

# Studentská Vědecká Konference 2012

## On-the-fly generalizace vícerozměrných dat

Radan Šuba<sup>1</sup>

### 1 Úvod

Práce se zabývá zjednodušováním map v reálném čase. Řeší volbu libovolného měřítka mapy, redundanci dat a strukturu pro ukládání výsledků generalizace. V úvodu předkládá problematiku zjednodušování map on-line, následně hledá možnou datovou strukturu pro generalizaci v reálném čase, tedy on-the-fly. Pracuje na topologickém strukturování dat na uzly, hrany, plochy a vytváří další mapová měřítka zjednodušováním měřítek základních. Těžištěm práce je implementace vybrané datové struktury s názvem tGAP (topological Generalized Area Partition) v databázovém systému Oracle.

### 2 On-the-fly generalizace

V tradičním přístupu ke generalizaci vícerozměrných (prostorových) dat jsou na začátku definovány konkrétní úrovně detailu (v případě prostorových dat měřítka) a všechny tyto úrovně jsou poté v databázi uloženy. Ve svém důsledku vede tento přístup ke vzniku duplicitních dat, což především z pohledu správy a aktualizace těchto vícerozměrných dat v databázi není žádoucí. Znamená to zvýšené časové nároky správců dat a i více diskového prostoru nutného pro uložení všech požadovaných úrovní detailu.

Ideálním stavem by byl mapový portál využívající uložená prostorová data pouze v jedné referenční vrstvě, přičemž referenční vrstva je nejpodrobnější. Navíc by měla umožňovat volbu libovolného měřítka a odvození (generalizování) všech mapových měřítek z nejpodrobnější, referenční vrstvy. Eliminace duplicitních dat rozhodně patří k dalším podstatným výhodám. Tento přístup je znám jako generalizace on-the-fly. Tato práce se snaží popsat generalizaci on-the-fly a poté vybrat a implementovat strukturu, s jejíž pomocí lze on-the-fly generalizaci realizovat.

Těžištěm práce je návrh efektivní implementace datové struktury tGAP (topological Generalized Area Partition). Tato datová struktura pracuje nad topologicky strukturovanými daty (uzly, hrany, stěny) a je implementována pomocí několika stromových struktur. Značnou výhodou tohoto přístupu je neduplicitní ukládání dat. Pro vlastní generalizaci hran byl zvolen Douglas-Peucker algoritmus. Tento algoritmus rekurzivním dělením linie na menší segmenty na základě vhodně zvolené geometrické podmínky „vyhlazuje“ průběh linie. Výsledky aplikace tohoto generalizačního algoritmu se ukládají do binární stromové struktury BLG-tree (Binary Line Generalization), který je součástí výsledné struktury tGAP. Další významnou součástí výsledného řešení je stromová struktura GAP-tree (Generalized Area Partion Edge Forest), která je použita pro proces zjednodušování hran jednotlivých stěn při slučování oblastí (agregaci) a vzniká jistým „spojováním“ dílčích BLG stromů. Implementace datové struktury tGAP vznikla na platformě Oracle 11g, s využitím prostorového rozšíření Spatial.

V práci je popsán celý proces zpracování prostorových dat, od jejich načtení do prostorové databáze, přípravu pro načtení do datové struktury tGAP, vlastní implementaci a naplnění struktury tGAP až po ukázkové generování generalizovaných výstupů na testovacích datech katastru nemovitostí. Tento proces je vysoce vhodný například při tvorbě map, kdy je mapy středních a malých měřítek možné vytvářet generalizací map velkých měřítek, které

---

<sup>1</sup> student magisterského studijního programu Geomatika, obor Geomatika, specializace Geodézie a geoinformační systémy, e-mail: radan.suba@volny.cz

obsahují potřebnou úroveň detailu. Pro úplnost je nutné dodat, že konkrétní proces tvorby map středních a malých měřítek pomocí generalizace nemůže v podstatě být plně automatizovaný, nicméně, lze se tomuto ideálnímu stavu užitím on-the-fly generalizace významně přiblížit.

### 3 Závěr

Přínos práce je v implementaci datové struktury pro generalizaci 2D prostorových dat s využitím prostorové databáze Oracle Spatial. Implementace je v současné době testována na reálných datech. Při nasazení do ostrého provozu tvůrců a zpracovatelů bází prostorových dat (např. základní báze goedat, spravované Zeměměřickým úřadem v Praze, se kterým již byla zahájena v této věci jednání) lze očekávat zpracování dat v objemu řádově stovky gigabajtů. Z tohoto důvodu je přínos práce i v tom, že data jsou ukládána pouze na jedné úrovni s nejvyšší mírou detailu a ke generování dalších úrovní s nižší úrovní detailu dochází na vyžádání (on-the-fly) a není nutné ukládat data explicitně pro každou zvolenou úroveň detailu. Představená implementace tak navíc poskytuje volnost při volbě zvolené úrovně detailu, což při klasických přístupech ke generalizaci vícerozměrných dat prostorových dat nebylo možné a bylo nutné dopředu definovat, jaké úrovně detailu se budou udržovat.

### Literatura

- MEIJERS, B. M. *Implementation and testing of variable scale topological data structures*. Delft: The Netherlands: 2006. Diplomová práce. Delft University of Technology, Section GIS Technology.
- ROBERT, W. a D. BURGHARDT. Generalization, On-the-Fly. In: SHARKKAR, S. a X. HUI. *Encyclopedia of GIS*. SpringerScience+Business Media LLC., 2008, s. 339-344. ISBN: 978-0-387-35973-1
- VAN OOSTEROM, P. a V. SCHENKELAARS. The Development of an Interactive Multi-Scale GIS. *International Journal of Geographical Information Systems*. The Hague: The Netherlands: 1993.
- VAN OOSTEROM, P. *Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest*. 2005. 331-346.