

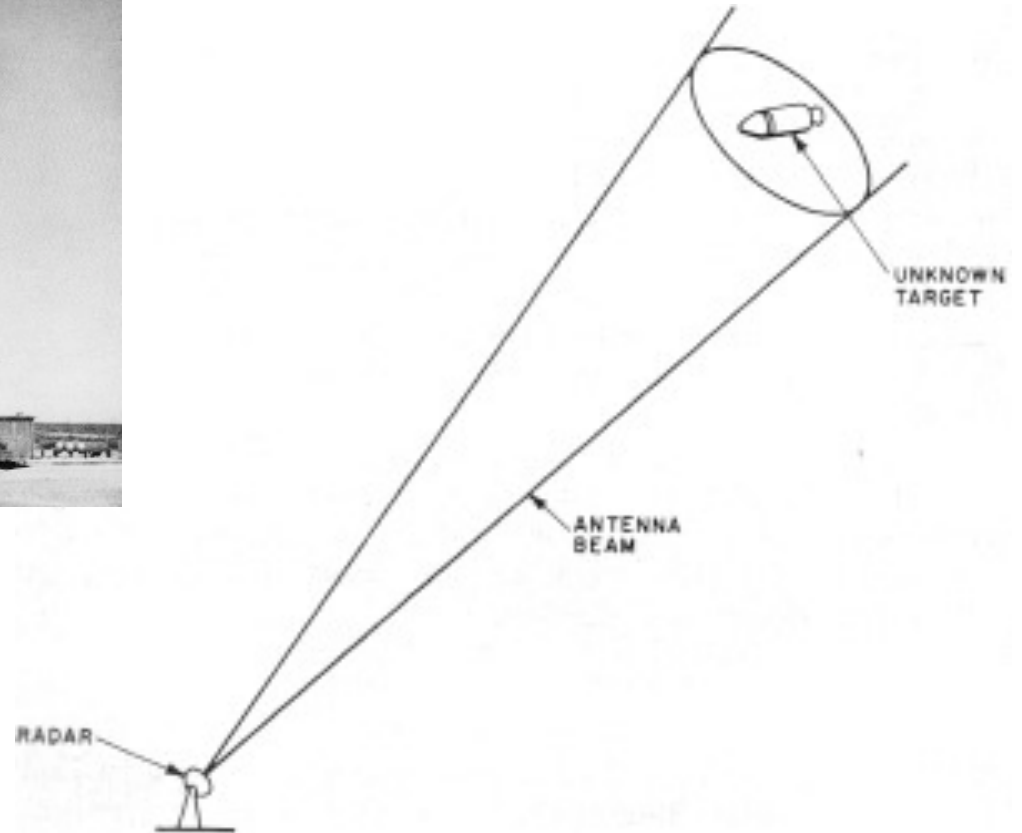


Fingerprinting Techniques for Network Forensics

Overview, Opportunities and Challenges

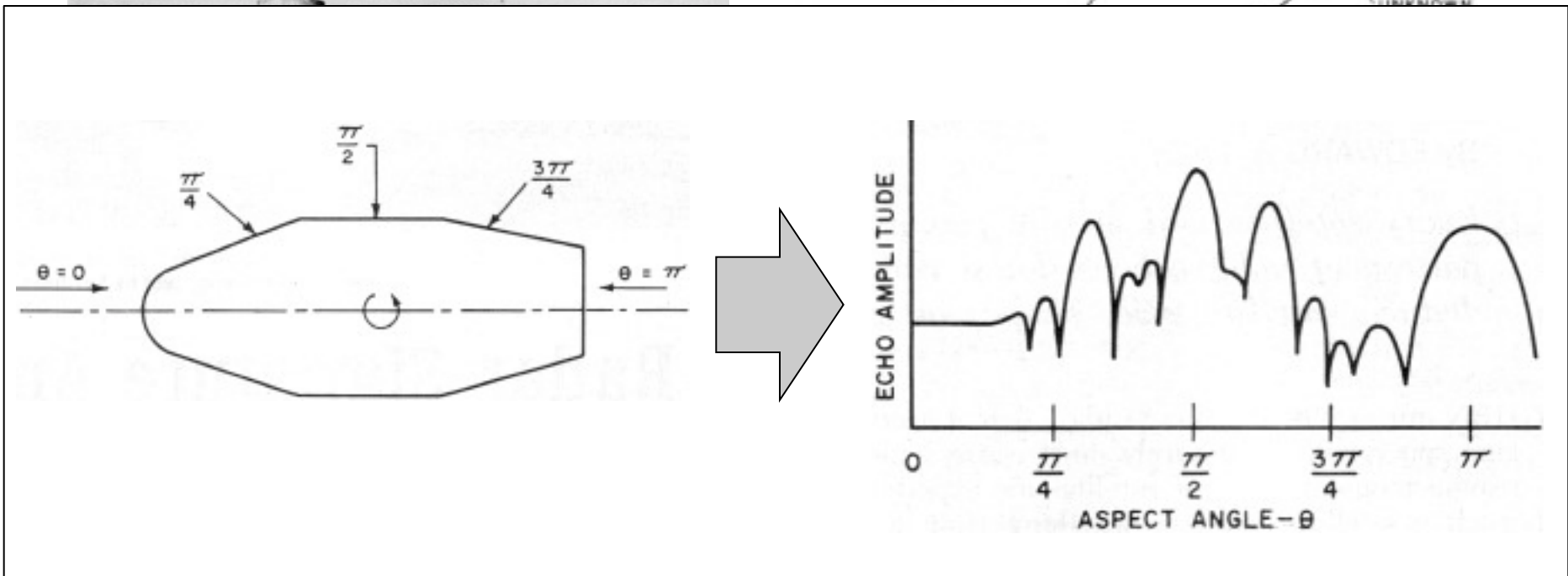
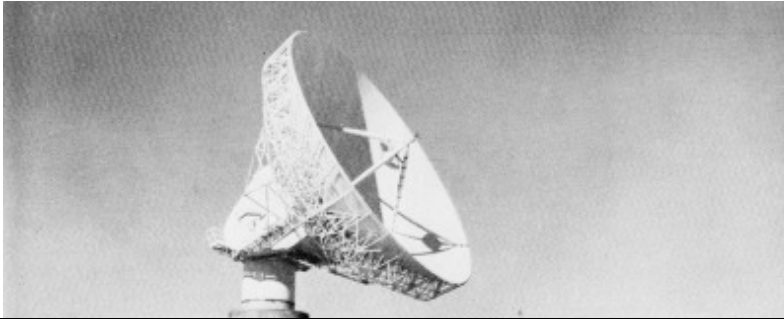
Dominik Herrmann

Fingerprinting Primer

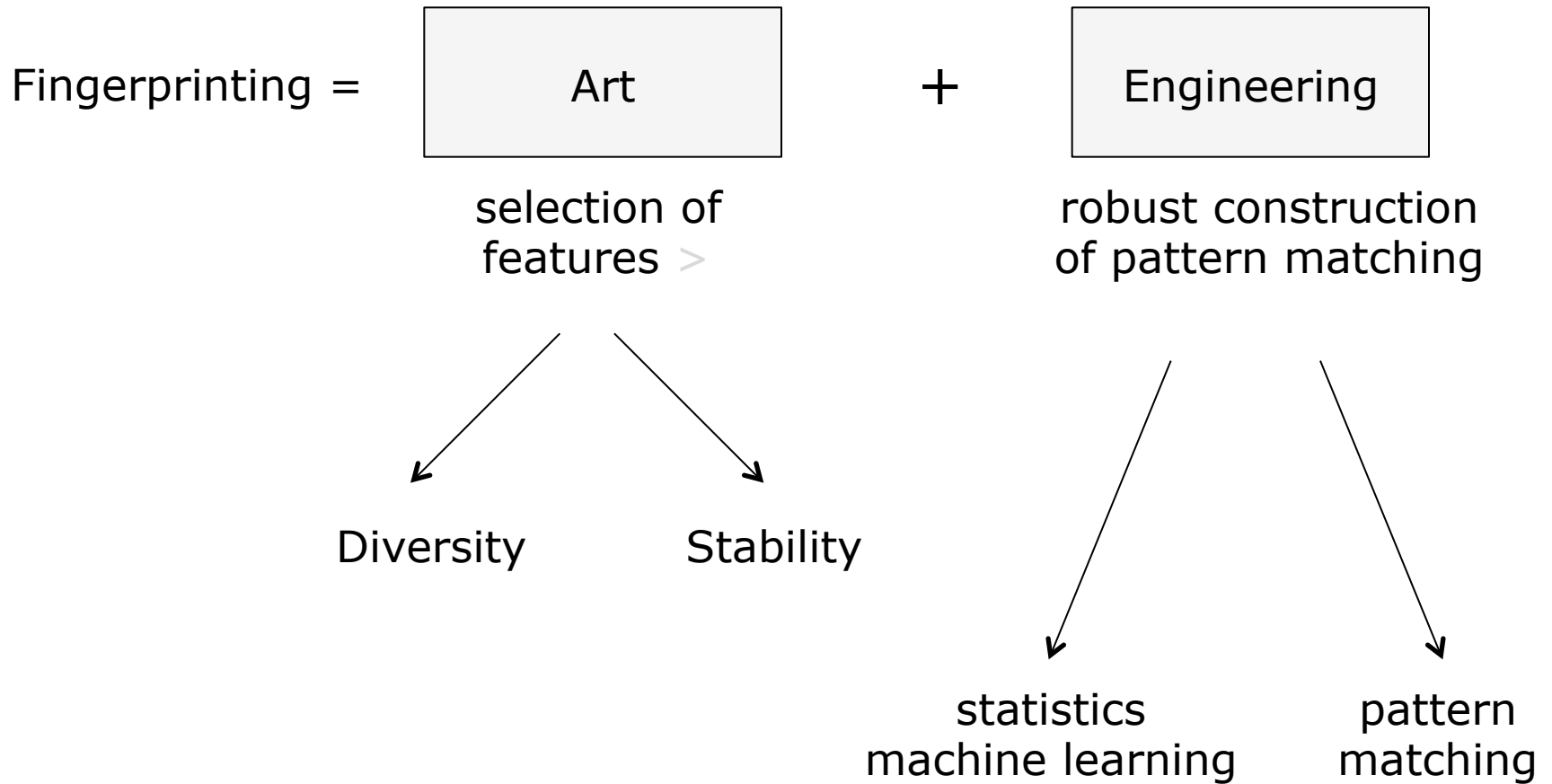


Origin:
Radar "fingerprints"

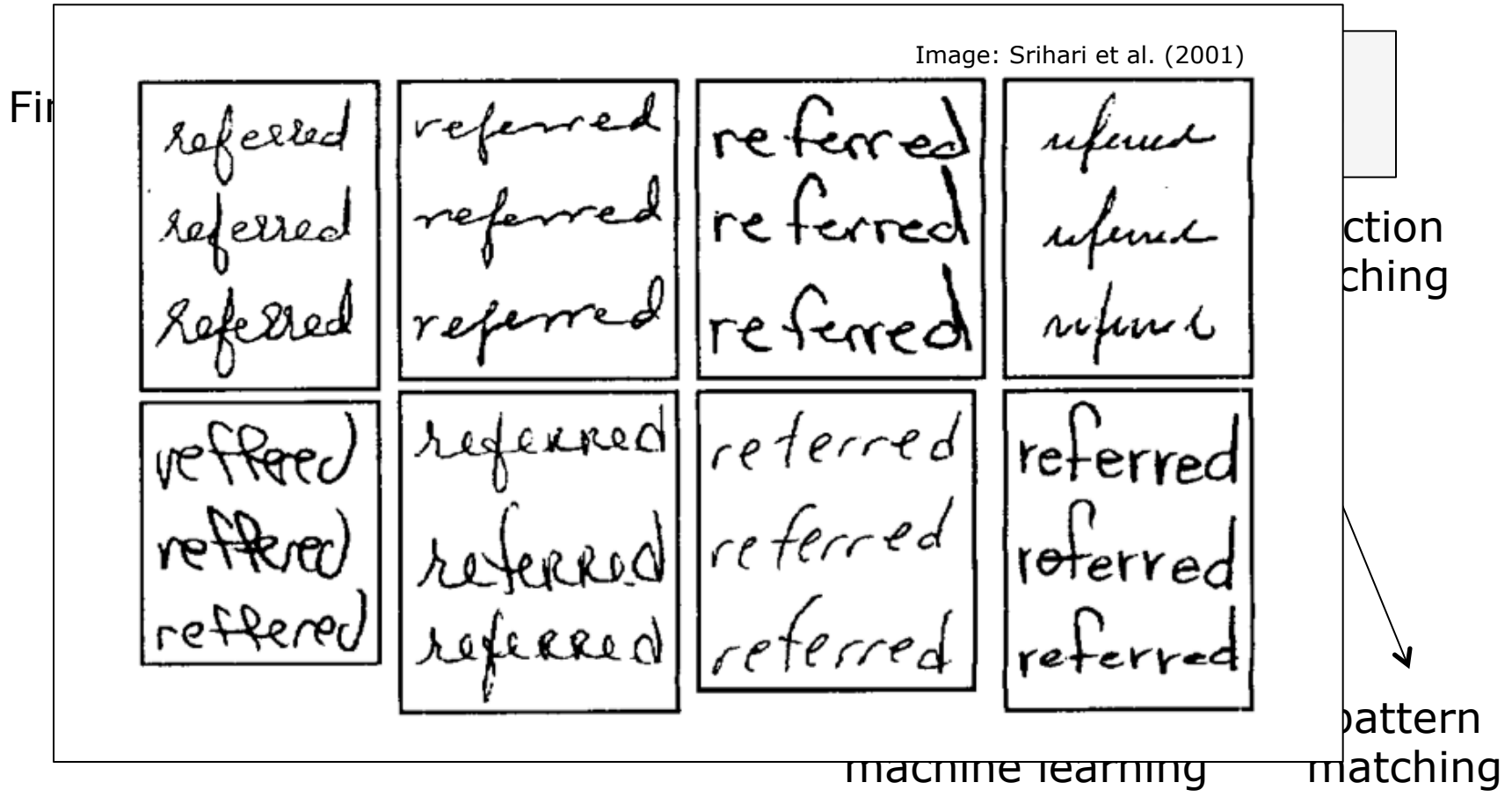
Fingerprinting Primer



Fingerprinting Primer



Fingerprinting Primer



Agenda

Fingerprinting Primer

From Computer Forensics to **Network Forensics**

Three Case Studies:

Website
Fingerprinting

Device/Software
Fingerprinting

Human Behavior
Fingerprinting

Fingerprinting for Forensics:
A new **promising** opportunity or a **dangerous** instrument?

The case for network forensics

Computer Forensics

- focus on HDD and RAM
- static dataset

Typical objectives

- deduce actions of a subject
- ascription of files/actions

However, some attacks do not leave suitable forensic traces.

We could look at network traffic to capture transient data and activities.

Network Forensics

- focus on network traffic
- transient dataset

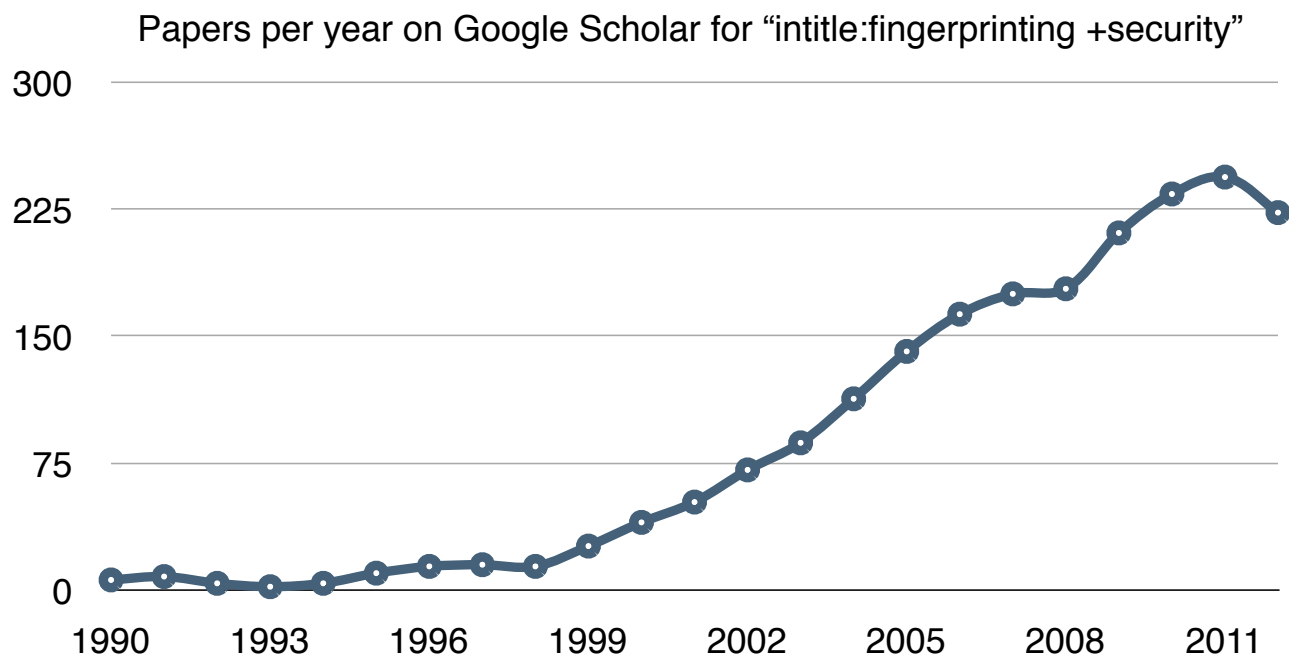
Typical objectives

- find source of criminal activity
- find evidence that a subject is involved in criminal activity

Challenges

- large volumes of traffic
difficult to analyze
 - cannot analyze content if it is encrypted before transmission
-

Rising interest in security-related fingerprinting lately



Can we leverage fingerprinting techniques for network forensics?

Yes!

1. Determine activities of a subject, even if traffic is encrypted
2. Find evidence for involvement in criminal activities



Objective 1:
Determining activities in encrypted traffic

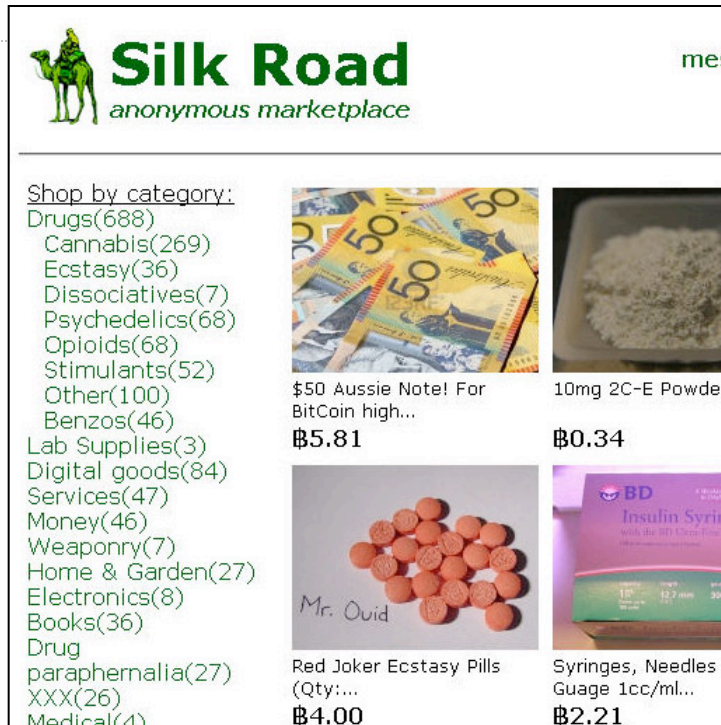
Case Study 1: Website Fingerprinting

Website Fingerprinting

- The Crime Scene
 - subject visits incriminating website
 - investigator has access to traffic
 - traffic is encrypted on network layer

