

Diplomová práce – posudek vedoucího

Stanislav Kovalčín: Slovníkové metody jako druhá fáze BWT

Předložená práce se zabývá možnostmi použití slovníkových metod LZC a LZSS jako druhé fáze po Burrows-Wheelerově transformaci (BWT), namísto klasické kombinace metod Move to Front (MTF) a Run Length Encoding (RLE). Celý výzkum se nezabývá jen kompresí nad abecedou písmen, jak bývá většinou zvykem, ale také nad abecedami slov a slabik.

V teoretické části práce autor nastudoval a velmi detailně popsal metody BWT, RLE, MTF, LZC a LZSS včetně jejich různých variant. Kromě slovního popisu nechybí ani názorné příklady a pseudokód. Všechny popsané metody byly dále autorem implementovány. Jedinou výjimkou je metoda BWT, jejíž implementace byla převzata se softwarového projektu XBW, jejímž členem byl i autor práce.

Velmi oceňuji, že práce je rozšířením již obhájeného projektu XBW. Výhodou pro autora bylo, že se nemusel starat o parsování textu na jednotlivé elementy (slabiky, slova, ale i XML značky), ani o samotnou BWT, ale mohl se naplno věnovat samotnému tématu práce.

Nejzajímavější částí práce jsou naměřené výsledky kompresního poměru při nastavení různých kombinací jednotlivých kompresních metod. Počet všech možných kombinací přesahuje množství, které by šlo v rámci práce publikovat. Autor byl tedy nucen vybrat na základě naměřených hodnot pouze některé zajímavé kombinace. Ke každé kapitole popisující nějakou variantu kompresní metody je přidána i tabulka s výběrem výsledků pro různé nastavení dalších parametrů.

Na závěr je pak tabulka shrnující nejlepší výsledky z jednotlivých dílčích tabulek. Bohužel je zde vždy uveden jen jeden parametr (ze které dílčí tabulky výsledek pochází) a ostatní parametry zde chybí, ale dají se dohledat v dílčích tabulkách.

Domnívám se, že jde o velmi zdařilou práci, ve které autor prokázal, že umí nastudovat netriviální algoritmy popsané v odborné literatuře a následně je efektivně implementovat. Experimenty pro srovnání algoritmů byly vhodně navrženy a získané výsledky jsou srozumitelně a přehledně vyhodnoceny.

Předloženou práci považuji po všech stránkách za práci splňující kritéria pro diplomové práce na MFF UK a doporučuji ji k obhajobě.

Praha, 17. 8. 2007

Mgr. Jan Lánský, KSI MFF UK

