

Zastarávanie geoinformácií a ich aktualizácia

Marián Adamják

Topografický ústav
974 01, Banská Bystrica, Slovensko
madamjak@topu.army.sk

Abstrakt. Geoinformatika a priestorové informácie sú kľúčové v rozhodovacích procesoch. Spoľahlivé informácie sú kľúčové k dosiahnutiu správnych rozhodnutí, najmä na úrovni riadenia krajiny a v krízovom manažmente. Článok popisuje a diskutuje niektoré dôležité pojmy v spojitosti so spoľahlivosťou priestorových informácií. Vek informácií je považovaný za základnú časť metaúdajov a je dávaný do súvislosti s pravdepodobnosťou nespoľahlivosti informácií. Pravdepodobnosť nespoľahlivosti údajov je m odelovaná jako funkcia závislá na čase. Aktualizačný proces, ktorý má zvyšovať spoľahlivosť geoinformácií, môže byť plánovaný na základe pravdepodobnosti nespoľahlivosti. Tým sa môže dosiahnuť optimalizácia a vysoká efektívnosť aktualizáčného procesu.

Kľúčová slova: priestorové informácie, aktualizácia, spoľahlivosť informácií, rozhodovacie procesy.

Abstract. Geoinformation (spatial information) is very important in decision processes. Reliability of spatial information is crucial to make good and correct decisions. Article describes and discusses several terms in connection with reliability of spatial information. Age of data is fundamental piece of information in metadata. Main term "probability of unreliability" is described based on age of data and characteristics of feature classes. Probability of unreliability is defined as function of time. Updating processes could be planed based on probability of unreliability to reach best of cost-effectiveness and optimal time consumption. Probability of unreliability is could use to predict fault in decision made by unreliable data..

Keywords: spatial information, updating, reliability of information, decision making

1 Úvod a ciele

Technologický pokrok, ktorý môžeme okolo nás vnímať, je spôsobený najmä rozvojom informačných technológií a ich možnosťami, ktoré priniesli ľuďom obrovské možnosti zvyšovania kvality života a rozvoja spoločnosti. Je prirodzené že možnosti informačných technológií výrazne ovplyvňuje ich samotný rozvoj. Tieto trendy zrejme platia všeobecne a týkajú sa teda aj informačných systémov, ktoré obsahujú priestorovú zložku.

Hlavnou črtou rozvoja v oblasti spracovania priestorových informácií je v poslednom období významný posun k spracovaniu a využívaniu vektorových typov modelov území. Údaje pre tieto typy modelov sú získavané rôznymi metódami a postupmi od presných geodetických meraní, cez metódy založené na fotogrametrickom spracovaní

až po vektorizácii mapových podkladov. Podstatná väčšina správne tvorených vektorových modelov územia obsahuje okrem primárnych údajov o objektoch v modelovanom území, aj súbor údajov o týchto primárnych údajoch – metaúdaje. Typickými predstaviteľmi týchto metaúdajov sú najmä údaje o geometrickej či atribútovej presnosti, údaje o metódach a zdrojoch z ktorých boli primárne údaje získané a údaje o čase (dobe) kedy boli tieto údaje získané, resp. údaje o dobe ku ktorej sa primárne údaje vzťahujú.

Metaúdaje o čase, kedy boli primárne údaje získané, sú mimoriadne dôležité pre posudzovanie aktuálnosti a spoľahlivosti primárnych údajov.

Cieľom tohoto príspevku je na jednej strane rozobrať pojem „aktuálnosť“ a na strane druhej načrtnúť niektoré možnosti, ako je možné zisťovať výskyt zmien vlastností objektov v krajine, zisťovať vznik nových resp. zánik existujúcich výskytu objektov. Okrem samotného pojmu aktuálnosť (vek informácie) je diskutovaný vzťah tohto činiteľa na spoľahlivosť rozhodovania založeného na údajoch, ktorých aktuálnosť sa hodnotí. V závere sú načrtnuté niektoré fakty súvisiace s vyhľadávaním zmien v krajine, čo má zásadný význam pre aktualizáciu údajov v informačných systémoch s priestorovou zložkou.

2 Vek údajov a čas poznania objektov

Aktuálnosť je možné v najjednoduchšej forme vyjadriť ako rozdiel času použitia informácie o objekte a času kedy bola informácia o objekte získaná, resp. zaznamenaná. Určenie času použitia informácie je obvykle triviálnou záležitosťou avšak určenie času kedy bola informácia skutočne získaná nemusí byť najmä pri nesprávnom zaznamenávaní metaúdajov o čase získania primárnej informácie o objekte jednoduchou záležitosťou.

Za čas získania informácie o objekte je potrebné vždy považovať okamžik, kedy bol naposledy objekt poznávaný v rozsahu, ktorý je určený parametrami vytváraného, prípadne aktualizovaného, vektorového modelu územia. V zásade je možné na účely vymedzenia času poznávania objektu vymedziť nasledovné skupiny postupov:

- a) kontaktné metódy, ktorými sú najmä geodetické metódy a metódy mapovania, zisťovania a overovania vlastností priamo v teréne,
- b) bezkontaktné metódy, ktorými sú najmä metódy diaľkového prieskumu zeme a digitálnej fotogrametrie,
- c) preberanie informácií z iných informačných systémov a zdrojov.

Časom poznania objektu pri kontaktných metódach je okamžik, kedy objekt bol týmito metódami naposledy meraný a poznávaný – tým sa však nemyslí len určenie geometrie objektu, ale najmä overenie samotnej existencie objektu či zisťovanie ďalších - obvykle negeometrických – vlastností.

Časom poznania objektu pri bezkontaktných metódach, je okamžik vyhotovenia záznamu na nosič obrazu. Pri leteckom meračskom snímkaní je to okamžik exponovania (osvetlenia) filmu. Časom poznania objektu zjavne nie je čas vyvolania filmu, či interpretácie snímok v prostredí digitálnej fotogrametrie.

Určenie času pri preberaní informácií z iných zdrojov je závislé na tom či je možné získať z informačného zdroja údaj o tom, kedy bol príslušný objekt naposledy

poznávaný. Jedným zo spôsobov preberania informácií z iných zdrojov je získavanie informácií z máp vektorizovaním ich obsahu, prípadne preberaním údajov o objektoch zobrazených v týchto mapách. Mapové podklady však obvykle obsahujú len údaje o redakčnej uzávierke, alebo údaj o dátume schválenia obsahu. Tieto časové údaje však neuvádzajú skutočný čas poznávania objektov. V týchto prípadoch skutočný čas poznávania objektov nie je možné spoľahlivo odvodiť. Ďalšími zdrojmi môžu byť rôzne databázy, katalógy a ročenky. Tieto zdroje buď vôbec neuvádzajú čas poznávania objektov či získania informácií, alebo ho uvádzajú len veľmi všeobecne. V týchto prípadoch je potrebné stanoviť čas poznania objektov odhadom a objektivizovať ho tak, aby takto odhadnutý čas poznávania objektov nebol neskorší ako sa v skutočnosti mohlo stať.

Bez ohľadu na metódy a postupy poznávania objektov sa ukazuje ako nevyhnutné stanoviť čas, kedy boli objektu skutočne poznávané. Úvahy uvedené v tomto článku vychádzajú z faktu, že čas poznávania objektu je stanovený. Pri použití skupín metód v ktorých sa informácie len preberajú sa predpokladá, že čas poznávania objektu bude vhodne odhadnutý a objektivizovaný.

3 Zložky informácie o priestorových objektoch a ich spoľahlivosť

Informácia o priestorovom objekte (entite ktorá patrí do určitej triedy objektov) je najčastejšie vyjadrená ako súbor hodnôt opisujúcich jeho vlastnosti (atribúty). Obvykle existujú atribúty viacerých typov – číselné, textové, logické, časové údaje, geometrie a pod. Spoľahlivosť pravdivosti každého atribútu sa v čase mení a klesá. Závislosť spoľahlivosti atribútu na čase je zrejme možné vyjadriť matematicky opísateľným vzťahom. Pre rozsiahle množiny objektov, bude mať tento vzťah pravdepodobnostný charakter. Tento vzťah bude potom vyjadrovať s akou pravdepodobnosťou v závislosti na čase, bude údaj (hodnota atribútu) ešte správny – pravdivý. Vyšetrením vzťahov a závislostí pravdepodobností pravdivosti jednotlivých atribútov by bolo možné kvantifikovať mieru spoľahlivosti na úrovni jednotlivých objektov.

Pravdepodobnosti pravdivosti jednotlivých údajov o jednom objekte sa v rámci tohto objektu navzájom ovplyvňujú. Vyšetrovanie týchto závislostí medzi funkciami vyjadrujúcimi pravdepodobnosť pravdivosti jednotlivých atribútov je veľmi dôležité a je možné ho demonštrovať na závislosti medzi atribútmi typu geometria a ostatnými atribútmi. Pri hodnotení pravdepodobnosti geometrických atribútov je možné vyjadrovať dve pravdepodobnostné charakteristiky pravdivosti a to – pravdepodobnosť existencie objektu (resp. jeho neexistencie) a pravdepodobnosť zmeny tvaru (objekt existuje ale zmenil tvar). V prípade že nastane zmena v existencii objektu implikuje nepravdivosť všetkých ostatných atribútov. V prípade zmeny tvaru objektu môže dochádzať u niektorých tried objektov k významným zmenám v pravdepodobnosti pravdivosti u všetkých vlastností a u niektorých tried len u niektorých vlastností. Zároveň je potrebné hodnotiť ako sa zmena existencie či vlastností jedného objektu prejaví na vlastnostiach iného objektu.