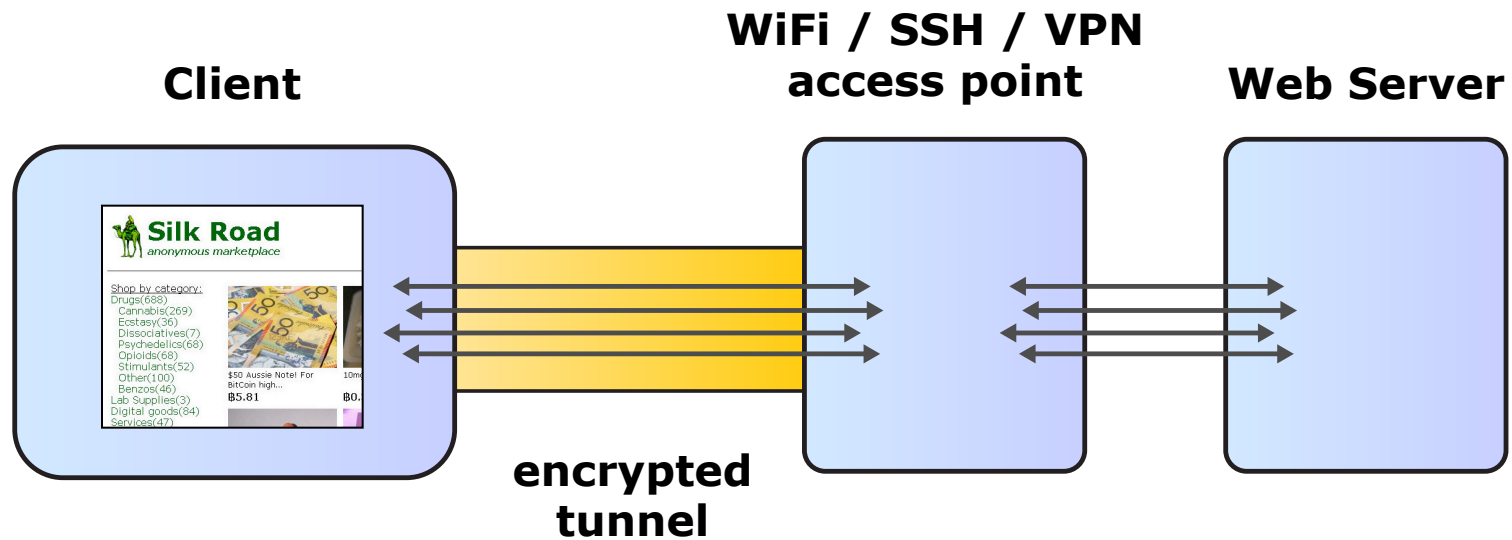
- Digital Forensics Objective
 - find corroborating evidence for specific incriminating activity

- Fingerprinting Approach
 - relies on metadata that is not encrypted (“**traffic analysis**”)
 - investigator **collects traffic samples** for interesting websites and extracts fingerprints (manually or via machine learning)
 - successful identification of site if recorded traffic of subject **matches** one of the known fingerprints



Technique 1: Characteristic Patterns in IP Packets

..., Herrmann, Wendolsky, and Federrath (2009), ...



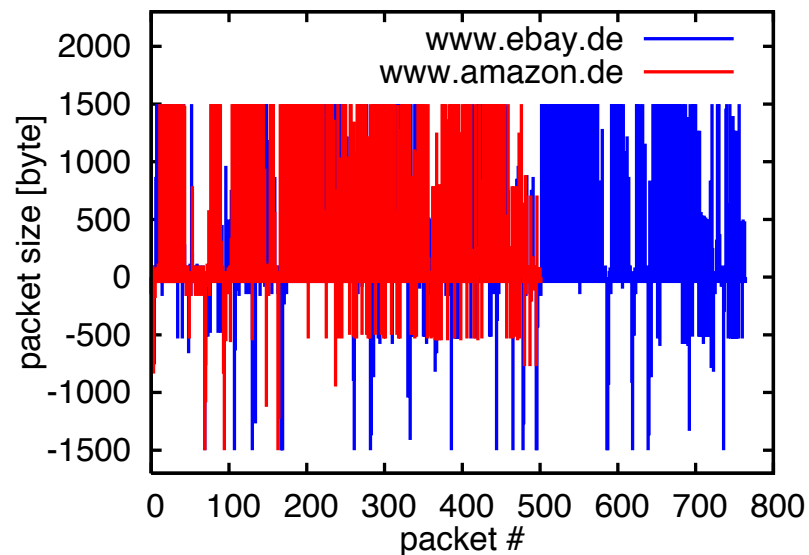
Investigator can only observe size, direction and timing of packets



Also applicable for anonymization services

Technique 1: Characteristic Patterns in IP Packets

..., Herrmann, Wendolsky, Federrath (2009), ...



Many websites cause characteristic patterns

Technique 1: Characteristic Patterns in IP Packets

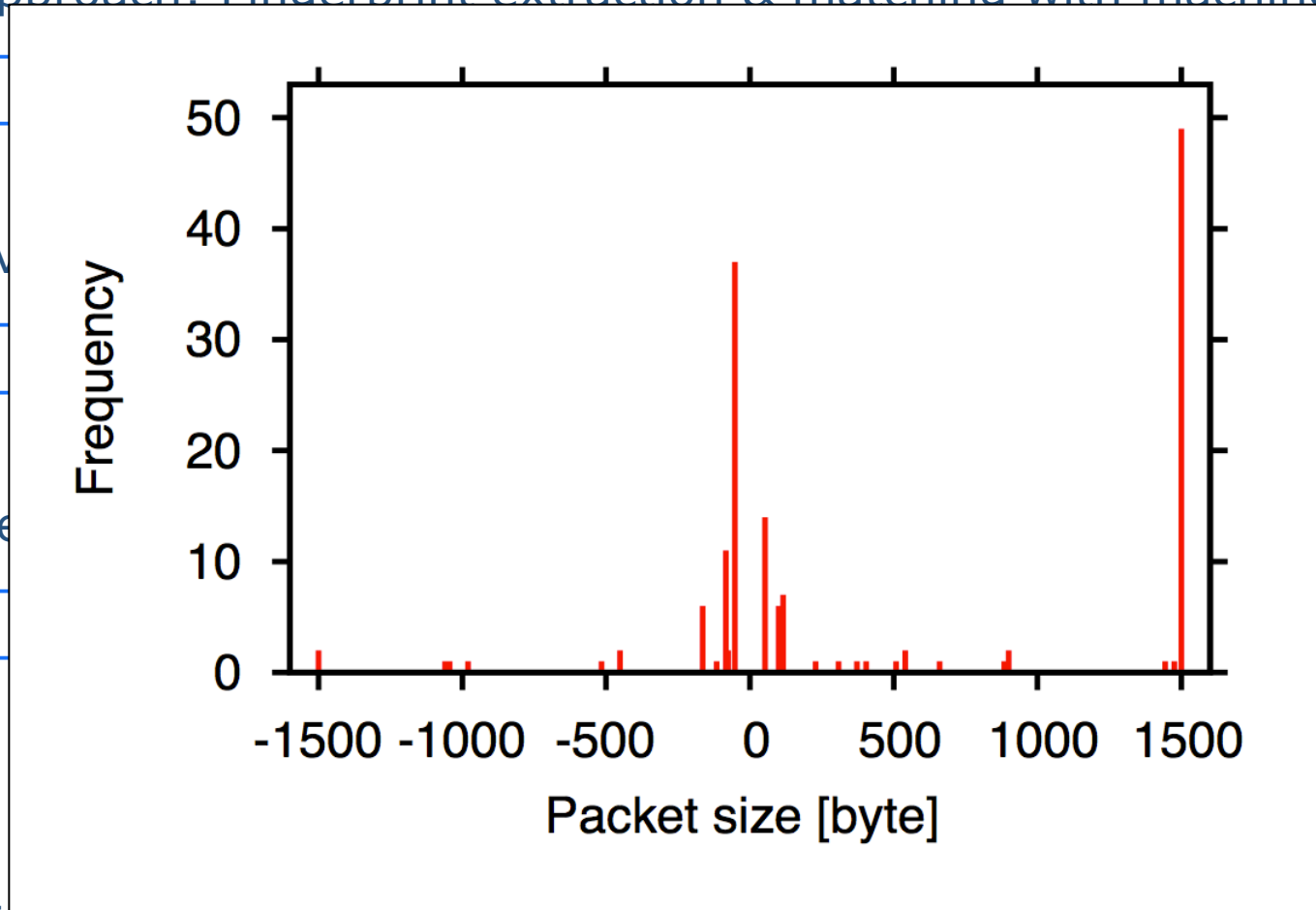
- Approach: Fingerprint extraction & matching with machine learning
 - features: histogram of packet sizes observed during download >
 - supervised learning technique: kNN and Naïve Bayes classifiers
- Evaluation
 - OpenSSL, stunnel, OpenVPN, IPsec, JonDonym, Tor
 - 775 popular sites from Alexa
- Results
 - accuracy > 95% for all systems (exception: Tor & JonDonym)
 - high efficiency: fingerprints keep for multiple days and a single training instance is sufficient

Next up:

Technique 2: Website fingerprinting via characteristic DNS queries

Technique 1: Characteristic Patterns in IP Packets

- Approach: Fingerprint extraction & matching with machine learning



download >
classifiers

- EV

- Re

(Anonym)
a single

Next up.

Technique 2: Website fingerprinting via characteristic DNS queries

unkrautex stahlrohr|

pesticide steel pipe

[webdesign by s@ndkes - Willkommen in unserem Forum über ...](#)

[www.schottlandforum.de](#) › [Highland Pub](#) ▾ [Translate this page](#)

Feb 7, 2009 - hallo macdubh, **Unkrautex** und Zucker im **Stahlrohr** war für mich immer der Inbegriff der Stümperei. Mein Sprengstoff basierte auf basis ...

[Thema: Feuerwerkszeugs | HalleSpektrum](#)

[hallespektrum.de](#) › [Foren](#) › [Halle \(Saale\)](#) ▾ [Translate this page](#)

Dec 27, 2012 - Bei uns hießen die Knaller aus **Unkrautex** und **Stahlrohr** "Eisenforze". Sicher etwas untertrieben, aber es war so ein herrliches Gafühl aus ...

["Pulver-Kurt" steht vor Gericht, Der Rentner mit dem Kriegsgerät ...](#)

[www.explorat.de/Forum/showthread.../page2](#) ▾ [Translate this page](#)

Jul 26, 2012 - 10 posts - 2 authors

Soviel zum Thema Basteln mit **Unkrautex**..... kenne ich auch,so ein fall....weitläufige verwandschaft.uex in ein **stahlrohr** und mit nem fäustel ...

[mosfetkiller.de • Thema anzeigen - Der Thread der "kleinen Fragen"](#)

[forum.mosfetkiller.de](#) › ... › [Sonstige Basteleien](#) ▾ [Translate this page](#)

Jan 5, 2013 - 15 posts - 9 authors

Jemand den ich kenne hat mit einer Mischung ein **Rohr** gefüllt und es ... parat: Kollege rzahlte wie er **Unkrautex** und Zucker im **Stahlrohr** mit ...

[More results from forum.mosfetkiller.de](#)

[!!!PYROTECHNIK VERSANDT!!! - Google Groups](#)

<https://groups.google.com/d/topic/z.../JFEZERlc4UE> ▾ [Translate this page](#)

Mar 18, 1998 - Ein Freund meinte, dass es in Bayern noch **Unkraut-Ex** gäbe. Z.B. haben ein paar Jugendliche aus einen **Stahlrohr** mal eine Kanone ge-

Ads ⓘ

[Stahlrohre aller Art](#)

[www.feld-eitorf.de/stahlrohre](#) ▾

★★★★★ 72 reviews for feld-eitorf.de

Stahlrohre auf Maß oder auf Gehrung - ganz einfach online bestellen!

