

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Petr Holeček
Oponent práce: RNDr. Lukáš Hermann
Název práce: Temporální data v grafové databázi v projektu Manta
Obor: Webové a softwarové inženýrství (magisterský)

Datum vytvoření: 19. 5. 2015

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
Komentář: Složitost zadání spočívala především v nutnosti seznámit se nejen s novou technologií, grafovou databází Titan, ale také její aplikací v komplexním projektu, a dále s problematikou temporality dat, která měla být nově aplikována jak pro tuto technologii, tak pro její konkrétní implementaci. Samotná implementační část práce ale na druhou stranu nebyla nijak rozsáhlá a jednalo se o rozšíření již existujících řešení.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
Komentář: Podle zadání se měla práce vypořádat s dvěma temporálními dimenzemi dat, ale procesní metadata nebyla v této práci vůbec analyzována. Na druhou stranu řešení historické dimenze tímto není omezeno a může být částečně využito i pro řešení procesní dimenze. V úvodu dále bylo přislíbeno a v závěru indikováno jako splněné, že se práce nad rámec zadání zabývá také výhodami a nevýhodami různých způsobů uložení temporálních dat v RDBMS a dále že obsahuje rešerši stávajících řešení problematiky temporality v grafických databázích. Obojí ale v samotném textu práce chybí.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Porovnejte rozsah předložené písemné zprávy s požadovaným rozsahem, viz Směrnice děkana č. 9/2011, článek 3. Pro hodnocení ZP je také důležité, zda všechny části písemné zprávy jsou informačně bohaté a pro práci nezbytné. Text ZP by neměl obsahovat zbytečné části.	
Komentář: Kapitoly o implementaci a testování obsahují zbytečně detailní Java kód, který by měl být nahrazen popisem nebo pseudokódem. Samotný kód by pak mohl být uveden v příloze. I bez tohoto kódu by ale práce splňovala požadovaný rozsah.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	65 (D)
Popis kritéria: Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.	

Komentář:

V části analýzy, kde se probírají jednotlivá řešení, se často uvádí, že některé řešení je pomalejší než jiné, ačkoliv toto není dostatečně argumentováno, případně prakticky ověřeno. Z analýzy je velmi špatně pochopitelný rozdíl mezi řešeními uváděnými v kapitolách 2.1.2 a 2.1.3. V kapitole implementace a benchmarking je zmíněno využití tzv. vertex-centric indexů, které urychlují řešení, bylo by vhodné lépe popsat rozdíl mezi těmito a standardními indexy.

Některé kapitoly jsou členěny nelogicky (viz kapitola 1.5 - Nástroj Manta pod kapitolou 1 - Temporalita obecně; nebo viz dělení kapitoly 2 - vše v podkapitole 2.1, kde sousedí například sekce Problémy diskretní časové osy a Verzování celé repository).

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

5. Formální úroveň práce

60 (D)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 9/2011, článek 3.

Komentář:

Práce obsahuje mnoho překlepů, některá slova jsou špatně rozdělena na nový řádek (zv-oleny, min-imální). V textu se vyskytuje špatné použití spojovníků a pomlček, v seznamech chybí interpunkční znaménka. Některé tabulky a některá slova v odstavcích přetékají okraje stránky. Text neobsahuje významnější pravopisné ani gramatické chyby. Autor se občas nejednotně uchyluje k použití první osoby jednotného čísla.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

70 (C)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

Student vycházel z relevantních zdrojů vzhledem k zaměření jeho práce a tyto správně citoval. Chybí širší rešerše přístupů k temporálním datům v grafových databázích, tedy není možné ověřit, zda už tento problém někdo neřešil. Na druhou stranu by vzhledem k vázanosti řešení na konkrétní projekt byla tato řešení pravděpodobně špatně aplikovatelná. Je dobře rozlišitelné, co již bylo implementováno v projektu a co je konkrétní přínos studenta.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

65 (D)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Hlavní cíl práce, tedy nalezení vhodné reprezentace uložení temporálních dat v grafové databázi a její implementace podle specifických potřeb projektu Manta, byl splněn. Kvalita implementace řešení odpovídá standardům kladeným na vývoj softwarových aplikací (funkční i výkonové testování, komentáře kódu, dokumentace). Pro lepší zhodnocení přínosu výsledné implementace by bylo vhodné provést výkonnostní testy vložení částečně změněného grafu do databáze (aktuálně jsou pouze pro stejný graf). Pro větší zobecnění práce pro další publikaci by bylo potřeba provést rešerši obecných přístupů k temporálním datům v grafových databázích a srovnat s nimi toto řešení.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uveďte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

Komentář:

Vzhledem k zaměření analýzy i implementace této práce na v praxi využívaný projekt jsou výsledky této práce přímo využitelné v praxi.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě

Popis kritéria:

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

- Můžete detailněji vysvětlit rozdíl mezi řešeními navrhovanými v kapitolách 2.1.2 a 2.1.3?
- Můžete vysvětlit rozdíl mezi normálním a vertex-centric indexem, které byly použity ve výkonostních testech?
- Můžete nastínit, jak by se lišila implementace temporální dimenze procesních metadat od historické dimenze?

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

65 (D)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Práce má nastavené přiměřené cíle, které jsou až na malé výjimky splněny. Práce má taktéž přímý aplikační přínos a její závěry je možné zobecnit. Na druhou stranu práci schází důležité rešerše, některé úvahy v analýze nejsou dostatečně vysvětleny a kazí ji typografické a lingvistické chyby. Vzhledem k průměrné obtížnosti zadání a uvedeným problémům při jeho zpracování navrhuji podprůměrné hodnocení.

Podpis oponenta práce: