

Lekce 10 – Analýzy prostorových dat

1. Cíle lekce	1
2. Základní funkce analýza prostorových dat	1
3. Organizace geografických dat pro analýzy	2
4. Údržba a analýza prostorových dat	2
5. Údržba a analýza popisných dat	3
6. Integrovaná analýza prostorových a popisných dat	3
6.1 Výběry, klasifikace, měření dat	3
6.2 Operace překryvů (overlay)	4
6.3 Operace se sousedními prvky	6
6.4 Incidenční operace	6
6.5 Prostorové statistiky	7
6.6 Prostorové interakce	7

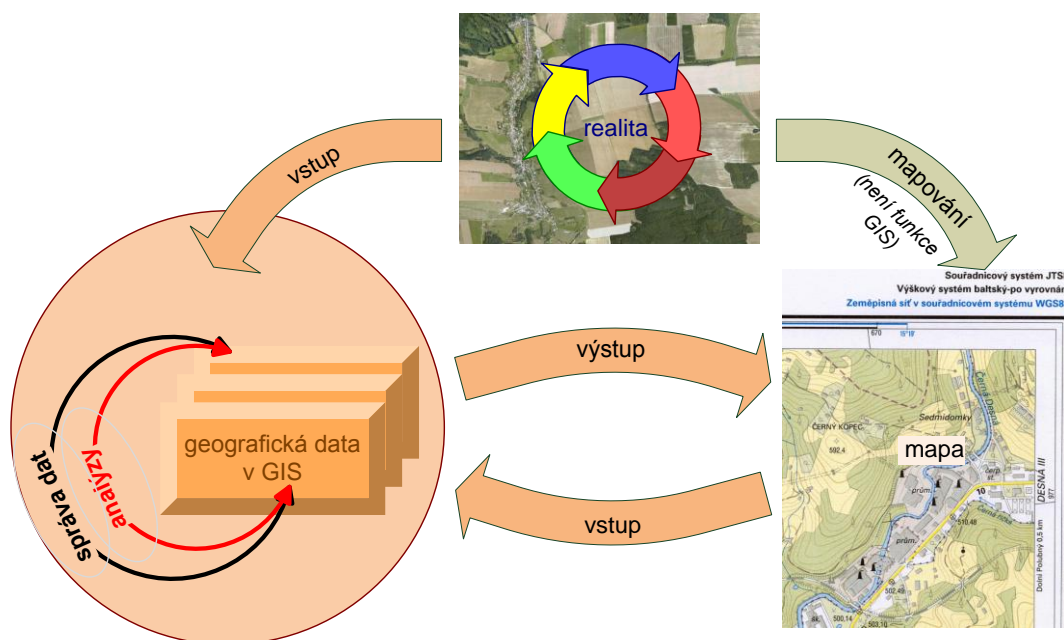
1 Cíle lekce

- uvést přehled analytických funkcí GIS
- ilustrovat některé funkce na příkladech

2 Základní funkce analýza prostorových dat

Mezi základní funkce GIS patří také analytické funkce:

- vstup
- management
- **analýza**
- výstup



Základní funkce GIS - analýza dat

V obecném pojetí se analytickými nazývají všechny funkce, při kterých vznikají nová data (výběrem, výpočtem, apod.).

Analytické funkce dávají odpovědi například na tyto typy dotazů:

- jaké objekty se nacházejí v zadaném území
- kde se nalézají objekty se zadanými vlastnostmi
- co se změnilo území za zvolené časové období
- co se stane, jestliže se změní zadaným způsobem geografická data (modelování)
- jaké jsou vztahy zadaných objektů k okolí

3 Organizace geografických dat pro analýzy

Při provádění analýz je nutné data organizovat tak, jak předpokládá programové vybavení, kterým se analýzy provádějí. Mezi základní stupně organizace dat, které používají analytické programové prostředky, patří:

- vrstvy
- výkresy

Vrstvy

Vrstva geografických dat je množina podobných objektů GIS (třída objektů GIS), která je definována seskupením geografických dat podle konvencí uživatele GIS a je vymezena v datovém modelu GIS. Obsah vrstvy může být závislý na podrobnosti dat, zaměření GIS a také, bohužel, SW možnostech GIS. Příklad: silnice a železnice mohou být v jedné vrstvě nebo v různých vrstvách podle charakteru úloh, které s těmito daty uživatel řeší.

Výkresy

Některé systémy vyžadují plošné rozdělení území na více částí - výkresy (mapové listy, dlaždice - tile). Rozdělení území vyžaduje složitou správu systému výkresů - proto jsou v současné době jednoznačně preferovány systémy založené na principech „bezešvé“ mapy.

Základní členění analytických funkcí

Analytické funkce lze dělit na:

- údržbu a analýzu prostorových dat
- údržbu a analýzu popisných dat
- integrovanou analýzu prostorových a popisných dat

4 Údržba a analýza prostorových dat

Funkce údržby a analýzy prostorových dat jsou ty funkce, které se týkají pouze prostorové složky geografických dat, popisná složka se nemění. K funkcím této kategorie patří:

- transformace formátu prostorových dat (konverze různých typů formátů prostorových dat za účelem přípravy vhodných dat pro další zpracování)
- geometrické transformace (lineární transformace, kvadratická transformace)
- transformace mezi kartografickými systémy (Cassini Soldner - S-JTSK)
- ztotožnění objektů na různých vrstvách systému (nalezení shodných objektů na několika vrstvách - tato situace může být způsobena například několikanásobnou grafickou reprezentací jednoho objektu na vstupu dat)
- ztotožnění objektů na hranicích mapových listů (slouží k propojování objektů, pokud jsou digitalizovány z mapových listů)
- editování grafických objektů (úprava průběhu a polohy prostorových dat)

- zjednodušování průběhu prostorových dat - geometrická generalizace (zmenšení počtu vrcholů lomené čáry, odstranění šumu v naskenovaných rastrových datech)

5 Údržba a analýza popisných dat

Údržba a analýza prostorových dat patří k běžným funkcím všech informačních systémů. Jedná se o tyto funkce:

- editace popisných dat (provádění změn v popisných datech)
- dotazy na popisná data (včetně vytváření nových popisných dat výběrem)

Pokud jsou popisná data uložena v relační databázi, využívají se pro jejich údržbu a analýzu služby RDBMS a jazyka SQL.

6 Integrovaná analýza prostorových a popisných dat

Integrovanou analýzou prostorových a popisných dat se liší GIS od ostatních informačních systémů, které mají implementovány pouze funkce operující nad popisnými daty, a od systémů CAD, které mají implementovány pouze (některé) funkce nad prostorovými daty.

Analytické funkce mohou být realizovány dvojím způsobem:

- funkce je realizována pomocí jedné úlohy (jediným procesem)
- pomocí několika navazujících úloh (procesů, modulů)

Výhodou prvního způsobu je snadnost obsluhy, výhodou druhého je širší spektrum možností, které poskytuje modulární systém. Způsobem implementace analytických funkcí se nadále nebudeme nadále zabývat.

6.1 Výběry, klasifikace, měření dat

V těchto funkcích jsou zpracovávány obě složky geografických dat (prostorová i popisná - fyzikální), mění se pouze popisná složka dat.

Výběry

Výběr dat, jejich zpracování a výstup dat beze změny prostorové složky dat. Pokud jsou data uložena v relační databázi, využívá se pro výběry jazyka SQL.

Klasifikace, generalizace klasifikací

Vytvoření nových vrstev zařazením prvků do tříd pomocí různých klasifikačních metod (jednorozměrné i vícerozměrné metody, faktorová analýza, shluková analýza, apod.)

Generalizovat klasifikací znamená vytvořit ze zadané klasifikace klasifikaci obecnější, méně detailní.

Měření

Měření délky čar (délka vedení potrubí, obvod areálu), měření ploch polygonů (výměra parcely, zastavěná plocha budov), výpočty zastavěných ploch pro daně, měření vzdálenosti mezi objekty, průměru polygonu, apod.

