Posudek oponenta bakalářské práce

Student: Chomo Tomáš
Téma: Identifikace aplikačních protokolů (id 20191)
Oponent: Ryšavý Ondřej, doc. Ing., Ph.D., UIFS FIT VUT

1. Náročnost zadání **obtížnější zadání**

Vzhledem k tomu, že se jedná o bakalářskou práci je zadání obtížnější, neboť vyžaduje pochopení pokročilých principů v oblasti identifikace vzorů a statistické klasifikace.

2. Splnění požadavků zadání **zadání splněno**

3. Rozsah technické zprávy **je v obvyklém rozmezí**

4. Prezentační úroveň předložené práce **60 b. (D)**

Práce obsahuje jednu kapitolu, která popisuje výsledky řešení. Ostatní kapitoly popisují známé algoritmy a metody. Význam kapitoly č.4 mi uniká, neboť informace v ní nejsou dále použity. Také informace v příloze A není příliš relevantní pro samotnou práci. Naopak kapitola 5 by mohla být rozdělena do dvou částí pro lepší pochopení řešení. Celkově je struktura práce nevyvážená, čímž se snižuje pochopitelnost pro čtenáře. Na druhou stranu je nakonec možné relevantní informace potřebné k pochopení s určitým úsilím dohledat.

5. Formální úprava technické zprávy **60 b. (D)**

Anglický abstrakt obsahuje gramatické chyby. Jazykovou stránku vzhledem k tomu, že je psána ve slovenském jazyce nejsem schopen plně posoudit. V textu se místy objevují překlepy a neobvykle formulované věty. Práce obsahuje také typografické chyby, například špatně vysázené odkazy na citované zdroje či přetečení řádku.

6. Práce s literaturou **75 b. (C)**

Literatura obsahuje relevantní zdroje vzhledem k řešenému tématu. Nejméně polovina se však týká ML algoritmů obecně. Citované zdroje z oblasti řešení jsou pak většinou 10+ let staré což znamená, že existují novější přístupy, které by mohly být příslibem lepších výsledků.

7. Realizační výstup **60 b. (D)**

Realizačním výstupem je extrakce nových vlastností pro klasifikátory. Přiložené zdrojové kódy obsahují implementaci extraktoru atributů a implementaci klasifikátorů. Některé zdrojové soubory mají hlavičku, ve které je uveden jako autor vedoucí práce. Kódy vytvořené (pravděpodobně) studentem hlavičku nemají a neobsahují ani komentáře či jinou dokumentaci. Toto může znamenat problém pro jejich další využití. V kódu se také objevuje množství TODO, z nichž některé označují místa, kde patrně dochází k problémům při zpracování. Z toho pohledu působí kód spíše rozpracovaným než dokončeným dojmem.

8. Využitelnost výsledků

Výsledkem práce je rozšíření stávajících metod pro klasifikaci provozu implementovaných v nástoji Netfox Detective o nové vlastnosti. Z výsledků je zřejmé, že rozšíření metody SPID přináší pro určité toky lepší výsledky. Tedy výsledky se zdají být částečně použitelné.

9. Otázky k obhajobě

Z uvedených výsledků a komentáře k nim není příliš zřejmé, jakého zlepšení jste dosáhl:

- Diskutujte prosím dosažené zlepšení pro obě uvedené metody
- Porovnejte prosím výsledky statistické metody s BN
- Diskutujte prosím vliv výběru atributů na klasifikaci, prováděl jste také analýzu atributů s cílem vyloučit nevhodné atributy pro klasifikaci?

10. Souhrnné hodnocení **65 b. uspokojivě (D)**

Práce se zabývala zajímavým tématem návrhu systému pro klasifikaci provozu v nástroji pro forenzní analýzu. Přestože zadání bylo splněno, jsou obě části zpracovány podprůměrně. V textové části jsou uvedené informace kompilátem známých informací, doplněné o popis rozšíření stávajících algoritmů o nové atributy. Vyhodnocení je pak nezbytné minimum pro splnění zadání. Praktická část je sice funkční, ale k použitelnosti vytvořeného řešení je potřeba upravit značnou část kódů, viz komentář k bodu 7.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 30. května 2017

.....
podpis