[Stahlrohrgroßhandel](#)

[www.kr-rohrsysteme.de/](#) ▾

Individuelle Lösungen für Großrohre z. B. 219,1 mm

📍 [Glockengießerwall 17, Hamburg](#)
040 609467100

[Stahlrohre div Qualitäten](#)

[www.stahlrohr.at/](#) ▾

Rohre neu - deklassiert - gebraucht
Dimensionen auf Anfrage

[Unkrautvernichter billig](#)

[unkrautvernichter.preisvergleich.de/](#)

★★★★★ 1,604 seller reviews

Riesenauswahl im PREISVERGLEICH.
Über 2.500 Händler & 8 Mio Produkte

[Unkrautex Online](#)

[unkrautex.living3000.de/](#) ▾

Unkrautex Ausverkauf
Viele Markenartikel bis -76%!

Technique 2: Website Fingerprinting via DNS Queries

Krishnan & Monroe (2010)

[webdesign by s@ndkes - Willkommen in unserem Forum über ...](#)

[www.schottlandforum.de](#) › Highland Pub ▾ [Translate this page](#)

Feb 7, 2009 - hallo macdubh, **Unkrautex** und Zucker im **Stahlrohr** war für mich immer der Inbegriff der Stü...

[Thema: Feuerwe](#)

[hallespektrum.de](#) ›

Dec 27, 2012 - Bei u
Sicher etwas untertr

["Pulver-Kurt" ste](#)

[www.explore.de/](#)

Jul 26, 2012 - 10 po
Soviel zum Thema B
verwandschaft.uex i

[mosfetkiller.de](#) •

[forum.mosfetkiller.](#)

Jan 5, 2013 - 15 pos
Jemand den ich ker
rzählte wie er Unkra
More results from fo

[!!!PYROTECHN](#)

<https://groups.google.com>

Mar 18, 1998 - Ein Freund meinte, dass es in Bayern noch **Unkraut-Ex** gabe. Z.B.
haben ein paar Jugendliche aus einen **Stahlrohr** mal eine Kanone ge-

Observable DNS queries due to prefetching (Firefox, Chrome, Safari):

- [www.schottlandforum.de](#)
- [hallespektrum.de](#)
- [www.explore.de](#)
- [forum.mosfetkiller.de](#)
- [groups.google.com](#)
- [www.feld-eitorf.de](#)
- [www.kr-rohrsysteme.de](#)
- [www.stahlrohr.at](#)
- [unkrautvernichter.preisvergleich.de](#)
- [unkrautex.living3000.de](#)

Ads ⓘ

[Stahlrohre aller Art](#)

[www.feld-eitorf.de/stahlrohre](#) ▾

★★★★★ 72 reviews for feld-eitorf.de
Stahlrohre auf Maß oder auf Gehrung
- ganz einfach online bestellen!

[Stahlrohrgroßhandel](#)

[www.kr-rohrsysteme.de/](#) ▾

Individuelle Lösungen für
Großrohre z. B. 219,1 mm
📍 [Glockengießerwall 17, Hamburg](#)
040 609467100

[Stahlrohre div Qualitäten](#)

[www.stahlrohr.at/](#) ▾

Rohre neu - deklassiert - gebraucht
Dimensionen auf Anfrage

[Unkrautvernichter billig](#)

[unkrautvernichter.preisvergleich.de/](#)

★★★★★ 1,604 seller reviews
Riesenauswahl im PREISVERGLEICH.c
Über 2.500 Händler & 8 Mio Produkte

[Unkrautex Online](#)

[unkrautex.living3000.de/](#) ▾

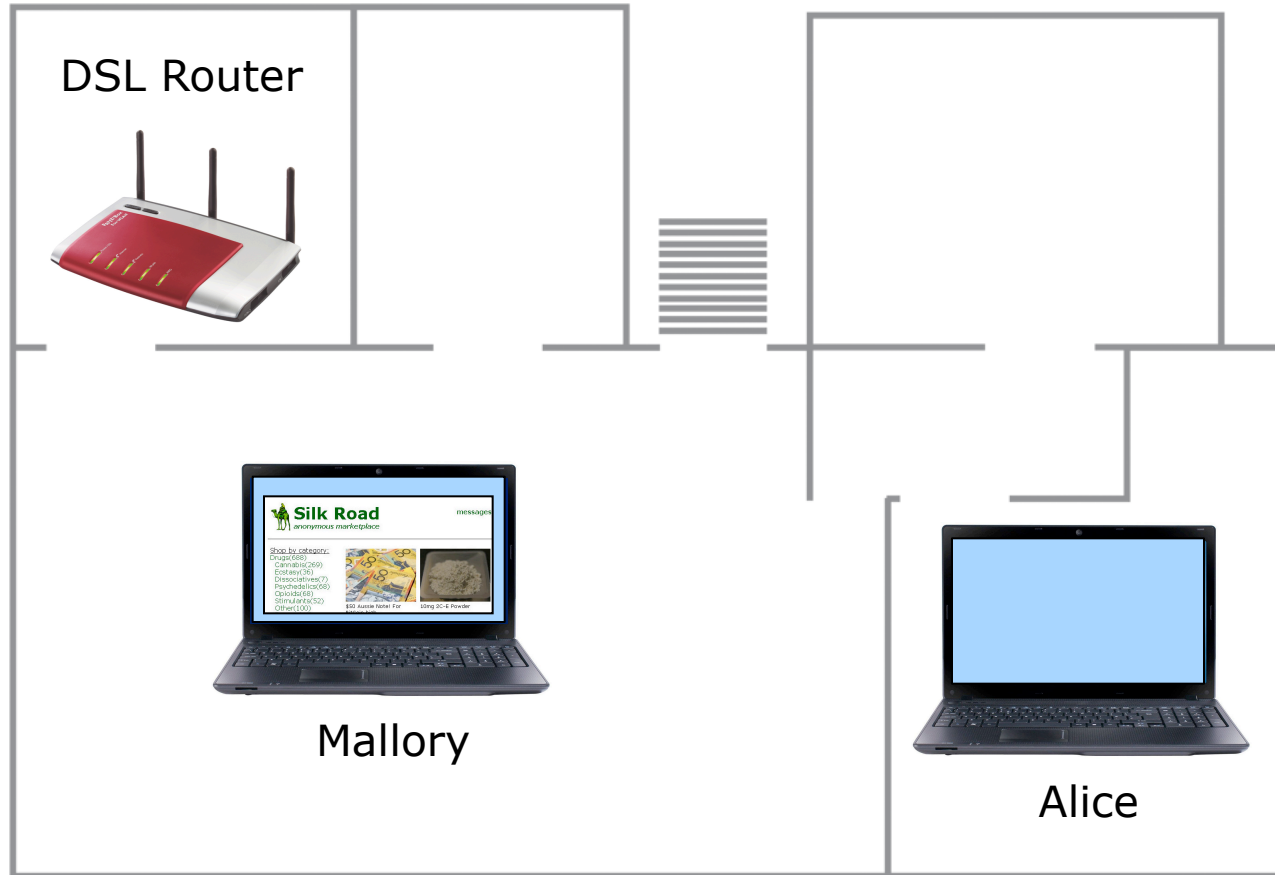
Unkrautex Ausverkauf
Viele Markenartikel bis -76%!



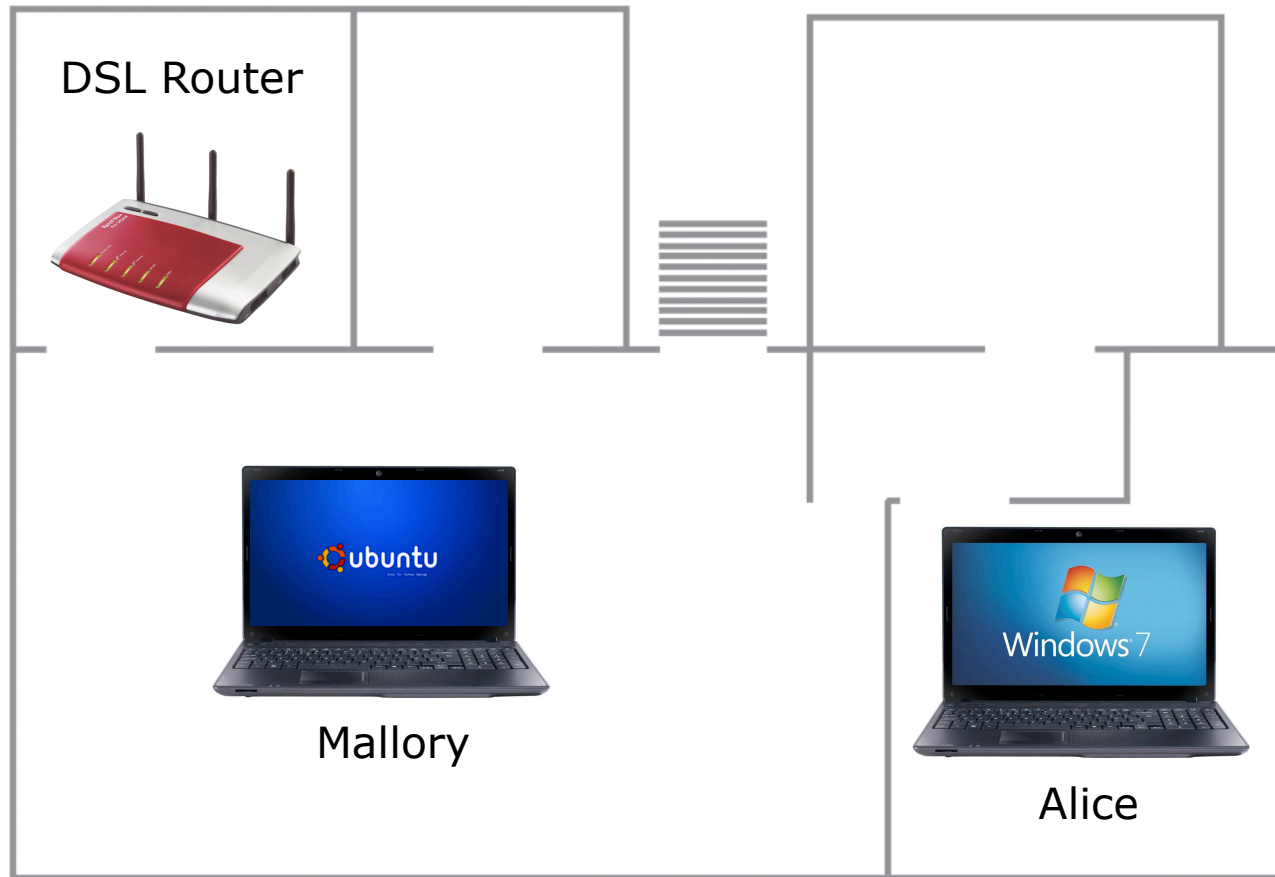
Objective 2:
Find evidence for involvement in criminal activities