Příklad - výběr hranic vyšších územních celků

Hranice obcí jsou v GIS udržovány jako lomené čáry v datovém modelu areálového grafu. Každý objekt typu lomená čára má v tabulce HRANOB několik atributů, dva z nich jsou kódy obcí (KODOB1, KODOB2), jejichž hranici lomená čára tvoří. Dále existuje tabulka OBCE příslušnosti obcí do okresů ve tvaru kód obce, kód okresu (KODOB, KODOK). Úkolem je vytvořit z hranic obcí hranice vyšších územních celků - okresů.

Příklad - výběr a zobrazení parcel v majetku města

Hranice parcel jsou v městském GIS udržovány podobně jako v předcházejícím případě v datovém modelu areálového grafu. Každý objekt typu lomená čára, který reprezentuje hranici parcely, má v tabulce HRANPARC atributy čísla parcel CISPARG1, CISPARG2. Dále existuje tabulka PARCELY s údaji o parcelách, jedním z údajů je číslo listu vlastnictví. Úkolem je zobrazit parcely, které jsou v majetku města - mají číslo listu vlastnictví rovno 1.

Příklad - zobrazení budov starších 100 let

Budovy jsou v GIS reprezentovány jako uzavřené lomené čáry představující půdorys zastavěné části. Objekty jsou uloženy ve špagetovém datovém modelu. Každý objekt typu lomená čára, který reprezentuje budovu, má v tabulce BUDOVSY atribut ROK_POSTAVENI, který obsahuje rok postavení (datum kolaudace) budovy. Úkolem je zobrazit budovy, které byly postaveny dříve než před sto lety.

Příklad - klasifikace okresů

Hranice okresů jsou v GIS udržovány jako lomené čáry v datovém modelu areálového grafu. Každý objekt typu lomená čára má v tabulce HRANOKR několik atributů, dva z nich jsou kódy okresů KODOK1, KODOK2), jejichž hranici lomená čára tvoří. Dále existuje tabulka OKRESY která má tyto atributy: vzdělanostní struktura, počet nezaměstnaných, věková struktura. Úkolem je zobrazit území s podobnými hodnotami uvedených atributů.

Postup:

- analýza (shluková, faktorová)
- doplnění typu okresu do atributu
- dále jako v předcházejících příkladech

6.2 Operace překryvů (overlay)

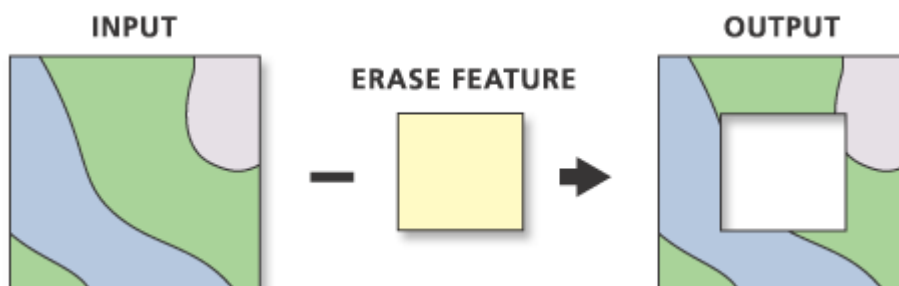
Do této třídy patří funkce, které řeší aritmetické nebo logické operace mezi vrstvami GIS. Systémy, ve kterých jsou implementovány tyto funkce, jsou často nazývány "opravdovými GIS".

- aritmetické operace - operace s prostorově lokalizovanou fyzikální složkou geografických dat vymezenou hodnotami pixelů v rastrovém tvaru nebo hranicemi polygonů ve vektorovém tvaru
- logické operace - vymezení území, kde jsou splněna nějaká kritéria

