Possibilities of Storage of Spatial Data of Real Estate Registry Information System

Karel JANEČKA and Václav ČADA, Czech Republic

Key words: information system designated for real estate registry, real estate register, model of spatial data, generalization, Oracle Spatial.

SUMMARY

An article describes an information system designated for the Real Estate Registry of the Czech Republic and the most significant periods of its extension and functionality. This information system was created in 1997. Due to the huge amount of data span and remote data access is necessary to optimize organization, management and storage of spatial data of the real estate registry. The need of an integrated architecture in database management systems for spatial data is very actual problem. We want to have the possibility to store non-spatial data with spatial data within one database. The main purpose of this article/ paper is to discuss the actual way to store digital cadastral maps in the information system of the real estate registry. We also propose the facilities of Oracle Spatial extension for management of selected spatial feature of digital cadastral map – of cadastral boundary. Further we present our experience with generalization of spatial data in the Oracle Spatial environment.

SUMMARY

Článek popisuje informační systém katastru nemovitostí České republiky budovaný od roku 1997 a důležité etapy rozšíření jeho funkcionality. Protože se jedná o rozsáhlý informační system jak po stránce objemu dat, tak i jeho využitelnosti dálkovým přístupem, je nezbytné optimalizovat organizaci, správu a uložení prostorových dat katastru nemovitostí. Potřeba integrované architektury v databázových systémech pro uložení geodat, kdy jsou v databázi uložena nejen data popisná (jméno, příjmení, adresa atd. s využitím standardních databázových typů), ale rovněž prostorová, je v současnosti zřejmá. Předložená stať se zabývá diskusí o aktuálním způsobu uložení prostorových dat digitálních katastrálních map a jsou představeny možnosti uložení vybraných prostorových dat katastru nemovitostí pomocí nadstavby databázového systému Oracle pro správu prostorových dat. Toto rozšíření nese označení Spatial. Rovněž je zmíněna je i problematika generalizace prostorových dat uložených s využitím možností nadstavby Spatial.

Possibilities of Storage of Spatial Data of Real Estate Registry Information System

Karel JANEČKA and Václav ČADA, Czech Republic

1. AN INFORMATION SYSTEM USED BY THE REAL ESTATE REGISTRY

An information system (RERIS) used by the real estate registry was created in 1997. Management of this system takes place within the Czech Office for Surveying, Mapping and Cadastre's (COSMC) Authority. Presently it is the largest (number of users, size of data storage) system in the Czech Republic. It supports the management of cadastre of real estates in the Czech Republic and it is a part of the information systems of public administration. RERIS ensures management, actualization and provides data on the local level and on the central level (the central database of RERIS) including the possibility of remote access to data via internet. The real estate registry represents the set of information about real estates in the Czech Republic which contain the list and descriptions including geometric and position determinations. The real estate registry also contains the evidence of ownerships and other rights connected with the estate. Real estate is registered according to the cadastre territories. The cadastre territory is the main unit of administration. The content of the Real estate registry is composed by the file of survey data and by the file of descriptive information. The service and the full functionality of RERIS has been currently classified as one of the conditions for the digitalization of cadastral documentation. The process of cadastral documentation digitalization has started in 1993. It was declare that the digital form of cadastre maps brings mainly these benefits:

- possibility to use modern digital technologies for the work with map at the activities related to the administration of real estate cadastre,
- possibility to get the agreement between the file of survey data and the file of descriptive information,
- facilitation of cadastral data interface among other information systems,
- possibility to get all important data of cadastre through the remote access via internet,
- possibility to use cadastral maps in other land information systems. It is necessary for building of the spatial information infrastructure.

The conception to digitalize cadastral maps was described/presented in the document "Rules of the cadastral office process of the survey data file digitalization in order to create a land oriented information systems" (COSMC reference number 2728/1994-22). It was strictly specified which activities must be done within the resort of COSMC and which activities can have the form of public contracts. The process of digitalization was controlled by the set of strict measures. Solving the problem of cadastral map digitalization was also postponed due to the fact that the main activities of COSMC was oriented on the administration of the real estate cadastre and on the inscription of data about legal relationships in the real estate cadastre. The resolution in the matter of separation of the digitalization process of the file of descriptive information and the file of survey data was doubtful. Due to this fact it was very hard to get the consent between the file of survey data and the file of descriptive information and in addition it was neccessary to do many activities repeatedly. The filling of the file of survey data depends on the process and results of cadastral map digitalization. Results of the digitalization of the file of survey data depend mainly on the quality of the actual analogue maps of real estates cadastre. More then 67% of actual cadastral maps in the Czech Republic have their origin in the field mapping for the stable cadastre (19th century). The procedure describing the digitalization of maps in fathom scale 1:2880 was created. [ČADA 1992]. This procedure was extended by complex technology designed for the creation of continous maps of cadastre of lands located in S-JTSK [ČADA 2001] through the usage of global transformation keys. This technology has been tested by COSMC since 2000. The seamless map of cadastre of lands as the base of digital cadastral maps is based on this proposed technology. The process of the digitalization of the file of survey data in the department of COSMC on the date of 31st of December 2006 and its progress describes table.1.

Situation in the year	Number of cadastral territories				Digitalized cadastral territories in %[1]
	DKM	KM-D	total in the year	total on 31.12.	
do 1997	436	1	437	437	3,4
1998	157	12	169	606	4,7
1999	239	94	333	939	7,2
2000	402	647	1049	1998	15,3
2001	180	260	440	2428	18,6
2002	305	148	453	2881	22,1
2003	475	68	543	3424	26,3
2004	380	3	383	3807	29,2
2005	314	0	314	4121	31,6
2006	279	0	279	4400	33,8
2007	406	0	406	4806	36,9

Table 1. Number of cadastral territories with digitalisation of the file of survey data in the department of The Czech Office for Surveying, Mapping and Cadastre.

Providers of data base management systems (DBMS) already have in their products integrated spatial data types accordance to OGC's (Open Geospatial Consortium) specifications [MEIJERS 2006]. In accordance with these specifications a spatial object can be represented in DBMS by two different logical vector models:

- geometrical data model,
- topological data model.

Both models have their pluses and minuses. Geometrical data model offers direct access to the object coordinates . Topological data model stores information about spatial relationships (e.g. adjacency). Due to the fact that the geometrical data model offers direct access to coordinates, common parts of adjacent polygons are stored for each polygon separataly which

[1] Total number of cadastral territories in the Czech Republic is today 13 027.

means the data redundancy. Topological data model eliminates data duplicities by using the references on unambiguously identified topology primitives.

1.1 Spatial data in RERIS

RERIS was built on the Oracle platform. It contains the file of survey data and the file of descriptive information. A very important part of the file of survey data is a digital cadastral map (DCM). DCM is a continous and seamless map on a scale 1:1000 in the Date of Unified Trigonometric Cadastral Network (DUTCN). [ČUZK 2004] describes the general way in which is DCM stored in RERIS. Based on this document we can say that the way of storage of DCM in RERIS is similar to the topological concept. There are stored references on the parcel on the left side and on the parcel on the right side for each edge. There are also stored references at the beginning point and at the end point for each edge. Basically there is a table for storing of each topological primitive in RERIS. In case we want to work with this data as with objects we get the problem becouse there are not stored references on consequential edges for any edge in RERIS. We can not use the structured query language (SQL) to solve this problem but when using some high-level language it is relatively easy to do so. On the other hand the data maintenance after the data editing is easier. It possible to use the minimum bounding rectangle (MBR) for each edge. MBR is used for example when we select data by the query window. But we must create and save MBR for each edge.

2.   STORAGE OF SPATIAL DATA IN ORACLE SPATIAL

2.1 Oracle Spatial

Oracle Spatial (further Spatial) is an integrated set of functions and procedures allowing storage, access and analysis of spatial data in rapid and affective way of the Oracle database system. Spatial consists of the following components:

- scheme describing the storage, the syntax and the semantics of supported geometrical data types,
- mechanism of spatial indexing,
- set of operators and functions for transaction of spatial queries and spatial analysis and other utilities in order to operate spatial data.

