

## **Analýza v GIS**

Prostorové analytické možnosti GIS tvoří jádro systému GIS, to GIS odlišuje od ostatních informačních systémů. Mezi otázky, na které nám GIS umožňuje odpovědět patří:

- Co se nachází na? Kde se nachází? Jaký je počet, které objekty spoňují? - statistické otázky.
- Co se změnilo od? Co je příčinou? Co když?

Analytické možnosti GIS můžeme rozdělit do následujících skupin:

- měřicí funkce,
- nástroje na prohledávání databáze (atributové i prostorové),
- topologické překrytí,
- mapová algebra,
- vzdálenostní analýzy,
- analýzy sítí,
- analýzy modelu reliéfu a dalších povrchů,
- statistické analýzy,
- analýzy obrazů, a další

### **Měřicí funkce**

GIS poskytují funkce na měření vzdáleností a ploch.

Při měření je třeba dát pozor na zobrazení:

- konformní - nedochází ke zkreslení úhlů,
- ekvivalentní - nedochází ke zkreslení ploch,
- ekvidistantní v určité soustavě křivek - nezkresluje délky ve směru určité soustavy křivek (většinou ve směru poledníků, rovnoběžek, případně jiných křivek),
- kompenzační - dochází k deformaci všech geometrických prvků (úhlů, délek i ploch), ale hodnoty deformací nejsou extrémní.

Poznámka: toto je důležité pro měření na mapách světa, kontinentů. Kartografická zobrazení ještě budeme řešit v teorii. Velká měřítká map, malý rozsah území – zkreslení zanedbatelné, lze měřit.

Poznámka: systémy umí velice efektivně používat různých délkových jednotek (stopy, cm, ...), případně mezi nimi i automaticky provádět převody.

Poznámka: při projektech v malém měřítku, a tedy většího plošného obsahu, má na měření vliv také zakřivení zemského povrchu, takže v poslední době umožňují GIS produkty uvažovat i tento faktor.

### **Dotazy na geografickou databázi**

Dotazováním se vybírají údaje, které odpovídají specifickému kritériu nebo podmínce.

Komponenty operace dotazu:

1. Specifikace údajů, kterých se týká.
2. Formulace podmínek, kterým musí údaje vyhovovat.
3. Instrukci, co se má na vybraných údajích vykonat.

Struktura dotazu: vyber z údajů typu T takové, které vyhovují podmínce P a vykonej na nich operaci O.

Dělení dotazů:

- Atributové - dotaz typu: "které geografické objekty (lokality) mají definovanou vlastnost".  
Například: "Vyber všechna města v ČR, která mají více jak 10 000 obyvatel"
- Prostorové - dotaz typu: "co se nachází na tomto místě, co se nachází v této oblasti".  
Například: "Vyber všechna města v ČR, která leží v Plzeňském kraji"
- Kombinované - dotaz typu: "které objekty splňují definovanou vlastnost a zároveň se nacházejí v nějaké oblasti"  
Například: "Vyber všechna města v ČR, která mají více jak 10 000 obyvatel a zároveň leží v Plzeňském kraji"

### **Atributové dotazy**

Atributové dotazy lze uskutečnit různými způsoby. První způsob spočívá v identifikaci jednotlivého

objektu na základě jeho jména, označení či jiného atributu.

Častěji se ale používá druhý způsob, který spočívá ve vyhledání všech objektů splňující intervalové a logické podmínky jednoho nebo více atributů.

Příklad atributového dotazu: vrstva silnice, máme dva atributy typ (I,II,D,o) a počet pruhů (1,2).

Zkusíme atributový dotaz, kde hledáme všechny silnice, které jsou typu I a mají dva jízdní pruhy: pomocí dotazovacího prostředí "Query By Example" (QBE):

vyplněný dialog ( [Trida\_sil] = "1") and ([J\_pruhy] = 2 )

pomocí SQL:

```
SELECT * FROM SILNICE WHERE TRIDA_SIL="1" and J_PRUHY=2
```

pro vyhledávání intervalových podmínek je možné použít operátorů <, >, =, <=, >=, <> (≠).

Intervalové podmínky jdou dále kombinovat pomocí logických operátorů (AND, OR, NOT) využívajících pravidel Booleovské logiky.

K vlastnímu definování podmínky se nejčastěji používá jazyk SQL nebo prostředí Query By Example.

