Z P R A C O V Á N Í D A T O L O K A L I Z O V A N Ý C H O B J E K T E C H

Jan S o k o l

Referát pojednává o strojovém zpracování dat, která jako jeden z podstatných atributů obsahují atribut lokalizace, tj. umístění ve vícerozměrném metrickém prostoru. Uvádí různé typy lokalizovaných objektů a podrobně si všímá základních algoritmů ukládání a výběru lokalizovaných dat. Na dvou příkladech je předvedeno zpracování bodových a liniových objektů v datové bázi.

1. Lokalizované objekty

Při strojovém zpracování dat obvykle předpokládáme, že zpracovávaná data se vztahují k jistým reálným objektům a že zachycují jejich důležité vlastnosti - atributy. Každý objekt je identifikován nějakým jménem či klíčem, který bývá také hlavním a nejčastějším výběrovým kritériem. Ve skladové evidenci to může být číslo materiálu, v personalistice jméno nebo osobní číslo zaměstnance. Vedle požadavku na výběr podle hlavního klíče se ovšem často setkáváme i s požadavky na výběr podle hodnot různých atributů, například podle množství ve skladu nebo podle výše platu. Výběrovým kritériem může přitom být jedna jediná hodnota určitého atributu, například plat 2800 Kčs nebo fakt, že se jedná o ženu. U vícehodnotových /tj. ne boolcovských/ atributů s více méně spojitou oblastí hodnot to často bývá interval hodnot, například plat v rozmezí 2600 až 3000 Kčs. Těmto problémům se v posledních létech věnovala značná pozornost v souvislosti s teorií datových bází a dotazových jazyků, která nakonec vedla k vytvoření relačního modelu dat jako zatím nejobecnějšího konceptu pro ukládání a výběr dat. Nicméně i v tomto konceptu se zpravidla předpokládá, že hodnoty všech atributů jsou skaláry, tj. boolcovské hodnoty, čísla nebo hodnoty na čísla převeditelné a tedy jednoznačně a lineárně uspořádatelné. Této vlastnosti klíčů a atributů se v technice hromadného zpracování dat hojně využívá, ať už při třídění datových souborů, při budování indexů či při vyhledávání a spojování v relačních tabulkách.

Lokalizovanými objekty míníme objekty, které mají z povahy věci mezi svými atributy atribut lokalizace, tj. umístění či souřadnic ve dvou- nebo vícerozměrném metrickém prostoru. Tento atribut, který bývá fakticky nejčastějším výběrovým kritériem, je tedy vektor /uspořádaná n-tice čísel/ a vykazuje výrazné topologické vlastnosti. Jednou z nejdůležitějších je vzdálenost, definovaná v obvyklé eukleidovské metrice, a na ní založený pojem okolí bodu. Také interval tu bude přirozeně definován jako interval vektorový, tj. uspořádaná n-tice číselných intervalů, která v příslušném prostoru vymezuje pravoúhlý rovnoběžník se stranami či stěnami rovnoběžnými s osami souřadnic. Při výběru či vyhledávání podle atributu lokalizace je intervalové zadání výběrového kritéria pravidlem, výběr podle jedné určité hodnoty se /mimo jiné vzhledem k neurčitosti zaměření bodů/ prakticky nevyskytuje. Základním problémem při zpracování dat o lokalizovaných objektech je tedy intervalový výběr podle vícehodnotového /vektorového/ atributu umístění. Protože hodnoty složek tohoto atributu, tj. hodnoty souřadnic, jsou více méně spojité, nelze tento výběr převést na obvyklé jednoznačné lineární uspořádání a nelze tedy použít obvyklých vyhledávacích algoritmů. Hodnoty souřadnic se sice v praxi vždy nějak zaokrouhlují a jsou pak přísně vzato diskrétní, nicméně v poměru k velikosti zpracovávaného univerza je musíme považovat za /téměř/ spojité.

Lokalizované objekty většinou souvisejí s nějakým dvou- nebo vícerozměrným zobrazením reality, například s mapou nebo plánem, kde právě zmíněné pojmy vzdálenosti a okolí hrají důležitou roli: mapa i plán je zobrazení, které zachovává vzdálenosti a okolí a v tom spočívá její hlavní smysl. Mezi lokalizované objekty v našem smyslu nebudeme tedy počítat zaměstnance, i když jeho bydliště je "lokalizováno" adresou nebo směrovacím číslem, ale zpravidla ani zboží ve skladu, i když jeho uložení může být pro automatizovaný sklad zapsáno v podobě vektoru souřadnic, protože s tímto atributem nebudeme pracovat jako se základním výběrovým kritériem. Typickými lokalizovanými objekty budou naproti tomu geodetická data jako výškopis, body trigonometrické sítě, parcely a podobně. Při zaměřování, zobrazování i zpracování lokalizovaných objektů vždy předpokládáme, že objekt je definován jedním nebo několika body, které jsou vlastními nositeli atributu lokalizace. Z hlediska topologic-