2.2 Object-relational model in Oracle Spatial

Spatial supports the object-relational model for the storage of spatial data. It actually means that each geometrical description of spatial object is stored in a cell within the table. This is possible because of to the usage of SDO_GEOMETRY object type of data. Then there are stored descriptive data together with spatial data in one table. One of the main facilities of object-relational way of modeling of spatial data is easiness of creation of spatial index. Spatial index enables the effective performance of spatial queries. On the other hand the object-relational model causes the data redundancy because an identical part of geometrical descriptions of adjacency spatial objects are stored twice.

2.2.1 SDO_GEOMETRY object data type

SDO_GEOMETRY object data type is defined as follows:

CREATE TYPE sdo_geometry **AS OBJECT** (
sdo_gtype **NUMBER**,
sdo_srid **NUMBER**,
sdo_point SDO_POINT_TYPE,
sdo_elem_info MDSYS.SDO_ELEM_INFO_ARRAY,
sdo_ordinates MDSYS.SDO_ORDINATE_ARRAY
);

SDO_POINT_TYPE type and SDO_ELEM_INFO_ARRAY type and SDO_ORDINATE_ARRAY type are defined as follows:

CREATE TYPE sdo_point_type **AS OBJECT** (
x **NUMBER**,
y **NUMBER**,
z **NUMBER**
);

CREATE TYPE sdo_elem_info_array **AS VARRAY** (1048576) **OF NUMBER**;

CREATE TYPE sdo_ordinate_array **AS VARRAY** (1048576) **OF NUMBER**;

The sdo_gtype attribute indicates a geometrical type of geometrical description of spatial feature. Value 2003 indicates polygon in two dimensions. The sdo_srid attribute indicates the spatial reference system associated with the spatial data. If value of this attribute is NULL then there is no spatial reference system associated with the spatial data. If we want to store point data in SDO_GEOMETRY the sdo_point attribute contains coordinates of this point. In another case it is recommended to set NULL value of this attribute. The sdo_elem_info attribute describes the way of interpretation of the sdo_ordinates attribute which contains coordinates of geometrical descriptions of spatial feature.

2.3 Use of SDO_GEOMETRY object type of data to model the cadastral boundary

The storage of spatial data by means of object type of data allows users/administrator to store the geometrical description of spatial feature in one cell within the table. We tried to store a cadastre boundary in object-relational way. We modeled this boundary as polygon. It means that the first couple of coordinates is the same as the last couple of coordinates of boundary geometrical description. The minus is if two cadastral territories share a part of boundary this way of modeling of spatial date leads to the origin of duplicity data. The definition of the table for storing of the geometrical descriptions of cadastral boundaries (the *boundary* attribute) looks as follows:

```
CREATE TABLE CADASTRE_BOUNDARY (
id                    NUMBER PRIMARY KEY,
cadastre_code         NUMBER(6),
boundary     SDO_GEOMETRY,
def_point    SDO_GEOMETRY
);
```

If we model the cadastre boundary as polygon we have to store the coordinates in counter-clockwise order. The special program for this was created. This program loads spatial data from RERIS and return the right geometrical description of particular cadastral territory. Next it is possible to store directly this description in SDO_GEOMETRY object type of data. Due to the usage of the application MapViewer by Oracle we were able to display data stored in SDO_GEOMETRY (vide Figure 1). Display of spatial data from the CADASTRE_BOUNDARY table run very quickly due to the fact that all geometrical descriptions are stored in one place. The speed of displaying spatial data is one of the most significant plus of this way of model the spatial data.

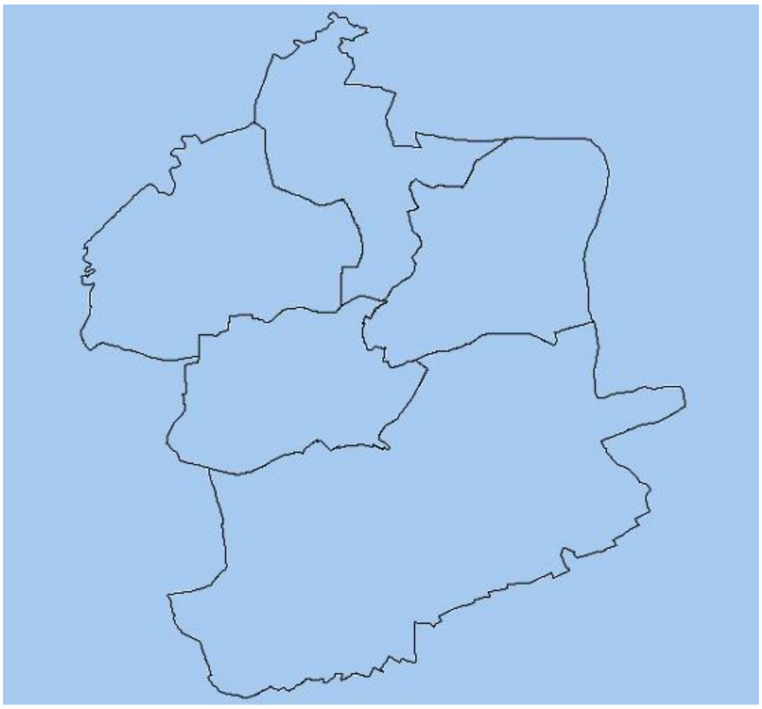

Figure 1. Plan of position map of cadastral territories.

It is presently possible to visualize data stored in SDO_GEOMETRY due to use of PL SQL Developer by Oracle.

2.4 Facilities of object-relational data model in Oracle Spatial

Without any doubt there are pluses in spatial data modeling by means of the SDO_GEOMETRY object type of data. For example it displays data fast and it is easy to build up a spatial index. Specifically there we can use the R-Tree structure to index spatial data in Spatial. Spatial index is important for effective evaluation of spatial query. The most negative part of object-relational concept that was tested are occurrence of data duplicity and difficult editing. If one cadastral boundary is changed then we must edit the boundary geometry of the adjacency cadastral territory. It is possible to reduce data duplicity by means of the topological concept - winged edge structure- we will discuss it later. Editing of spatial data should be then also easier.

3. DATA GENERALIZATION IN ORACLE SPATIAL

We tested available Spatial's functions over spatial data stored as objects. We can use these functions written in PL/SQL programming language for the purpose of generalization. The generalization of spatial data is very actual problem. The main purpose of data generalization is to produce data with less spatial details from the very detailed data.. Very important request regarding the function generalization it is preservation of all spatial relationships. We want to keep adjacency after simplification of origin geometrical descriptions in case of cadastral boundaries. All tested functions of generalization need a geometrical input of SDO_GEOMETRY object type of data. Functions were tested with various values of input parameters. More details about this testing are in [JANEČKA 2005]

3.1 REMOVE_DUPLICATE_VERTICES Function

The main purpose of this function from SDO_UTIL package is to eliminate duplicate vertices from the geometrical description of spatial feature. If there are no duplicate vertices in the geometrical description we can use this function for simplification of geometrical description by the usage of suitably choice of values of input paramenters. All points within distance ε will be removed, vide Figure 2.

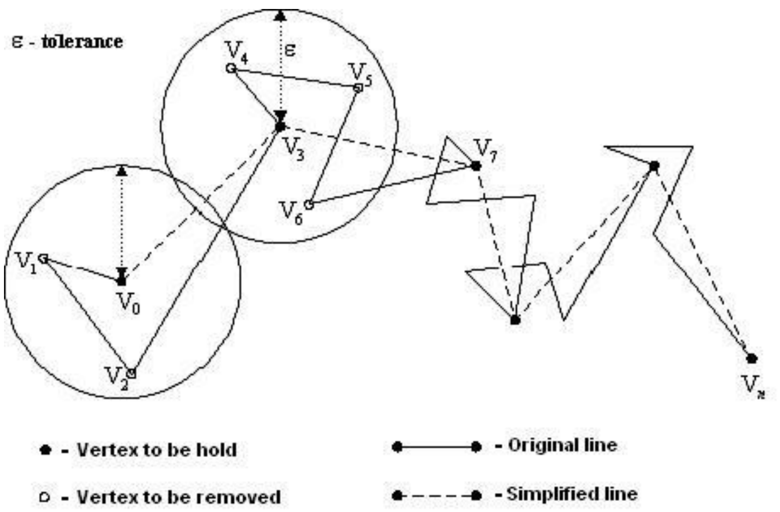

Figure 2. Principle of the REMOVE_DUPLICATE_VERTICES function.

In the Figure 3 we can see the result after applying the REMOVE_DUPLICATE_VERTICES function on geometrical descriptions of cadastral boundaries stored in the CADASTRE_BOUNDARY table. We can see that geometrical descriptions were simplified but the adjacency relationship between adjacency territories was damaged. The joint part of cadastre boundary is not identical for both adjacency territories.

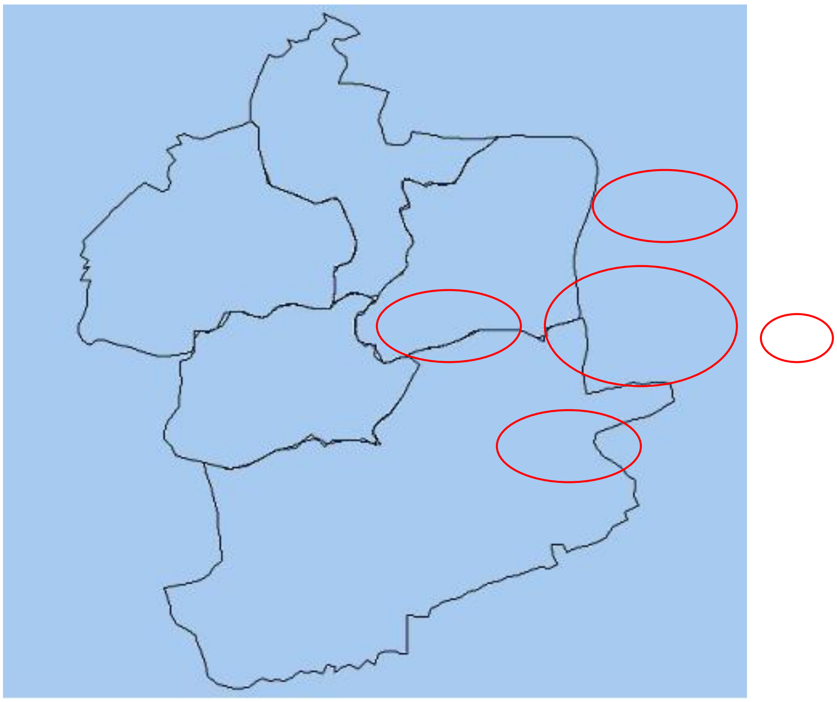

Figure 3. Plan position map after applying of REMOVE_DUPLICATE_VERTICES function.

3.2 SIMPLIFY function

The main purpose of function SIMPLIFY is to simplify the geometrical description of spatial feature by the usage of the Douglas-Peucker's algorithm. Detailed description of this algorithm can be found in [DOUGLAS 1973]. Principle of Douglas-Peucker's algorithm is showed in the Figure 4.

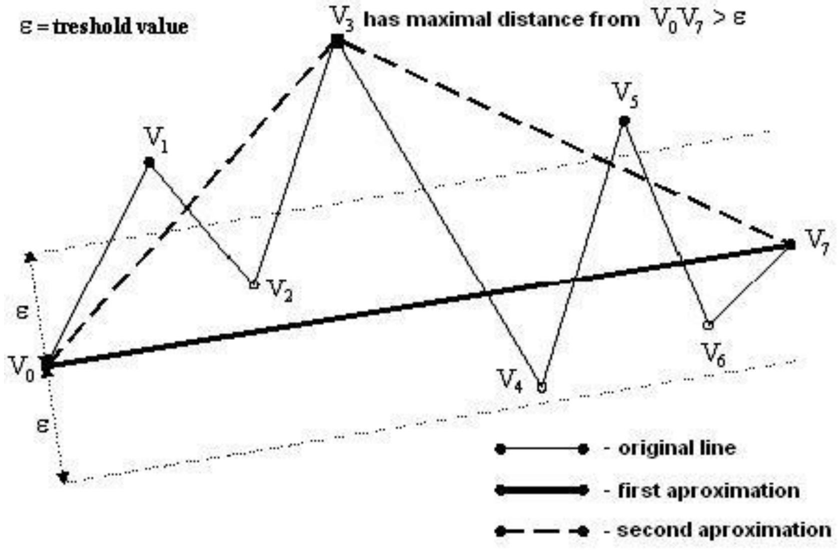

Figure 4. Principle of the SIMPLIFY function.

The function SIMPLIFIY is one of the two functions in Oracle Spatial which main purpose is the generalization of geometrical descriptions. After applying this function a geometrical description of spatial object can be reduced for example from polygon data type to line data type. Very important fact is that after applying this function origin topological relationships are not preserved. This fact is illustrated in the Figure 5.

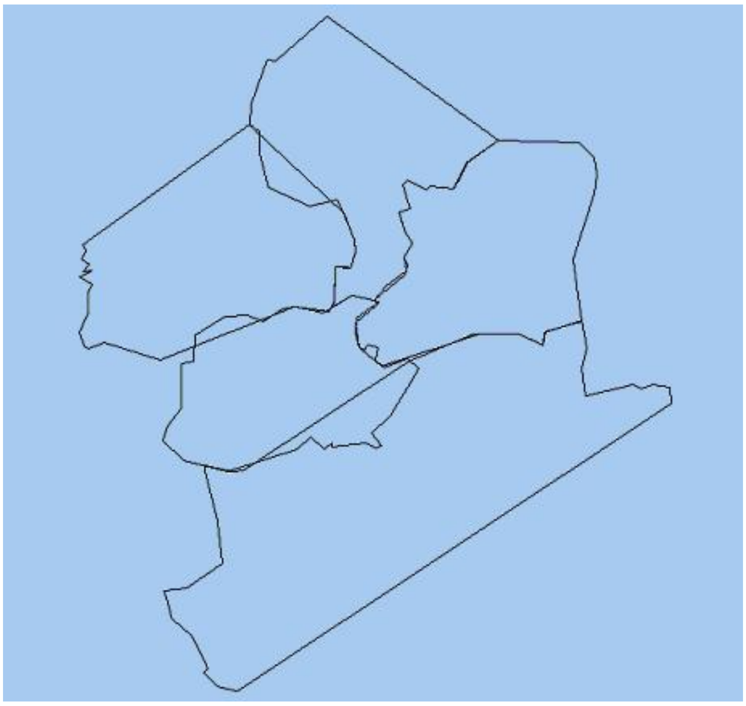

Figure 5. Plan position map after applying function SIMPLIFY.

3.3 SIMPLIFY_GEOMETRY function

This function is a part of the SDO_SAM package. The main purpose of this function is to simplify geometrical description of spatial feature. We compare the area of origin polygon to the area of simplified polygon. This ratio should not be bigger then the user given threshold value in percentage. This function is not well documented in the official Spatial's documentation. Only the name of this function which is SIMPLIFY function in SDO_UTIL package. The result - after applying this function - is presented in Figure 6. It is evident that the adjacency between simplified geometrical descriptions was not preserved. The ratio was set on the value of 5 percent.