Príkladom takéhoto posudzovania nech sú triedy objektov „budova“, „cestný úsek“ a „mostná konštrukcia“. V prípade zmeny geometrie cestného úseku zjavne došlo k prestavbe celého objektu a jeho vlastnosti (napr. šírka, materiálové zloženie povrchu) je potrebné považovať za nespoľahlivé – pravdepodobnosť pravdivosti sa blíži nule. Naproti tomu zmena geometrie budovy bude mať zrejme odlišný dopad na spoľahlivosť atribútu „účel využitia budovy“ a odlišný dopad na spoľahlivosť atribútu „výška budovy“. Zmena existencie objektu triedy „mostná konštrukcia“ zásadne ovplyvní spoľahlivosť niektorých vlastností objektov v bezprostrednom okolí. Vybudovaním mostnej konštrukcie (premostenia) nad cestným úsekom spôsobí zníženie pravdepodobnosti pravdivosti jeho atribútu „prejazdna výška“ na nulu, keďže premostenie vybudované nad cestou môže obmedzovať (a často aj obmedzí) maximálnu výšku vozidla, ktoré môže popod mostnú konštrukciu prejsť.

Z vyššie uvedených príkladov sa dá usúdiť, že by bolo možné nájsť vzťahy medzi spoľahlivosťami jednotlivých atribútov v rámci jedného objektu i medzi viacerými objektmi a zároveň, že by bolo možné kvantifikovať spoľahlivosť jednotlivých atribútov a aj objektov ako celok.

Kvantifikácia nespoľahlivosti jednotlivých atribútov a následne aj informácií o objektoch ako takých umožní zaviesť pojem riziko nespoľahlivej informácie.

4 Riziko nespoľahlivej informácie a jeho obmedzovanie

4.1 Odhad nespoľahlivosti informácií

Informačný systém, jeho údajový sklad, databáza i informácie v tejto databáze „žijú svoj životný cyklus“. Tento životný cyklus je možné v prípade informácií o priestorových objektoch opísať nasledovne:

- a) prvotné poznávanie objektu a získanie prvotných informácií o objekte
- b) existencia informácie o objekte bez zmeny tejto informácie bez ohľadu na skutočný stav objektu – postupný pokles spoľahlivosti informácie o objekte
- c) pravidelné preverovanie vlastností objektu a ich prípadná aktualizácia
- d) existencia informácie o objekte bez zmeny tejto informácie bez ohľadu na skutočný stav objektu – postupný pokles spoľahlivosti informácie o objekte – po istom čase návrat k pravidelnému preverovaniu vlastností (predchádzajúci bod),
- e) zánik informácie o objekte spolu so zánikom objektu

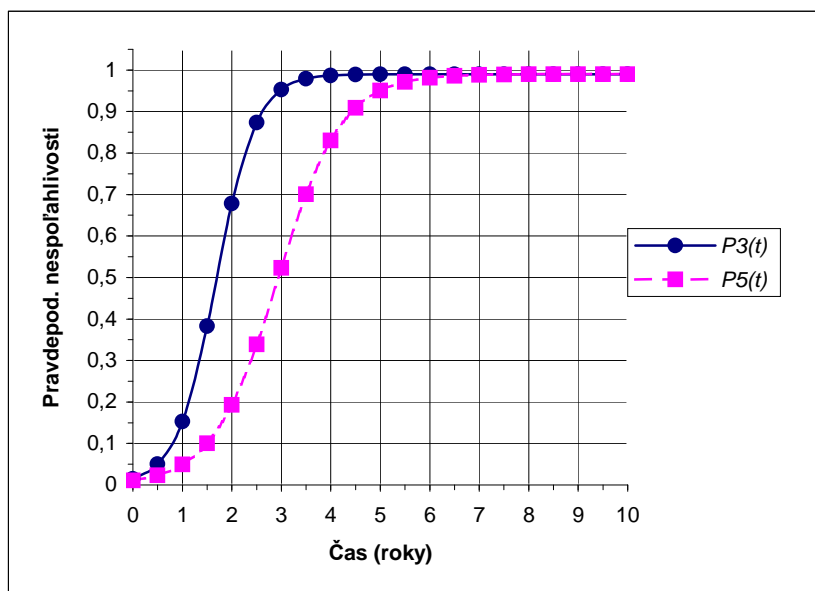
Po každej etape životného cyklu, v ktorej sa poznávajú vlastnosti objektu, dochádza k zvýšeniu spoľahlivosti informácie o objekte, tzn. k zvýšeniu pravdepodobnosti, že informácia o objekte je pravdivá. Miera úspešnosti zvýšenia pravdepodobnosti pravdivosti je závislá od účinnosti metódy poznávania a overovania vlastností jednotlivých objektov. Túto úspešnosť je možné tiež kvantifikovať a to ako pravdepodobnosť, že zvolená metóda poznávania zistí všetky zmeny existencie a vlastností objektu. Príkladom rozdielných účinností je napr. použitie metód geodetického merania a poznávania v teréne oproti a bezkontaktným metódam kedy sa napr. porovná existujúci vektorový model územia a ortogonalizovaná letecká snímka a kvalifikovane sa odhadne miera zmien objektov. V prípade geodetických

metód je účinnosť poznania zmien objektov a teda zvýšenie pravdepodobnosti pravdivosti informácie o objekte veľmi vysoké. Naproti tomu bezkontaktné metódu budú mať výrazne nižšiu účinnosť zisťovania zmien vlastností objektov. Čiastkovým záverom, ktorý môžeme vyvodit' z predchádzajúcich úvah je, že pravdepodobnosť pravdivosti informácie o objekte zachytená vo vektorovom modeli územia je vždy nižšia ako jedna – informácie o objektoch nie sú absolútne spoľahlivé.

Modely územia vrátane vektorových slúžia najmä na podporu rozhodovania ľudí pri riadení rozvoja krajiny, pri ochrane životného prostredia, pri riešení krízových situácií a v mnohých ďalších prípadoch. Rozhodovanie týchto ľudí – autorít – má obvykle zásadný dopad na životy a osudy ďalších ľudí, na ekonomiku a rozvoj krajiny. Pri použití nespoľahlivej informácie v procesoch vyššie spomínaného rozhodovania, vzniká nebezpečenstvo, že výsledky a závery odvodené na základe takejto informácie budú nesprávne a ohrozia správnosť ďalších rozhodnutí a následne zvolených postupov. Mieru rizika, že bude použitá nepravdivá informácia je možné kvantifikovať tak, že sa pokúsime určiť pravdepodobnosť, že niektorý z údajov o priestorovom objekte je nepravdivý. Potom miera nebezpečenstva nesprávneho rozhodnutia je priamo úmerná tejto pravdepodobnosti. Pre stanovenie odhadu pravdepodobnosti, že údaj je nepravdivý (pravdepodobnosť nespoľahlivosti) sa navrhuje použiť napr. funkciu obmedzeného rastu:

$$P(t) = \frac{ML}{M + (L - M)e^{-Lrt}} \quad (1)$$

kde $P(t)$ je pravdepodobnosť, že údaj nie je spoľahlivý v čase t . Hodnota M reprezentuje počiatočnú nespoľahlivosť, L reprezentuje koncovú pravdepodobnosť nespoľahlivosti a r je koeficient reprezentujúci rýchlosť „starnutia“ informácie o objekte.



obr. 1

Na obrázku 1 sú demonštrované dva možné priebehy starnutia údaju o objekte. Horizontálna os reprezentuje čas a vertikálna os reprezentuje pravdepodobnosť že údaj o objekte je nepravdivý.

Ak je údaj považovaný v horizonte 3 rokov za zastaralý a zároveň predpokladáme,

1. že pravdepodobnosť zastarania údaju za 0,5 roka je menšia ako 0,05
2. a zároveň, že údaj je zastaraný po 2,5 roku s pravdepodobnosťou viac ako 0,95 potom

budú mať koeficienty približne nasledovné hodnoty:

$$M \approx 0,015 \quad L \approx 0,99 \quad \text{a} \quad r \approx 2,5$$

Ak je údaj považovaný v horizonte 5 rokov za zastaralý a zároveň predpokladáme,

1. že pravdepodobnosť zastarania údaju za 1 rok je menšia ako 0,05
2. a zároveň že údaj je zastaraný po 4 rokoch s pravdepodobnosťou viac ako 0,95 potom

budú mať koeficienty približne nasledovné hodnoty:

$$M \approx 0,011 \quad L \approx 0,99 \quad \text{a} \quad r \approx 1,55$$

Na obr. 1 sú zobrazené grafy oboch funkcií. Funkcia $P3(t)$ prislúcha prípadu, kedy uvažujeme že údaj je po 3 rokoch nespoľahlivý a funkcia $P5(t)$ demonštruje nespoľahlivosť po 5 rokoch. Hodnoty koeficientov boli určené odhadom a sú určené len na demonštráciu. Pre každú triedu objektov a metódu zisťovania údajov o objektoch tejto triedy, bude potrebné kvalifikovane odhadnúť jednotlivé hodnoty koeficientov. Je zrejmé že koeficient L bude rovnaký pre všetky triedy, keďže sa jedná o maximálnu očakávanú pravdepodobnosť nespoľahlivosti údajov o objektoch. Koeficient M reprezentuje zvyškovú nespoľahlivosť, ktorej sa pri zisťovaní údajov o objektoch objektívne alebo subjektívne dopúšťame. Tento koeficient bude zrejme odlišný pre kontaktné metódy zisťovania vlastností objektov (geodetické merania, mapovanie v teréne) a pre bezkontaktné metódy (fotogrametria). Pre kontaktné metódy sa zrejme bude blížiť nule, čo vyjadrí ich vyššiu spoľahlivosť. Koeficient r reprezentuje rýchlosť straty spoľahlivosti údajov o objekte. Táto rýchlosť je ovplyvnená najmä nasledovnými faktormi:

- a) povahou objektu – niektoré objekty sa v čase vyvíjajú iné zostávajú prakticky bez zmeny. Typickým príkladom je lesný porast ktorý rastie a teda sa plynulo v čase mení. Naproti tomu cesta, budova, linka elektrického vedenia či iné človekom vytvorené objekty sa v čase menia omnoho pomalšie než je tomu u porastov,
- b) vplyvmi človeka alebo prírody samotnej na charakter krajiny – niektoré oblasti krajiny podliehajú v príslušnej časovej perióde veľkým zmenám, u iných je ich vývoj pomalý. Typickým príkladom je budovanie veľkých priemyselných oblastí, ktoré majú významný vplyv na široké okolie a objekty typu budova alebo cesta tu budú podliehať zmenám rýchlejšie než v iných oblastiach.

Autor pripúšťa, že pre určenie pravdepodobnosti je možné použiť aj iné druhy funkcií prípadne metód. V ďalšom texte sa pridržiava len faktu, že pravdepodobnosť nepravdivosti údaju je určená dostatočne vhodnou a spoľahlivou metódou. Parametre týchto funkcií či metód by zrejme tiež mali byť súčasťou technickej dokumentácie príslušného vektorového modelu územia (metaúdajov o tomto modeli).

4.2 Obmedzovanie nespoľahlivosti informácií

Model určovania pravdepodobnosti nespoľahlivosti údajov demonštrovaný v predchádzajúcej časti vytvára nástroj na posudzovanie spoľahlivosti jednotlivých údajov, ktoré opisujú konkrétny objekt. Bude teda zrejme možné v ľubovoľnom čase stanoviť aké spoľahlivé údaje o jednotlivých objektoch sú vo vektorovom modeli územia zaznamenané. Toto vytvára predpoklad pre efektívne obmedzovanie nespoľahlivosti údajov vo vektorových modeloch územia.

Obmedzovanie nespoľahlivosti údajov je možné zabezpečiť:

- a) pravidelnou aktualizáciou týchto údajov, alebo
- b) vytváraním takých údajových štruktúr, ktoré len odkazujú na zdroje údajov, tzn. na také údajové štruktúry do ktorých sú údaje primárne zaznamenané (napr. kataster nehnuteľností, informačné systémy odbornej štátnej správy a pod.).

Riešenie uvedené v bode b) je už v súčasnosti známe aj keď ešte nie je rozšírené a zároveň nie je riešením pre všetky druhy údajov a modelov územia. Preto sa v ďalšom budeme zaoberať len aktualizáčnym procesom (spomenutom v bode a)).

Aktualizáciu je možné vykonávať rôznymi metódami a postupmi. Tieto metódy a postupy budú mať rôznu úspešnosť pri vyhľadávaní (identifikácii) zmien vlastností objektov a zároveň budú mať rozdielnu schopnosť poskytovať aktuálne údaje v očakávanej kvalite. Pri plánovaní aktualizácie sa bude teda možné oprieť:

- a) o určenie pravdepodobnosti s akou údaje o jednotlivých objektoch zastarávajú v konkrétnom čase a vyberať len tie, ktorých miera nespoľahlivosti prekročila určenú medzu a zároveň
- b) poznajúc účinnosť aktualizáčnych metód vybrať optimálnu z nich, ktorá efektívne znížila úroveň nespoľahlivosti údajov o vybraných objektoch.

