ZAJIŠTĚNÍ KONZISTENCE PROSTOROVÝCH DAT V INFORMAČNÍM SYSTÉMU KATASTRU NEMOVITOSTÍ

Karel Janečka[1]

[1]Katedra matematiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni,
Univerzitní 22,
306 14, Plzeň, Česká republika
kjanecka@kma.zcu.cz

Abstrakt. V České republice je primárním zdrojem prostorových dat informační systém katastru nemovitostí, spravovaný Českým úřadem zeměměřickým a katastrálním. V tomto systému jsou uložena nejen data atributová, ale také data prostorová, především v podobě digitální katastrální mapy. Pro uložení prostorových dat katastru nemovitostí je využito nadstavby databázového systému Oracle - Oracle Spatial. Důležitým problémem v oblasti využití vektorových prostorových dat je zajištění konzistence prostorových vztahů, které se mezi geometrickými reprezentacemi jednotlivých prostorových prvků vyskytují. Obecné požadavky na topologicky čistá data jsou všeobecně známy. V článku se zaměřuji na specifické požadavky na prostorová data katastru nemovitostí a především na zajištění konzistence těchto prostorových dat. Míra využití a funkcionalita nadstavby Spatial není taková, aby mohly být složitější postupy vedoucí k nalezení nekonzistencí v prostorových datech opřeny právě o funkčnost samotné nadstavby Spatial. Mnou navržené řešení spočívá v aplikování vhodných algoritmů z oblasti výpočetní geometrie použitých nad prostorovými daty ISKN.

Klíčová slova: informační systém katastru nemovitostí, konzistence prostorových dat, výpočetní geometrie.

Abstract. Ensuring of consistence of spatial data in the information system of cadastre. There is the main source of spatial data in the Czech Republic – the Information System of Cadastre (ISC), maintained by the Czech Office for Survaying, Mapping and Cadastre. There is stored attribute data as well as spatial data in form of the digital cadastral map in the ISC. Oracle Spatial – a part of the Oracle Database Enterprise Edition, is used for the storage of spatial data in the ISC. I would like to mention the fact we have to take into account during the processing and utilization of spatial data from the ISC – the inconsistency of spatial relationships. General requirements on the topological valid data are well known. There I concentrate on specific demands on spatial data of cadastre in the article. Consistency assuring of spatial relationships between geometries of spatial features is very important for an effective usage of data. The functionality and the utilisation rate of Oracle Spatial in the ISC does not allow to find all inconsistencies in spatial data only by using the means of Spatial. I suggested a solution based on the application of convenient algorithms from field of computational geometry.

Keywords: information system of cadastre, consistence of spatial data, computational geometry.

1 Prostorová data v informačním systému katastru nemovitostí

Informační systém katastru nemovitostí (ISKN) vytvořený nad databázovým systémem firmy Oracle obsahuje kromě dat souboru popisných informací (SPI) rovněž prostorová data, především v podobě digitální katastrální mapy (DKM). Digitální katastrální mapa, vedená jako spojitá a bezešvá mapa, se zpracovává pro jednotlivá katastrální území nebo jejich část v S-JTSK ve vztažném měřítku 1:1000. O struktuře uložení DKM ve zdrojové databázi ISKN lze získat nejvíce informací z dokumentu [2], který úzce souvisí s novým výměnným formátem (NVF).

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. Dodavatelé databázových systémů mají ve svých produktech již integrované prostorové datové typy více či méně odpovídající specifikacím OGC (Open Geospatial Consortium [7]) [6]. Podle těchto specifikací může být prostorový objekt reprezentován v systému řízení báze dat dvěma rozdílnými logickými vektorovými modely:

