Posudek na diplomovou práci Davida Matouška
„Metody sledování chování procesů na Windows"

Řešený projekt, který se předložená práce snaží popsat, i struktura této práce (až na kapitolu 7 popisující vlastní projekt) jsou v pořádku, ale...
Celá práce je napsána velmi obecným, neexaktním jazykem – často se jedná o obecně platná tvrzení, která se, mimo jiné, možná vztahují i na popisované problémy a jejich řešení. Způsob vyjadřování a nedokonalá čeština nepřispívají k dobrému dojmu z práce. U mnoha vět je těžké uhádnout jaká byla původní autorova myšlenka – to připomíná spíše mnohé „odborné" blogy na Internetu než diplomovou práci. U mnoha tvrzení chybí alespoň explicitně uvedený odkaz buď na přílohu, nebo na spolehlivý zdroj exaktních doplňujících informací.

str. 5 – „rekapituluje vybrané metody"
– jak kompletní je ten výběr?
– vlastní rekapitulace je nesystematická a nekonkrétní

str. 5 – „... v přesně definovaných stavech jsou na základě aktuálního stavu paměti vytvářeny elementární události sledovacího systému. Tyto události lze dále analyzovat ..."
– ale v celé práci nejsou tyto „přesně definované stavy" definovány vůbec, natož přesně;
– a kde jsou definovány ty „elementární události";
– a kde je ta analýza?

str. 6 – kde je ta slibovaná podpora pro současné verze Windows, když celý systém je „ušit na míru" pro Windows XP SP 3 (která už nejsou z profesionálního hlediska detekce malwareu zajímavá) a nikde není ani zmínka např. o User Account Control nebo Code Access Security nebo .NET Framework Runtime Security Policy.

str. 8 – Object Manager – odstavec srozumitelně vysvětlí význam handle-identifikátorů objektů, ale nevysvětlí, proč je Object Manager pro účely předložené práce důležitý, k čemu a jak se používá.

str. 8 – objekt procesu vs. proces – datová struktura vs. proces, který vykonává nějaké činnosti a jehož vlastnosti jsou zachyceny v té datové struktuře

str. 8 – „podpora HW pro exekuci kódu v jiném režimu než kódu jádra systému"
– jaký je význam této fráze?

str. 9 – „... libovolný program spuštěný uživatelem není omezen svými právy."
– neomezují „práva" ale „omezení";
– opět se jedna o nekompletně vyslovenou myšlenku, je ponecháno na čtenáři, aby si domyslel, co tím chtěl autor vlastně říci;
– přístup k datům uživatele je něco jiného než např. přístup k libovolným datům na disku, špatné nastavení bývá „vynuceno" tím, že některé starší programy zapisovaly individuální uživatelská data do sytémových adresářů (Windows, Program Files,...) a klíčů v registry (HLM) a uživatel (ani jeho Internetoví poradci) nedokáže určit správné nastavení „minimálních potřeb" aplikace, kterou chce používat (viz např. Compatibility Mode ve Windows 7).

str. 9 – věta „S výjimkou sdíleného mapování paměti nebo mapování souboru nemůže kód procesu v uživatelském módu nijak ohrozit bezpečnost systému nebo uživatelská data, ovlivnit chování systému nebo jiných procesů a pro práci se zmíněným mapováním paměti musí také nejdříve využít volání systémových služeb."
– tato věta by si zasložila nejméně celý jeden odstavec, aby se z hlediska bezpečnosti vysvětlil rozdíl mezi

- prací s namapovanou pamětí či souborem,
- vlastní akcí „mapování" a
- mapováním souborů nebo mapováním paměti, což jsou dvě zcela odlišné technologie.

– význam termínu „mapování paměti" – manipulace s tabulkami stránek (jiného) procesu?
– jaká se předpokládá „spolupráce" uživatele se sledovaným kódem pro akce, pro které jsou vyžadována nadstandardní privilegia:
LPVOID *ptr* = **VirtualAllocEx** (*hProcess, adresa, velikost, flag, ochrana*)
– requires PROCESS_VM_OPERATION
– **WriteProcessMemory**
– requires PROCESS_VM_WRITE
– mapování souborů do takto alokované paměti?