Príkladom môže byť na jednej strane porovnávanie materiálov diaľkového prieskumu zeme (vrátane leteckého meračského snímkovania) ktoré vznikli v rozdielnych časoch a na druhej strane preberanie údajov zo stavebného úradu. Metódy porovnania materiálov DPZ a LMS z rôznych časových období môžu indikovať rôzne zmeny vo vegetačnom kryte krajiny, rozsiahlu novú zástavbu atď. na veľkých častiach územia avšak nemôžu podať spoľahlivú a úplnú informáciu o charaktere týchto zmien. Naproti tomu údaje zo stavebného úradu umožnia rýchlo nájsť nové budovy, opravované budovy či demolácie. Takto identifikované budovy sú spoľahlivo vymedzenou podmnožinou zo všetkých budov, ktoré sú zmenené. Informácie zo stavebného úradu môžu zároveň obsahovať aj ďalšie spoľahlivé údaje vhodné pre aktualizáciu informácií o týchto budovách. Oproti materiálom DPZ a LMS je však zo stavebného úradu získať len informácie týkajúce sa obmedzenej množiny objektov z niektorých tried.

Obdobne, tak ako bola kvantifikovaná miera nespoľahlivosti údajov, bude možné odhadnúť mieru účinnosti jednotlivých aktualizáčnych metód a postupov. Nech účinnosť U je vyjadrená ako pravdepodobnosť, že nespoľahlivý údaj bude identifikovaný a opravený, hodnota $P(t)$ nech je pravdepodobnosť nespoľahlivosti údaju (určená v (1)), tak potom hodnota RI (rizikový index):

$$RI = P(t)(1 - U) \quad (2)$$

udáva mieru rizika s akým je možné údaj po aktualizácii použiť. Ak účinnosť U bude vyjadrená v intervale $(0,1)$ potom rizikový index RI nadobudne hodnotu taktiež v

intervale (0,1), kde 0 znamená žiadne riziko a 1 znamená maximálne riziko (údaj je nespoľahlivý).

Ak neaplikovanie žiadnej aktualizáčnej metódy má účinnosť $U = 0$ a zároveň ak teoretická metóda s účinnosťou $U = 1$ spoľahlivo zistí všetky zmeny a opraví údaje potom je možné hodnotu RI považovať za pravdepodobnosť že údaj aj po aplikovaní určitej metódy zostal nespoľahlivý.

Ak zoberieme do úvahy fakt, že koeficient M je počiatočnou neistotou v zistení údajov o objekte a RI je rizikový index, ktorý reprezentuje zvyškovú pravdepodobnosť, že údaj aj po aktualizácii metódou s účinnosťou U zostal nespoľahlivý, potom po aktualizácii bude koeficient M rovný indexu RI.

4.3 Závislosti a zovšeobecňovanie nespoľahlivosti údajov

Ako už bolo spomenuté existujú vzájomné väzby medzi zmenami vlastností v rámci jednotlivého objektu i medzi viacerými objektmi. Vyšetrovanie týchto vzájomných závislostí bude zložité a náročné avšak potrebné.

Pre úspešné zvládnutie týchto závislostí bude potrebné dopracovať rozbor určenia pravdepodobnosti, že daný údaj je nespoľahlivý v kontexte zmien údajov o iných objektoch a zároveň posúdiť vplyv zmien jednotlivých údajov podľa účelu ku ktorému sa údaje budú využívať.

Určovanie pravdepodobnosti nespoľahlivosti a rizikových indexov by mali mať za následok, že bude možné:

- a) spoľahlivo vyberať objekty u ktorých je potrebné preveriť spoľahlivosť údajov o nich (a prípadne ich opraviť),
- b) spoľahlivo vymedzovať oblasti v ktorých sa bude potrebné zamerať najmä na zisťovanie nových objektov resp. zisťovať zaniknuté objekty,
- c) spoľahlivo určovať dôsledky zistených zmien v údajoch vybraných objektov, vznik nových objektov a zánik objektov na ostatné objekty,
- d) využívať aj neaktualizované vektorové modely územia a pritom zohľadniť nespoľahlivosť údajov o objektoch. V rozhodovacích procesoch, ktoré budú založené na neaktualizovaných vektorových modeloch územia, bude možné efektívne zohľadniť mieru nespoľahlivosti jednotlivých konkrétnych údajov o objektoch.

Z vyššie uvedeného je možné vyvodiť čiastkový záver, že zrejme bude možné optimalizovať proces aktualizácie, bezpečnejšie voliť vhodné metódy vyhľadávania a poznávania zmien objektov. Táto optimalizácia bude dôležitá najmä pri aktualizácii vektorových modelov územia veľkého rozsahu, ktoré zaznamenávajú informácie o obrovskom množstve objektov a vo veľkom rozsahu, u ktorých je spoľahlivosť údajov mimoriadne dôležitá.

