

Posudek oponenta diplomové práce

Autor/autorka práce: **David Hrbáček**

Název práce: **Zpracování časových údajů pro jejich vizualizaci**

Obsah práce

Předmětem práce je průzkum algoritmů pro zpracování grafových dat a nástrojů pro zpracování a reprezentaci grafů. Dále výběr algoritmu vhodného pro ohodnocení významnosti uzlů grafu historických událostí, osob, míst a dalších entit, návrh, implementace a otestování nástroje pro zpracování grafů tohoto typu. Obsahově práce pokrývá nutné aspekty práce srozumitelným textem a logicky navazující strukturou kapitol. Přesto existují výhrady k této stránce práce. První je hned úvod práce, z kterého není dostatečně patrná motivace pro její vytvoření ani konkrétní cíle. Zmínka o tom, že práce přispívá k větší skupině diplomových a bakalářských prací, které se spojují ve větší aplikaci, následuje až dále v textu práce a důvody pro vývoj této aplikace a její celkové funkce nejsou zmíněny vůbec. Po úvodu text práce začíná poněkud prudce a pouští se přímo do konkrétnějších aspektů grafových algoritmů s ohledem konkrétní použití pro graf historických událostí. To vyžaduje od čtenáře určitou míru znalosti samotné teorie grafů a později i matematických postupů a pro člověka v těchto oborech méně znalého může způsobit nesrozumitelnost práce. Při rozsahu 57 stran od úvodu po závěr by jistě byl na místě krátký obecný úvod do těchto problematik a vysvětlení později používaných pojmů jako $E(G)$ a hloubka okolí, které v tomto stavu práce berou jako obecně známé. Další drobnou konkrétní výtkou je pak porovnání existujících knihoven pro implementaci grafů. I když jejich popis v textu je dostačující a nevýhody, které brání jejich použití, jsou dobře zdůrazněny, přesto by pro názornost a celkové shrnutí přispěla jejich tabulka s potřebnými kritérii, která splňují, případně nespĺňují. Obecnějším problémem je pak absence popisu obsahu přiloženého CD, návodu ke spuštění a použití aplikace (který je pouze stručně na CD ve formě textového souboru) a rovněž seznamu obrázků, tabulek a rovnic, kterých je v práci nemalé množství. V závěru pak chybí kritické zhodnocení práce a komentář ke splnění bodů zadání.

Kvalita řešení a dosažených výsledků

Co se týče implementované aplikace a jejího popisu, je opět v obecném měřítku vše v pořádku a výtky jsou spíše dílčího charakteru. Popis implementace v textu je srozumitelný a popisuje zásadní principy a části programu. Implementace samotná je velmi rozsáhlá (63 .java souborů tříd, rozhraní a JUnit testů), její dekompozice na třídy a segmentace do metod je více než uspokojivá a jeví se logicky správná a CD obsahuje všechny části nutné k zprovoznění a používání aplikace, včetně těch přejatých z paralelních prací. Aplikace je funkční a splňuje požadavky na ní kladené. Komentování kódu je částečně rozporuplné, protože v některých případech jsou třídy komentovány velmi důsledně včetně členských proměnných, private metod a komentářů jiných než dokumentačních, zatímco v jiných třídách je míra i úroveň komentování výrazně slabší. První případ ale výrazně převládá. V popisu implementace je jediný drobný nedostatek v podobě nekonzistence popisu atributů uzlů a hran grafu. Konkrétně atributu properties, který je zdá se nepovinný, ale na rozdíl od předchozích atributů to u něj není explicitně řečeno. V implementaci samotné je jedinou drobnou výhradou pojmenování tříd pro reprezentaci uzlů a hran. Třídy představující logickou strukturu prvků jsou pojmenovány Vertex a Edge, zatímco třídy nesoucí data těchto prvků mají názvy Node a Bond. Ač oddělení těchto dvou aspektů uzlů hran do 2 skupin tříd je logické a správně použité, samotná terminologie může být na první pohled matoucí. Jasnější by bylo použití stejných základů názvů s patřičnými postfixy

určujících konkrétní účel tříd. Poslední připomínka se vztahuje k testování programu. Aplikace disponuje unit testy a byla testována i její časová a paměťová náročnost, srovnány výsledky různých variant algoritmu PageRank a prozkoumána ideální hodnota zastavovací podmínky. Popis testů ale postrádá konfiguraci hardwaru a softwaru, na které bylo testování prováděno, a která může výsledky časové a paměťové náročnosti ovlivnit ne ve smyslu jejího charakteru, ale konkrétních dosažených časů a spotřeby paměti. Dalším užitečným prvkem by bylo porovnání výkonnosti implementovaných algoritmů s již existujícími knihovny, popisovanými v textu práce. To ale nebylo součástí zadání.

Formální úroveň

Po formální a jazykové stránce je práce kvalitní a výhrady jsou opět pouze dílčí. Použité vyjadřovací prostředky jsou adekvátní stejně jako formátování textu a obrázků. Výjimkou jsou jen grafy z kapitoly o testování, kde použité barvy v černobílé verzi nelze rozeznat. Frekvence překlepů a interpunkčních chyb je zanedbatelná, za zmínku stojí snad jen opakující se chybějící tečka, či čárka za použitou rovníčkou. Zvláště působí nevysvětlení některých zkratk (API, URL, HTTP) a to ani při prvním výskytu v textu, ani v seznamu zkratk. Jedná se sice o pojmy v oboru obecně známé, nicméně, těmi jsou i pojmy URI, XML a REST a ty vysvětleny jsou. Poslední drobnou vadou jsou někdy až příliš dlouhé popisky obrázků (viz obrázek 7.1), jejichž většina textu se měla spíše objevit v samotném textu práce.

Práce s literaturou

Práce cituje 18 očividně validních a věcně odpovídajících zdrojů, z nichž 8 je elektronických.

Splnění zadání

Splněno bez výhrad

Dotazy k práci

Odkud, kým a kdy byla stanovena struktura grafu historických událostí?

Jsou všechny knihovny existující před započítím práce open-source? Jestli ne, které jsou?

Jaké byly alternativy knihovny pro práci s maticemi a vektory oJAlgo a jaké nevýhody je diskriminovali oproti knihovně zvolené?

Navrhuji hodnocení známkou ~~výborně~~ / **velmi dobře** / ~~dobře~~ a práci doporučuji k obhajobě.

V Plzni 27.8.2015

Ing. Petr Pícha