Case Study 2: Device/Software Fingerprinting

Device and software fingerprinting

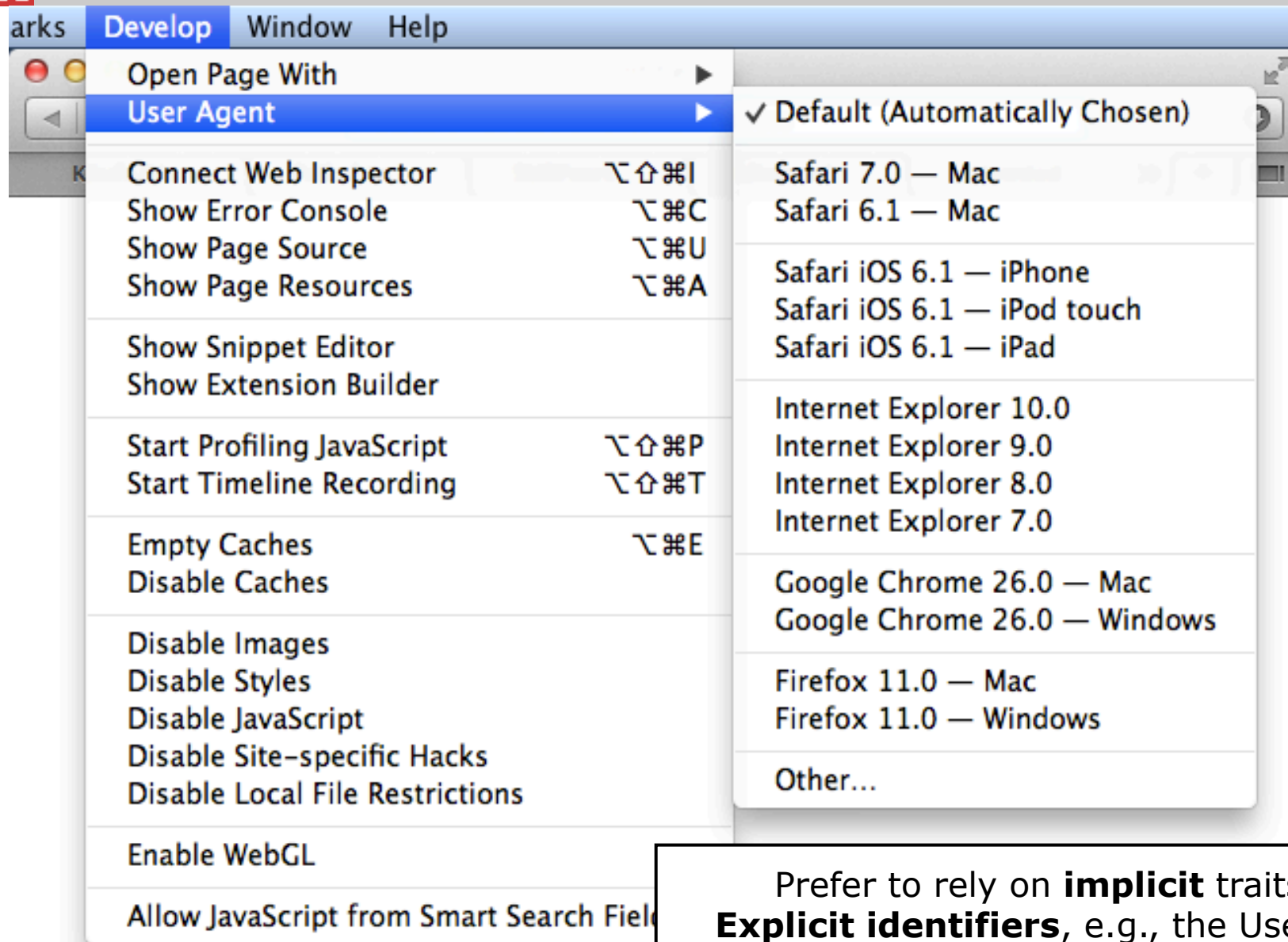


Device and software fingerprinting



Device and software fingerprinting

- The Crime Scene
 - subject carries out criminal activity on the network
 - investigator has access to the traffic of the subject
 - an investigation is launched and all hardware is seized
 - the subject denies any involvement and incriminates a flat mate
 - no traces of the activity can be found on any of the machines
- Digital Forensics Objective
 - **ascription/association**: find corroborating evidence that one of the machines was in fact used for the criminal activity
- Fingerprinting Approach
 - relies on **differing implicit behavior** of devices/software
 - build a **corpus**: investigator collects behavioral samples of network traffic for various systems
 - infer **system architecture, operating system, browser**, etc. by matching recorded traffic of suspect to patterns from corpus



Prefer to rely on **implicit** traits.
Explicit identifiers, e.g., the User Agent header, can be **forged** easily.

Various Device Fingerprinting Techniques

- Operating system fingerprinting
 - characteristics in **TCP stack**, Comer&Lin (1994)
 - now readily available in tools, e.g. **p0f & nmap**
- Device fingerprinting
 - **Skew of real-time clock** is characteristic, Kohno et al. (2005) >
 - Runtime of **JavaScript** code is characteristic for browser, operating system and CPU architecture, Mowery et al. (2011)
 - Text rendering in **HTML5** <canvas>, Mowery&Shacham (2012)>
- Browser fingerprinting
 - Characteristic **TCP flows** allow identification, Yen et al. (2009)
 - EFF **Panoptlick**: plugins, fonts, etc., Eckersley (2011)>

Note: **class** characteristics vs. **individual** characteristics

Various De

- Operating
 - chara
 - now r

- Device fi
 - **Skew**
 - Runti
 - opera
 - Text

- Browser
 - Chara
 - EFF P

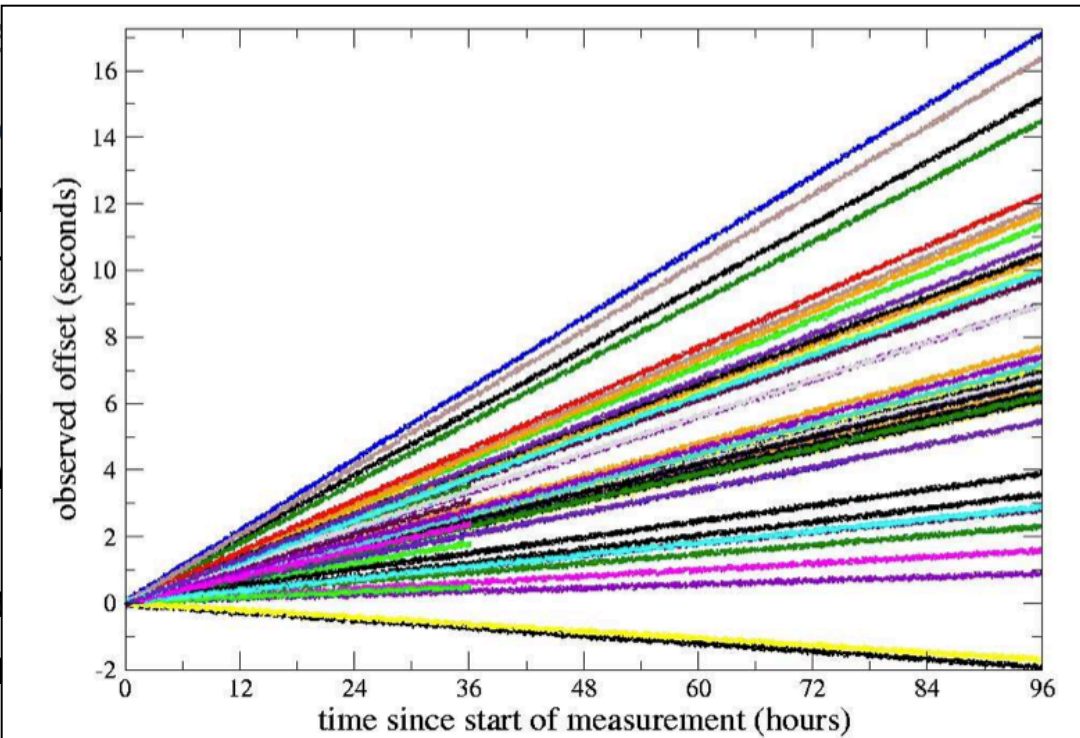


Figure 4. TSopt clock offset-sets for 69 Micron 448MHz Pentium II machines running Windows XP Professional SP1. Trace recorded on host2, three hops away, 2004-09-10 08:30PDT to 2004-09-14 08:30PDT.

et al. (2005) >
 browser,
 et al. (2011)
 acham (2012)>

et al. (2009)
 (2011)>

Note: **class** characteristics vs. **individual** characteristics

Various Device Fingerprinting Techniques

- Operating system fingerprinting
 - characteristics in **TCP stack**, Comer&Lin (1994)
 - now readily available in tools, e.g. **p0f & nmap**

How quickly daft jumping zebras vex. (Also, punctuation: &/c.)

How quickly daft jumping zebras vex. (Also, punctuation: &/c.)

How quickly daft jumping zebras vex. (Also, punctuation: &/c.)

- Browser fingerprinting
 - Characteristic **TCP flows** allow identification, Yen et al. (2009)
 - EFF **Panoptlick**: plugins, fonts, etc., Eckersley (2011)>

Note: **class** characteristics vs. **individual** characteristics

Your browser fingerprint appears to be unique among the 3,628,476 tested so far. Currently, we estimate that your browser has a fingerprint that conveys at least **21.79 bits of identifying information**.