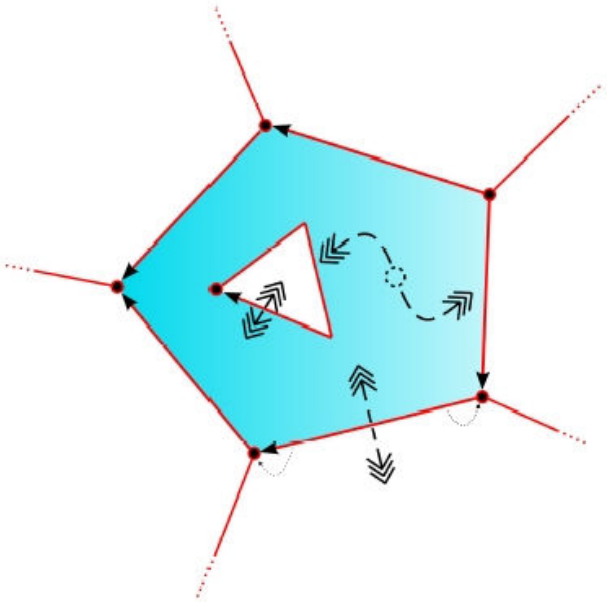
- Geometrický datový model.
- Topologický datový model.

Oba modely mají své silné a slabé stránky. Geometrický přístup nabízí přímý přístup k souřadnicím objektů, topologický přístup naproti tomu ukládá informace o prostorových vztazích (např. sousednost). Vzhledem k tomu, že geometrický model poskytuje přímý přístup k souřadnicím, společné části sousedních polygonů jsou uloženy vícekrát, což vede k redundanci dat. Topologický přístup eliminuje tyto duplicity v datech použitím *odkazů* (*references*) na jednoznačně identifikovatelná topologická primitiva.

1.1 Současný způsob uložení prostorových dat v ISKN

Zaměřím se na stručný popis současného způsobu uložení prostorových dat v ISKN, konkrétně na digitální katastrální mapu. Jak bylo zmíněno výše, lze vektorová prostorová data v systémech řízení báze dat ukládat podle geometrického, či topologického modelu. Způsob uložení DKM se nejvíce blíží tzv. „left-right“ topologii s ukládáním odkazů na hrany, která se řadí mezi topologické modely.

„Left right“ topologie s ukládáním odkazů na hrany je založena na ukládání informací o stěně (v našem případě parcele) vlevo a vpravo pro každou orientovanou hranu (v našem případě úsek hranice parcely). Na základě této informace můžeme získat kolekci hran náležející jedné stěně a z nich stěnu zrekonstruovat. Všechny hrany jsou orientované a každá hrana má počáteční a koncový uzel. Pro každou stěnu (parcelu) je uložen odkaz na jednu hranu, která je součástí dané stěny. Při hledání hranice konkrétní stěny (parcely) musí mezi vybranými hranami proběhnout operace vyhledání hran, které lze dočasně vzájemně „navázat“, tzn. mající společný uzel. Výhodou tohoto přístupu je, že se do databáze neukládá příliš mnoho odkazů. Nevýhodou je naproti tomu nutná operace vyhledání sousedních (navazujících) hran při vyhledávání hranice stěny. Schematicky je tento způsob ukládání vektorových prostorových dat zachycen na obrázku 1.



Obr. 1. „left-right“ topologie s ukládáním odkazů na hrany [6].

2 Řešení nekonzistencí v prostorových datech ISKN

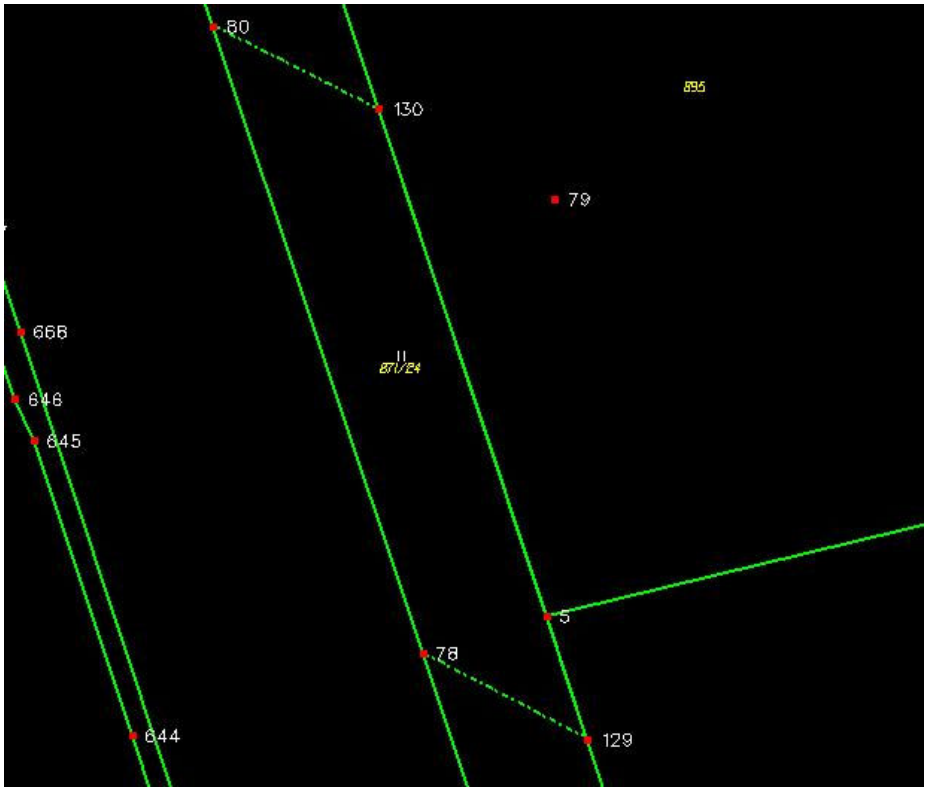
V rámci spolupráce se Sekcí centrální databáze katastru nemovitostí dochází k definování a identifikaci chyb, nacházejících se v prostorových datech ISKN. Cílem je data zkontrolovat, následně opravit a mít je ve stavu, kdy budou chyby negativně ovlivňující například automatickou generalizaci, eliminovány. Data po zobrazení v klientu (Kokeš, Microstation,…) mohou být ve zvolené úrovni podrobnosti vizualizována uspokojivě, ovšem při dostatečně velkém přiblížení se na část dat lze pozorovat nedostatky v prostorových vztazích mezi prvky, například hranicemi parcel. Tento způsob identifikace chyb je pro celou oblast, na které se digitální katastrální mapa vyskytuje, samozřejmě nereálný a především kategoricky špatný. Efektivním řešením je vyhotovení jednotlivých funkcí či aplikací, které dokáží podle předem definovaných pravidel konkrétní chybu odhalit. Rád bych uvedl typy chyb, které se v prostorové složce ISKN vyskytují a postupy, který dokáží tyto chyby identifikovat. V práci [5] jsou vyjmenovány charakteristiky topologicky konzistentní množiny dat:

- Každá hrana musí být ohraničena dvěma uzly (počáteční a koncový uzel).
- Každá hrana má stěnu vlevo a stěnu vpravo.
- Každá stěna má uzavřenou hranici sestávající ze střídavé posloupnosti uzlů a hran.
- Kolem každého uzlu je střídavá uzavřená sekvence hran a stěn.
- Hrany se vzájemně neprotínají, pouze v uzlech.

2.1 Prázdné polygony

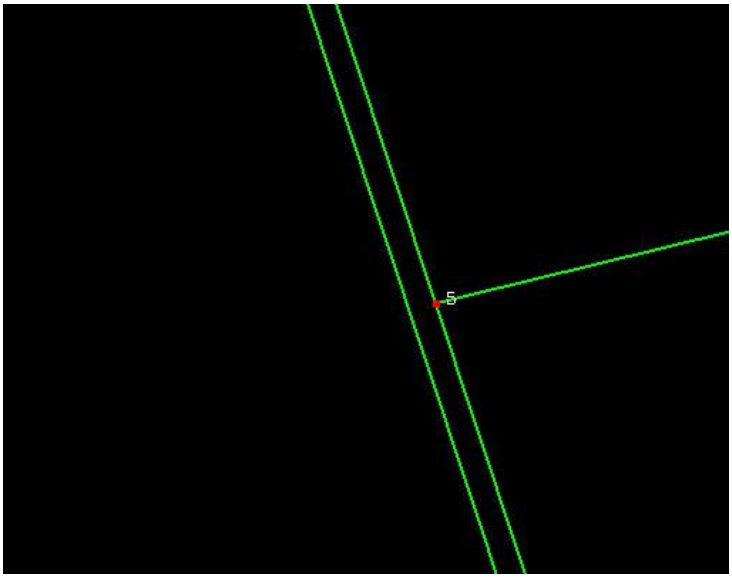
Příkladem chyby, která se v digitální katastrální mapě vyskytuje, může být prázdný polygon. Prázdným polygonem se rozumí uzavřená posloupnost vrcholů na sebe navazujících hraničních úseků parcel, která ovšem nepředstavuje hranici parcely. Prázdný polygon typicky vzniká ve chvíli, kdy dochází ke vzniku nové parcely a při vložení nového bodu hranice

parcely nedojde k jeho přesnému umístění na část hraničního úseku, který se v datech již vyskytuje. Geometricky může situace vypadat například jako na obrázku 2. Všimněme si spojnic bodů 130 – 5, 5 – 129 a 129 – 130.