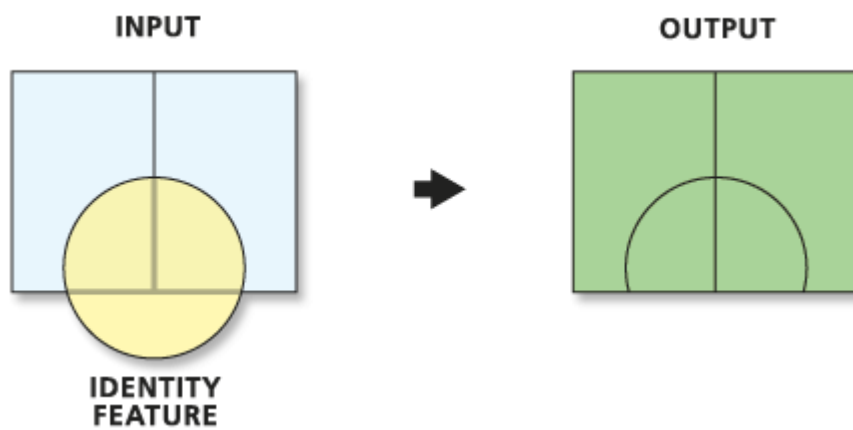
Na způsob provádění operace překryvů má značný vliv typ základního datového modelu (raster x vektor), z důvodu snadnějšího provádění překryvů vznikly raster-vektorové GIS.

Typy operací překryvů (převzato z <http://webhelp.esri.com/arcgisdesktop/9.3>)

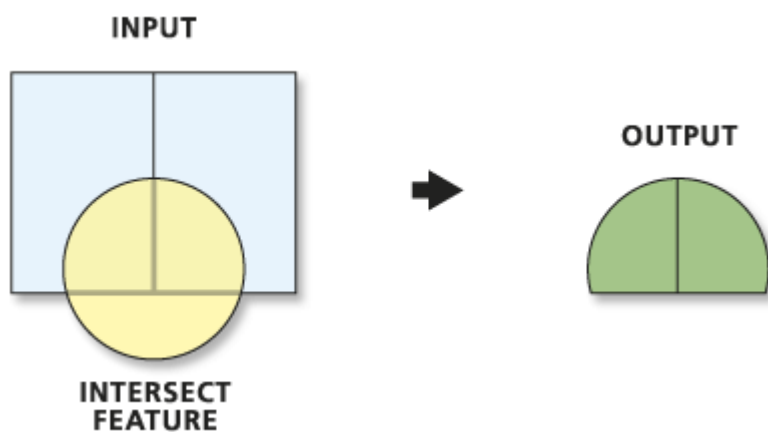
a) výmaz



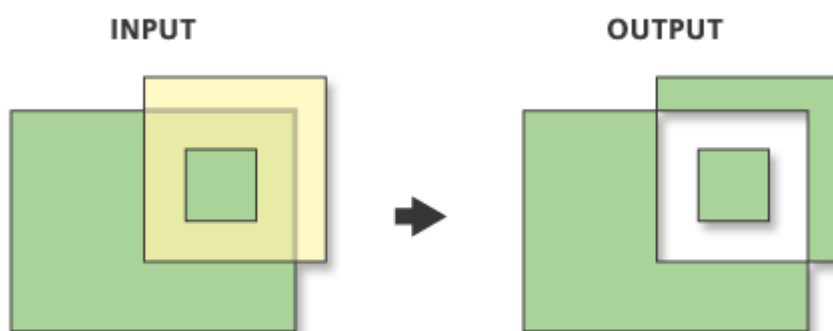
b) Identifikace



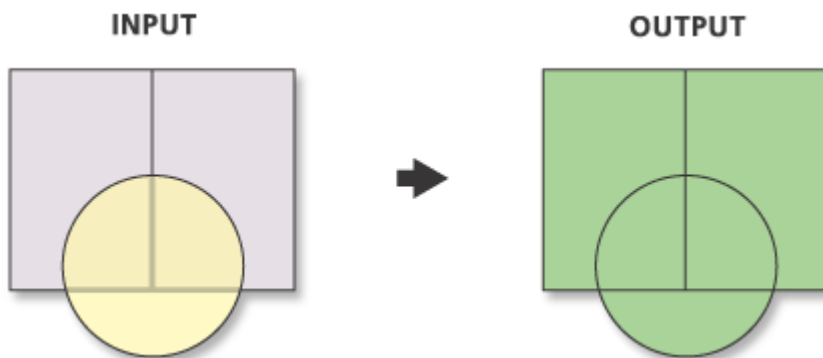
c) Průnik



d) Symetrická diference



e) Sjednocení



Příklad - výpočet měsíčního množství srážek

V jednotlivých vrstvách GIS je uchováno množství denních srážek. Vypočítejte množství srážek za měsíc. Předpokládejte:

- data jsou uložena v rastrovém modelu ve vrstvách 1 - 31 podle dnů v měsíci, v každé vrstvě je v každém pixelu uchováno denní množství srážek
- data o srážkách jsou za každý den uložena ve vektorovém tvaru v areálovém grafu, hrany grafu tvoří hranice oblastí, které jsou charakterizovány shodným množstvím denních srážek

Řešení prvního případu spočívá v prostém součtu hodnot pixelů: vytvoří se nová prázdná vrstva, a do každého jejího pixelu se sečtou hodnoty stejně položených pixelů ve vrstvách 1 až 31.

Pro řešení případu (b) je nutné zvolit komplikovanější postup, například:

- výpočet překryvů (průniků) všech 31 grafů, u vznikajících oblastí je nutné uchovat, ze kterých oblastí vznikly
- výpočet měsíčního množství srážek nových oblastí podle množství srážek ve zdrojových oblastech
- sloučení nových oblastí se shodným měsíčním množstvím srážek

6.3 Operace se sousedními prvky

Operace se sousedními prvky vyhodnocují charakteristiky území v okolí zadaného prvku. Operace se sousedními prvky je zadána

- specifikací prvku, v jehož okolí bude operace probíhat
- specifikace okolí
- specifikace výběrové funkce (jakým způsobem budou prvky v okolí vybírány)

Příklady:

- průměrná cena budov v okolí hasičské stanice
- seznam nemocnic v okolí místa dopravní nehody, apod.

Operace se sousedními prvky využívají operace s prostorovými daty: bod v polygonu, čára v polygonu, atd., které byly specifikovány v kapitole 2.

Speciálními operacemi této třídy jsou *interpolace*, což je doplnění hodnot pixelů podle hodnot sousedních pixelů v rastrovém modelu, a *generování linií* z bodů, které linie reprezentují (vrstevnice, hranice areálů, průběh inženýrské sítě z naměřených bodů na potrubí, apod.).

6.4 Incidenční operace

Operace této třídy zpracovávají prvky, které spolu incidují. Při zpracování jsou vytvářeny nové prvky. Mezi tyto operace patří:

- operace vycházející ze spojitosti prvků (generování souvislých prvků se shodnými vlastnostmi - řetězení čar, seskupování pixelů, apod.)
- operace vycházející z blízkosti prvků (bufferování - generování zón okolo bodů, linií, ploch)

- operace aplikované na síťové modely dat (optimalizace cesty, umístění zdrojů do sítě, apod.)

Příklad - GIS správce kanalizační sítě

Kanalizační síť tvoří orientovaný síťový graf. Každý objekt typu šachta je uzlem sítě a jeho vlastnosti jsou popsány v tabulce SACHTA, která mimo jiné atributy obsahuje také primární klíč CISLO_SACHTY. Dále každý objekt typu úsek kanalizační stoky je hranou sítě a jeho vlastnosti jsou popsány v tabulce STOKA, která obsahuje mimo jiné atributy CISLO_VSTUPNI_SACHTY (číslo šachty, ze které úsek stoky teče) a CISLO_VYSTUPNI_SACHTY, (číslo šachty, do které úsek stoky ústí). Dále víme, že mezi dvěma šachtami vede nanejvýš jeden úsek stoky a každý úsek stoky začíná a končí v nějaké šachtě.

Úkolem je určit všechny šachty, do kterých teče kanalizace ze zadané šachty CISLO_SACHTY.

6.5 Prostorové statistiky

Distribuční funkce a hustota dvou (nebo více) náhodných proměnných, které mohou reprezentovat spojité nebo diskrétní jevy v E2 (nebo ve vícerozměrném prostoru). Tvar prostorových statistik lze vizualizovat pomocí isolinií hustoty, tečkových diagramů a dalších metod tématické kartografie.

6.6 Prostorové interakce

Newtonův gravitační model ($F = m(1) * m(2) / (d * d)$). Slouží k odhadům frekvence dopravy, množství toku informací, potenciálu ekonomického růstu.

Připomínky a dotazy k obsahu lekce posílejte, prosím, na adresu:

Rudolf Richter, richter@fi.muni.cz