COMPUTER IDENTIFICATION USING TIME INFORMATION

Jakub Jirásek
Master Degree Programme (3), FIT BUT
E-mail: xjiras02@stud.fit.vutbr.cz

Supervised by: Libor Polčák
E-mail: ipolcak@fit.vutbr.cz

Abstract: This work deals with the identification of a remote computer by monitoring TCP timestamps of the tracked device. It is possible to determine computer's clock skew from these timestamps as the clock skew is unique for every device. Passive data capturing ensures that the identification process is invisible to the fingerprinted computer. It is necessary that the network communication of fingerprinted computer is visible to the observing device.

Keywords: Computer identification, device's clock, clock skew, TCP timestamps

1 ÚVOD

Všechna síťová zařízení mají vlastní vnitřní hodiny, jejichž takt udává oscilátor řízený krystalem. Vnitřní perioda těchto hodin však není úplně přesná, vliv na ni mají okolní teplota, vlhkost a především samotný krystal. Vnitřní hodiny tak vykazují drobný, ale měřitelný posun oproti reálnému času [1]. Za předpokladu, že je tento posun v rámci jednoho počítače s časem konstantní a současně je tato hodnota dostatečně rozdílná v rámci různých počítačů, mohl by být posun vnitřních hodin počítače využit k jeho identifikaci [2]. Nezáleží přitom na tom, kde se počítač nachází, jakou má sítovou adresu, zda je připojený za směrovačem a využívá překladu síťových adres (NAT) nebo jaký operační systém na daném počítači běží.

2 ZÍSKÁNÍ ČASOVÝCH INFORMACÍ

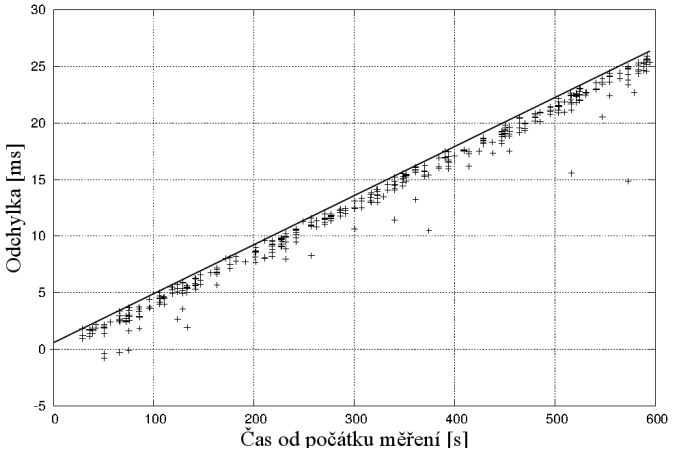
Jako nejvhodnější zdroj časových informací pro získání odrazu vnitřních hodin sledovaného počítače je možné použít časová razítka TCP, která jsou součástí hlavičky TCP. Hodnota *TCP timestamp* není ovlivněna časovou synchronizací protokolem NTP [2], je generována přímo z vnitřních hodin počítače a pouhým pasivním odposlechem umožňuje získání velkého množství informací z komunikace nad protokolem TCP. K čemu slouží a jak s ním může být nakládáno popisuje RFC 1323.

3 URČENÍ POSUNU HODIN Z ČASOVÝCH RAZÍTEK

K určení posunu vnitřních hodin sledovaného počítače nejprve potřebujeme získat množinu paketů TCP obsahujících časová razítka podle doporučení RFC 1323. Z této množiny vytvoříme posloupnost, seřazenou podle času přijetí jednotlivých paketů. Tuto posloupnost nazveme T. Nechť t_i je čas, kdy sledující počítač zachytil i-tý paket z posloupnosti T, platí tedy formule (1).

$$\forall i \in \{1, \ldots, |T|\} . t_i \geq t_{i-1} \tag{1}$$

Na posun vnitřních hodin se také můžeme dívat jako na rychlost změny odchylky. Tato změna může být jak kladná (hodiny jdou napřed), tak záporná (hodiny se opožďují). Pro každé získané časové razítko určíme jeho odchylku od reálného času. Jedinou komplikací je zde rozdílná frekvence generování časových razítek u různých operačních systémů, abychom mohli pracovat s odchylkami



Obrázek 1: Graf posloupností odchylek časových razítek TCP. Plnou čárou je znázorněn hledaný posun hodin.

v sekundách (resp. milisekundách), potřebujeme tuto frekvenci vypočítat (2) a časová razítka získanou hodnotou podělit.

$$f = \frac{\mathrm{d}T}{\mathrm{d}t}[Hz] \tag{2}$$

Dostaneme tak posloupnost odchylek O_T, náležící k získané posloupnosti paketů T a seřazenou podle času přijetí

$$O_T = ((x_i, y_i) : i \in \{1, \ldots, |T|\}), \tag{3}$$

kde x_i je relativní čas přijetí paketu a $y_i = (T_i - T_1)/f - x_i$ jeho odchylka.

Pro výpočet posunu hodin může být použita metoda založená na lineárním programování [1]. Vychází z faktu, že posun dvou porovnávaných hodin je s časem konstantní. Zpoždění jednotlivých paketů se sice může měnit (dojde k zahlcení linky, jinému směrování apod.), při dostatečně dlouhém měření je ale právě lineární změna naměřených odchylek známkou konstantního posunu obou hodin. Chceme-li zjistit rychlost změny odchylek, proložíme body přímkou a změříme úhel, který tato přímka svírá s osou x (obrázek 1).

K proložení grafu je použita lineární funkce, shora ohraničující množinu všech odchylek. Výsledná hodnota není zatížena chybami způsobenými zpožděnými pakety a kopíruje nejrevelantnější hodnoty

$$\alpha \cdot x_i + \beta \geq y_i, i \in \{1, \ldots, |T|\}, \tag{4}$$

kde α určuje sklon přímky a β posun na ose y.

Druhou důležitou vlastnost, kterou funkce musí splňovat, je minimální vertikální vzdálenost všech bodů posloupnosti odchylek O_T od této funkce [2]

$$min\{\sum_{i=1}^{|T|}(\alpha \cdot x_i + \beta - y_i)\}. \tag{5}$$

Hledaný posun vnitřních hodin sledovaného počítače je pak hodnota α, určující sklon této lineární funkce. Na obrázku 1 je zobrazena přerušovanou čárou.

4 IMPLEMENTACE NÁSTROJE

Implementovaný nástroj je ve formě démona, naslouchajícím na síťovém rozhraní. Po získání dostatečného množství dat od sledovaného počítače je nástroj schopen se tento počítač tzv. naučit a při

budoucí komunikaci rozpoznat.

4.1 ODPOSLECH DAT

Základem nástroje je sledování síťového provozu na zvoleném rozhraní. K tomuto účelu je využita knihovna pcap, konkrétně její implementace *libpcap*. Knihovna *libpcap* pracuje na úrovni jádra operačního systému, díky čemuž má přímý přístup k veškerým datům procházejícím přes síťové rozhraní, včetně údaje o přesném čase zachycení paketu. Z přijatých dat jsou získávána časová razítka a časy přijetí paketů. Tyto hodnoty jsou ukládány do paměti pomocí obousměrně vázaných seznamů, jeden seznam pro každý sledovaný počítač.

4.2 VÝPOČET POSUNU HODIN

Pro určení posunu hodin sledovaného počítače (a tedy hodnoty identifikující počítač) je potřeba získat rovnici přímky, která shora co nejtěsněji ohraničuje hodnoty odchylek. Pro nalezení všech přímek, které shora ohraničují odchylky, je využito principu nalezení konvexní obálky — konkrétně upraveného *Graham scan* algoritmu [3]. Z těchto přímek je vybrána ta, která splňuje nejmenší vzdálenost od všech bodů. Úhel, který přímka svírá s osou x, je hledaný posun. Protože je výpočet posunu hodin časově poměrně náročná operace, není počítán po každém přijatém paketu, ale až po získání předem daného počtu paketů. Současně je automaticky generován graf (viz obrázek 1), který umožňuje sledovat vývoj posunu hodin s časem v grafické podobě.

5 ZÁVĚR

Doposud provedené testy při měření pěti různých počítačů v reálném prostředí (avšak s uměle generovaným datovým tokem) ukazují, že pokud sledovaný počítač používá časová razítka TCP způsobem popsaným v RFC 1323, je nástrojem rozpoznán. Nezáleží při tom na tom, zda je měření prováděno v rámci lokální sítě nebo pomocí internetu. Důležitějším faktorem než samotný počet přijatých paketů je pak doba měření. Při paketech rovnoměrně rozložených v čase se naměřený posun může očekávané hodnotě přiblížit již po dvou minutách, přesnějších hodnot však dosáhneme až při délce měření v řádu desítek minut. Rozdílnost naměřených posunů hodin u testovaných počítačů byla poměrně velká, pro stanovení závěrů tak bude potřeba provést měření s mnohem větším počtem zařízení. Plánované experimenty s nástrojem dále zahrnují rozpoznání počítačů s různými operačními systémy, rozlišení počítačů se stejnou konfigurací, virtualizované systémy apod.

PODĚKOVÁNÍ

Tato práce je součástí projektu VG20102015022 podporovaného ministerstvem vnitra České republiky. Tato práce vznikla za podpory projektu MŠMT CZ.1.07/2.3.00/09.0067 TeamIT – Budování konkurenceschopných výzkumných týmů pro IT.

REFERENCE

[1] Moon, S. B., Skelly, P., Towsley, D.: Estimation and removal of clock skew from network delay measurements. In: INFOCOM ’99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 1999, s. 227-234

[2] Kohno, T., Broido, A., Claffy, K. C.: Remote physical device fingerprinting. In: Dependable and Secure Computing, IEEE Transactions on, 2005, s. 93-108

[3] Graham, R. L.: An efficient algorithm for determining the convex hull of a finite planar set. In: Information Processing Letters, 1972, s. 132-133