

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Dominik Smrž
Název práce Feature extraction from Android application packages and its usage in machine learning for malware classification
Rok odevzdání 2017
Studijní program Informatika **Studijní obor** Obecná informatika
Autor posudku Jan Kofroň **Role** Oponent
Pracoviště KDSS

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Komentář Zadání považuji za nešťastně zformulované. Jedná se o jednu větu (souvětí), jejíž význam je i pro mě poněkud nejasný. Celkově tedy nedokážu s určitostí říct, jestli je zadání splněno nebo ne. Práce neobsahuje žádné přílohy (implementaci), ani tabulky, seznamy, či výsledky nad rámec toho, co je uvedeno v textu.				

Textová část práce	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Analýza	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Text je velmi stručný a uvedené výsledky velmi zběžně uvedeny. Autor v závěru uvádí, že „we have discussed several ways to extract and refine API calls...“. Můj dojem po přečtení práce je, že zde uvádí jen jeden způsob (který ještě není nijak zvlášť popsán, extrahovaná volání jsou pak následně zpracovávány statisticky. Práce v závěru uvádí, že hlavní cíl práce bylo vyvinout algoritmus pro detekci malwaru pro systém Android. I když mi nepřipadá, že by tohle byl hlavní bod zadání, v textu práce je algoritmus popsán pouze slovně a to ještě příliš stručně na to, aby bylo možné jej pochopit (to jsem výše označil za nedostatečnou vývojovou dokumentaci, i když to v pravém slova smyslu dokumentace není). Text práce je psán v angličtině, obsahuje lehce nadpřůměrné množství gramatických chyb a překlepů, což ale nesnižuje jeho čitelnost. Práce je prostá jakékoliv analýzy, nebo jsem ji v textu nerozpoznal. Na několika málo místech autor srovnává různé přístupy, ale srovnání je nesystematické a jen slovní. Autor často argumentuje přílišnou časovou nebo paměťovou náročností některých přístupů, aniž uvádí, co to vlastně znamená. Myslím, že v tomhle případě nějaká ta hodina nebo gigabyte skutečně není nepřekonatelná překážka. V práci jsou použité termíny, jako třeba „logistic regression“, „ROC curve“ bez referencí – myslím, že typický čtenář práce termíny znát nemusí. Z toho vyplývá i špatná srozumitelnost některých grafů, například význam obrázků 6.2 a 8.1 mi zůstal utajen.</p>				

Implementační část práce	lepší	OK	horší	nevyhovuje
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Tato část se v práci nenachází, implementace jako taková by asi nebyla případná, ale nějaká příloha s daty by práci doprovázet měla.</p>				

Celkové hodnocení Neprospěl(a) Choose an item.

Práci navrhuji na zvláštní ocenění Ne

Datum 23. srpna 2017

Podpis