kých vlastností objektů samých můžeme lokalizované objekty rozdělit na bodové /např. triangulační body, geologické vrty/, liniové a grafové /např. dopravní cesty, podzemní či nadzemní vedení/ a plošné /např. parcely nebo katastry/. U bodových objektů je veškerá informace vztažena přímo k bodům, takže objekt a bod často splývá. U liniových a grafových /tj. větvených/ objektů k tomu přistupují informace o spojnicích mezi těmito body, o jejich pořadí a orientaci v rámci objektu a konečně u objektů plošných informace o plochách, omezených body a jejich spojnicemi.

V krajním případě může být atribut lokalizace jediným atributem bodu, případně objektu. Zpravidla jsou však lokalizované objekty všech typů nositeli dalších atributů, které sice nemají povahu lokalizace, ale které se většinou váží ke zcela určitým složkám lokalizovaných objektů: některé k objektu jako celku, jiné k určitým plochám, liniím a bodům. Obecně tedy můžeme atributy lokalizovaných objektů rozdělit do dvou velkých skupin: atributy topologické čili kresebné a zbývající atributy popisné, které se na mapě či kresbě neobjeví. Obecná struktura lokalizovaného objektu se tedy skládá z topologické složky, tvořené objekty, plochami, čarami a body, přičemž každý prvek topologické složky může být nositelem dalších atributů, které patří do popisné složky.

Typické aplikace strojového zpracování dat o lokalizovaných objektech se vyznačují následujícími charakteristickými vlastnostmi:
- velký počet objektů
- převaha vyhledávání podle lokalizace /umístění/
- poměrně jednoduché a stálé vazby mezi prvky topologické složky /plochami, čarami a body/
- složitá vnitřní struktura popisné složky vzhledem k rozmanitosti popisovaných objektů
- těsná návaznost na metody a prostředky grafického zpracování /např. digitalizace a kresba/.

V dalším textu si všimneme některých obecně zajímavých problémů sběru, vyhledávání a aktualizace dat o lokalizovaných objektech různých typů, popíšeme základní algoritmy a zmíníme se o příkladech implementací. Problematika vlastního grafického zpracování, byť velice zajímavá, přesahuje už rámec tohoto referátu.

2. Bodové objekty

Nejjednodušším typem lokalizovaných objektů jsou objekty bodové: každý objekt je reprezentován jediným bodem, k němuž se váže veškerá informace o objektu. Lze si představit aplikace, kde by to byla pouze informace topologická, v praxi je však většinou provázena ještě více či méně složitou informací popisného charakteru.

Dotazy či požadavky na výběr dat se zpravidla týkají nepatrného zlomku uložených dat a výběrovým kritériem je právě umístění, zadané intervaly souřadnic. Základním problémem je tedy takové uložení dat, které by umožnilo účinný intervalový výběr podle souřadnic bez prohledávání celé datové báze. I když údaje souřadnic nejsou úplně spojité - rastr zaměření bývá u geodetických aplikací asi 10 cm, u stavebních aplikací kolem 1 cm - nelze použít prosté setřídění podle jedné souřadnice nebo podle součtu souřadnic; takové uspořádání nezachovává okolí. Pro ukládání a výběr se často používá tzv. algoritmus čtverců v různých modifikacích. Celá oblast, zobrazovaná v datové bázi /"univerzum" datové báze/ se rozdělí na pravidelné rovnoběžníky tak, že se v každé souřadnici zvolí jistý krok. Je-li rozměr univerza v souřadnici x roven m-násobku kroku x a rozměr v souřadnici y roven n-násobku kroku y, vznikne síť o m krát n stejných obdélnících, které jsou pak organizační jednotkou /např. vlastnickým segmentem/ pro ukládání a výběr. Je-li jako výběrové kritérium zadán interval $/x_1,y_1 - x_2,y_2/$, určí se snadno indexy čtverců, v nichž mohou bodové objekty ze zadaného intervalu ležet. V rámci takto vybraných čtverců /primárního výběru/ se pak další vyhledávání provádí sekvenčně /sekundární výběr/. Síť, provádějící rozdělení na čtverce, může být zvolena libovolně, ovšem za předpokladu, že není ani příliš hustá /většina čtverců je prázdná/ ani příliš řídká /sekundární výběr se zpravidla provádí na zbytečně mnoha datech/. Rozumná velikost čtverce by se měla blížit velikosti typického výběrového intervalu, který lze v rámci určité aplikace dobře odhadnout.

