A Report on

802.11 User Fingerprinting

- Authors -

Jeffrey Pang, Carnegie Mellon Univ

Ben Greenstein, Intel Corp (Seattle)

Ramakrishna Gummadi, Univ of So Cal

Srinivasan Seshan, Carnegie Mellon Univ

David Wetherall, Univ of Washington

Published at *MobiCom '07*, Montréal, Canada

9-14 Sept 07

Presented by James Gaskell

7 Nov 11

CS 577, Advanced Computer Networks

Prof Kinicki, WPI, Fall 2011

Overview

1. **What is Fingerprinting**
2. Experimental Methodology
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. Three Scenarios
6. Automated Data Reduction
7. Results
8. Summary
9. Questions & Comments

Define “Fingerprinting”

1. Not a real “fingerprint.”

2. Characteristics of a fingerprint:

    a. Each human has a unique fingerprint (?).

    b. Finding a (clear) fingerprint at a crime scene can uniquely identify someone at the crime scene, but does not necessarily tie that identity to a particular person.

        i. Needs a database of fingerprints.

        ii. If you can obtain the fingerprint of a suspect, that might be able to tie that suspect to the crime scene.

“Fingerprinting” here.

1. Individual wireless adapters using the network are called USERs.

2. Passive only sensing used.

   USER has no way of knowing s/he is being “investigated.”

A. Detect a USER

   Identify definable characteristics of each and every USER that separate one the USER from other USERS.

B. Detect an INDIVIDUAL

   Perhaps identify characteristics that tie a USER’s identity to a particular PERSON.

C. Opposite of a FINGERPRINT is:

ANONYMOUS

Overview

1. What is Fingerprinting
2. **Methodology**
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. Three Scenarios
6. Automated Data Reduction
7. Results
8. Summary
9. Questions & Comments

Methodology

1. Use easily obtainable monitoring tools.

2. Be completely passive (receive only).

3. Monitoring devices may be placed in more than one location.

   If “close together” these multiple sensors can use signal strength as a determining factor.

4. Different types of environments:

   a. Public networks (difficulty of key distribution)

   b. Home networks (link layer encryption)

   c. Corporate networks

Methodology 2

5. MAC address identity was not used. However, it was assumed that the Pseudonym address was used for an hour before being changed.

    This means that information gathered over one hour from a single address could be combined in an effort to fingerprint that User.

6. “Training” sessions are when a particular User is isolated and the Adversary is able to get traces known to be from that User.

    These sessions can occur either before or after the fact.

Methodology 3

7. For the experiment, one half of the full trace was used as the “Training” session and the other half as the “Validating” session.

“Profiled” Users must exist in both sessions.

The Experiment Goal

Answer the following 2 Questions:

a. Is this traffic from User U?

b. Was User U here today?

Aside #1

Lack of communications even 30 years ago much more prevalent than today.

Hardwired phones, often with no answering machines.

No e-mail, only snail mail.

Low tech ID cards.

Virtually no surveillance cameras.

No RFID devices.

No facial recognition.

Old folks got used to being able to travel in cognito.

Overview

1. What is Fingerprinting
2. Methodology
3. **Obvious Fingerprints**
4. Four less obvious Fingerprints
5. Three Scenarios
6. Automated Data Reduction
7. Results
8. Summary
9. Questions & Comments

Obvious Fingerprints

1. MAC Address

    a. Globally unique
    b. Persistent
    c. Defense by Pseudonyms

2. SSID Probes

    a. Default action of various OSs
    b. Defense is to change default operation.

3. Personal IP Addresses

    a. E-mail server
    b. Bookmarked locations
    c. Defense by encrypting.

802.11i

Best Practices for Securing Networks

1. User Authentication
2. Service Authentication
3. Data Confidentiality
4. Data Integrity

Does not guarantee ANONYMITY!

AKA, LOCATION PRIVACY.

Overview

1. What is Fingerprinting
2. Methodology
3. Obvious Fingerprints
4. **Four less obvious Fingerprints**
5. Three Scenarios
6. Automated Data Reduction
7. Results
8. Summary
9. Questions & Comments

Less Obvious Fingerprints

These were used in the experiment.

1. netdests: IP address + port

2. SSID probes

3. Broadcast packet sizes

   NetBIOS naming advertising
   M/S Office advertising
   Filemaker advertising

4. MAC Protocol Fields

   OS, driver, hardware characteristics

Aside #2

Acceptance of Loss of Anonymity

1. Surveillance cameras.
2. RFID, Easy Pass, etc
3. Cell phone tower tracking
4. Built-in GPS
5. Facial recognition
6. Facebook

Overview

1. What is Fingerprinting
2. Methodology
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. **Three Scenarios**
6. Automated Data Reduction
7. Results
8. Summary
9. Questions & Comments

Three Scenarios

1. Sigcomm

    a. 37 hrs training, 54 validating
    b. 1974/3391 samples, 377/412 users

2. UCSD (Univ of Calif @ San Diego)

    a. 10 hrs training, 11 validating
    b. 587/1240 samples, 225/371 users

3. Apartment Building

    a. 119 hrs training, 345 validating
    b. 1638/1473 samples, 97/196 users

Overview

1. What is Fingerprinting
2. Methodology
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. Three Scenarios
6. **Automated Data Reduction**
7. Results
8. Summary
9. Questions & Comments

Automated Data Reduction

[Read the Paper!]

Overview

1. What is Fingerprinting
2. Methodology
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. Three Scenarios
6. Automated Data Reduction
7. **Results**
8. Summary
9. Questions & Comments

Results

Surprise! There are so many distinguishing idiosyncrasies among Users that it was often possible to:

1. Probabilistically identify Distinct Users among All Users.

2. At a different session, recognize that a particular Distinct User might probably have returned.

Overview

1. What is Fingerprinting
2. Methodology
3. Obvious Fingerprints
4. Four less obvious Fingerprints
5. Three Scenarios
6. Automated Data Reduction
7. Results
8. **Summary**
9. Questions & Comments

Summary

What's the problem that we're trying to solve?

In the electronic, camera, GPS world that keeps adding more and more of these devices that have more and more accuracy and extend into more and more locations:

What are our expectations of privacy?

What amount of privacy is needed?

What can be solved by legal means?

What protections can be added?

Lastly: Here I am, come get me!

Questions?

Comments.

Thank You!

Jim Gaskell