COMPUTER IDENTIFICATION BASED ON PACKET'S TIMESTAMPS

Jan Novotný

Bachelor Degree Programme (4), FIT BUT

E-mail: xnovot38@stud.fit.vutbr.cz

Supervised by: Jan Kaštil

E-mail: ikastil@fit.vutbr.cz

Abstract: This work describes perspective identification technique called computer identification based on packet's timestamps of TCP. This technique describes the calculation of the clock skew of the computer's clock. Finally, we will use this technique to identify a number of computers in the real computer network and evaluate the results and possible extensions.

Keywords: Computer identification, timestamps, clock skew.

1. ÚVOD

Tento příspěvek se zabývá identifikací počítače pomocí časových značek paketů a to konkrétně pomocí časových značek transportního protokolu TCP [3]. Cílem práce je implementovat níže zmíněnou techniku pro identifikaci počítače, ověřit funkci implementované aplikace v reálné počítačové síti, vyhodnotit dosažené výsledky a uvést možná rozšíření.

Technika vzdáleného snímání počítače v síti označovaná jako *remote physical device fingerprinting* [1] byla představena v roce 2005 (Kohno, Broido, Claffy). Pomocí této techniky můžeme sledovat zařízení, které může být jakkoli vzdálené od našeho měřícího zařízení a bez toho aniž by o tom snímané zařízení vědělo nebo jakkoliv spolupracovalo s naším měřícím zařízením. Technika se zakládá na výpočtu zkreslení hodin počítače. Zkreslení hodin počítače je mikroskopická odchylka v hardwaru počítače. Vypočítané zkreslení hodin nám slouží jako unikátní identifikátor počítače.

2. URČENÍ ZKRESLENÍ HODIN POČÍTAČE

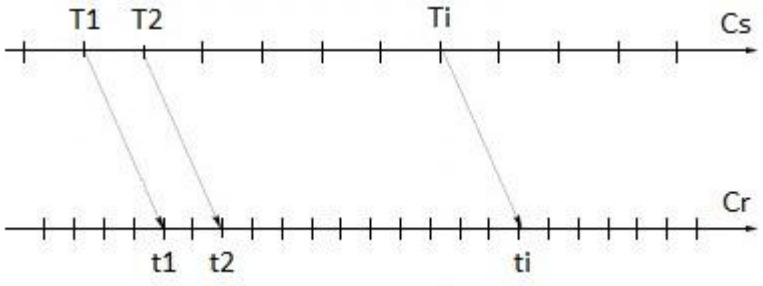
V této sekci uvedeme, jak spočítat zkreslení hodin v reálné počítačové síti mezi dvěma počítači. Tyto dva počítače si označíme jako, příjemce (náš počítač, ze kterého provádíme měření) a odesílatel (měřený počítač). Každý z těchto počítačů obsahuje své vlastní hodiny, které si označíme C_s jako hodiny odesílatele a C_r jako hodiny příjemce, kde hodnoty hodin odesílatele reprezentují časové značky protokolu TCP $C_s(t)$, a hodnoty hodin příjemce zase aktuální dobu přijetí paketu od odesílatele $C_r(t)$. Pokud tedy získáme záznam síťového provozu mezi dvěma počítači, obsahující časové značky protokolu TCP [3], jsme schopni z tohoto záznamu získat hodnoty časových značek TCP protokolu a hodnoty časů přijetí paketů.

2.1. VÝPOČET FREKVENCE HODIN

Proto, abychom mohli vypočítat zkreslení hodin, musíme nejprve znát frekvenci hodin odesílatele. Frekvence hodin je první derivace funkce hodin v čase $C_s'(t)$ [2]. Funkce hodin je lineární funkce ve tvaru $y = ax + b$, kde y odpovídají hodnoty časových razítek protokolu TCP T_i (v_i) a x odpovídají hodnoty aktuální doby přijetí paketu t_i (x_i). První derivací této funkce tedy dostaneme tvar $y' = a$, kde a [Hz] odpovídá frekvenci hodin odesílatele f_s.

2.2. VÝPOČET ZKRESLENÍ HODIN

Obrázek 1 znázorňuje, jak odesílatel generuje časová razítka T_i. Časová razítka vkládá do paketů, konkrétně do volitelné položky TCP segmentu, které následně odešle příjemci.



Obrázek 1: Znázornění hodin příjemce a odesílatele [2].

- $\mathbf{C_s}$: hodiny odesílatele (sledovaného zařízení).
- $\mathbf{C_r}$: hodiny příjemce (sledujícího zařízení).
- **N**: počet paketů, které přijme příjemce od odesílatele.
- **|N|**: počet paketů obsahující časové značky, které přijme příjemce od odesílatele.
- $\mathbf{t_i}$: je čas v sekundách, kdy příjemce obdržel i-tý paket od odesílatele, kde $i \in \{1,\ldots, |N|\}$.
- $\mathbf{T_i}$: je časové razítko protokolu TCP vložené odesílatelem do i-tého paketu, kde $i \in \{1,\ldots, |N|\}$.

Kohno, Broido a Claffy [1] představili následující postup pro výpočet zkreslení hodin pomocí časových značek paketů protokolu TCP:

$$x_i = t_i - t_1 \qquad (1)$$
$$v_i = T_i - T_1 \qquad (2)$$
$$w_i = v_i / f_s \qquad (3)$$
$$y_i = w_i - x_i \qquad (4)$$
$$\sigma = \{(x_i, y_i) : i \in \{1,\ldots, |N|\}\} \qquad (5)$$

Postupným aplikováním vzorců (1), (2), (3) a (4) získáme množinu bodů (5), která obsahuje čas zachycení paketu (x_i) a hodnotu rozdílu (offset) hodin (y_i). Pokud těmito body proložíme přímku a získáme směrnici vzniklé přímky, potom směrnici přímky odpovídá zkreslení (skew) hodin. Zkreslení hodin je tedy náš unikátní identifikátor sledovaného zařízení (počítače).

3. DOSAŽENÉ VÝSLEDKY

V tabulce 1 jsou zobrazené výsledky, kdy jsme testovali tři různé počítače. Daný počítač byl vždy propojen s testovacím notebookem a to třemi různými způsoby zapojení. Testy probíhaly tak, že z daného počítače jsme stahovali vždy stejný soubor. Tuto síťovou komunikaci jsme zaznamenávali pomocí programu tcpdump a na zachycený soubor v PCAP formátu jsme spustili námi vytvořený program pro výpočet zkreslení hodin počítače.

Pokud se podíváme na výsledky měření pro PC1 a PC2, tak pro různá zapojení došlo ke změně zkreslení hodin v rozsahu max. 4,5 %. Testy pro PC3, které mělo výrazně nižší frekvenci hodin, nevyšly příliš shodně. Při zachycení mnohem většího množství paketů, kdy jsme testovali vliv množství zachycených paketů na zkreslení hodin, se zkreslení hodin pro hodiny s frekvencí 100Hz od milionu zachycených paketů měnilo v rozsahu max. několika desetin PPM. Pro hodiny s frekvencí 10Hz se zkreslení hodin ustálilo až po zachycení více jak pěti milionů paketů.

číslo testu	počítač	zapojení	počet paketů obsahující časové značky	frekvence hodin odesilatele [Hz]	zkreslení [PPM]
1	windows 7 - PC1	přímo síťovým kabelem	1 122 349	100	75.3972
2	windows 7 - PC1	PC - router - NB	1 171 080	100	75.4674
3	windows 7 - PC1	PC - router - router - NB	1 171 083	100	72.8160
4	windows 7 - PC2	přímo síťovým kabelem	1 122 348	100	50.9337
5	windows 7 - PC2	PC - router - NB	1 122 347	100	50.4108
6	windows 7 - PC2	PC - router - router - NB	1 171 082	100	50.4794
7	windows XP - PC3	přímo síťovým kabelem	1 169 352	10	153.6755
8	windows XP - PC3	PC - router - NB	1 171 083	10	147.6430
9	windows XP - PC3	PC - router - router - NB	1 171 083	10	161.9558
10	windows XP - PC3	přímo síťovým kabelem	5 057 551	10	91.6276
11	windows XP - PC3	PC - router - NB	5 061 733	10	89.4187
12	windows XP - PC3	PC - router - router - NB	5 063 623	10	89.1852

Tabulka 1: Hodnoty zkreslení hodin pro různé počítače.

4. ZÁVĚR

Vzhledem k dosud dosaženým výsledkům, lze pomocí této techniky identifikovat počítač v síti. Ovšem uvedené testy jsou pouze předběžné a v rámci pokračování práce je třeba provést měření na více počítačích. Technika identifikace počítače by se dále mohla využít pro zjištění počtu počítačů nacházejícími se za NATem. Dále by se tato technika mohla použít pro ověření, zda daný tok přichází z daného počítače nebo se jedná o podvrh.

PODĚKOVÁNÍ

Chtěl bych poděkovat svému vedoucímu, panu Ing. Janu Kaštilovi, za jeho vůli a ochotu diskutovat se mnou všechny problémy, na které jsem při vypracovávání narazil.

REFERENCE

[1] KOHNO, Tadayoshi, Andre BROIDO a K. C. CLAFFY. *Remote physical device fingerprinting*. San Diego, May 2005. IEEE Symposium on Security and Privacy 2005. Department of Computer Science & Engineering, University of California.

[2] MOON, Sue B., Paul SKELLY a Don TOWSLEY. *Estimation and Removal of Clock Skew from Network Delay Measurements*. Amherst, 1998. Technical Report 98-43. Department of Computer Science University of Massachusetts.

[3] JACOBSON, V., R. BRADEN a D. BORMAN. *TCP extensions for high performance*. RFC 1323, May 1992.