Browser Characteristic	bits of identifying information	one in x browsers have this value	value
User Agent	13.07	8618.71	Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:25.0) Gecko/20100101 Firefox/25.0
HTTP_ACCEPT Headers	16.79	113389.88	text/html, */* gzip, deflate en-us,en;q=0.8,de;q=0.5,de-de;q=0.3
Browser Plugin Details	21.79+	3628476	Plugin 0: Google Talk Plugin Video Renderer; Version 4.9.1.16010; o1dbrowserplugin.plugin; (Google Talk Plugin Video Renderer; application/o1d; o1d). Plugin 1: Java Applet Plug-in;... Shockwave Flash 11.9 r900; Flash Player.plugin; (Shockwave Flash; application/x-shockwave-flash; swf) (FutureSplash Player; application/futuresplash; spl). Plugin 4: iPhotoPhotocast; iPhoto6; iPhotoPhotocast.plugin; (iPhoto 700; application/photo;).
Time Zone	2.64	6.23	-60
Screen Size and Color Depth	11.95	3965.55	1120x700x24
System Fonts	21.79+	3628476	Adobe Caslon Pro Bold, Adobe Caslon Pro Bold Italic, Adobe Caslon Pro Italic, [300 more fonts], Yuppy TC Regular, Zapf Dingbats, Zapfino (via Flash)
Are Cookies Enabled?	0.43	1.35	Yes
Limited supercookie test	0.95	1.93	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No

OS Fingerprinting based on DNS Queries

au.download.windowsupdate.com
watson.microsoft.com ipv6.msftncsi.com
gadgets.live.com weather.service.msn.com
money.service.msn.com

Windows 7

swscan.apple.com swdist.apple.com
swcdnlocator.apple.com su.itunes.apple.com
time.euro.apple.com radarsubmissions.apple.com
internalcheck.apple.com identity.apple.com
configuration.apple.com init.ess.apple.com [init-p\[x\]md.apple.com](http://init-p[x]md.apple.com)
[p\[x\]-contacts.icloud.com](http://p[x]-contacts.icloud.com) [p\[x\]-caldav.icloud.com](http://p[x]-caldav.icloud.com)
[p\[x\]-imap.mail.me.com](http://p[x]-imap.mail.me.com) [\[x\].guzzoni-apple.com](http://[x].guzzoni-apple.com)
akadns.net ax.init.itunes.apple.com
[a\[x\].phobos.apple.com](http://a[x].phobos.apple.com) keyvalueservice.icloud.com

MacOS X 10.8.5

au.v4.download.windowsupdate.com ds.download.windowsupdate.com
bg.v4.emdl.ws.microsoft.com definitionupdates.microsoft.com
spynet2.microsoft.com watson.telemetry.microsoft.com
sqm.telemetry.microsoft.com clientconfig.passport.net ssw.live.com
client.wns.windows.com appexbingfinance.trafficmanager.net
appexbingweather.trafficmanager.net appexsports.trafficmanager.net
[appexdb\[x\].stb.s-msn.com](http://appexdb[x].stb.s-msn.com) de-de.appex-rf.msn.com
finance.services.appex.bing.com [financeweur\[x\].blob.appex.bing.com](http://financeweur[x].blob.appex.bing.com)
weather.tile.appex.bing.com

Windows 8

similar for iOS, Windows Phone and Android OS

mirrorlist.centos.org
[\[x\].centos.pool.ntp.org](http://[x].centos.pool.ntp.org)

CentOS 6

changelogs.ubuntu.com ntp.ubuntu.com geoip.ubuntu.com
daisy.ubuntu.com _https._tcp.fs.one.ubuntu.com [fs-\[x\].one.ubuntu.com](http://fs-[x].one.ubuntu.com)

Ubuntu 12.04

OS Fing

au.download
watson.micr
gadgets.live
money.servi

Window

swdist.apple.com su.itunes.apple.com
time.euro.apple.com internalcheck.apple.com
identity.apple.com configuration.apple.com p[x]-
contacts.icloud.com p[x]-caldav.icloud.com
[x].guzzoni-apple.com.akadns.net
keyvalueservice.icloud.com

MacOS X 10.0.5

au.v4.download.windowsupdate.com
definitionupdates.microsoft.com
spynet2.microsoft.com
watson.telemetry.microsoft.com
clientconfig.passport.net
ssw.live.com
client.wns.windows.com
appexbingweather.trafficmanager.net ...

*similar for iOS, Windows
Phone and Android OS*

mirrorlist.centos.org
[x].centos.pool.ntp.org

CentOS 6

Browser Fingerprinting based on DNS Queries

*aus3.mozilla.org download.cdn.mozilla.net fhr.data.mozilla.com
 services.addons.mozilla.org versioncheck-bg.addons.mozilla.org
 versioncheck.addons.mozilla.org addons.mozilla.org cache.pack.google.com
 download.mozilla.org [x].pack.google.com safebrowsing-cache.google.com
 safebrowsing.clients.google.com tools.google.com*

Firefox

*safebrowsing.google.com translate.googleapis.com [xxxxxxxxxx].
 [domain] apis.google.com cache.pack.google.com clients[x].google.com
 [x].pack.google.com safebrowsing-cache.google.com
 safebrowsing.clients.google.com ssl.gstatic.com tools.google.com
 www.google.com www.google.de www.gstatic.com*

Chrome

*apis.google.com clients.l.google.com clients1.google.com
 safebrowsing-cache.google.com
 safebrowsing.clients.google.com ssl.gstatic.com
 www.google.com www.google.de www.gstatic.com*

Safari

*ctldl.windowsupdate.com iecvlist.microsoft.com
 t.urs.microsoft.com*

Internet Explorer

DNS leaks information about setup & environment

Environmental fingerprinting?

1278194041.274	134.100.15.31	www.cnn.com A +
1278194041.278	132.100.15.31	ad-emea.doubleclick.net A +
1278219213.110	132.100.15.31	download.windowsupdate.com A +
1278221941.040	132.100.15.31	fbidc2008a.informatik.uni-hamburg.de SRV +

Installed Applications

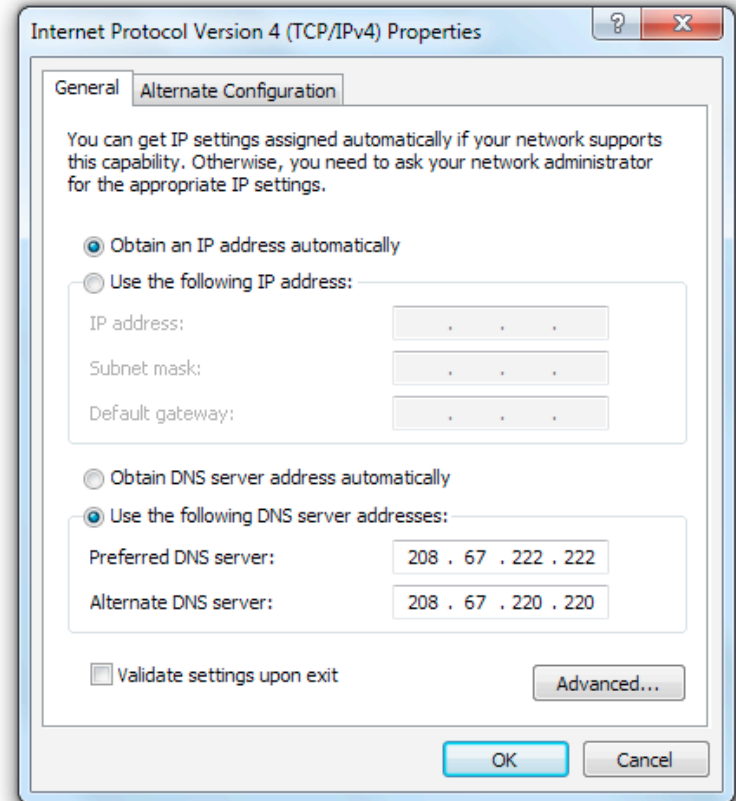
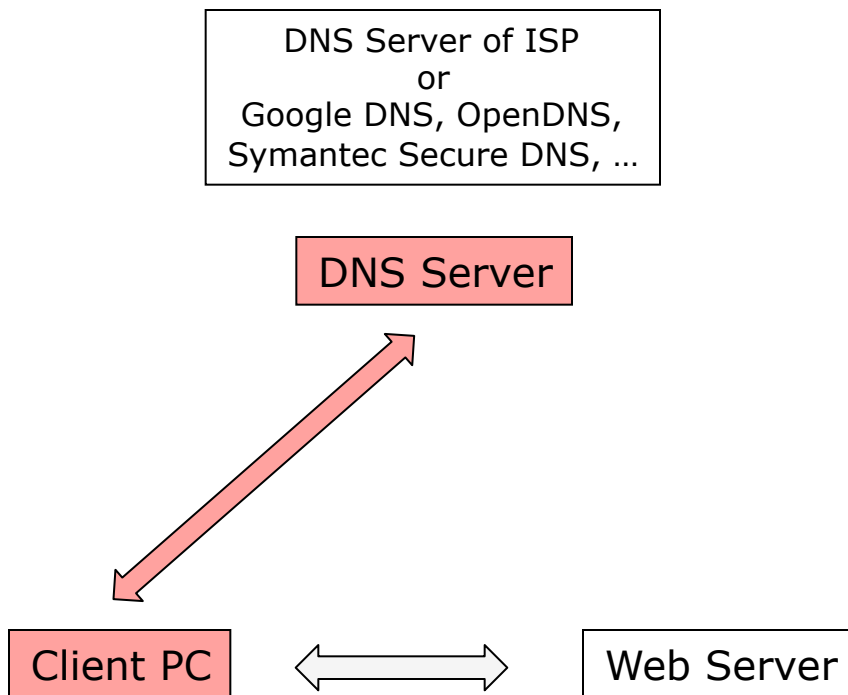
Neighboring machines

1279552941.192	87.2.55.11	FRITZ!NAS.fritz.box A +
1279553021.142	87.2.55.11	personal.avira-update.com A +
1279823365.030	87.2.55.11	ui.skype.com A +
1279553010.891	87.2.55.11	PAULSPC-16K2966SDJJ.fritz.box A +

Own hostname

Local domain suffix

Where can DNS data be observed or confiscated?





