

Posudek vedoucího diplomové práce

Peter Irikovský: Graph algorithms in text retrieval

Původní zadání této práce bylo zadáno velmi široce jako využití grafových algoritmů v dokumentografických informačních systémech. Musím zde vyzdvihnout velmi aktivní přístup autora k této práci. Během studia podkladů pro diplomovou práci autor absorboval velmi obsáhlou kolekci současných článků, které se tématu týkají a využitím grafových algoritmů pro zlepšení kvality odezvy se zabývají. Spolu s tím, jak se rozšiřovala oblast témat, kterými se diplomant zabýval bylo zřejmé, že bude nutné práci konkrétněji zacílit. Nakonec padla volba v první řadě na možnosti, které nabízejí algoritmy pro práci s grafem Webu, tedy algoritmus PageRank, jeho modifikace.

Problematikou výpočtu PageRanku se zabývá také implementační část práce, jejímž cílem bylo implementace paměťově i časově efektivního výpočtu PageRanku pro vyhledávač Egothor. Implementaci je věnována především pátá kapitola, která popisuje implementaci včetně výsledků, kterých bylo dosaženo pro snapshoty fakultní a katedrální domény.

Příloženo je několik variant implementace lišících se způsobem práce s pamětí, případně dalšími parametry. Dosažené výsledky jsou shrnuty v textu práce.

Práce má však i svoji přehledovou část, která je širěji pojatá, a obsahuje přehled řady dalších algoritmů. Pro tuto část práce byly z původně nastudovaných algoritmů vybrány především algoritmy příbuzné zvolenému cíli implementace, respektive ty, které by bylo možné dále kombinovat s výpočtem hodnot PageRank. Této přehledové problematice je věnována část páté kapitoly a kapitoly následující. Úvodní kapitoly potom obsahují nezbytné základní seznámení s principy vyhledávání v textech a s některými dále používanými grafovými algoritmy.

Text práce je napsán anglicky. Dle mého názoru je angličtina na poměrně vysoké úrovni a text práce je dobře čitelný a strukturovaný. Přehledová část seznamuje čtenáře s řadou současných výsledků použití grafových algoritmů pro vyhledávání v textech. Implementační část rozšiřuje schopnosti na fakultě vyvíjeného vyhledávače. Celkově lze říci, že práce splnila zadané cíle. Vzhledem k tomu, že k práci nemám žádné vážnější připomínky, doporučuji ji uznat jako práci diplomovou.

V Praze dne 14. 5. 2007

RNDr. Michal Kopecký, Ph.D.
KSI MFF UK

