model has some limitations in maintaining topological relationships: complex geometric computations leading inefficiency in maintaining topological consistency, unclear connectivity relationship, and big data volume. Then network-based topological data model based on graph theory was raised to overcome these limitations. In this data model, because topological relationships in complex 3D space is described as a simple network structure using nodes and edges, computational complexity to define adjacency and connectivity relationships is expected to be reduced. For this reason, the network-based topological data model has been believed in more efficient than feature-based data models for 3D spatial analysis with no practical tests yet. In this paper, to verify this general assumption, we perform comparative analysis about efficiencies on spatial queries between two data models through a practical implementation.

1. INTRODUCTION

Many scholars have developed spatial information technologies since 1980. As a result, the analysis tools in the two-dimensional space have considerably improved. But because the analysis methods in the three-dimensional space involve more complicated operations as compared with the analysis methods in the two-dimensional space, more researches and developments are necessary for the three-dimensional space. As large-scale and complex indoor living spaces increase gradually, there are increasing demands on analysis techniques in the three-dimensional indoor space to provide better location-based services to people in normal or emergency situations. The analysis technologies in the three-dimensional indoor space have become important especially since the September 11 attacks in 2001 and the London bombings in 2005.

Topological relationships among spatial objects have to be figured out effectively to search the spot where the accident took place, analyze and understand the emergency situation. That means it is important to decide how to define topological relationships among spatial objects to do efficient spatial queries

Feature-based data model using topological primitives based on boundary representation (B-rep) has been widely used to express topological relationships among spatial objects. But, feature-based data model has the weak points as follows. First, it is inefficient in maintaining consistency of topological relationships because geometric and topological elements of spatial object are expressed together and it leads to complex geometric computations. Second, it has a problem with finding routes because connectivity relationships are not represented clearly. Accordingly, it is difficult to find the fastest route depending on circumstance. Third, it needs big storage space by reason of the complicated data storage structure (Lee 2004). Therefore, network-based topology data model based on a graph theory has been raised to make improvements. This model represents topological relationships among real world's spatial objects using nodes (spatial objects) and edges (relationships), the topological relationships of the complicated real world space are described as very simple network structures.(Lee 2005) So this model became well known as very efficient data model to find out topological relationships such as adjacency, connectivity. But actually experimental evidences to support these common assumptions are not yet defined.

So in this paper, we perform comparative efficiency analyses on time complexity of spatial queries and data volume between two data models through practical implementations. The first criterion is the volume of data storage. After spatial data sets are constructed based on two data models, two volumes of data

* Corresponding author.

storage are compared each other. The second criterion is the time complexity of query processing. For searching spatial objects and their neighborhoods, we compare the times required for query processing using different two data models. We conduct the comparative efficiency analyses using not only 2D data but also 3D data.

This paper is organized as following. In the next section, we review the research related works. In following section, we introduce comparative analysis methods between two data models. In section 4, we do an experimental implementation on efficiency verification. Finally, we describe the further research topics in conclusion.

2. RELATED WORKS

In the last 20 years, various topological data models have been developed (Lee 2005). 3D Formal Data Structure (3D FDS) (Molennar 1990), TEtrahedral Network (TEN) (Pilouk, 1996), Simplified Spatial Model (SSM) (Zlatanova 2000) are the representative feature-based data modes based on B-rep. 3D FDS considers integration of geometric and thematic properties (Zlatanova 2002). TEN has four primitive (tetrahedron, triangle, arc, node primitive), and useful for representing spatial object that has indiscernible boundaries. SSM is a developed data model from 3D FDS. It is useful for the visualization. These data models have been proposed according to each purpose. And Rikkers (1994) stored 3D FDS' conceptual data in relational databases, and tested the some topologic quires. This is the first study about how topologic queries can be conducted in relational database. Lee (2005) proposed combinatorial data model (CDM) that is a logical data model and network structure. CDM is very useful to find out the topologic relationships (adjacency, connectivity) among the spatial objects.

There are many topological data models, and many studies have been carried out. But there is a lack of study on efficiency of spatial queries. The efficiency of spatial queries should depend on topologic data models. So, the purpose of our study is to analyze efficiencies of spatial queries according to two topologic data models; one is a B-rep based data model, the other is a network-based topologic data model.

