ANALYSIS OF AUTOMATED GENERATION OF SIGNATURES USING HONEYPOTS

Lukáš Bláha

Master Degree Programme (2), FIT BUT

E-mail: xblaha04@stud.fit.vutbr.cz

Supervised by: Michal Drozd

E-mail: idrozd@fit.vutbr.cz

Abstract: One of the biggest security issues in today's computer world is the rapid-growing number of newly created malicious software. The process of updating security mechanisms based on signatures stands on the manual creation of these patterns which extends the reaction time to defend against newly discovered attacks. This paper proposes an analysis, implementation and testing of existing methods of automated generation of signatures using deception systems called Honeypots. The main objective of this work is to suggest improvements to the existing method of automatic generation of signatures which was created under scientific project AIPS at FIT BUT.

Keywords: malware, automatic generation of signatures, honeypot, IDS, signature

1. ÚVOD

V dnešní době jsou počítačové systémy používány ve všech odvětvích komerční i osobní sféry a je na nich uloženo velké množství citlivých dat. Tato data jsou velmi lákavá pro počítačové zločince, kteří za účelem získání důvěrných dat denně vyprodukují obrovské množství nových škodlivých kódů. Tyto kódy jsou den ode dne sofistikovanější a díky síti Internet dokáží infikovat počítačové systémy po celém světě během několika minut [1].

Současný postup aktualizace anti-malwarových systémů založených na signaturách je postaven na manuální tvorbě těchto signatur pro každou nově nalezenou zranitelnost. Tento proces je při tak velkém množství nového škodlivého software značně pomalý. Malware dokáže v časovém intervalu od objevení zranitelnosti po vydání aktualizace bezpečnostních komponent kompromitovat tisíce až miliony počítačových systémů [1].

Tato práce se věnuje metodám automatického generování signatur (detekčních profilů) s využitím honeypotu. Vysvětlení pojmu honeypot je k nalezení v [2]. Automatizované generování signatur by mělo zajistit rychlejší doručení aktualizací do bezpečnostních zařízení, čímž by měla být snížena reakční doba na útok a počet systémů infikovaných novým škodlivým kódem by měl klesnout.

2. POPIS EXISTUJÍCÍCH METOD

V rámci této práce bylo analyzováno 13 existujících metod automatické tvorby signatur. Při získávání informací o metodách bylo čerpáno především z [3]. U každé metody byly popsány vstupní parametry a výstup, který daná metoda generuje. Detailně byl analyzován především samotný mechanizmus použitý pro generování signatury. Vybrané metody byly rozděleny do následujících kategorií podle toho, jaké informace daná metoda používá k tvorbě signatur.

2.1. CONTENT-BASED SIGNATURY

Content-based signatury jsou široce používány systémy pro automatické generování signatur. Signatura má následující formát:

```
(číslo IP protokolu, cílový port, sekvence bajtů)
```

Hlavním úkolem je identifikace charakteristického řetězce, který se nachází v těle škodlivých síťových paketů. Detekční systémy poté pouze porovnají každý procházející paket s řetězcem uvedeným v signatuře, a pokud je nalezena shoda, je paket označen jako škodlivý. Tento způsob je často používán díky své jednoduché implementaci a rychlosti. Hlavním problémem tohoto přístupu je jeho statická povaha. Pokud bude v těle paketu pozměněn nějaký znak, nebude detekční systém schopen tento paket identifikovat jako škodlivý.

V této kategorii byly analyzovány metody Honeycomb, Polygraph, Earlybird, Autograph a Taint-Check.

2.2. FLEXIBILNĚJŠÍ CONTENT-BASED SIGNATURY

Tyto metody pracují s řetězci bajtů podobně jako metody popsané v předchozí podkapitole. Jsou však flexibilnější, jelikož se nepokouší pouze porovnat řetězce nebo podřetězce v příchozích paketech. Signatury generované těmito technikami popisují vzory, které berou v úvahu uspořádání škodlivých bajtů v paketu. Některé techniky používají například regulární výrazy.

Do této kategorie patří analyzované metody PAYL a PADS.

2.3. METODY POUŽÍVAJÍCÍ KONTEXT A SÉMANTIKU APLIKACE

Metody uvedené doposud vytvářely signatury na základě analýzy řetězce bajtů. Techniky zařazené do této kategorie využívají sofistikovanější postupy. K tvorbě signatur používají znalosti aplikačních protokolů a dokáží tak rozhodnout, v jakém stavu se musí aplikace nacházet, aby byl škodlivý kód odhalen.

Tato kategorie obsahuje analyzované metody Nemean a COVERS.

2.4. OSTATNÍ PŘÍSTUPY

Metody zahrnuté do této kategorie nejsou klasifikovány jako systémy pro automatické generování signatur. Přináší však zajímavé nápady a postupy, které by mohly být použity pro optimalizaci metod pro automatické generování signatur.