Rozdíly vektorové a rastrové reprezentace:

- U vektorové reprezentace se zpracovávají údaje atributových tabulek připojených k jednotlivým vektorovým objektům.
- U rastrové se zpravidla zpracovávají údaje uložené v buňkách jednotlivých vrstev (není to ale podmínka, i u rastrových reprezentací je možné mít připojené atributové tabulky).

### **Prostorové dotazy**

Prostorové dotazy jdou uskutečnit opět dvěma základními způsoby.

První je jednoduchý a spočívá v identifikaci geografického objektu na základě jeho souřadnic, a to buď ručně (zadáním souřadnic) nebo interaktivně (ukázáním na objekt myší), viz. obrázek .

Druhý způsob spočívá v prohledávání prostoru různých geometrických tvarů (obdélníky, kružnice, polygony, linie).

Rozdíly rastrové a vektorové reprezentace:

- U vektorů je vždy vybrán celý objekt, u rastrů je vybírána vždy konkrétní buňka či skupina buněk.

### **Kombinované dotazy**

Kombinované dotazy umožňují práci i s více vrstvami (či množinami objektů), je zde opět možnost propojovat je pomocí operátorů Booleovské logiky, podobně jako u atributových dotazů.

Kombinované dotazy také zčásti mohou využívat topologické překrývání vrstev, viz níže.

Příklad:

Najdi všechna sídla, která jsou ve vzdálenosti do 5 km od dálnice.

Je vidět, že nejdříve musela být provedena analýza atributová (vyber dálnice) a poté prostorová (najdi sídla ve vzdálenosti do 5 km od této dálnice).

Lze vytvářet složitější dotazy - najdi sídla, která jsou ve vzdálenosti do 5 km od dálnice, mají méně než 50000 obyvatel a protíná je železniční trať, ...

### **Topologické překrytí**

Obecně dotazování dvou nebo více informačních vrstev se označuje jako topologické překrytí (overlay) těchto vrstev.

Klasicky se tento problém řešil překrytím dvou tématických map na průhledných fóliích. Stejný problém se v GIS řeší pomocí základních algoritmů počítačové grafiky (test bodu v polygonu, hledání průsečíku dvou objektů, ořezávání). Výsledkem postupu je pak identifikace nových objektů, které mají kombinace vlastností objektů ze zdrojových informačních vrstev.

Pro kombinaci vstupních objektů se opět používají pravidla Booleovské logiky. Systémy obvykle nabízejí:

- INTERSECT (AND - průnik),
- UNION (OR - sjednocení),
- IDENTITY (AND NOT - přiřazení na základě prostorového umístění).

Při těchto operacích dochází k řešení vztahů bod, linie nebo polygon v polygonu (výjimkou je sjednocení, které mohou provádět pouze mezi dvěma polygonovými vrstvami). Názorné obrázky naleznete v helpu v ArcGIS.

Z procesu topologického překrytí vznikají nové objekty (vrstvy), kterým jsou přiřazeny také atributy. Tím se topologické překrytí liší od prostorových dotazů, kde žádné nové vrstvy nevznikají. Vstupní vrstva může být jak bodová, liniová, tak polygonová, druhá vrstva (ta co provádí topologické překrytí) musí být vždy polygonová.

Speciálními případy topologických operací jsou: CLIP a ERASE, UPDATE a SPLIT, atributy nejsou spojovány, ale přejímány ze vstupní vrstvy (jedna vrstva je vždy vstupní a druhá na ní provádí výše uvedené operace). Tyto funkce je možné zařadit i do kategorie restrukturalizace dat.

- CLIP - ořízne vstupní vrstvu pomocí definovaných polygonů v druhé vrstvě.
- ERASE - opak CLIP odstraní části vstupní vrstvy pomocí polygonů definovaných v druhé vrstvě.
- SPLIT - rozdělí vstupní vrstvu na části pomocí hranic definovaných polygony ve druhé vrstvě.
- UPDATE - vyjme tu část vstupní vrstvy, která bude aktualizovaná druhou vrstvou a místo ní vloží prvky z druhé vrstvy.

Do topologických operací je možné zařadit i úlohy typu DISSOLVE ("rozpuštění" hranic objektů na základě stejného atributu) MERGE (spojení dvou vrstev do jedné a odstranění hranic mezi objekty se stejnými atributy), ty ale spíše patří do kategorie restrukturalizace dat.

Příklady:

*Intersect:* Zkusím najít všechny úseky železničních tratí, které procházejí lesem.

Pokus o nalezení všech úseků železnic procházejících lesem pomocí prostorového dotazu – nedá požadovaný výsledek, výsledkem jsou všechny tratě jako celky, u nichž alespoň část prochází lesem. Pomocí topologického překrytí (INTERSECT tratě a lesy) dostaneme výsledek, kde jako atributy jsou přiřazeny ke každému úseku jak atributy tratě (číslo tratě, ...), tak i lesa (typ lesa, stáří lesa, ...), jak byly obsaženy ve vstupních vrstvách.

Další příklad (intersect polygon - polygon) může být hledání všech částí parcel, na kterých je přítomen les.

*Union:*

(Union polygon - polygon): budu hledat vlastníky zalesněných ploch a kolik procent z celkové výměry jejich parcely tato zalesněná plocha činí. Vezmu parcely (polygonovou vrstvu) a lesy (polygon). Po provedení sjednocení dostanu vrstvu, kde budou jak parcely s lesem, tak i část bez lesa, ze kterých je pak snadné udělat požadovanou statistiku.

*Identity:*

chci rozdělit všechny železniční tratě na probíhající lesem a neprobíhající lesem. Tak udělám identity železnic s lesy. Výsledkem je vrstva, kde budu mít jak úseky probíhající s lesem, tak i úseky neprobíhající lesem s tím, že atributy úseků v lese budou opět obsahovat typ lesa (tak jako u průniku) a úseky mimo les nebudou mít tyto atributy naplněné.

Rozdíl oproti průniku je jasný. Rozdíl oproti sjednocení je takový, že ve výstupní vrstvě jsou pouze objekty vstupní vrstvy s přiřazenými atributy podle polygonů vrstvy, která prováděla identifikace.

Poznámka: u rastrů je analýza překrytí obdobná, říkáme jí mapová algebra

## Mapová algebra

U rastrových reprezentací se místo topologického překrytí používá nástroj zvaný mapová algebra. Ten je určen výhradně pro ně a umožňuje kombinovat rastrové vrstvy pomocí různých

matematických operací.

Tyto matematické operace se vykonávají buď na jedné nebo na dvou (i více) vrstvách a jejich výstupem je vždy nová vrstva, kterou je samozřejmě možné používat v dalších analýzách. To vytváří z mapové algebry mocný prostředek pro prostorové modelování a analyzování. více lze najít v hlepu k nástrojům, které funkce mapové algebry umožňují (arcINFO, GRASS,...) nebo v literatuře.

### Vzdálenostní analýzy

Na první pohled jednoduchá problematika, která se však ve spolupráci s mapovou algebrou (pro rastrovou reprezentaci), či analýzami sítí (pro vektorovou reprezentaci) může rozvinout ve velice složitý a sofistikovaný nástroj.

Bezesporu nejpoužívanějším nástrojem vzdálenostních analýz je **tvorba obálky (bufferu)**, což spočívá ve vektorové reprezentaci ve vytvoření polygonů v určené vzdálenosti kolem bodů, linií či polygonů.

Vytvořené polygony jsou uloženy jako standardní vrstva s definovanou topologií, tudíž je možné je používat v dalších analýzách topologického překrytí.

Příklad: vzdálenostní analýzy kombinované s topologickým překrytím:

Zjistí plochu lesů, které jsou do 3 km od dálnice.

Vstupní vrstvy: silnice, lesy.

1. krok: tvorba bufferu kolem dálnice ve vzdálenosti 3km ,

2. krok: průnik lesů s vytvořenou obálkou. Výsledkem je vrstva lesů do 3km od dálnice, nad kterou je již snadné provést dotaz na jejich celkovou plochu.

Z výš uvedeného vyplývá, jakým způsobem jsou prováděny vzdálenostní analýzy ve vektorovém datovém modelu.

U rastrové reprezentace je tvorba obálky opět jen otázka metriky. Všechny buňky, které jsou od daného objektu v menší vzdálenosti než definovaná budou označeny.

Velice často se ale v rastrové datové struktuře dělá analýza okolí tak, že se spočítá vzdálenost každé buňky od požadovaného objektu a ta vzdálenost se uloží do nové vrstvy, kterou je možné dále reklasifikovat (např. buňky s hodnotou menší než 3km - hodnota 1, jinak hodnota "NO\_DATA").

Tato analýza se často nazývá nalezení vzdálenosti (Find Distance).

### Analýzy sousedství (Proximity analysis)

Mezi vzdálenostní analýzy také patří takzvané analýzy sousedství (proximity analysis).

Spočívají v tvorbě "individuální plochy" kolem každého ze vstupních bodů, které definují příslušnost dané lokality k nejbližším z objektů.

Pro vlastní výpočet se používá metody **Thiessenových polygonů** nebo Voronoi diagramy, což jsou duální funkce k triangulaci. Funkci proximity analysis je možné provádět jak v rastrové, tak ve vektorové podobě.

Jako příklad analýzy sousedství uveďme vyhledání prostorové příslušnosti k jednotlivým železničním zastávkám.

Další vzdálenostní analýzou je tzv. **vážená vzdálenost** (weighted distance).

Vážená vzdálenost si všímá jedné podstatné vlastnosti, a to, že při běžných vzdálenostních analýzách se vůbec neuvažují vlivy okolí, vše je měřeno vzdušnou čarou za ideálních podmínek. V reálném světě ale tento model zdaleka neodpovídá skutečnosti.

Reálná vzdálenost však neodpovídá vzdálenosti "vzdušnou čarou - má na ni vliv převýšení terénu, do kopce se jde hůře než z kopce, když proti mně bude foukat vítr, tak také spotřebuji více pohonných hmot než po větru a další. Tyto faktory lze do analýzy zahrnout právě pomocí vážené vzdálenosti.

Více informací lze najít v literatuře, problematika není triviální.

### Analýzy nad vektorovou sítí

Další významnou oblastí jsou analýzy sítí. V podstatě se jedná opět o hledání nejkratší vzdálenosti, ale s tím rozdílem, že sítě jsou vektorovou reprezentací.

Síť tvoří (orientovaný) ohodnocený graf, skládající se z uzlů (průsečíků) a hran (linií).

Postup tvorby sítě:

- Je třeba získat liniovou vrstvu, nad kterou budou analýzy prováděny (ulice, rozvody, kanalizace).
- Tato data musí být topologicky čistá (hlavně musí splňovat konektivitu a znalost směru).
- Následně lze síti přiřadit pravidla, která určují, jak je možné se pohybovat mezi jednotlivými uzly:

Vlastní analýzy

- Hledání konektivity - hledání všech propojených prvků s daným uzlem.  
Příklad: nalezení všech vedení a odběratelů postižených vyhořeným transformátorem.
- Modelování zatížení sítě - analýza transportu vody/splavenin ve vodních tocích, pohyb plynu v potrubích (na základě objemu, průřezu, sklonu, tlaku).  
Příklad: připojím nové odběratele k plynovému potrubí a po analýze zjistím, jaký tlak budou mít na přípojce, o kolik se tlak sníží původním odběratelům.  
Nebo v případě energetických rozvodů může vyhořet transformátor a je nutné provést nouzové propojení přes jiný transformátor. Nové propojení však musí být dimenzováno podle odběratelů.  
Hledání optimální trasy - jde o vyhledání optimální trasy mezi dvěma nebo více body (ve stanoveném pořadí nebo bez) na základě ceny cesty (vzdálenost, čas, ...). Analýza umí produkovat i pokyny o cestě pro řidiče
- Hledání cesty do nejbližšího zařízení - drobná modifikace předchozí analýzy. Jde o vyhledání optimální trasy do nejbližšího (optimálního) zařízení.  
Příklad: Hromadná dopravní nehoda ve velkém městě. Jde o to, nalézt co nejrychlejší způsob, jak se k nehodě dostat sanitkou. Řešení je nalezení optimální cesty od optimálního zařízení k nehodě.  
Tím to ale nekončí, jelikož je možné ještě hledat optimální cestu od nehody do nejbližší nemocnice. Tyto cesty totiž vzhledem ke konfiguraci sítě (jednosměrky) či vzhledem k času (ucpané ulice v určitém v důsledku nehody) nemusí být stejné.
- Alokace zdrojů - další možnost aplikace analýzy sítí. Vyhledání všech lokalit, které jsou od vybraného objektu vzdáleny nějakou cenou cesty.  
Příklad: vzdálenost do 30 minut od vyhlášené restaurace. Jak je vidět, je to analýza podobná vytváření obálek (buffers), ale bere v úvahu cenu cesty definovanou pomocí sítě (není to jen vzdálenost vzdušnou čarou). Výsledkem této analýzy jsou tzv. izochrony, což jsou čáry spojující body se stejným časem k dosažení výchozího bodu.