3. METHODS

This research uses two data models to identify the more efficient data model for spatial queries. One of them is a data model to comply with ISO 19107 spatial schemas (Herring 2001). ISO 19107 is the international standard schema that defines how to express objects geometrically and topologically. And it expresses objects based on B-rep using boundaries and co-boundaries. The other is a network-based topologic data model. It describes spatial relationships among objects using network structures that consist of nodes and edges. So, topological relationships can be expressed more simple and efficient. Because the data model based on a network expresses only topological relationships, this research uses Oracle Geometry type to represent geometries of 3D objects

3.1 ISO 19107 Spatial Schema

ISO19107 Spatial Schema uses topology primitives (Node, Edge, Face, and Solid) to define spatial objects. Solid Primitive is expressed in combination of Face boundaries. Face Primitive is expressed in combination of Edges. Coordinates that are used to represent geometry location are defined in Node Primitive. Likewise, ISO19107 Spatial Schema expresses spatial objects as B-rep structure.

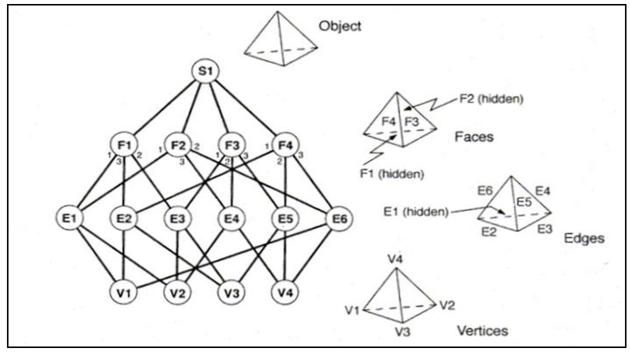

Figure 1. 3D B-rep Structure

Fig.1 shows how to define the 3D solid objects according to ISO 19107 data model based upon B-rep. This method includes geometric and topologic representations, which so far is used in many applications widely.

In this paper, data based on ISO 19107 data model are stored in Relational Database. Fig.2 shows the structure of relational database of 3D B-rep. Each Node, Edge, Face, Solid tables have 'many-to-many' relationships. Solid table refers to Room Type table to link solid object's semantic information. And Face table refers to Door table that has information about whether each face has a door or not. It is used to find connectivity relationships among Solid Objects (rooms).

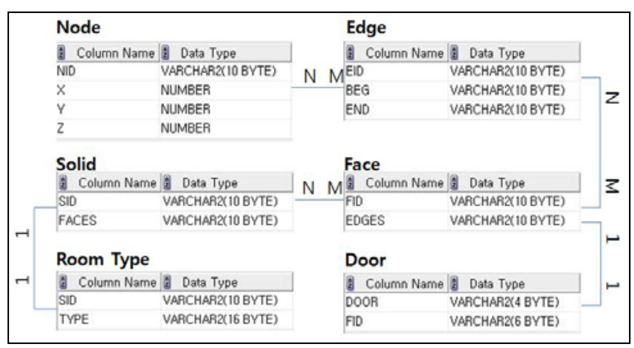

Figure 2. 3D B-rep Structure of the Relational Database

3.2 Network-based Topological Data Model

In network-based topologic data model, spatial objects and their topological relationships are represented as nodes and edges. So it is clear and intuitive to grasp topological relationships. Therefore, this data model has been known as more efficient than B-rep based data model.

Fig.3 shows an example of network structure that expresses topological relationships among spatial objects. 3D solids such as rooms (Faces in 2D) are represented as nodes, and their relationships (adjacency in Fig.3) are represented as edges. Network-based topologic data model represents only topological relationships except geometric representations, which requires dual data structures. Therefore in this model, geometry representation is complemented by Oracle Spatial 11g's SDO_Geometry data types (In contrast, ISO 19107 data mode has topologic and geometric representation).

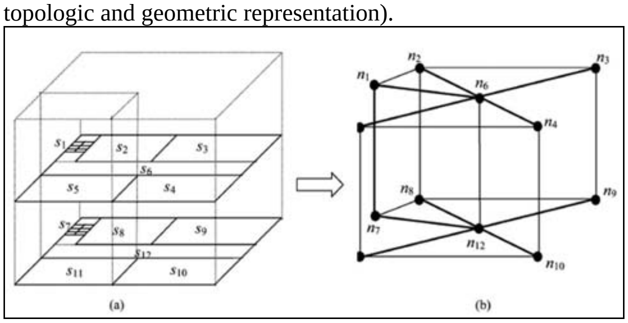

Figure 3. Network Structure (Lee 2005)

Oracle's Geometry data type is based on Simple Features Geometry (OGC) and Solid primitive is added to store 3D spatial objects.