Z praktických důvodů sběru a identifikace bodů a jejich souřadnic je výhodné, je-li příslušnost bodu k základnímu čtverci zahrnuta už v jeho identifikátoru, který můžeme v jistých mezích

zvolit rovněž libovolně. Zaměřování se běžně dělá po mapových listech např. měřítka 1:1000 a čísla bodů se přiřazují celkem nahodile rovněž v rámci těchto listů. Ztotožníme-li síť čtverců s těmito mapovými listy a zavedeme-li jako identifikátor bodu číslo složené z čísla /indexu/ mapového listu a vlastního čísla bodu na listu, zjednoduší se podstatně ukládání dat. Objekty a body nejsou na sledovaném území rozloženy rovnoměrně, často se vyskytují ve shlucích a jinde vznikají prázdné oblasti. Proto je výhodné základní síť čtverců podle potřeby zjemňovat. Pro toto zjemňování lze navrhnout různé algoritmy, nejlépe se osvědčilo prosté půlení v okamžiku, kdy je určitý čtverec "hodně" naplněn. Složitější algoritmy, které volí dělicí čáry s ohledem na rozložení bodů, jsou sice o něco účinnější, ke čtvercům se však musí ukládat více pomocných informací. Ještě vážnější je však skutečnost, že při plnění datové báze se z povahy věci postupuje obvykle systematicky, např. zleva doprava a shora dolů. Čtverce se tedy plní nerovnoměrně a dělení, které se v určitém okamžiku zdá být optimální, se dalšími daty zase vyvede z rovnováhy. Metoda prostého půlení je naproti tomu stabilní a jednoduchá jak při ukládání, tak při výběru. Problém bodů ležících na dělicí čáře lze obejít buď tak, že dělicí čáry leží mimo zaokrouhlovací rastr na souřadnicích, které se nemohou vyskytnout, nebo tak, že se např. levý a dolní okraj přiřadí k mapovému listu a pravý a horní z něho vyloučí.

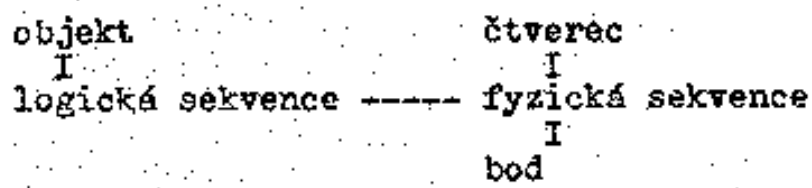
Jako příklad datové báze s bodovými objekty uvedeme evidenci stavebních geologických vrtů na území města Prahy. Báze obsahuje řádově desetitisíce jednobodových objektů nerovnoměrně rozložených po celém území města. Topologická složka objektu je velice jednoduchá /souřadnice a číslo bodu/. Popisná složka je složitě hierarchicky členěna a obsahuje jak celkové charakteristiky vrtu /hloubka, úroveň spodní vody, datum, firma atd./, tak geologické, statické a další údaje o jednotlivých vrstvách v místě vrtu a konečně o jednotlivých vzorcích, odebraných v rámci určité vrstvy. Výběrovým kritériem primárního výběru je buď identifikační číslo vrtu /bodu/ nebo interval souřadnic. Výstupy jsou tiskové nebo souborové pro další zpracování, mezi něž patří vykreslení profilu vrtu nebo interpolace průběhu geologických vrstev v oblasti mezi několika vrty. Báze vznikla ve spolupráci PÚDIS Praha a VÚMS, n.ú., Praha a byla uvedena do provozu v roce 1982.