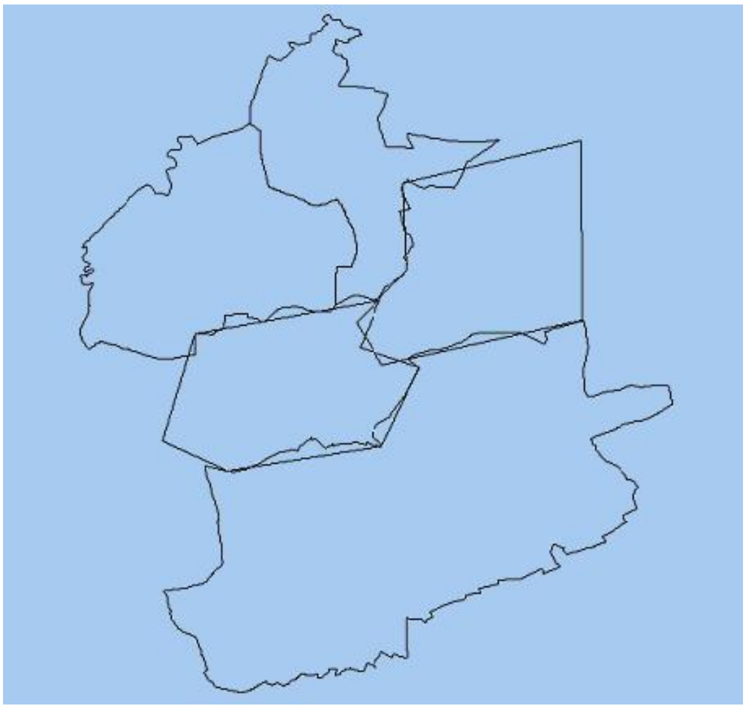

Figure 6. Position plan map after applying of SIMPLIFY_GEOMETRY function.

4. SPATIAL DATA MODELING BY THE USAGE OF A TOPOLOGIC DATA MODEL

To store spatial data by some topological data model for face tables, edges and nodes must be created. It means that we need at least three tables. The structure of tables depends on specific topology of data model. When the description of geometry faces is build up through reference between topological primitives an implementation of function building the face geometry from these primitives can be realised. Some operations over the topological structured data are difficult. Otherwise the topology data model has generally good properties:

- remove duplicity
- easier asurance of data consistency after editing,
- evaluation efficiency of some spatial queries (e.g. find all neighbours).

4.1 Topology of Oracle Spatial data model

The topology of Oracle Spatial data model corresponds with the winged-edge structure. The start node and the end point determine the edge orientation. For each edge we store reference on the face on the left side and on the face on the right side. Next we store references on four edges, vide Figure 7.

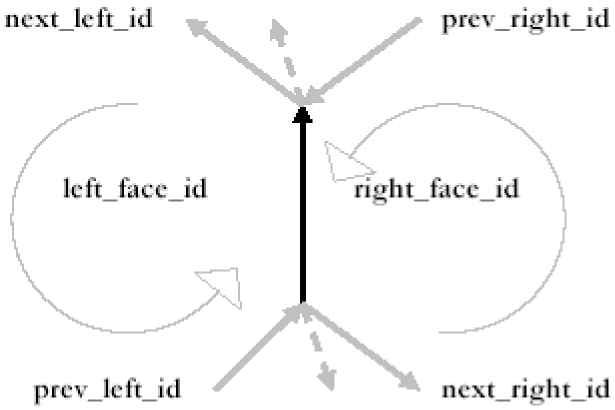

Figure 7. Winged–edge structure in Oracle Spatial [MURRAY 2003].

The model of the spatial data created by winged-edge structure has also some minuses. In [PENNINGA 2005] - there is mentioned that the model of the spatial data in the topological way requires more data space for its storage. The reason is - vast amount of references. The problem is also to get the geometry of particular face. Is is not possible to declare a "navigation" access to data in SQL – follow the next reference until we reach the start node. It is preferable to use some high level language to solve this problem. It is possible to use built–in (standard) function to solve this problem too in Oracle Spatial 10g. The usage of topological structure to represent the cadastral data we can find for example in Dutch cadastre.

4.2 Hierarchical data model in Oracle Spatial

Oracle Spatial 10g offers very interesting facility for data modeling. It is possible to create several hierarchical layers for one data set. The higher layer level consists of the lower layer level within the frame of hierarchical data model. This can be very useful for cadastre because the regional division has apparently a hierarchical character. At the beginnig we have data which represents boundaries of plats.We can get layers from this data with boundaries of cadastral territories and so on without duplicities.