Attribute	Type
SDO_GTYPE	NUMBER
SDO_SRID	NUMBER
SDO_POINT	SDO_POINT_TYPE
SDO_ELEM_INFO	SDO_ELEM_INFO_ARRAY
SDO_ORDINATES	SDO_ORDINATE_ARRAY

Table 1. SDO_GEOMTRY Data Type in Oracle (Kothuri, 2009)

SDO_Geometry data type has five attributes (Table 1). SDO_GYTPE specifies the form (e.g. point, line, polygon. etc) of spatial object, SDO_SRID specifies coordinate system information, SDO_POINT specifies coordinates (if geometry is a point), and SDO_ELEM_INFO specifies where in the SDO_ORDINATES array a new element starts and kinds of geometric primitives (Kothuri, 2009).

Fig.4 shows the structure of relational database of network structure. Network is composed of nodes (representing rooms). Room Type table has semantic information, and Geometry table stores the geometry information

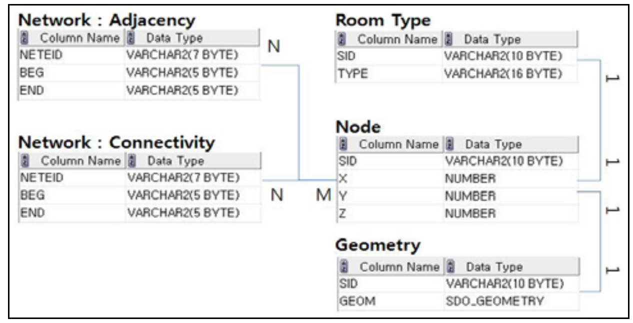

Figure 4. Network-based Data Structure of the Relational Database

3.3 Efficiency Verification

To test efficiencies of spatial queries according to two data models, criteria are shown as bellow.

1) Data Volume
2) Time Complexity for Adjacency Queries
3) Time Complexity for Connectivity Queries

The first criterion is data storage volume. Because ISO19107 data mode is based on B-rep, geometric and topologic representations are stored together. On the other hand, the network-based topologic data model represents only topologic representations, so geometric representation should be stored separately (in this paper, used Oracle Spatial geometry type). Because the methods of representing/storing geometry and topology are different, it is necessary to compare data volumes. The next is time complexity of adjacency/connectivity query. The ways to define the adjacency and connectivity relationships are different depending on the models. So, time complexity will be different. And, this verification is conducted using both 2D data and 3D data.

4. IMPLEMENTATIONS

In order to exam the efficiency of spatial queries, an implementation was conducted with real data. All data (2D and 3D data) were stored at database, and all experiments were performed on database management system. Efficiency verifications can be performed on not only database but also application; but in this study only on database.

The 21 Century building located at University of Seoul, South Korea was used as a study site. Three floors of the building were used (fig.5) and the total number of rooms is 138 (each floors have 46 rooms). In this study, 'rooms' include all types such as 'hall', 'lecture room', 'professor office', 'elevator', 'stairway', 'toilet'.

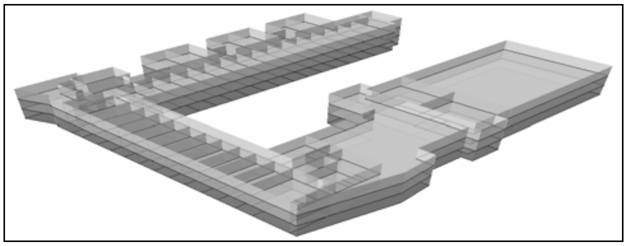

Figure 5. 21 Century Building, University of Seoul

4.1 Data preparations

Fig.6 shows that the data is stored in a database according to ISO 19107 data model

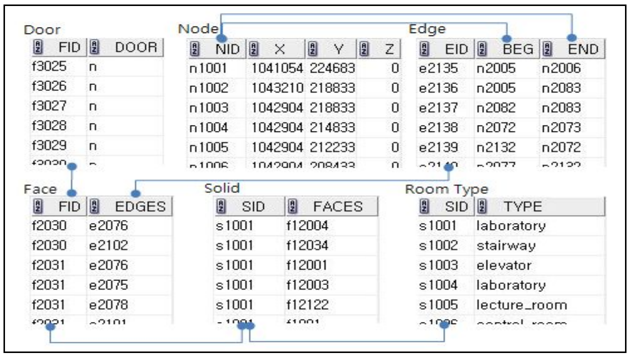

Figure 6. Data of Database according to ISO 19107 data model

Fig.7 shows spatial data of a study area represented by the network-based topologic data model; there are three types of data including geometric data, adjacency network data and connectivity network data.

Fig.8 shows that the real data is stored in a database according to the network-based topologic data model.

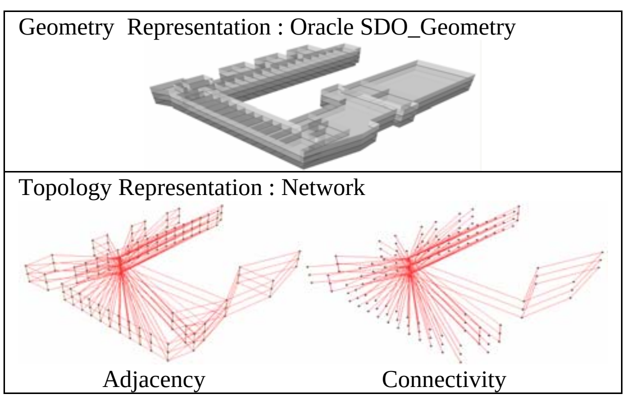

Figure 7. Data of Database according to ISO 19107

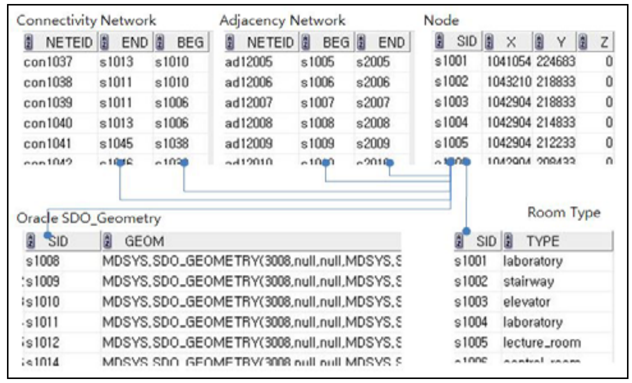

Figure 8. Data of Database according to Network-based topologic model

4.2 SQL Queries

In our work, we implement two types of spatial queries; one is adjacency query, and another is connectivity query. The spatial queries are to find spatial neighbourhood objects in the order relationships shown in Fig.9.

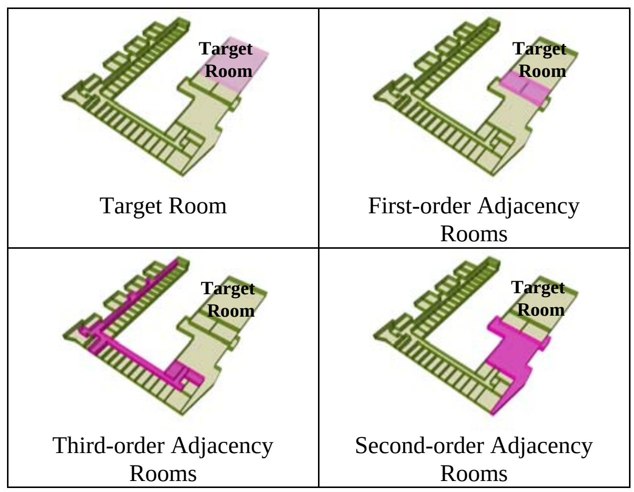

Figure 9. Spatial Query of Finding Adjacency Rooms in the Order

Table 2 shows the SQL query statements that find the adjacent neighbourhood objects represented in 3D B-rep based on ISO19107 data model. The process of the SQL queries is as following.

1) First, select a room (solid) randomly and find the faces that compose the solid.
2) Find the adjacent rooms that share the faces found in the first step: First order adjacent rooms are detected
3) Find the faces that compose the first adjacent faces.
4) Find the second order adjacent rooms that share the faces found in the third step: Second adjacent face detected.
5) Repeat the above process to find the next order adjacent rooms.

```
DECLARE
  vsid VARCHAR2(10);
  vtype VARCHAR2(16);
  ransid VARCHAR2(10) :='s'||TO_CHAR(ROUND(DBMS_RANDOM.VALUE(1001,
1046),0));

  CURSOR c1 IS
  SELECT sid, TYPE FROM f123_3d_room_type
  WHERE sid IN (SELECT DISTINCT SID FROM f123_3d_solid
                      WHERE faces IN (SELECT faces FROM f123_3d_solid
                                           WHERE sid in (ransid)) and
            sid <> all (ransid));

  CURSOR c2 IS
  SELECT sid, TYPE FROM f123_3d_room_type
  WHERE sid IN (SELECT DISTINCT SID FROM f123_3d_solid
                      WHERE faces IN (SELECT faces FROM f123_3d_solid
                                           WHERE sid in (select sid from order_1)) and
            sid <> all (select sid from order_1 ));

<omitted>

BEGIN

  FOR x IN 1..6 LOOP

  IF (x=1) THEN

    OPEN c1;
    LOOP
      FETCH c1 INTO vsid, vtype;
      EXIT WHEN c1%notfound;
      insert into order_1 (sid, "ORDER") values (vsid, x);
    END LOOP;
    CLOSE c1;
    insert into order_0 (sid, "ORDER") values (ransid, 0);

  ELSIF (x=2) THEN
end loop;

<omitted>
```

Table 2. SQL for Adjacency query based on ISO 19107 data model in 3D

Table 3 shows the SQL query statements that find the adjacent neighbourhood objects represented in 3D network on network-based data model. The process of the SQL query is as following

1) First, select a room (solid) randomly and find the adjacent rooms using the adjacency network table: First order adjacent rooms are detected.
2) Find the second adjacent rooms using the adjacency network table: Second order adjacent room detected
3) Repeat the above process to find the next order connected rooms.

```
DECLARE
  vsid VARCHAR2(10);
  vtype VARCHAR2(16);
  ransid   VARCHAR2(10)   :='s'||TO_CHAR(ROUND(DBMS_RANDOM.VALUE(1001,
1046),0));

  CURSOR c1 IS
  SELECT SID, TYPE FROM F123_3D_ROOM_TYPE WHERE (SID) IN
  (SELECT BEG FROM F123_ADJACENCY_INFO2 WHERE END = ransid);

  CURSOR c2 IS
  SELECT SID, TYPE FROM F123_3D_ROOM_TYPE WHERE (SID) IN
  (SELECT BEG FROM F123_ADJACENCY_INFO2 WHERE END in (select sid from
order_1));

                                   <omitted>

BEGIN

  FOR x IN 1..6 LOOP

  IF (x=1) THEN

    OPEN c1;
    LOOP
      FETCH c1 INTO vsid, vtype;
      EXIT WHEN c1%notfound;
      insert into order_1 (sid, "ORDER", type) values (vsid, x, vtype);
    END LOOP;
    CLOSE c1;
    insert into order_0 (sid, "ORDER") values (ransid, 0);

  ELSIF (x=2) THEN
                                   <omitted>
```

Table 3. SQL for Adjacency query based on network-based topologic data model in 3D

Table 4 shows the SQL query statements that find the connected neighbourhood objects that have connectivity relationships represented in 3D B-rep based on ISO19107 data model. The process of the SQL query is as follows.

1) First, select a target room (solid), and find the faces that compose the room and include a door.
2) Find the room that share the faces: First order connective rooms are detected.
3) Find the faces that include doors and compose the solids found in the second step.
4) Find the solids that share the faces found in the third step: Second order connective face detected.
5) Repeat the above process to find the next order connected objects.

```
  DECLARE

    vsid VARCHAR2(10);
    ransid  VARCHAR2(10)  :='s'||TO_CHAR(ROUND(DBMS_RANDOM.VALUE(1001,
1046),0));

    CURSOR c1 IS
    select SID from f123_3D_SOLID
    where FACES IN
    (select   fid   from   f123_door,   F123_3D_SOLID   where   f123_door.fid   =
F123_3D_SOLID.faces and sid = 's1001' and door = 'y')
    and  sid <> 's1001';

    CURSOR c2 IS
    select DISTINCT SID from f123_3D_SOLID
    where FACES IN ((select fid from f123_door, F123_3D_SOLID where f123_door.fid =
F123_3D_SOLID.faces and
    sid in (select sid from order_con_1) and door = 'y'))
    and  sid <> all(select sid from order_con_1);

                                   <omitted>

  BEGIN

    FOR x IN 1..9 LOOP

    IF (x=1) THEN

      OPEN c1;
      LOOP
        FETCH c1 INTO vsid;
        EXIT WHEN c1%notfound;
        insert into order_con_1 (sid, "ORDER") values (vsid, x);
      END LOOP;
      CLOSE c1;
      insert into order_con_0 (sid, "ORDER") values ('s1001', 0);


    ELSIF (x=2) THEN

                                   <omitted>
   /
```

Table 4. SQL for Connectivity query Based on ISO 19107 data model in 3D

In 3D space, the SQL query statements that find the objects which have connectivity relationship in network-based topologic data model are similar to the 'adjacent query of Table 3'. Merely, the adjacent table is replaced with the connectivity table in the SQL query statements.

.

4.3 Results

In terms of data volume, first of all, 2D and 3D data volume of dual data model is bigger than B-rep data volume. A data volume of 2D B-rep is 196Kbytes that includes topology and geometry representations together. On the other hand, a data volume of 2D dual data model is 328Kbytes that includes topology (196Kbytes) and geometry (131 Kbytes) representations separately. The 3D data volume of dual data model's topology part is equal to 2D data volume of dual data model's topology part. But there is quite a big difference between 2D and 3D data volume of geometry part (Fig.10).

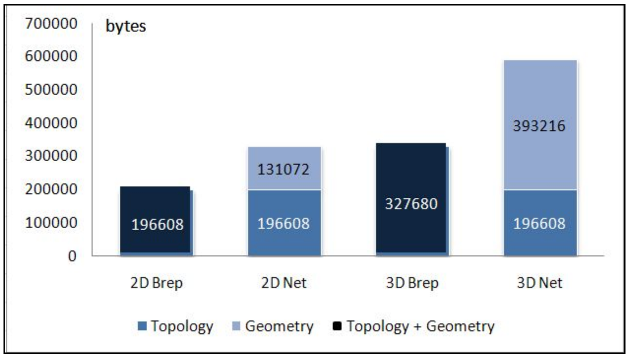

Figure 10. Data of Database according to ISO 19107

Fig.11 is the result of the adjacency query in 2D space using two data models. The x-axis of the graph denotes the counts of the query. One count means to query from a target room to adjacent rooms by first though fourth orders. The y-axis of the graph denotes the time taken. The query times rarely seem to have differences between the two data models until 300 counts. When 600 times adjacency queries are run, it takes 24 seconds using B-rep data while taking 15 seconds using network-based data. It means that the data structured as network is much more efficient than the data structure as B-rep for adjacency queries.

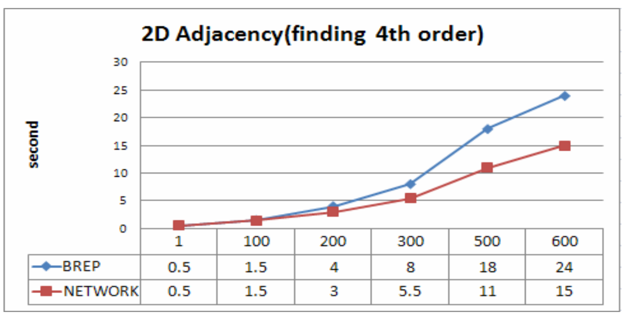

Figure 11. Data of Database according to ISO 19107

Fig.12 is the result of the connectivity query in 2D space using the two data models. As the same as the result of the adjacency query, the data structured as network is much more efficient than the data structure as B-rep.

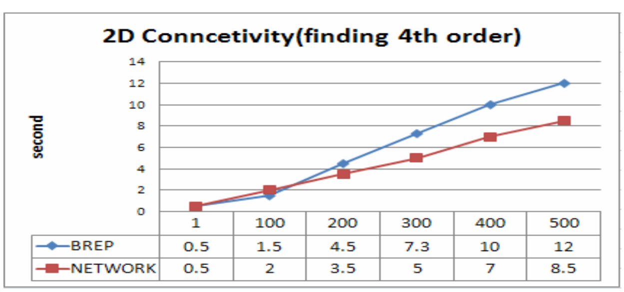

Figure 12. Data of Database according to ISO 19107

Fig.13 is the result of the adjacency query in 3D space using two data models. The query times rarely seem to have differences between the two data models until 20 counts. When 60 times adjacency queries are run, the amount of time using B-rep data takes about 40 times more than the amount of time using network data. The result of the connectivity query is similar to the result of the adjacency query.