Objective 2:
Find evidence for involvement in criminal activities

Case Study 3: Human Behavior Fingerprinting

The problem of linking activities of a user over time

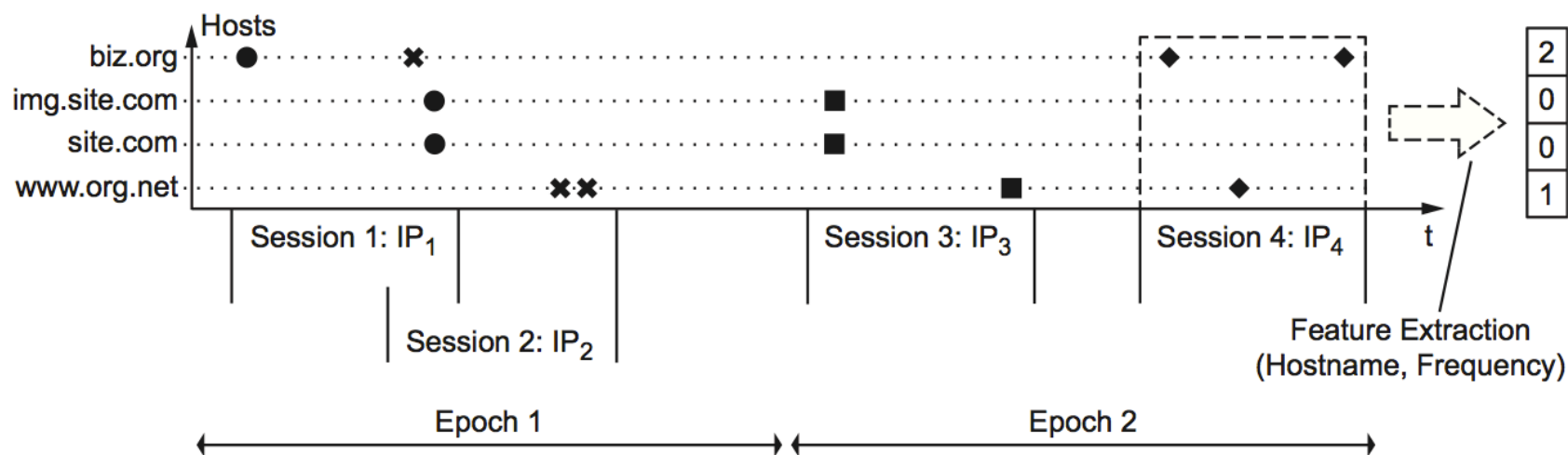


Behavioral fingerprints of users

- The Crime Scene
 - Day 1: subject carries out criminal activity on the network
 - Day 2: subject identifies himself during online shopping
 - investigator has access to network traffic on Day 1 and Day 2
- Digital Forensics Objective
 - **ascription/association:** find corroborating evidence that the subject identified on Day 2 is the same as the subject that was involved in criminal activity on Day 1
- Fingerprinting Approach
 - relies on characteristic **behavior** of humans
 - train a classifier: investigator collects traffic samples of multiple users on Day 1 and uses machine learning to extract fingerprints
 - classifier is used to determine whether the session of the suspect on Day 2 matches the behavioral fingerprint from Day 1

Behavior-based linking of sessions of a subject

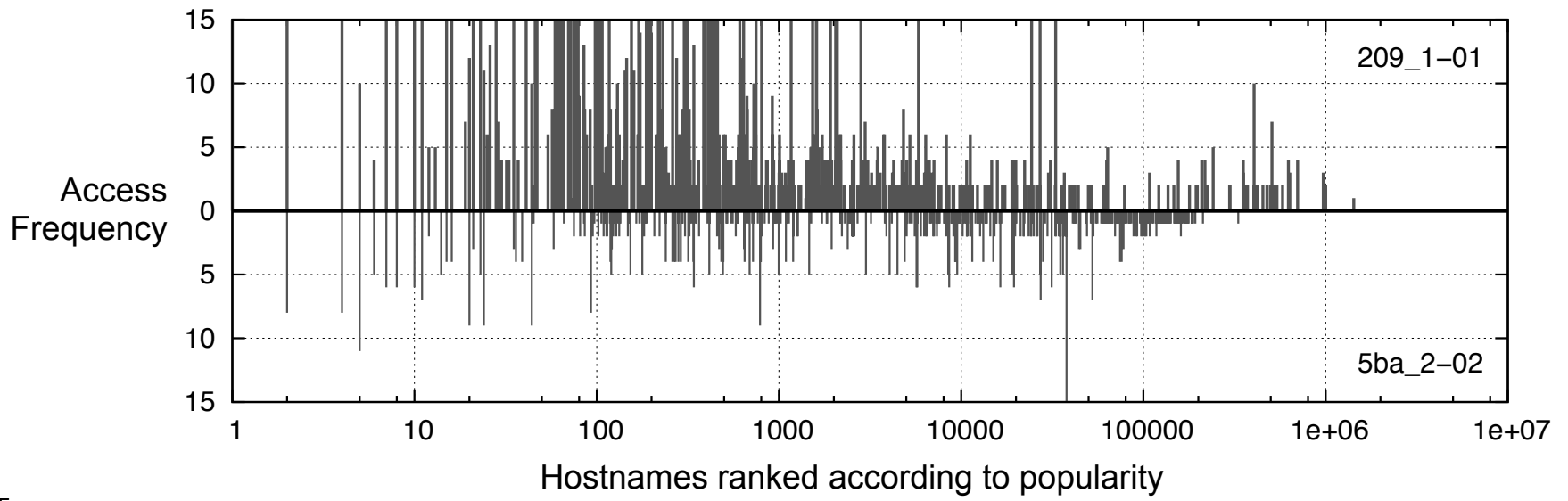
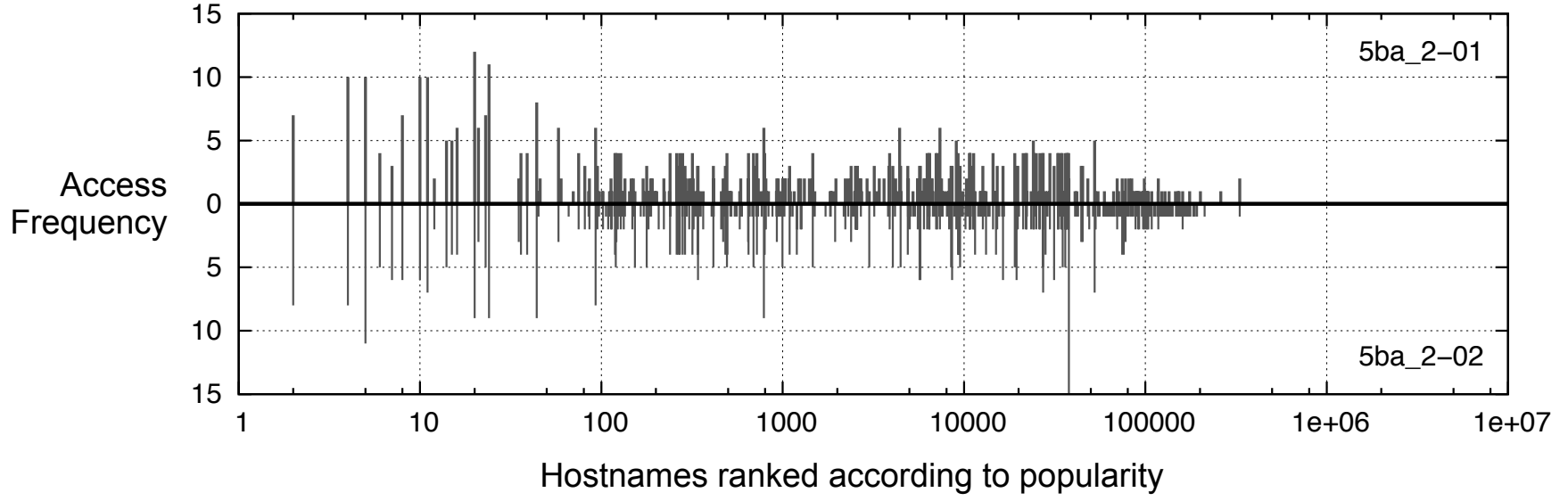
..., Herrmann, Banse, and Federrath (2013), ...



- Fingerprinting approach
 - profile: hostnames in DNS queries, number of queries per name
 - all queries of a user within a session grouped by source IP

(approach not limited to DNS traffic)

Is behavior-based fingerprinting feasible?



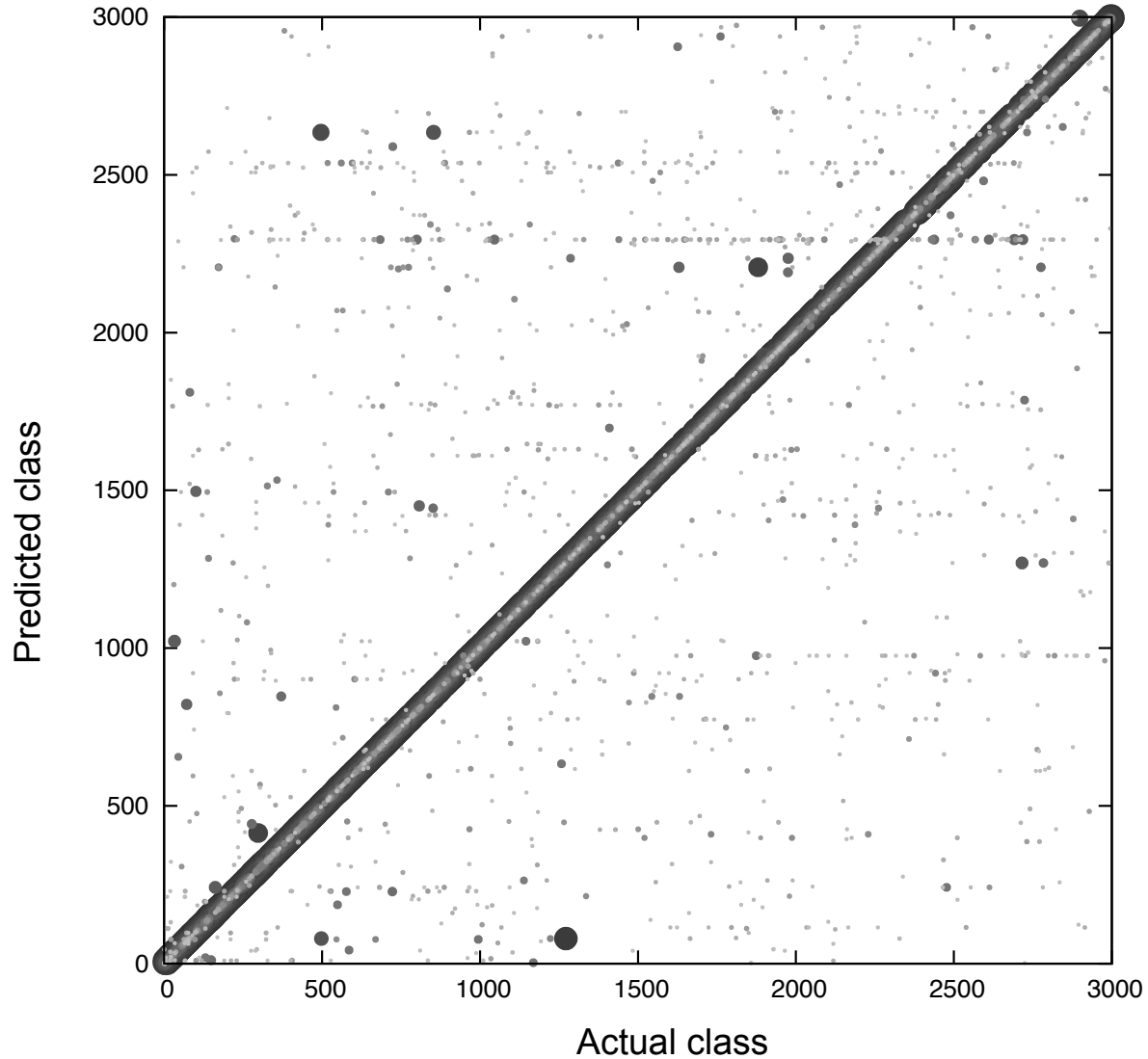
Behavior-based linking of sessions of a subject