str. 9 – nepřesné a příliš neurčité pojednání o knihovnách funkcí – např. adresy požadovaných funkcí se do tabulek ukládají při vytváření procesu jen vyjímečně, např. pokud nelze při mapování DLL knihovny uspokojit její požadavek na bázovou adresu

str. 10 – „**Může** obcházet veškerá zabezpečení definovaná systémem práv ..." – jak se „systém práv" kódem běžícím na ringu 0 obchází?

str. 10 – protože Windows standardně používaji jen ring 0 a 3, chybí vysvětlení, proč je zde zmínka i o Ringu 1 a 2.

str. 10 – zmínka o možnosti vykonávat kód v režimu jádra i bez ovladače je natolik silným tvrzením, že nutně vyžaduje alespoň odkaz na konkrétní popis

str. 10 – „Některé aplikace vyžadují přístup k jádru systému a systém jim to umožňuje prostřednictvím ovladačů."
– co je „přístup k jádru systému"? – ten je snad umožněn voláním API funkcí
– pomocí jakých ovladačů jim to systém umožňuje?

str. 11 – „nebo se jinak propracoval k režimu jádra"
– upřesnit jak jinak

str. 11 – „jen velmi málo legitimních programů vyžaduje práci s ovladačem"
– opět volná terminologie
– to chce rozlišit „manipulaci s ovladačem" a „využívání služeb ovladače"

str. 11 – čeština: „chování kódu běžícím v režimu jádra"

str. 32 – chybí reference na bezpečnostní systémy, které se snažily neúspěšně o sandboxing, a na analýzy a důkazy této neúspěšnosti

str. 32 – systémové omezení práv procesu – co to skutečně znamená a jak se to dělá?

str. 43 – Základní vlastnosti systému – objekty jádra – je někde uveden konkrétní přehled objektů které jou sledovány a **co** a **proč** je na nich sledováno? Způsob výběru těchto objektů a seznam vybraných objektů by měly tvořit skutečné jádro této práce.

str. 44 – detekce virtuálního prostředí – pouze marketingový popis – pro které způsoby detekce virtualizace je systém PABOV neviditelný a které by ho naopak odhalily

str. 44 – co je stav operačního systému uložený v externím souboru (datbáze stavu) – jaké informace obsahuje – viz str. 46 – jaký byl způsob a postup při vytváření databáze

str. 45 – sledování 130 adres – o jaké adresy se jedná, jak byl proveden výběr

str. 45 – „Na rozdíl od tradičního přístupu, nelze v případě PABOVu dynamicky rozpoznat adresy v neznámém jádře ani lokace důležitých parametrů."
– co tím chtěl autor říct?
– proč to nelze?

str. 46 – chybí přesnější určení (identifikátory, způsob jejich vyhledání) tabulek, o kterých je řeč, popř. grafické znázornění jejich struktury

str. 46 – „PABOV má k dispozici kompletní seznam všech událostí, které ve virtualizovaném systému nastaly od jeho startu."
– opět marketinově přehnané nekonkrétní tvrzení
– kde PABOV ten seznam vzal?

str. 48 – zobrazení množství informací o registrační databázi by mělo být ponecháno na uživateli

str. 48 – „Aktuální pokrytí na výstupu"
– tomuhle titulku nerozumím, jedná se o část nějaké fráze vytrženou z kontextu

str. 49 – slabiny konkurenčních systémů
– zcela nekonkrétní komerční odsouzení konkurence,
– chybí jakékoliv konkrétní údaje nebo alespoň odkazy na specifické přílohy
– některé nedostatky vytýkané konkurenčním produktům má ale PABOV také

str. 49 – struktura výstupu – text je pěkným doplňujícím vysvětlením nějaké stručné a přehledné tabulky, která ale chybí

str. 50 – chybí odkaz na přílohu, je ponecháno na čtenáři, jestli ho napadne ji vůbec hledat

str. 75 – Srovnání sledovacích systémů
– chybí shrnutí ve formě přehledné tabulky

Jaký je vztah PABOVu ke skutečně profesionálním produktům (antiviry, antimalwary, ...), které buď také provádějí takové analýzy lokálně, včetně vyhodnocení, nebo využívají centrální databáze výsledků podrobných analýz,resp. obojí dohromady, a v případě podezření provedou nezbytná opatření?

Doporučuji předloženou práci k obhajobě s výhradou, že bude doplněna o přílohu – tabulku - s přehledem sledovaných funkcí systému.

.....

V Praze, 4.1.2010

RNDr. Vojtěch J. Jákl