5 Možnosti a metódy aktualizácie vektorových modelov územia veľkého rozsahu

Aktualizácia rozsiahlych vektorových modelov územia, akými je napr. Centrálna priestorová databáza Vojenského informačného systému o území Ozbrojených síl SR (CPD VISÚ) resp. Základná báza údajov pre GIS Slovenskej republiky (ZB GIS), vyžaduje vzhľadom k množstvu registrovaných objektov, aby tento proces bol efektívne organizovaný a riadený.

V proces prípravy a plánovania aktualizácie bude potrebné spoľahlivo identifikovať zmeny, odhadnúť ďalšie rizikové údajové množiny a na základe týchto údajov optimalizovať celý aktualizčný proces. Pravdepodobnosti nespoľahlivosti jednotlivých údajov a miery účinnosti aktualizčných metód, ktoré boli spomenuté v predchádzajúcom texte budú práve nástrojom na optimalizáciu celého aktualizčného procesu.

V nasledujúcom texte bude v krátkosti pojednané o niektorých metódach, ktoré zrejme v procese aktualizácie rozsiahlych vektorových modelov územia budú použité. Sú to tieto skupiny metód:

- a) detekcia zmien porovnávaním obrazov,
- b) opakované mapovanie a poznávanie objektov,
- c) preberanie údajov.

5.1 Detekcia zmien porovnávaním obrazu

Tieto metódy sú založené na automatickej, alebo poloautomatickej identifikácii rozdielov medzi obrazmi, ktoré boli zaznamenané v rozdielom čase. Pod pojmom obrazy sa na tomto mieste rozumejú najmä letecké meračské snímky alebo obrazové materiály diaľkového prieskumu Zeme (satelitné snímky). V nasledujúcom texte bude použitý termín snímka, ktorý zahŕňa rôzne druhy vyššie spomenutých obrazových materiálov.

Metódu je možné realizovať v dvoch nasledovných modifikáciách:

- a) porovnávanie dvoch snímok resp. ich ortogonalizovaných verzií, ktoré boli získané v rozdielom čase a obe zobrazujú skúmané územie,
- b) porovnávanie obrazu snímky s vektorovým modelom územia, ktorý bol vytvorený v čase pred zaznamenaním snímky.

Metóda porovnávaním dvoch snímok je použiteľná za predpokladu, že spôsob záznamu snímok a ich geometrické parametre sú prakticky zhodné. Metóda je úspešne použiteľná najmä ak sa jedná o multispektrálne satelitné snímky a javy ktoré sa majú identifikovať sú dostatočne veľké a výrazne sa prejavajú na snímkach. Príkladom je zisťovanie zmien v krajine v oblasti Reno a jazera Tahoe v Nevade, ktoré bolo realizované Geologickou službou USA [2]. Porovnávanie dvoch ortogonalizovaných leteckých meračských snímok naproti tomu nemusí vždy priniesť uspokojivé výsledky. Dôvodom je najmä skutočnosť, že nie je možné zaručiť rovnaké svetelné podmienky pri ktorých sa snímkovanie realizuje a zároveň veľkosť javov, ktoré majú byť identifikované obvykle nie je dostatočne veľká na to, aby významné rozdiely medzi obrazmi nezakryli v rozdieloch zapríčinených inými faktormi.

Identifikáciu zmien väčšieho rozsahu na leteckých meračských snímkach úspešne použil vo svojej štúdií vývoja vysokohorskej krajiny Martin Boltížiar [1].

Metóda porovnávania obrazu snímky zaznamenatej neskôr ako vznikol vektorový model územia s týmto modelom, môže u niektorých tried objektov priniesť očakávané výsledky. Metóda je založená na vyhľadávaní extrémnych hodnôt zložiek spektra jednotlivých pixelov v rámci veľkého súboru pixelov, ktoré by mali reprezentovať objekty jednej triedy. U niektorých objektov jednej triedy sa dá predpokladať, že hodnoty zložiek spektra budú homogénne. Metóda bude založená na štatistickom posúdení súborov pixelov, ktoré budú „prekryté“ entitami vektorového modelu územia patriaceho do jednej triedy.

Obe z vyššie spomenutých metód dokážu rôznou mierou spoľahlivosti určiť, či v skúmanej oblasti došlo, alebo nedošlo k významným zmenám, prípadne určiť ktoré triedy objektov zastúpené vo vektorovom modeli územia sú najviac zmenami dotknuté. Tieto metódy môžu upresniť hodnoty pravdepodobnosti nespoľahlivosti údajov o objektoch avšak nedokážu túto nespoľahlivosť efektívne znižovať. Majú však svoj význam v procese prípravy a plánovania aktualizáčného procesu, kedy môžu spresniť rozsah požadovanej aktualizácie.

