



Posudek oponenta závěrečné práce

Student: Bc. Lucie Svitáková
Oponent práce: RNDr. Lukáš Hermann
Název práce: Query Analysis on a Distributed Graph Database
Obor: Webové a softwarové inženýrství

Datum vytvoření: 27. 1. 2019

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. 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 dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
Komentář: Všechny body zadání byly splněny v popsaném rozsahu.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	100 (A)
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
Komentář: Všechny části práce obsahují dostatečné množství informací pro pochopení zadání, návrhu, implementace i konkrétních výsledků. Práce neobsahuje faktické chyby, provedené rešerše obsahují informace platné k době vzniku práce. Práce je logicky rozdělená do sedmi kapitol - první kapitola seznamuje s teorií grafů, druhá s problémem rozdělení grafu, třetí obsahuje rešerši souvisejících technologií, čtvrtá analýzu požadavků a návrh, pátá popis realizace, šestá výstupy testování na vybraných vstupech a poslední návrhy na další vylepšení. Jednotlivé kapitoly na sebe logicky navazují a postupně uvádí čtenáře do problematiky a jejího řešení. Práce neobsahuje pravopisné či gramatické chyby, typografické zpracování je taktéž na vysoké úrovni. Práce obsahuje všechny náležitosti - abstrakty v obou jazycích, klíčová slova, obsah, seznamy obrázků, tabulek a literatury, závěrečné zhodnocení výsledků, přiložené CD se zdrojovými kódy a dokumentací. Všechny zdroje jsou řádně citovány a z textu je zřejmé, které části práce jsou implementovány studentem. Velká část referencí je ve formě odkazů na webové stránky a je u nich uvedena doba, kdy byly informace z těchto zdrojů čerpány. Veškeré knihovny třetích stran byly využity v souladu s jejich licencí.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	90 (A)
Popis kritéria: Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů	
Komentář: Práce obsahuje prototyp implementace navrženého řešení. První částí prototypu je rozšíření knihovny TinkerPop o logování spuštěných dotazů v různých formátech, druhou částí pak implementace navrženého redistribučního algoritmu. Obě části slouží především k ověření proveditelnosti a efektivnosti navrženého řešení na různých sadách dat. Z tohoto důvodu neobsahují žádné unit testy. Způsob konfigurace a exekuce aplikace je dostatečně popsán v dokumentaci. Samotný kód odpovídá kvalitě prototypu, využívá dostupné knihovny, v některých částech však není dostatečně okomentovaný.	

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
4. Hodnocení výsledků, jejich využitelnost	90 (A)
<i>Popis kritéria:</i> Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.	
<i>Komentář:</i> Práce nevytváří kompletně nový algoritmus, ale navrhuje úpravu stávajícího tak, aby byly využity jeho přednosti a částečně odstraněny jeho nedostatky. Navržená úprava distribučního algoritmu by se znalostí její limitace (malá dysbalance na uzlech clusteru) mohla být implementována v praxi, jelikož na rozdíl od podobně efektivního algoritmu vyžaduje výrazně menší výpočetní prostředky, což je nezbytný předpoklad pro nasazení v systémech zpracování velkého množství dat, kde může být distribuované nasazení grafových databází nezbytné.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – nehodnotí se</i>
5. Otázky k obhajobě	
<i>Popis kritéria:</i> Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).	
<i>Otázky:</i> 1. Výsledkem práce je testovací prototyp, vytvářející návrh nového umístění uzlů. Můžete v obecnosti popsat, jaké další kroky by musely být provedeny pro implementaci samotné redistribuce uzlů ve vámi zvolené databázi? 2. V kombinaci s vámi navrženým algoritmem pro redistribuci uzlů v grafu, jaký myslíte, že by byl nejlepší způsob udržení dobré balance na uzlech clusteru při průběžném přidávání nových uzlů a hran? 3. O jaké další operace byste rozšířila vámi navržené logování v rozhraní TinkerPop, abyste dostatečně zachytila dopad distribuovanosti databáze na její využití?	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
6. Celkové hodnocení	95 (A)
<i>Popis kritéria:</i> Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.	
<i>Text hodnocení:</i> Předložená závěrečná práce obsahuje všechny podstatné části pro analýzu, návrh a prototypovou implementaci algoritmu pro dobré rozmístění uzlů v distribuované grafové databázi. Všechny části práce jsou zpracovány ve vysoké kvalitě z hlediska předložených informací, formálních náležitostí, použitých zdrojů i prototypové implementace. Samotná prototypová implementace by mohla být širší, ale možnosti dalšího rozšíření jsou v práci dostatečně popsány.	

Podpis oponenta práce: