PIITracker: Automatic Tracking of Personally Identifiable Information in Windows

Meisam Navaki Arefi (mnavaki@unm.edu)

Geoffrey Alexander

Jedidiah R. Crandall

THE UNIVERSITY of NEW MEXICO

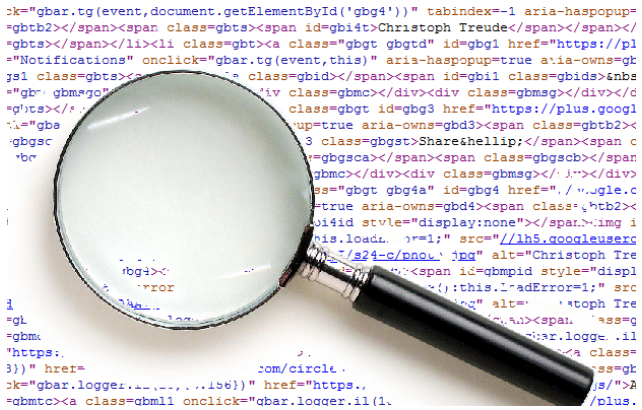
What is PII?

- **Personally Identifiable Information (PII)** is information that can be used to distinguish or trace an individual's identity.

- Examples of PII:
  - Name, Address, Phone number, SSN.
  - MAC Address, Hard drive serial number, IP address.

PII Tracking

- Needs considerable effort to reverse engineer an application.



- To automate PII Tracking process and save reverse engineers substantial time and effort, we propose **PIITracker**.

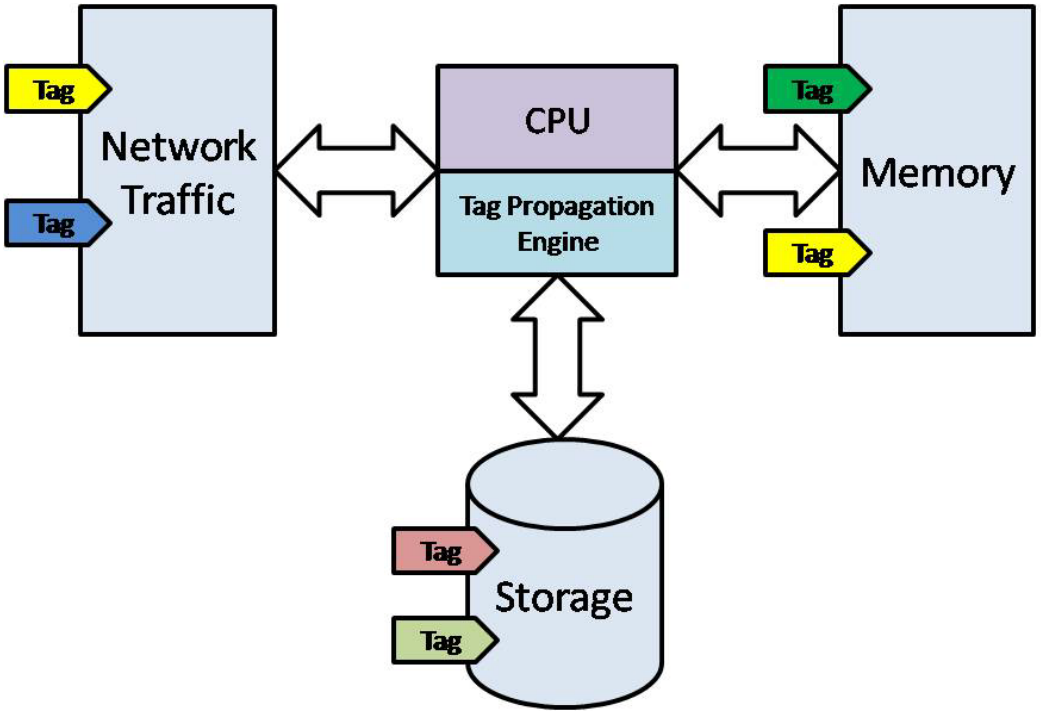
Motivation

- Applications that send PII over the network pose a threat to user privacy and anonymity.

- No other tools track PII in an automatic fashion specifically for Windows.

Background - DIFT

- **Dynamic information flow tracking (DIFT)**, aka Dynamic Taint Analysis, is a promising technology for making systems transparent.



Our Approach

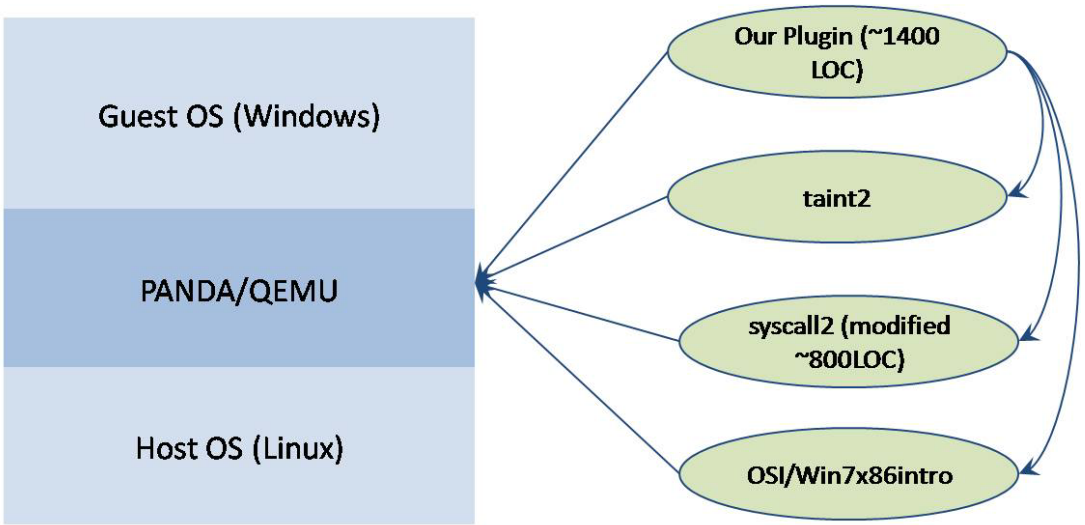
- PIITracker is based on Dynamic Information Flow Tracking (DIFT).
- **PIITracker**:
  1. Monitors reading PII (by monitoring specific function and system calls).
  2. Taint PII with unique tags and track them (using taint2 plugin in PANDA).
  3. Monitors out-going network traffic for tainted bytes (i.e. PII).

PII Data points

- The PII that we have investigated in this paper are:
  - MAC address
  - Hard drive serial number
  - Hard drive model name
  - Volume serial number
  - Host name
  - Computer name
  - Security identifier number (SID)
  - CPU model
  - Windows version and build

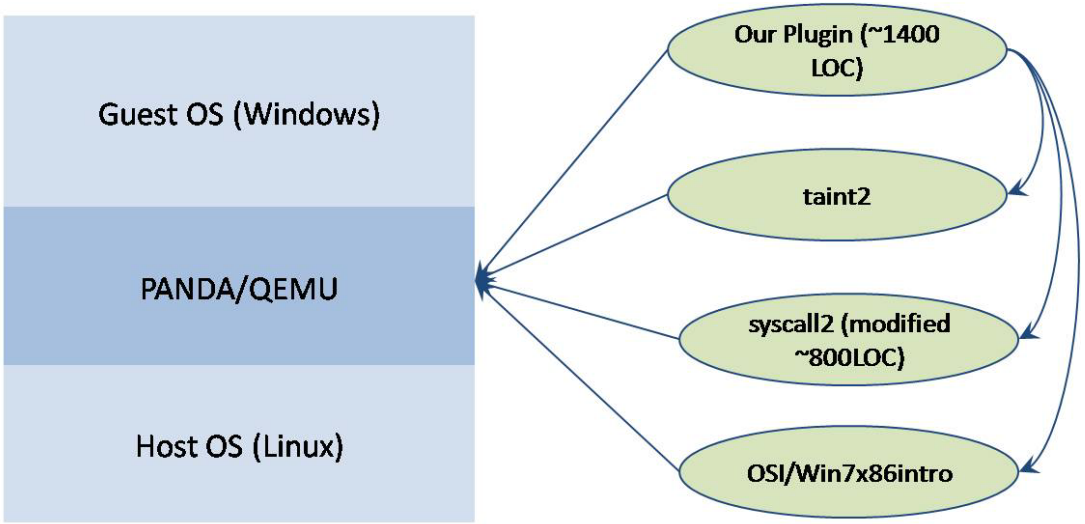
System Architecture

- PIITracker is implemented as a plugin to PANDA whole-system dynamic analysis framework.
- Supports Windows 7 as the guest OS.
- Runs on top of Linux as the host OS.



System Architecture

- PIITracker interacts with other plugins:
  - **Taint2:** whole-system taint analysis engine
  - **Syscalls2:** Callbacks whenever system calls invoked
  - **OSI/Win7x86intro:** Callbacks whenever process-related events happen.



Placing Hooks

- PIITracker utilizes Windows API function calls and system calls as hooks.

- Once a specific function or system call occurs, we get the memory address of the desired argument, and taint that memory location using the taint2 plugin API.

Placing Hooks

- List of functions used to place hooks for each PII data point.

PII data point	Monitored function and system calls
MAC address	GetAdaptersInfo, GetAdaptersAddresses, NtDeviceIoControlFile
Hard Drive Serial Number	NtDeviceIoControlFile, ZwDeviceIoControlFile, DeviceIoControlFile
Hard Drive Model Name	NtDeviceIoControlFile, ZwDeviceIoControlFile, DeviceIoControlFile
Volume Serial Number	GetVolumeInformation, GetVolumeInformationByHandle,
Host Name	gethostname
Computer Name	GetComputerName, GetComputerNameEx
Security Identifier Number (SID)	LookupAccountName, LookupAccountNameLocal
CPU Model	GetSystemInfo
Windows Version and Build	GetVersion, GetVersionEx

Query

- To monitor the outgoing network traffic, PIITracker uses the NtDeviceIoControlFile system call.

- We query the memory address of every byte in the outgoing network traffic to determine if it has any tags.

Results: Analyzing Popular Windows Applications

- We have investigated **15** popular Windows applications, mostly chat applications and web browsers.

- We determined that **12** of these applications collect some form of PII, meaning that they send PII over the network.

Results: Analyzing Popular Windows Applications

PII / Application	Network MAC Address	Hard Drive Serial Number	Hard Drive Model Name	Volume Serial Number	Host Name	Computer Name	Security Identifier Number	CPU Model	Windows Version and Build	Destination address
[1] Baidu Browser V7.6.100.2089	Yes	Yes	Yes	Yes	Read	No	Yes	Yes	Yes	*.br.baidu.com
[2] QQ Browser V9.2.5748.400	Yes	Yes	Yes	No	Yes	No	Read	Yes	Yes	wup.imtt.qq.com
[3] UC Browser V5.5.10106.5	No	Yes	Read	Yes	Read	Yes	Read	Yes	Yes	uc.ucweb.com
[4] 360 Secure Browser V9.1.0.358	No	Yes	Yes	No	No	Read	Yes	Read	Read	dd.browser.360.cn
[5] WeChat V2.6.0.56	Yes	No	No	No	Yes	Yes	No	No	Yes	qq.com 183.232.96.107
[6] Tencent QQ International V2.11	Yes	Read	Yes	No	Read	No	No	Read	Read	203.205.144.238
[7] Viber V7.5.0.97	No	No	No	No	No	Read	No	Read	Yes	content.cdn.viber.com
[8] Line V5.4.2.1560	No	Yes	Read	No	No	No	No	No	Yes	webmaster.naver.com
[9] Telegram V1.1.23	No	No	No	No	No	No	No	No	No	NA
[10] IMO V1.1.2	No	No	No	Read	No	No	Yes	Yes	Yes	192.12.31.77 38.90.96.67(PageBites Inc.)
[11] KakaoTalk V2.6.3.1672	Read	Read	Read	Read	No	Read	Yes	Read	Read	app.pc.kakao.com
[12] Firefox V57.0.3	No	Read	Yes	Read	No	Read	Yes	Read	Yes	184.51.0.249 (detect-portal.firefox.com)
[13] Internet Explorer V8.0.7601	No	No	No	No	No	No	Yes	Yes	Yes	204.79.197.200 (Microsoft)
[14] Chromium V63.0.3239.108	No	Read	Read	Yes	No	Read	Yes	Read	Yes	172.217.12.3 (Google Inc.)
[15] Mullvad	No	No	No	No	No	No	No	Read	Read	NA

Results: Analyzing Popular Windows Applications

- The chat applications that we could not find any serious PII-related privacy issues were **Telegram** and **Viber**.

- **All** Chinese chat and web browser applications that we investigated collect some form of PII.

- **Firefox** and **Chromimum** also collect some form of PII.

False Positive and False Negative Analysis

- Comparison with previous works.
  - Using PIITracker, we could verify the results of other researchers.

- Evaluating PIITracker via our own developed test applications.
  - Worked as expected.

Performance Evaluation

- Whole-system information flow tracking is intrinsically heavyweight.
  - Performance has not been a priority for PIITracker.

- PIITracker exhibited a 67X slowdown on average compared to PANDA replay.

Related Works

- TaintDroid
  - Detects data leakage of Android applications.
- TaintEraser
  - Detects leakage of sensitive data such as password and credit card numbers in Windows.
  - Requires users to manually specify what actually is a password or credit card number.
- None of them are able to track PII in an automatic way in Windows.

Conclusions

- Presented PIITracker, a novel tool for tracking personally identifiable information (PII) in Windows.
- Analyzed 15 popular Windows applications
  - Majority of these applications collect some form of PII.
- PIITracker:
  - Saves reverse engineers substantial time and effort in practice.
  - Provides valuable information including the relevant memory addresses of leaked PII, as well as network socket info.
- PIITracker is available for public download
  - https://github.com/mnavaki/PIITracker

Thank you!

- Contact: Meisam Navaki Arefi
  - mnavaki@unm.edu

- Download PIITracker:
  - https://github.com/mnavaki/PIITracker