DETECTION OF OPERATION SYSTEMS IN NETWORK TRAFFIC USING IPFIX

Martin Vymlátil

Bachelor Degree Programme (3), FIT BUT

E-mail: xvymla01@stud.fit.vutbr.cz

Supervised by: Petr Matoušek

E-mail: matousp@fit.vutbr.cz

Abstract: This task deal with detection of operation system in network traffic using IPFIX. The main idea of this task is based on the fingeprinting, when we use information from IP and TCP headers to determine operation system.

Keywords: detection, operation system, fingerprinting.

1. ÚVOD

V dnešní době, kdy neustále roste počet operačních systémů, ale také počet chytrých telefonů, není určitě na škodu, např. pro administrátora sítě, vědět, kdo a s jakým operačním systémem pracuje. Tuto informaci můžeme poté využít nejen při rozvoji sítě, ale také při nákupu nového softwaru, při správě počítačů v síti, např. pro nasazení nového OS nebo přechodu na novější verzi, nebo můžeme detekovat neautorizované a nebezpečné zařízení v naší síti, atd. Cílem práce je plugin pro sondu FlowMon společnosti INVEA-TECH, který bude rozšiřovat její funkčnost o detekci OS, a s tím související rozšíření IPFIX o informaci o operačním systému.

2. DETEKCE OPERAČNÍCH SYSTÉMŮ

Detekce operačních systémů v síti probíhá na základě tzv. OS fingerprintingu, což je proces, který pomocí kombinace parametrů a informací zjištěných v síťovém provozu dokáže určit příslušný OS. Téměř všechny techniky fingerprintingu jsou založeny na detekci odlišností v paketech zasílaných různými OS. Tyto techniky běžně analyzují IP datagram (konkrétně hodnoty TTL a ID), TCP protokol (Window size a příznaky SYN a SYN+ACK), DHCP request a ICMP request. Některé techniky pak využívají běžicí služby nebo otevřené porty.

Rozlišujeme aktivní a pasivní fingerprinting [1] [2].

Pasivní fingerprinting, jehož zástupce je například nástroj p0f, monitoruje síťový provoz bez vytváření a zasílání speciálních paketů a analyzuje pouze zachycené pakety cílového hostitele. Jak hlavička IP datagramu, tak TCP protokolu nemá konkrétní implicitní hodnoty, tyto hodnoty jsou většinou doporučené. Výhodou pasivních metod je, že jsou prakticky nezjistitelné, nejeví žádnou aktivitu, pouze analyzují hlavičky průchozích paketů. Jejich nevýhoda spočívá v tom, že nemusí mít všechny potřebné informace pro určení operačního systému, z tohoto důvodu je u těchto metod nutné zavést určité heuristiky.

Ovšem, různé OS se projevují odlišnostmi v určitých hodnotách IP a TCP hlavičky.

2.1. VLASTNÍ DETEKCE OPERAČNÍCH SYSTÉMŮ

Detekce OS probíhá na základě pasivního fingerprintingu, který využívá hodnot z IP a TCP protokolu. TCP protokol se objevuje v síťovém provozu poměrně často, ovšem je nutné, aby měli tyto pakety příznaky SYN nebo SYN+ACK.

Kromě typických hodnot používaných při pasivním fingerprintingu můžeme k určení operačního systému dále využít také velikost SYN paketu[3] a hodnot z TCP Options – počet NOP bytů (NO operation)[3] a při dalším zkoumání hodnot v poli Options si můžeme všimnout, že operační systémy se liší také v hodnotě Window Scaling.

Kombinací příslušných hodnot z IP a TCP protokolu můžeme určit, o jaký se jedná operační systém.

2.2. ODLIŠNÉ HODNOTY PŘI SKUTEČNÉM MĚŘENÍ

Bohužel nelze explicitně říci, jaké hodnoty má která položka v IP a TCP protokolu pro konkrétní operační systém. Ve skutečnosti se tyto hodnoty mohou lišit na základě verze OS nebo konkrétní distribuce. Z toho důvodu, nemůže OS určit pouze na základě jedné nebo dvou hodnot. Tučně označené hodnoty v tabulce 1 jsou ty, které jsou odlišné od hodnot udávaných v literatuře a zdrojích.

	Windows XP	Windows 7 Home	Windows 8 Prof	Ubuntu, Redhat	FreeBSD	MAC OS
Time to live	128	**64**	128	64	64	64
SYN packet size	48	52	52	60	60	64
Maximum segment size	**1460**	**1452, 1460**	**1460**	1460	1460	1460
Window size	65535	8192	8192	**14600**	65535	65535
Selective acknowledge	Set	Set	Set	Set	**Set**	Not set
No operation	2x	3x	3x	1x	1x	3x
Window scale	None	2	8	3	7	4

Tabulka 1: Zjištěné skutečné hodnoty z IP a TCP hlaviček.

