

Review of Master's Thesis

Student: Kubov Peter, Bc.
Title: Scalable Binary Executable File Similarity (id 23736)
Reviewer: Kolář Dušan, doc. Dr. Ing., DIFS FIT BUT

- 1. Assignment complexity** **average assignment**
Zadání je relativně široké a vyžaduje škálovatelné řešení, ale při jeho řešení bylo možno, a bylo využito řadu existujících postupů i řešení. Nad rámec zadání student navrhl vlastní způsob kódování/uložení rovnoměrně rozložených dat tak, aby byla uložena úsporně a přitom rychle dosažitelná.
- 2. Completeness of assignment requirements** **assignment fulfilled with minor reservations**
Zadání považuji za splněné, ale zpráva má jisté mezery, k té mám zásadní výhrady, viz níže.
- 3. Length of technical report** **in usual extent**
Rozsah je na obvyklém rozmezí, když si odmyslíme pár zbytečných, či velkých obrázků, nebo text, který není podstatný, dostaneme se k dolní hranici.
- 4. Presentation level of technical report** **75 p. (C)**
Skvělý úvod, dále pak stále snižující se kvalita, jak po obsahové stránce, tak po stránce srozumitelnosti a pochopitelnosti. Např. popis implementace zahrnuje vždy výčet technologií, k tomu reklamní obrázek, snímek obrazovky. Toto opravdu není popis implementace, ani tvrzení, že implementace byla popsána v návrhu - nebyla, tam o tom, jak se to implementovalo, nic není. Podobně, návrh metody uložení není popsán dostatečně, v algoritmech (jediné dva, str. 42) jsou symboly typu "černá skříňka" - nikdo netuší, co to je, popis chybí. Formalizace celého postupu je proto asi zbytečná, když nelze bez dovysvětlení pochopit, co a jak.
- 5. Formal aspects of technical report** **70 p. (C)**
Úvodní části, až do strany strany cca 31 obsahují jen 3 vážnější chyby, potom se kvalita rapidně zhoršuje. Nepřiměřeně velké (např. 4.5, 4.6, 4.7) obrázky a zbytečnosti (např. všechny obrázky v kapitole 5). Informačně slabé, vše působí jako vycpávka na dostatečný počet stran.
- 6. Literature usage** **95 p. (A)**
Volba pramenů pokrývá danou oblast slušně.
- 7. Implementation results** **50 p. (E)**
Zdrojové kódy nekomentované, původně do IS i na SD kartu odevzdány nekompletní, později (31. 5. 2021 18:07) do IS přidány další zdrojové kódy. Jelikož kódy neobsahují komentáře, takže nelze určit, co je převzaté, co je upravené, co je zcela původní. Tak jak jsou, nejdou přeložit, i když by asi velké části nic moc nebránilo. Nicméně celý kód je funkční, i když rozsahem nijak bohatý, ale do šířky s velkým záběrem. Nízké hodnocení je díky špatné odevzdávce, nekomentování kódů (zdůvodněno firemní politikou, která ale pro DP nemá žádný význam), nejasností o původu.
- 8. Utilizability of results**
Výsledky běží a jsou testovány v rámci fy Avast.
- 9. Questions for defence**
-
- 10. Total assessment** **70 p. good (C)**
Na práci kladně hodnotím nový způsob ukládání, šířku záběru práce, je nasazeno. Text práce i samotná předložená implementace, pozdější zásahy do ní jsou však velkým zklamáním.

In Brno 7 June 2021

Kolář Dušan, doc. Dr. Ing.
reviewer