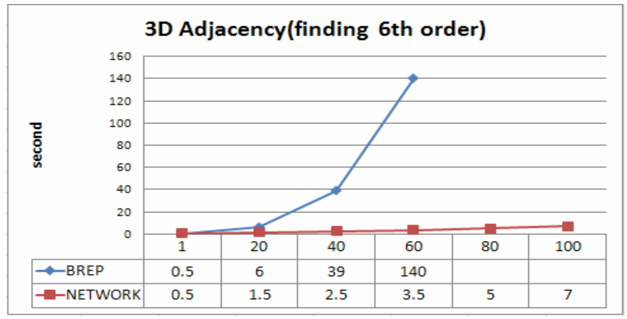

Figure 13. Data of Database according to ISO 19107

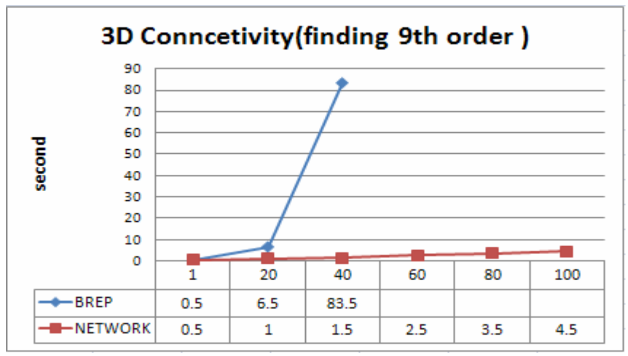

Figure 14. Data of Database according to ISO 19107

5. CONCLUSIONS

This study conducted the verification tests of efficient spatial queries based on two data models; ISO 19107 data model based on B-rep, and network-based topologic data model with Oracle Spatial Geometry. As results, 1) Data volume of ISO 19107 data model's data is smaller than the volume of network-based topologic data model, and 2)Time complexity of spatial queries (adjacency, connectivity) based on network-based topologic data model is better than ISO 19107 data model's time complexity.

The query processing on network-based topologic data model's data is relatively simple, because objects that have topological relationships with other objects can be found directly using linked information from edge data. But query processing on B-rep based data model's data is relatively complex. To find objects that have topological relationships with a target object, first, boundaries of the target object should be defined. Then those all boundaries should be checked one by one to find neighbourhood objects that have topological relationships. In 3D space, because one object is composed of more boundaries than in 2D space, the time complexity become worse.

In conclusion, total data volume of network-based topologic model would be bigger than B-rep data, but it has good performances on spatial queries. Especially in 3D space, time complexity of spatial queries is much better. In the future, because many large-scale and complex indoor spaces should be handled, further researches that improve spatial query more quickly and accurately are required continuously.

REFERENCES

J. R. Herring, The OpenGIS Abstract Specification, Topic1: Featuare Geometry (ISO 19107 Spatial Schema), OpenGeospatial Consortium, Inc., 2001, version 5, OGC Doc. No. 01-101.

Lee, J. (2004). "A Spatial Access Oriented Implementation of a Topological Data Model for 3D Urban Entities." GeoInformatica 8(3): 235-262.

Lee, J. and M-P. Kwan. (2005). "A Combinatorial Data Model for Representing Topological Relationships between 3-D Geographic Entities." International Journal of Geographical Information Sciences, 19(10): 1039-1056.

Molenaar, M.,1990, "A formal data structure for 3D vector maps", In: Proceedings of EGIS'90

Pilouk, M., 1996, "Integrated modelling for 3D GIS", Ph.D. Dissertation, ITC.

Ravi Kothuri, Albert Godfrind, and Euro Beinat, 2009, Pro Oracle Spatial for Oracle Database 11g.pdf, pp.60-61

Zlatanova, S., 2000, "On 3D topological relationships ", In: proceedings of the 11th International Workshop on Database ad Expert System Applications, Greenwich, London, UK, pp913-919

ACKNOWLEDGEMENTS

"This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(20100012941)"

"This work was researched by the supporting project to educate GIS experts"