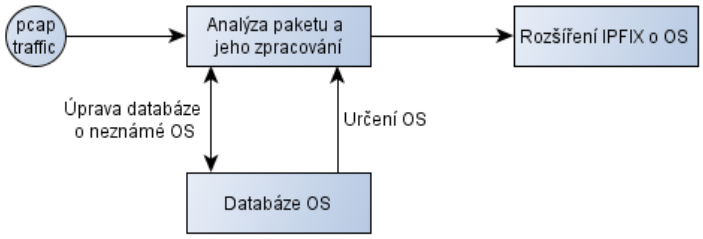
2.3. NÁVRH DETEKCE OPERAČNÍCH SYSTÉMŮ

Při návrhu detekce OS bylo nutné dodržet technologii a architekturu sondy FlowMon. Tato sonda je neviditelná pro ostatní zařízení na L2 a L3 vrstvě a proto je použit princip pasivního fingerprintingu pro zjišťování OS. Návrh je pak rozdělen do dvou částí. První část se skládá ze získání potřebných informací a to probíhá na úrovni vstupního pluginu. V druhé části, která se děje na úrovni procesního pluginu, dochází ke zpracování dat a vlastnímu určení operačního systému. Proto nelze využít žádný z dostupných nástrojů pro detekci OS. Poté dochází k exportování záznamů ze sondy na kolektor. Je také nutné rozšířit IPFIX, konkrétně šablonu pro ipfixcol. Ipfixcol sice přijme takto rozšířený záznam, ale neví, jak s ním dále pracovat.

Detekce operačních systémů může probíhat ze souboru (pcap soubory) nebo na základě monitorování provozu na síťovém rozhraní. Po zachycení paketu dojde k jeho analýze, kdy získáme potřebné informace pro určení OS. Potřebujeme především tyto informace:

- IP: Time to live, SYN packet size.
- TCP: Window size, Maximum segment size, Window scale, No operation, Selective akcnowlidge.

Získané informace jsou následně porovnány s databází operačních systémů. Pro správné určení OS je nutné, aby se získané informace shodovali s databází minimálně v 5 hodnotách. Pokud ovšem neznáme OS, i přes dostatek dat, můžeme získat informaci o OS např. z položky User agent protokolu HTTP a rozšířit tak databázi operačních systémů. IPFIX je rozšířen pouze o konkrétní operační systém.



Obrázek 1: Zjednodušené schéma detekce OS.

2.4. TESTOVACÍ PROGRAM

Pro zjednodušení práce a lepší zpracování zdrojových souborů vznikl program, který zpracovává pcap soubory. Paket po paketu prochází hlavičky a získává potřebné informace a následně určí, o jaký se jedná OS. Operační systém vždy přiřadí k IP adrese. Pokud nemá dostatek informací, nebo se zjištěná data neshodují s žádným operačním systémem v databázi, program určí jako OS Unknown, tedy neznámý OS.

```
IP: 173.194.113.68      OS: Unknown     ttl: 50  df: 2  synsize 0       win size 0
IP: 173.194.70.125      OS: Unknown     ttl: 44  df: 2  synsize 0       win size 0
IP: 192.168.20.1        OS: MAC OS      ttl: 64  df: 1  synsize 64      win size 65535
IP: 31.13.81.128        OS: Windows 8   ttl: 82  df: 1  synsize 60      win size 14480
IP: 81.91.84.180        OS: Free BSD    ttl: 53  df: 1  synsize 60      win size 5792
IP: 173.194.113.78      OS: Unknown     ttl: 50  df: 2  synsize 60      win size 42540
IP: 68.232.35.139       OS: Free BSD    ttl: 50  df: 1  synsize 60      win size 14480
IP: 185.31.17.185       OS: Free BSD    ttl: 47  df: 1  synsize 60      win size 14280
IP: 173.194.70.84       OS: Unknown     ttl: 43  df: 2  synsize 60      win size 42540
IP: 8.37.70.21  OS: Free BSD    ttl: 45  df: 1  synsize 60      win size 14480  max ss
IP: 23.209.127.139      OS: Free BSD    ttl: 52  df: 1  synsize 60      win size 14480
IP: 23.62.237.104       OS: Free BSD    ttl: 54  df: 1  synsize 60      win size 14480
IP: 23.209.114.110      OS: Free BSD    ttl: 52  df: 1  synsize 60      win size 14480
```

Obrázek 2: Ukázka výstupu z testovacího programu.

3. ZÁVĚR

Pro určení operačního systému musíme znát minimálně 5 hodnot z tabulky 1, v ideálním případě, všech 7 hodnot. V případě, že nemůžeme rozhodnout, o jaký se jedná OS, poslouží nám hodnota Window scale, která se liší u všech sledovaných OS.

REFERENCE

[1] Sanders Ch.: Practical Packet Analysis using Wireshark to solve real-word network problems. No Starch Press, druhé vydání, červenec 2011: s. 194-197, ISBN: 978-1-59327-266-1.

[2] Forensic Wiki: OS fingerprinting. [online], [cit. 2013-10-22]. URL http://www.forensicswiki.org/wiki/OS_fingerprinting

[3] Krmíček, V.: Hardware-Acceleratd Anomaly Detection in High-Speed Networks. [online], 2011, [cit. 2013-12-17].