

## Reprezentace XML dat pomocí C-store

Předkládaná diplomová práce se zabývá možnostmi využití metod ukládání relačních dat po sloupcích (tzv. column oriented storage) pro správu a dotazování XML dokumentů.

Práce je rozdělena do několika částí. V první části (kapitolách 1 a 2) je stručně popsán úvod do problematiky, použítá implementace C-store a kritéria pro srovnávání navržených algoritmů. Ve druhé části (kapitolách 3 – 5) jsou popsány dva navržené přístupy – ukládání se znalostí schématu a bez něj – a implementace navržených přístupů. Třetí část (kapitola 6) práci hodnotí a popisuje její možná rozšíření.

Hlavními klady práce jsou především:

- Zajímavá a neotřelá problematika
- Navržení dvou přístupů k ukládání XML dat – s a bez schématu – a jejich srovnání
- Velmi pěkný formální popis algoritmů včetně analýzy jejich složitosti
- Minimální množství překlepů a gramatických chyb

Z hlediska důležitých připomínek je možné uvést následující:

- V práci zcela schází analýza existujících přístupů ke správě (nejen) XML dat v relačních databázích, především pak pomocí přístupu ukládání po sloupcích.
- V částech popisujících navržené přístupy k ukládání dat není v podstatě žádná diskuse možností. Je pouze uveden zvolený přístup, ale není vysvětleno proč byla tato řešení zvolena, zda existují alternativy, v čem jsou lepší/horší apod.
- Samotné experimenty jsou poměrně triviální. Pokud již implementace existuje, bylo by vhodné srovnat algoritmy na různých typech dokumentů (co do velikosti, struktury apod.).

K méně zásadním připomínkám patří tyto:

- Struktura práce je poněkud nepřehledná, členění do kapitol je matoucí.
- V práci zcela schází kapitola, která by alespoň krátce přiblížila použité technologie.
- Není jasné, proč nejsou výsledky experimentů uvedeny přímo v textu práce, zvláště když se jedná o necelé 2 strany textu.
- V referencích jsou uvedeny pouze 3 citace. Zcela chybí odkazy na zmiňované technologie (např. XML, relační model), metody a přístupy (např. DTD graf, DF algoritmus), tvrzení (např. XML data obvykle schémata nemají) apod.

I přes výše uvedené připomínky, práce splňuje původní zadání i cíl. Studentka dostatečně pronikla do zadané problematiky a prokázala schopnost detailně zanalyzovat a řešit zadaný problém. Práce Zuzany Částkové podle mého názoru splňuje podmínky na diplomovou práci kladené na obhajobě.

V Praze, 12.5.2009

RNDr. Irena Mlýnková, Ph.D.  
KSI MFF UK