V této kategorii byly analyzovány metody Paid, Vigilante, DOME a HoneyStat.

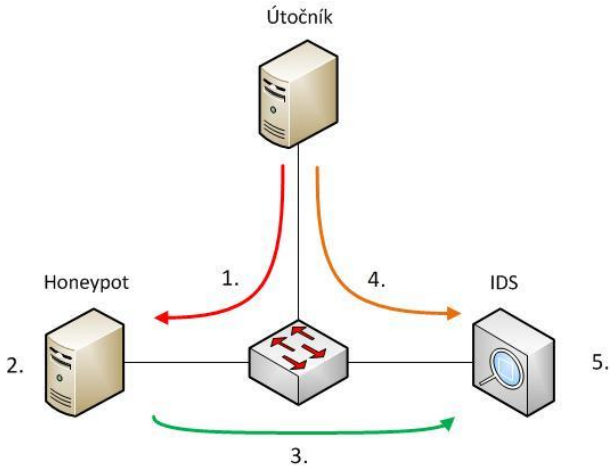
3. NÁVRH ANALÝZY METOD

Pro potřeby testování existujících metod bylo vytvořeno virtualizované testovací prostředí. Při tvorbě tohoto prostředí byl brán ohled především na jeho jednoduchost, proto byly vybrány pouze tři nezbytně nutné systémy. Vytvořené testovací prostředí je zobrazeno na obrázku 1.

Jako první krok testování je proveden útok ze stanice útočník na stanici honeypot. Zde musí být škodlivý provoz detekován a předán subsystému, který automaticky vygeneruje jeho signaturu. Vygenerovaný vzorek je poté v kroku 3 distribuován na systém IDS (Intrusion Detection System). K tomuto účelu byl vytvořen skript, který přenos signatury k detekčnímu systému automatizuje. Poté na systému IDS dojde k aktualizaci databáze signatur. Ve čtvrtém kroku je proveden stejný útok jako v kroku jedna, tentokrát však vedený ze stanice útočník na stanici IDS. V posledním kroku je vyhodnoceno, zda IDS systém dokázal pomocí vytvořené signatury detekovat provedený útok.

K útokům na jednotlivé systémy byl použit Metasploit Framework (http://www.metasploit.com), který obsahuje rozsáhlou databázi exploitů, což umožňuje vyzkoušet velké množství útoků v poměrně krátkém časovém horizontu a je tak usnadněno celé testování. Jako IDS systém poslouží open source řešení Snort (http://www.snort.org), které je hojně využíváno i v produkčních prostředích.

Kvalita generovaných signatur je kromě počtu úspěšně detekovaných útoků dána také nízkým (nejlépe nulovým) počtem falešných poplachů (false positives). Pro otestování této vlastnosti bude na systém IDS zasílán legitimní provoz a bude zaznamenán počet vzniklých falešných poplachů.



Obrázek 1: Schéma navrženého testovacího systému.

Při posuzování kvality otestovaných metod automatického generování signatur bude nejvíce záležet na počtu úspěšně detekovaných útoků a na co nejnižším počtu vygenerovaných false positives. Jako další kritérium bude brán čas, za který metoda dokáže vygenerovat výslednou signaturu.

4. ZÁVĚR

V této práci byla provedena analýza existujících metod automatického generování signatur pro škodlivé kódy. Vybrané metody byly implementovány do vytvořeného testovacího prostředí. Po dokončení testovací fáze budou zpracovány získané výsledky a budou vyhodnoceny nejefektivnější metody automatické tvorby signatur. Na základě výsledků z analýzy vybraných metod budou navrženy optimalizace pro existující metodu vytvořenou pro projekt Automatizované zpracování útoků (http://www.fit.vutbr.cz/research/view_project.php.cs?id=465), případně bude navržen zcela nový způsob automatického generování signatur.

PODĚKOVÁNÍ

Příspěvek vznikl v rámci výzkumného záměru Výzkum Informačních technologií z hlediska bezpečnosti (MSM0021630528). Práce byla podporována operačním programem Výzkum a vývoj pro inovace v rámci projektu Centrum excelence IT4Innovations (CZ.1.05/1.1.00/02.0070) a projektem Pokročilé bezpečné, spolehlivé a adaptivní IT (FIT-S-11-1).

REFERENCE

[1] Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S., Weaver, N.: *The Spread of the Sapphire/Slammer Worm.* Dostupné na URL: `http://www.caida.org/publications/papers/2003/sapphire/sapphire.html`

[2] Provos, N., Holz, T.: *Virtual Honeypots: From Botnet Tracking to Intrusion Detection.* Addison Wesley, 2007, ISBN: 0-321-33632-1

[3] Waraich, R.: *Automated Attack Signature Generation: A Survey*, ETH Zurich, 2005, Technical Report