Posudek oponenta bakalářské práce

Student: Dižová Natália
Téma: Systém pro testování YARA pravidel (id 23046)
Oponent: Křivka Zbyněk, Ing., Ph.D., UIFS FIT VUT

1. **Náročnost zadání** — **průměrně obtížné zadání**
   Samotné zadání považuji za průměrně náročné, ale jistě s potenciálem na různá rozšíření.
2. **Splnění požadavků zadání** — **zadání splněno**
3. **Rozsah technické zprávy** — **je v obvyklém rozmezí**
   Textová část práce má 52 normostran včetně schémat a autorských obrázků.
4. **Prezentační úroveň předložené práce** — **75 b. (C)**
   Využít formalismus pro popis jazyka pravidel nástrojů Yara a Yarka.
   Občas byla matoucí snaha o stylistickou různorodost, kterou považuji v technických textech spíše za nežádoucí (např. označování kódu jednou jako softvér, jednou jako malware a jindy jako kód). Vyšší technické úrovně by bylo možné dosáhnout i formálnějším popisem jazyka Yara, který údajně není k dispozici ani v interních dokumentacích. Některé pojmy a souvislosti mohly být lépe vysvětleny (např. vztah entity pro "scan" a pro "scan point", dále dynamický sken).
5. **Formální úprava technické zprávy** — **90 b. (A)**
   Typograficky i jazykově je práce na velmi dobré úrovni, obsahuje minimum překlepů a pravopisných chyb.
6. **Práce s literaturou** — **90 b. (A)**
   Práce s literaturou byla na výnikající úrovni. U některých poznámek pod čarou bych raději zvolil klasickou referenci do Literatury (viz str. 8).
7. **Realizační výstup** — **83 b. (B)**
   Implementace je funkční a oceňuji i snahu o její zrychlení. Překvapilo mě, že nebyla implementována také alespoň základní vyrovnávací paměť (knihovny existují), která by značně omezila nutnost pokaždé stahovat celé vzorky ze vzdálených serverů. Zdrojový kód se skládá především z doplnění tříd a metod do systému Yarka v jazyce Python 3, a to v rozsahu cca 4500 řádků (včetně komentářů).
8. **Využitelnost výsledků**
   Dle studentky se již úpravená verze nástroje Yarka používá ve firmě Avast, kde ji využívají analytici při údržbě pravidel sloužících pro detekci malware.
9. **Otázky k obhajobě**
   - Architektura vašeho distribuovaného systému je založená na zprávách (viz str. 16). Proč nebyla využita některé existující architektura, která se zabývá distribuovaným zpracováním nezávislých úloh?
10. **Souhrnné hodnocení** — **82 b. velmi dobře (B)**
    Potenciál tématu mohl být vytěžen ještě více. Návrh rovnou modifikuje nástroj Yarka. Text je velmi pěkný, ale stále se najdou místa, kde bylo vysvětlování těžkopádné nebo zbytečně zdlouhavé. Navrhuji **velmi dobře** (B).

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

V Brně dne: 22. června 2020

Křivka Zbyněk, Ing., Ph.D.
oponent