Posudek oponenta bakalářské práce

Autorka	Jindra Marvalová
Název práce	Správa a vizualizace časoprostorových bodových dat
Studijní obor	Geomatika
Oponent práce	Ing. Jiří Skála, Ph.D.

Splnění cílů práce	velmi dobře
Odborný přínos práce	netradiční postupy, zpracování výsledků z různých zdrojů
Matematická (odborná úroveň)	velmi dobrá
Věcné chyby	téměř žádné
Grafická, jazyková a formální úroveň	vynikající

Práce Jindry Marvalové se zabývá aktuálním problémem generalizace a vizualizace kartografických bodových dat. Vychází při tom z nejnovějších výsledků, cituje práce z roku 2012. Cíl práce je v úvodu jasně stanoven. Kapitola 2 uvádí přehled shlukovacích algoritmů, jejich výhody a nevýhody pro následné použití na kartografická data. V kap. 2.2 na str. 6 by bylo vhodné uvést citaci článku MacQueen, 1967. Dále v popisu algoritmu není přesně uvedeno, jak se „přepočítají těžiště". Kapitola 3 dává přehled kartografických generalizačních metod. Kapitola 4 popisuje metody kartografické vizualizace s ohledem na zobrazování shluků.

Kapitolou 5 začíná praktický přínos práce; popisuje způsob uložení dat v relační databázi. Kapitola 6 se zabývá generalizací bodových dat metodami popsanými v teoretické části. Tabulka 6.1 na str. 28 navrhuje počty shluků pro různá měřítka mapy. Bylo by vhodné uvést, z čeho při tom autorka vycházela, případně jak se navržené počty osvědčily v praxi. Na str. 30 je popsáno, že algoritmus k-means nedává dobré výsledky pro větší měřítka mapy. Pro případnou další práci doporučuji zkusit náhodnou volbu počátečních těžišť, případně použít algoritmus pro jejich odhad, např. Bradley, Fayyad: *Refining initial points for k-means clustering*, 1998; nebo Wang, Zhu: *Research on selecting initial points for k-means clustering*, 2008.

Oceňuji kvalitní porovnání shlukovacích algoritmů, kde autorka nejprve stanoví hodnotící kritéria s ohledem na cíl práce. Jako nejvhodnější je zvolena metoda k-means. Vlastnosti všech algoritmů jsou shrnuty v tabulce. Kapitola 7 navrhuje tři metody vizualizace generalizovaných dat. Proč jsou metody 1 a 2 demonstrovány na datech z algoritmu SnapToGrid, přestože za nejlepší byl zvolen k-means? V kapitole 8 je navrženo možné pokračování práce zohledněním časové složky dat. Kapitola 9 shrnuje výsledky práce.

V práci jsem našel několik nepodstatných chyb. Na str. 15 odst. 3 místo „kartografické vizualizace" asi má být „kartografické generalizace". Na str. 29 je v popisu vzorce uvedeno, že proměnné *min*, *max* jsou pole. Správně jsou to minimální a maximální indexy vstupního pole bodů. Obrázek 6.3 neodpovídá popisu na str. 37. Na téže straně k popisu shlukování dat v Plzni by bylo vhodné připojit obrázek.

Bakalářská práce je logicky uspořádaná, poměr teoretické a praktické části je vyvážený. Stylistická i typografická úroveň je bezvadná. Použité termíny jsou v textu řádně vysvětleny (databázový systém, PHP, agregační funkce). U metod (shlukování, generalizace) je nejprve stručně popsáno, co dělají, k čemu slouží. Pak následují konkrétní algoritmy a detaily. Oceňuji, že se práce drží svého oboru, a tak např. metodu MarkerCluster kriticky zkoumá z hlediska kartografické vizualizace.

K práci mám následující dotazy:

Umožňují uvedené generalizační metody přidávat nové body do existujících dat?

Jak by se dala vizualizovat data s respektováním typů bodů (keší)?

Práci Jindry Marvalové považuji za zdařilou, doporučuji uznat ji jako kvalifikační a navrhuji hodnocení *výborně*.

16. června 2013

Ing. Jiří Skála, Ph.D.