5.2 Opakované mapovanie a poznávanie objektov

Pod pojmom opakované mapovanie a poznávanie objektov sa rozumie použitie zhodných alebo veľmi podobných metód získavania údajov o objektoch a poznávania územia aké boli použité v predchádzajúcej etape získavania údajov o objektoch.

Výhodou tohto riešenia je že účinnosť aktualizáčnej metódy je rovnaká ako v predchádzajúcej etape, tzn. zvyšková neistota (parameter M rovnice (1)) bude rovnaký a teda vektorový model územia bude mať stále rovnaké charakteristiky spoľahlivosti. Nevýhodou metódy je, že časová a ekonomická náročnosť opakovaného zisťovania údajov o objektoch bude zrejme vysoká. V prípade vektorových modelov územia veľkého rozsahu akými je napr. ZB GIS je použitie opakovaného mapovania prakticky nerealizovateľné.

Riešením je výber tých tried objektov a oblastí, u ktorých sa predpokladá najväčšie zníženie spoľahlivosti údajov a u týchto tried separátne vykonať overovanie údajov a ich následná aktualizáciu. Ďalším riešením je overenie spoľahlivosti údajov vo vybratej oblasti štatistickými metódami a na základe toho prijať rozhodnutie či vektorový model územia skúmanej oblasti je potrebné aktualizovať a v akom rozsahu.

5.3 Preberanie údajov

Pod pojmom preberanie údajov sa rozumie získavanie údajov o objektoch z takých údajových zdrojov, kde sa údaje o týchto objektoch primárne evidujú. Takýmito zdrojmi sú najmä kataster nehnuteľností a databázy odborných orgánov štátnej správy – databázy správcov.

Táto metóda však vyžaduje aby boli splnené nasledovné predpoklady:

- a) údaje o objektoch vo vektorovom modeli územia je možné spoľahlivo stotožniť s údajmi v databáze správcu z ktorej sa majú údaje prebrať,

- b) vektorový model územia i databáza z ktorej sa majú údaje prevziať musí obsahovať časové údaje na základe ktorých je možné identifikovať či údaje z databázy správcu sú použiteľné pre aktualizáciu vektorového modelu územia,
- c) charakteristiky geometrickej i atribútovej presnosti a vernosti údajov o objektoch v databáze správcu sú zhodné alebo lepšie ako vo vektorovom modeli územia, ktorý sa má aktualizovať,
- d) domény vlastností sú zhodné, alebo navzájom prevoditeľné medzi databázou správcu a vektorovým modelom územia.

Využívanie databáz správcov je z pohľadu efektívnosti aktualizácie a účinnosti aktualizáčnej metódy zrejme najvhodnejšou metódou. Je však potrebné pripomenúť, že údaje ktoré sú súčasťou vektorového modelu územia akým sú ZB GIS či CPD VISÚ nie sú v databázach správcov dostupné v požadovanej štruktúre a kvalite alebo nie sú dostupné vôbec.

6 Záver

V článku sú naznačené niektoré možnosti posudzovania a odhadovania spoľahlivosti údajov o objektoch vo vektorových modeloch územia. Nie len z článku je zrejmé, že údaje vo vektorových modeloch územia starnú bez ohľadu na spôsob vytvorenia týchto modelov. Vektorové modely územia veľkého rozsahu, ktoré pokrývajú celé územie štátu a ktoré majú slúžiť najmä na podporu rozhodovania pri riadení štátu, regiónov a krajiny musia byť spoľahlivé, alebo poskytovať dostatok informácií o spoľahlivosti údajov v nich obsiahnutých.

Posudzovanie spoľahlivosti údajov vo vektorových modeloch územia zrejme nebude samoúčelné a okrem poskytnutia základnej informácie o tejto spoľahlivosti, môže veľmi účinne napomôcť pri hľadaní optimálnych metód a postupov aktualizácie vektorových modelov územia.

Referencie

1. Boltiziar M., Spálenisko pod Slavkovským štítom – zmeny krajinnej štruktúry v rokoch 1949 - 2003 s využitím výsledkov GIS a DPZ, *Geoinformation – elektronický časopis 1/2005*, Univerzita Konštantína Filozofa Fakulta prírodných vied Katedra geografie a regionálneho rozvoja, Nitra (http://www.fpv.ukf.sk/kg/geoinfo/archiv/2005_text_1.pdf)
2. Saylor K., Southwest U.S. Change Detection Images from the EROS Data Center Reno and Lake Tahoe, Nevada, U.S. Geological Survey (<http://geochange.er.usgs.gov/sw/changes/natural/reno-tahoe/>)