3. Liniové objekty

Liniové objekty se od bodových liší tím, že objekt je reprezentován větším počtem bodů, většinou spojených v jistém pořadí čarami. Objekt tedy nelze ztotožnit s bodem, je charakterizován umístěním a jistou rozlohou /maximem a minimem jednotlivých souřadnic/ a k topologickým informacím o bodech přistupují informace o jejich spojnicích, jako je tvar čáry /např. přímka, kružnice, interpolovaná křivka n-tého stupně/, typ čáry /tloušťka, druh kresby/ a další. Vzhledem k tomu, že objekty mají vlastní rozměry, musíme algoritmus zjemňovaných čtverců dále modifikovat: řada objektů bude přecházet přes hranice čtverců tak, jak jsme je popsali v předchozím odstavci. Myšlenka této modifikace "hierarchických čtverců" navazuje na myšlenku zjemňování. Základní sítí mohou opět tvořit hranice mapových listů. Dojde-li při algoritmu zjemňování k půlení čtverce, je původní čtverec nahrazen dvojicí menších. Algoritmus hierarchických čtverců se liší tím, že původní čtverce v tomto případě zůstávají v platnosti pro objekty, které přecházejí přes novou dělicí čáru, a naopak se nad čtvercí základní sítě vybudují podle potřeby nadřazené, např. dvojnásobné čtverce pro uložení objektů, které se do původních čtverců nevešly. Nakonec tak vznikne hierarchie s vrcholovým čtvercem, jehož plocha je /teoreticky/ rovna 2^n-násobku základního čtverce /mapového listu/ a který /prakticky/ pokrývá celé univerzum datové báze a binární strom proměnné hloubky /nejméně však n/, jehož uzly odpovídají půleným čtvercům vyšší úrovně. Uzly n-té úrovně odpovídají původním čtvercům - mapovým listům. Každý objekt ukládáme do nejmenšího čtverce, který ho beze zbytku obsáhne. Hierarchicky nejvyšší čtverec bude obsahovat objekty velmi rozsáhlé, ale i všecky další, které překračují dělicí čáru druhé úrovně a tak dále. Názorně si lze algoritmus hierarchických čtverců představit jako soustavu stále jemnějších sít, kterými objekt "propadává", dokud se nezarazí o nějakou dělicí čáru. Při každém primárním výběru musíme ovšem vybrat nejen určitý počet nejjemnějších čtverců, ale i všecky čtverce, které je obsahují /nadřazené v přímé linii/. Hierarchické čtverce se dobře hodí i pro ukládání takových "liniových objektů", jako je název ulice, který je "umístěn" po celé její délce a přitom se v každém konkrétním výběru, který jakoukoli část zahrnuje, má vyskytnout a vykreslit celý.

Většina liniových objektů, s nimiž se v praxi setkáváme, netvoří pouze lomené čáry, ale různě se větví, takže tvoří obecné, zpravidla orientované a souvislé grafy, v některých případech acyklické /např. kanalizace/. Z důvodů, které uvedeme níže, je výhodné větvené liniové objekty v bodech větvení /tzv. uzlové body/ dělit na sekvence, které musí být souvislé a nevětvené. Algoritmem hierarchických čtverců pak ukládáme jednotlivé sekvence a objekt chápeme jako posloupnost sekvencí . Ukázalo se z uživatelského hlediska velice žádoucí, aby pořadí bodů v sekvenci i pořadí sekvencí v rámci objektu určoval uživatel báze a aby se toto pořadí ve všech případech zachovávalo i při výběru z báze. Běžné algoritmy linearizace grafu totiž vedou k pořadí, které je v určitých případech pro uživatele velmi nepřehledné a nepřirozené. Zachovávání pořadí sekvencí znamená sice v datové bázi jistou komplikaci, je však řešitelné bez dalšího průchodu daty.

U větvených liniových objektů je běžné, že dva nebo více objektů v určitém úseku /sekvenci/ sdílí tutéž trasu, tj. probíhá týmiž body. Příkladem může být mnohažilový kabel, kolektorová šachta nebo úsek železniční trati. Z kresebných i jiných důvodů je užitečné, aby taková sekvence a její body byly v datové bázi uloženy pouze jednou a účastnily se v několika objektech. Fakt sdílení je totiž dán v popisované realitě samé a nevzniká jen nahodilou shodou souřadnic. Sdílení se ovšem týká jen topologické složky sekvence a bodů, kdežto popisné údaje /evidenční, správní, technologické/ musí zůstat pro každý objekt zvlášť. Zavedení sdílených sekvencí si tedy vynucuje rozlišení mezi logickou a fyzickou sekvencí, případně logickým a fyzickým bodem. Fyzická sekvence se skládá z fyzických bodů, může být sdílena více objekty a je zásadně nositelem topologických a kresebných atributů. Logická sekvence je jednoznačně přiřazena objektu, je nositelem popisných atributů a na druhé straně se jednoznačně vztahuje k určité sekvenci fyzické - nanejvýš s opačnou orientací. Obecné schéma větvených liniových objektů můžeme tedy znázornit takto:



Toto obecné schéma předpokládá přímý /primární/ výběr podle identifikace objektu /případně třídy objektů/ a podle umístění /čtverce/. Databázové schéma, které bude složitější, by mělo dovolit obě kritéria kombinovat v rámci primárního výběru, aby zadání např. třídy objektů omezilo objem primárně vybraných dat a tedy urychlilo výběr.