5. CONCLUSION

If we want to keep a topologically structured spatial data in database we must consider carefully pluses and minuses of it.. We must especially consider demands for the data maintenance and the functionality of the application working with these data. Fast evaluation of the most frequent usage of the spatial queries is preferred very often. Possibility of the usage of the Oracle Spatial and the winged-edge structure for spatial data of the Real Estate Registry information system is the topic for our next research. We want to focus mainly on the usage of the hierarchical data model. Selecting data by query window might be the best used spatial query on spatial data of cadastre. The modification of data by the winged–edge structure makes the evaluation of this query slower. We plan to use the mechanism of function based indices in order to reduce this negative effect. Advantage of the usage of topologically structured data and the hierarchical data model can be possible when producing hierarchical layers for the one data set. These layers can be used as a plan of position maps.

REFERENCES

[ČADA 1992] ČADA, V.: *Návrh postupu digitalizace pozemkové mapy 1:2880*. KGKS, Plzeň 1992.

[ČADA 2001] ČADA, V.: *Návod pro obnovu katastrálního operátu přepracováním ze systému stabilního katastru.* ČÚZK, Praha 2001.

[ČUZK 2004] Czech Office for Surveying, Mapping and Cadastre. (ČÚZK). *Struktura výměnného formátu informačního systému katastru nemovitostí České republiky*. Číslo jednací 6665/2004-24. Praha, 2004.

[DOUGLAS 1973] DOUGLAS, D., PEUCKER, T.: *Algorithms for the reduction of the number of points requiered to represent a digitised line or its caricature.* The Canadian Cartographer, 1973.

[JANEČKA 2005] JANEČKA, K.: *Modelování geoprostorové báze dat na úrovni datového modelu KN.* Master's thesis. Plzeň, 2005.

[MEIJERS 2006] MEIJERS, M.: *Implementation and testing of variable scale topological data structures.* Master's thesis. TU Delft, 2006. 114 p.

[MURRAY 2003] Murray, Ch.: *Oracle Spatial User's Guide and Reference, 10g, Release 1*. 2003.
[ONLINE]
http://www.oracle.com/technology/products/spatial/spatial_10g_doc_index.html

[PENNINGA 2005] PENNINGA, F., QUAK, W., TIJSSEN, T., OOSTEROM, P.: *Storage and Querying of Topological Structures in Oracle Spatial.* Topology and Spatial Databases Workshop. Glasgow, 2005.

BIOGRAPHICAL NOTES

Karel Janečka graduated at the University of West Bohemia (WBU) in Pilsen. Since 2005 he is a Ph.D. student of the Faculty of Applied Sciences at WBU. Within the scope of his doctoral thesis he proposes and implements algorithms identifying non-valid parts of spatial data from the Real Estate Registry Information System (RERIS). Further he examines capabilities of Oracle Spatial for the storage and the administration of spatial data from RERIS. He has successfully published own science findings in several international conferences.

Václav Čada graduated at the Czech Technical University in Prague. From 2004 he is the docent. His main research topics are oriented on surveying, computer cartography and geographical information systems. Nowadays he is the head of the Geomatic Section at the Department of Mathematics at the University of West Bohemia. He has published many scientific articles for international symposiums and technical papers. He is a member of the Czech Association of Surveying and Cartography, of the Czech Association for Geoinformation, of the Cartographic Association of the Czech Republic and of the Nemoforum.

CONTACTS

Karel Janečka
University of West Bohemia
Univerzitní 8
Plzeň
CZECH REPUBLIC
Tel. +420377632691
Email: kjanecka@kma.zcu.cz
Web site: http://www.kma.zcu.cz/janecka

Václav Čada
University of West Bohemia
Univerzitní 8
Plzeň
CZECH REPUBLIC
Tel. +420377632678
Email: cada@kma.zcu.cz
Web site: http://www.kma.zcu.cz/cada