

Nově vznikající proudové systémy mají vůči dotazovacím enginům pro grafové databáze výhodu při agregaci dat (části Group by a Order by), protože jim stačí uchovávat pouze agregované prvky, ale zase nedokáží provádět vyhledávání vzoru.

V této práci jsme vytvořili statickou grafovou databázi s Labeled-property datovým modelem a pro ni dotazovací engine, který agreguje data klasickým přístupem až po dokončení vyhledávání vzoru. Dotazovací engine jsme následně upravili po vzoru proudových systémů tak, aby prováděl agregace již v průběhu vyhledávání vzoru. Cílem této práce bylo zjistit, zda danou úpravou dokážeme urychlit zpracování dotazů.

Pro upravený i původní přístup jsme navrhli a implementovali několik jednovláknových i paralelních řešení. Řešení jsme porovnali v rychlosti zpracování dotazů na reálných grafech s uměle vygenerovanými hodnotami vlastností. Zjistili jsme, že existují situace, kdy zmíněnou úpravou dochází k urychlení zpracování dotazů. Konkrétně se tak stalo pro upravené paralelní řešení části Order by, které předčily původní řešení při třídění pomocí hodnot vlastností. Rovněž jsme dokázali navrhnout jednovláknové řešení části Group by, které bylo rychlejší než všechna původní řešení. Dále jsme zjistili, že nastalo značné urychlení pro upravené jednovláknové i paralelní řešení Single group Group by (dotaz obsahuje agregační funkce a nemá část Group by).