Sběr dat o lokalizovaných objektech probíhá obvykle na několika oddělených místech, v různých dobách a je výsledkem různých profesních činností. Topologické údaje /kromě souřadnic/ zadává např. geodet na základě zaměření, mapy či plánu. Údaje o souřadnicích bodů vznikají strojovou digitalizací mapy nebo fotogrammetrického snímku, kdežto popisné údaje musí vyplnit různí odborní pracovníci na základě jiných podkladů nebo zjištění. Tuto skutečnost je třeba respektovat: projekt, který by trval na úplnosti dat, není příliš realistický. Velmi vážným problémem je naopak kontrola vstupních dat. Pro topologickou složku je nejlepší kontrolou kresba, jinak je možno zavést do vstupních dat jisté redundance, které zajistí aspoň nejdůležitější podmínky konzistence. Tak např. číslo bodu musí být vyplněno jak ve vlastní topologické složce, tak v datech z digitalizace, při vhodně zvoleném číslování bodů lze kontrolovat číslo mapového listu proti souřadnicím atd.

Další praktický problém se týká aktualizace dat. Protože se jedná o údaje složitě strukturované a protože o vnitřní struktuře dat neví odborní pracovníci uživatele zpravidla téměř nic, nelze aktualizovat na základě jednoznačných identifikátorů, jak tomu bývá při zpracování nelokalizovaných dat. Aktualizaci bude dále provádět několik různých pracovníků a bude podléhat různým schvalovacím a kontrolním procedurám. Dáváme proto přednost aktualizaci na datech, zapsaných ve tvaru velmi blízkém prvotním datům, která z datové báze vzniknou na základě požadavku jako tzv. nabídkový soubor. Aktualizace pak probíhá v několika krocích:

1. požadavek na aktualizaci určitého objektu a vytvoření nabídkového souboru,
2. vlastní úprava dat v nabídkovém souboru,
3. kontroly a ověřování upraveného souboru,
4. definitivní schválení nového /upraveného/ stavu.

Kroky 2. a 3. se přitom mohou vícekrát opakovat, součástí kontrol bude zpravidla i kontrolní kresba. Nabídkový soubor se jednak vydává, ale zároveň vzniká i v datové bázi a vlastní opravování musí technicky probíhat přímo v datové bázi, aby se daly vyloučit

konfliktní aktualizace i když bude "cyklus" aktualizace probíhat poměrně dlouhou dobu. V průběhu tohoto cyklu vzniká v datové bázi paralelně ke starému stavu postupně nový, který lze kdykoli zrušit. Po dobu aktualizace je pro běžné výběry k dispozici stále starý stav, nanejvýš s příznakem, že "na datech se pracuje". Teprve v okamžiku definitivního schválení se starý stav zruší a nahradí novým. Toto řešení je sice precnější, ukazuje se však, že je nezbytné a při vhodné realizaci může být dokonce i úspornější než okamžitá aktualizace bod po bodu přímo na místě.

Předkládané názory a zkušenosti vznikly v průběhu práce na programovém vybavení evidence podzemních vedení v rámci městského informačního systému města Prahy. Toto programové vybavení představuje poměrně univerzální systém ukládání, výběru a aktualizace liniových lokalizovaných dat. První aplikace, které se má uvést do provozu v nejbližších měsících, bude sloužit k evidenci vodovodních, kanalizačních, silových a dalších veřejných podzemních vedení na území hlavního města. Předpokládaný cílový objem činí řádově deset milionů bodů. Evidence podzemních vedení navazuje na další datové báze /mapa velkého měřítka aj./ a měla by nahradit rychle zastarávající technickou mapu města. Užitečnost takového soustředění informací může posoudit každý, kdo se ucházel o stavební povolení a musel obejít deset a více různých "správců" podzemních vedení na území města. Celý projekt, který opět vzniká ve spolupráci PÚDIS Praha, VÚMS,k.ú.o.,Praha a dalších organizací, je naprogramován v jazyce PASCAL, používá databázový systém DBS-25 v rámci operačního systému DOS-4/EC a měl by se provozovat na počítači EC 1045. Z dalších výhledových aplikací /kromě zavádění v dalších městech/ bych rád uvedl evidenci vedení v technologicky složitých budovách, která je příkladem trojrozměrné aplikace. Možnosti aplikace na plošné objekty jsou zatím ve stadiu úvah a studií.