Obr. 2. Ukázka prostorových dat ISKN.

Že je zde spojnice 129 – 130 navíc? Pokud se vizualizací přiblížíme k bodu 5, je vidět, že dochází ke vzniku prázdného polygonu mezi body 130, 5 a 129 (viz. obr. 3) a tudíž se v datech objevuje i spojnice 129 - 130.



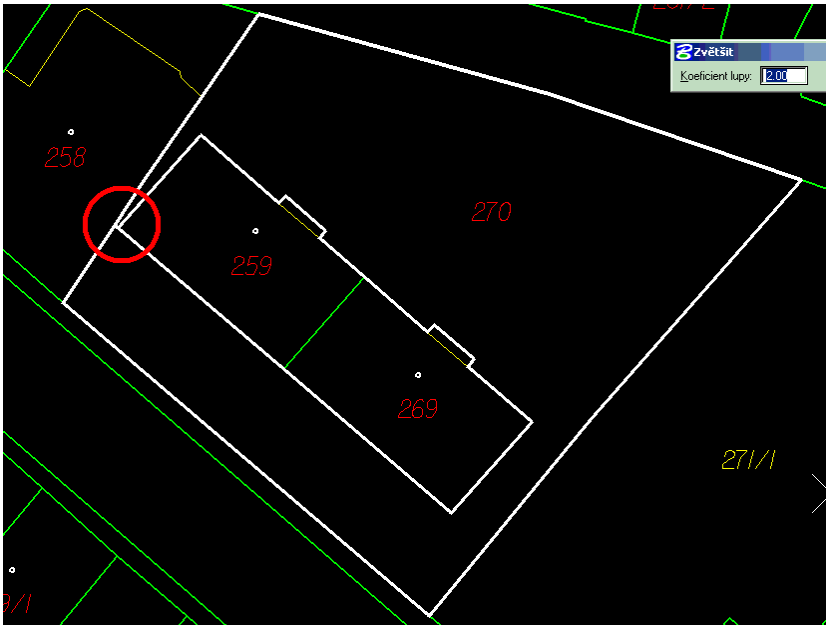
Obr. 3. Detail prázdného polygonu.

Algoritmus pro nalezení prázdných polygonů jsem publikoval v [4].

2.2 „Šťopky“

Ze způsobu uložení hranic parcel v ISKN vyplývá, že je pro každou hranu (segment hranice parcely) uložen v odpovídajícím záznamu identifikátor parcely „vlevo“ a identifikátor parcely „vpravo“. Pokud neexistuje digitální katastrální mapa po obou stranách daného hraničního úseku, je příslušný identifikátor roven hodnotě NULL.

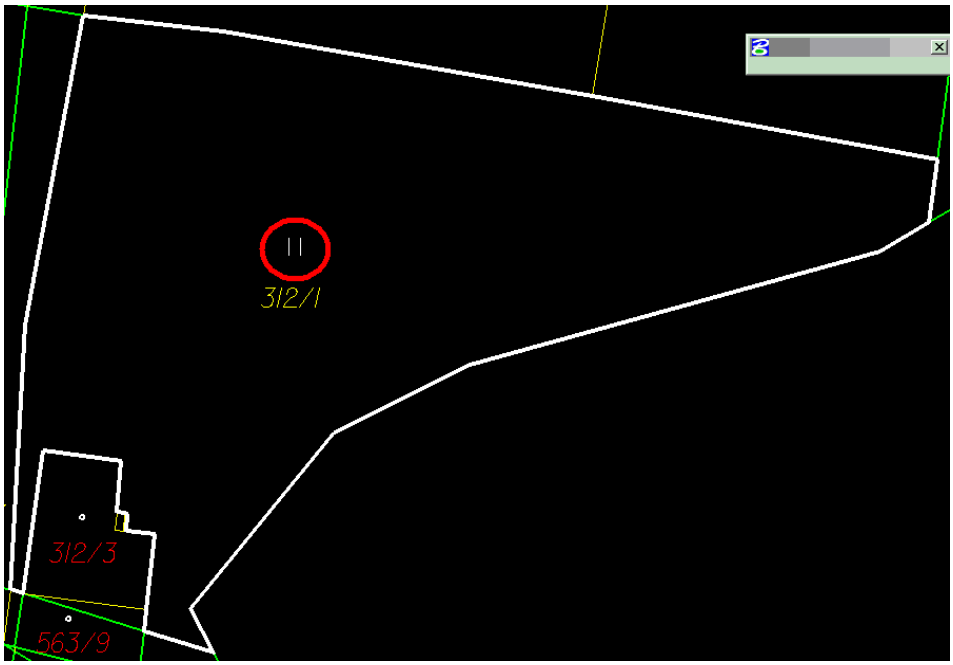
V případě, kdy hrana reprezentuje segment vlastnické hranice oddělující dvě různé parcely (za předpokladu existence DKM po obou stranách hrany), je vyžadováno, aby pro tuto hranu byly uloženy odkazy na parcely po obou stranách. Můžete se stát, že oba dva uložené identifikátory jsou shodné, resp. odkazují na identickou parcelu. Pak hovoříme o tzv. „šťopkách“. Bohužel, není mi znám žádný odborný ekvivalent. Příklad „šťopky“ ilustruje obrázek 4. Pro automatizovanou lokalizaci této chyby byl vytvořen dotaz v jazyce SQL, který ve všech lokalitách pokrytých DKM vyhledá tento typ chyby.



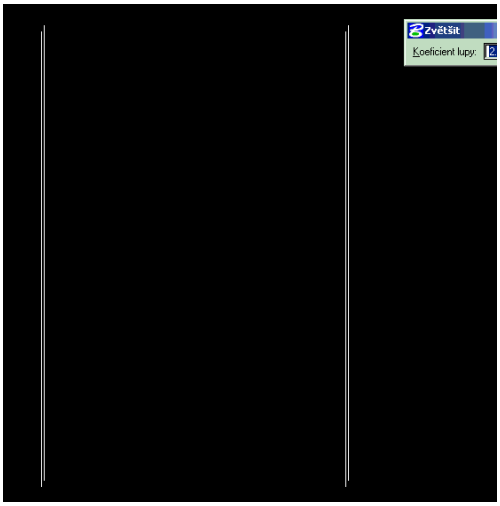
Obr. 4. „Šťopka“.

2.3 Duplicitní mapové značky

Při grafickém určení značky druhu pozemku může dojít k tomu, že bude značka zanesena dvakrát (případně vícekrát) ve smyslu, že vztažné body značek budou ve vzdálenosti menší než určená prahová hodnota. Řešení tohoto problému opět vyžaduje automatizovaný přístup. Implementoval jsem v programovacím jazyce Java algoritmus využívající techniky sweep line [3] známé z oblasti výpočetní geometrie, který prohledá vstupní množinu dat (vztažné body mapových značek vyjadřujících druh pozemku), nalezne dva nejbližší body a na základě prahové hodnoty rozhodne o tom, zda jsou dva body identické. Pokud jsou dva nejbližší body od sebe ve vzdálenosti větší než prahová hodnota, algoritmus dále nepokračuje. V opačném případě najde další dvojici nejbližších bodů a provede nové porovnání vůči prahové hodnotě. Na obrázcích 5 a 6 je znázorněna duplicitní mapová značka „Trvalý travní porost“ (pozn.: Algoritmus prozatím nebyl testován nad provozními daty ISKN, jedná se o situaci nasimulovanou nad lokální testovací databází).



Obr. 5. Duplicitní mapová značka – duplicita ovšem není patrná.



Obr. 6. Duplicitní mapová značka - detail.

2.4 Hraniční úseky parcel se shodnými koncovými body

Hraniční úseky parcel jsou v ISKN určeny pomocí dvou koncových bodů a případných mezilehlých bodů. V datech se vyskytuje případ, kdy hraniční úsek sestává ze dvou bodů o shodných Y, X souřadnicích. Vytvořil jsem dotaz v jazyce SQL, který všechny takovéto úseky v datech identifikuje.

2.5 Nespočtené průsečíky křížení linií vnitřní kresby

Digitální katastrální mapa obsahuje i tzv. vnitřní kresbu, kterou se znázorňují například hranice budov. Je žádoucí, aby pro všechna křížení linií vnitřní kresby existovaly spočtené průsečíky a údaje o nich byly uloženy v databázi. Prozatím tomu tak být nemusí, takže je

žádoucí, mít k dispozici například nějaký program, který by tyto nespočtené průsečíky v prostorových datech ISKN našel.

Pro řešení problému se opět nabízí možnost, sáhnout po vhodném algoritmu z oblasti výpočetní geometrie. Algoritmus, jehož popis lze najít v [1] a který je založen na již zmíněné *sweep-line* technice, přesně splňuje kritéria na takový algoritmus kladená – aby byly nalezeny ve vstupní množině dat (liniové úseky vnitřní kresby) všechny průsečíky efektivním způsobem, tzn. ne hrubou silou, čili testováním každého úseku s každým. Algoritmus jsem implementoval v programovacím jazyce Java s využitím vhodných datových struktur. V současné době probíhá testování implementovaného algoritmu.

3 Závěr

Z důvodu historie vývoje ISKN se v současné době může v prostorových datech vyskytnout několik typů chyb, které mohou později negativně ovlivnit například automatickou generalizaci. Hlavním důvodem, proč se snažit o odstranění všech topologických nekonzistencí v prostorových datech ISKN, je chystaný vznik centrální databáze katastru nemovitostí. Je proto potřeba tyto chyby definovat a následně pomocí automatizovaných postupů identifikovat.

Problematika konzistence prostorových dat katastru nemovitostí je velice široká. V práci jsem se pokusil nastínit možné metody, které vedou k identifikaci a následnému odstranění vybraných typů chyb (nekonzistencí) v prostorových datech ISKN. Některé z nich jsou již součástí standardních GIS, např. hledání nespočtených průsečíků. Přínos mého řešení vidím v tom, že je možné kontrolu konkrétní chyby provést nad daty celé ČR, respektive tam, kde je k dispozici digitální katastrální mapa (DKM) najednou. V softwaru typu MicroGEOS Nautil je možné pracovat po katastrálních územích a pro každé katastrální území spouštět požadované a dostupné kontroly zvlášť. Mnou navržené řešení spočívá v aplikování vhodných algoritmů z oblasti výpočetní geometrie použitých nad prostorovými daty ISKN. Data do algoritmů budou vstupovat v rozsahu všech digitálních katastrálních map. Výstupem implementovaných postupů je identifikace chyb způsobem, aby bylo možné výstup „roztřídit“ podle jednotlivých katastrálních pracovišť. Vlastní odstranění chyb proběhne na příslušném pracovišti odpovědným pracovníkem. Je důležité si uvědomit, že navržené postupy mají identifikační charakter, tzn. říkají, kde je jaký typ chyby. Většina algoritmů, naprogramovaných v jazyce Java, již byla úspěšně otestována a použita v praxi. Nalezené chyby budou opraveny na příslušných katastrálních pracovištích.

Reference

1. M. de Berg, M. van Kreveld, M. Overmans, O. Schwarzkopf. *Computational Geometry: Algorithms and Applications.* Springer-Verlag, Berlin, 1997.

2. Český úřad zeměměřický a katastrální. *Struktura výměnného formátu informačního systému katastru nemovitostí České republiky.* Číslo jednací 6665/2004-24. Praha, 2004.

3. Closest Pair in the Plane with (2,4) – trees.
http://www.cs.mcgill.ca/~cs251/ClosestPair/ClosestPairPS.html

4. Janečka, K. *Konzistence prostorových dat ISKN.* Sborník 9. Odborné konference doktorského studia JUNIORSTAV. Brno, 2007. ISBN: 978-80-214-3337-3.

5. Kainz, W.: *Geographic Information Science, Technology, and Infrastructure*. http://homepage.univie.ac.at/wolfgang.kainz/Lehrveranstaltungen/GIS/GIS-V2.3.pdf

6. Meijers, M.: *Implementation and testing of variable scale topological data structures.* Master's thesis. TU Delft, 2006.

7. Open Geospatial Consortium. http://www.opengeospatial.org