- Evaluation approach
 - obtained a DNS log of University of Regensburg
 - 2 months, 3860 users, 431 mn. queries, 5 mn. hostnames
 - implement linking technique with 1NN and Naïve-Bayes classifier

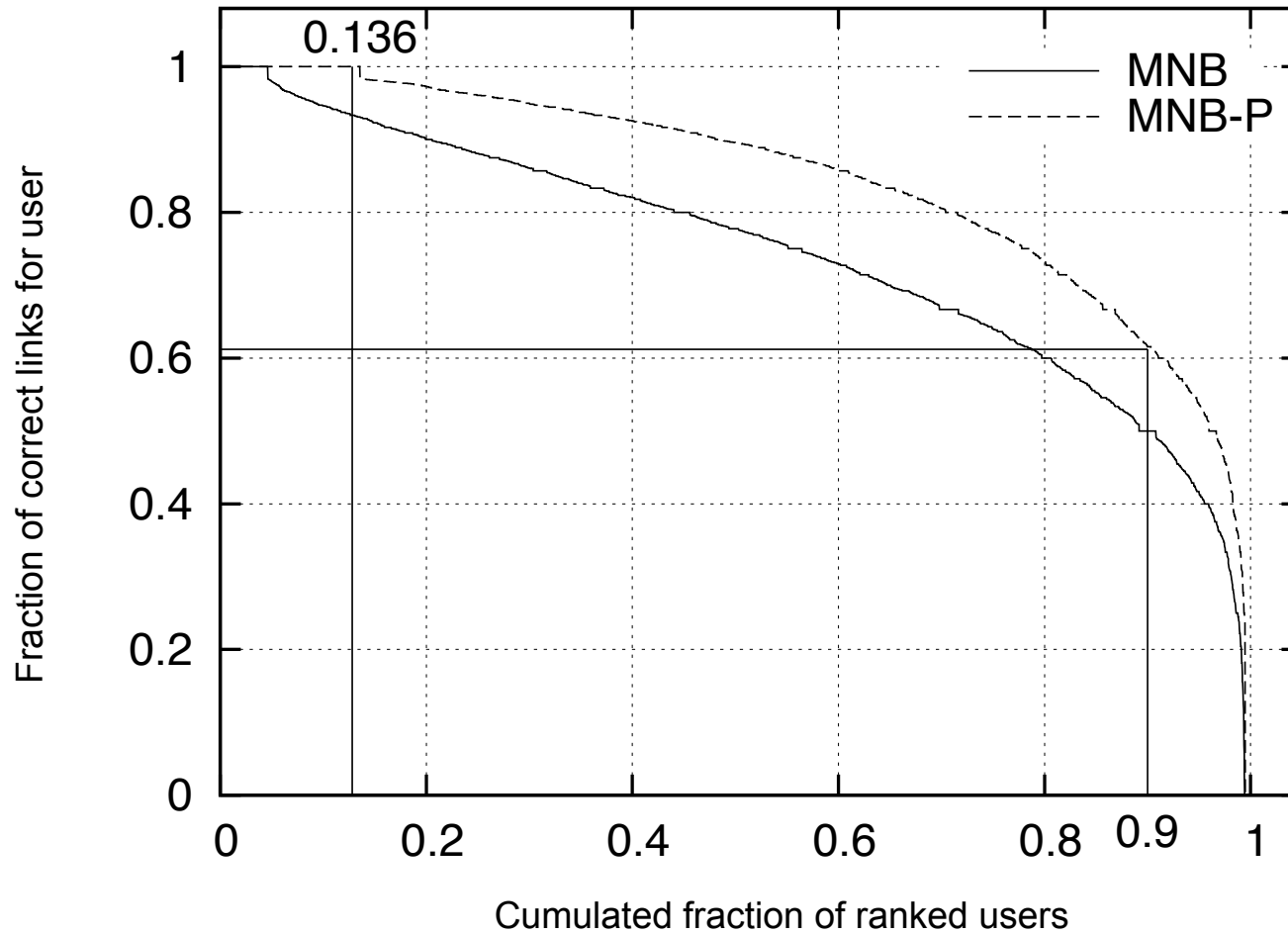
Apache Hadoop Cluster 18 quadcore desktop machines



Result: on average 86 % of day-to-day sessions linked correctly



Result: most users re-identified correctly most of the time



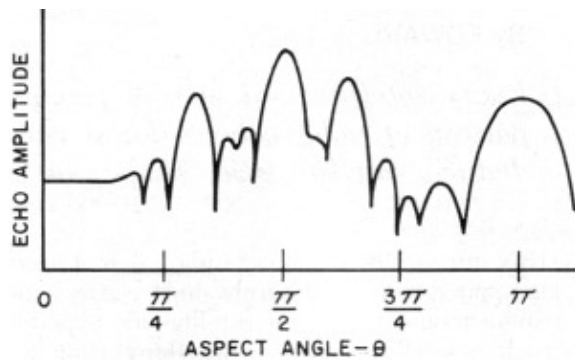
Fingerprinting for Forensics:
A new **promising** opportunity or a **dangerous** instrument?

Opportunities for Fingerprinting in Network Forensics

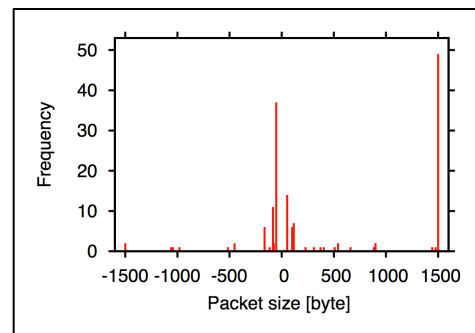
- Use cases
 - infer actions even when communication is encrypted
 - ascription of criminal actions, association/involvement of devices
- Utility for **blanket surveillance** and dragnet investigations
 - trace back potentially incriminating activities to the source to determine what should be investigated in detail (“leads”)
- Utility as corroborating evidence **in court**
 - implicit characteristics are unavoidable, difficult to forge (?)
- Utility of fingerprints for defense: to **disprove false accusations?**
 - should users pre-emptively keep a log of their own activities to provide counterevidence?

Challenges and Risks

- Unclear probative value
 - poor explainability of the decision of a machine learning system
 - required accuracy? robust evaluation (via standard corpora)?
- Future work: active fingerprinting via labeling/watermarking?
- Will feasible techniques lead to calls for pre-emptive surveillance?
- Identity theft vs. fingerprint theft
 - fingerprints can be stolen and re-injected
 - easier than with fingerprints of physical devices (?)



VS



Summary

Fingerprinting: diversity and stability of characteristics

Determine activities of a subject, **even if traffic is encrypted**
Infer associations: evidence for involvement in criminal activities

Three Case Studies:

Website
Fingerprinting

Device/Software
Fingerprinting

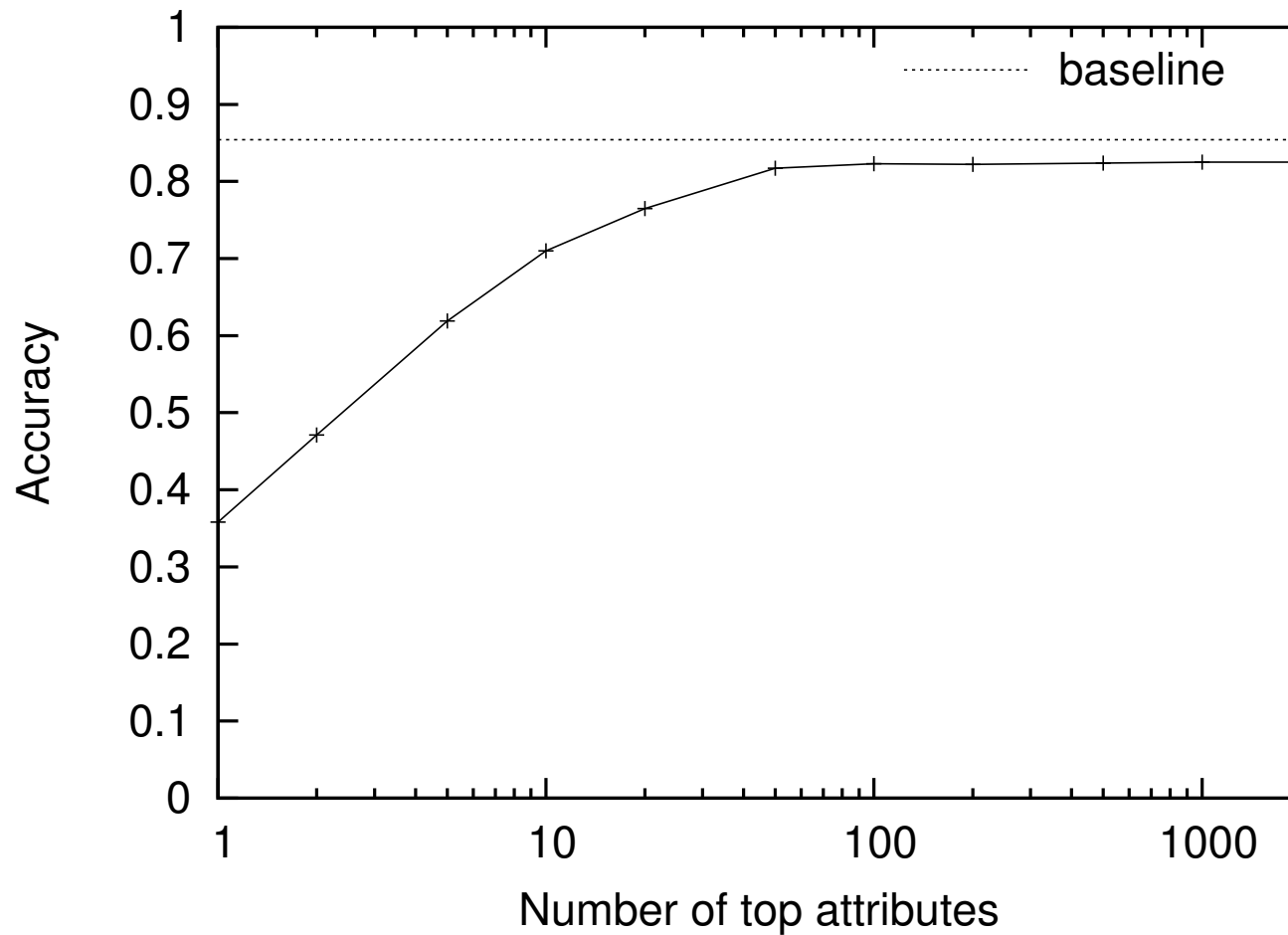
Human Behavior
Fingerprinting

Fingerprinting for Forensics:
A new **promising** opportunity or a **dangerous** instrument?

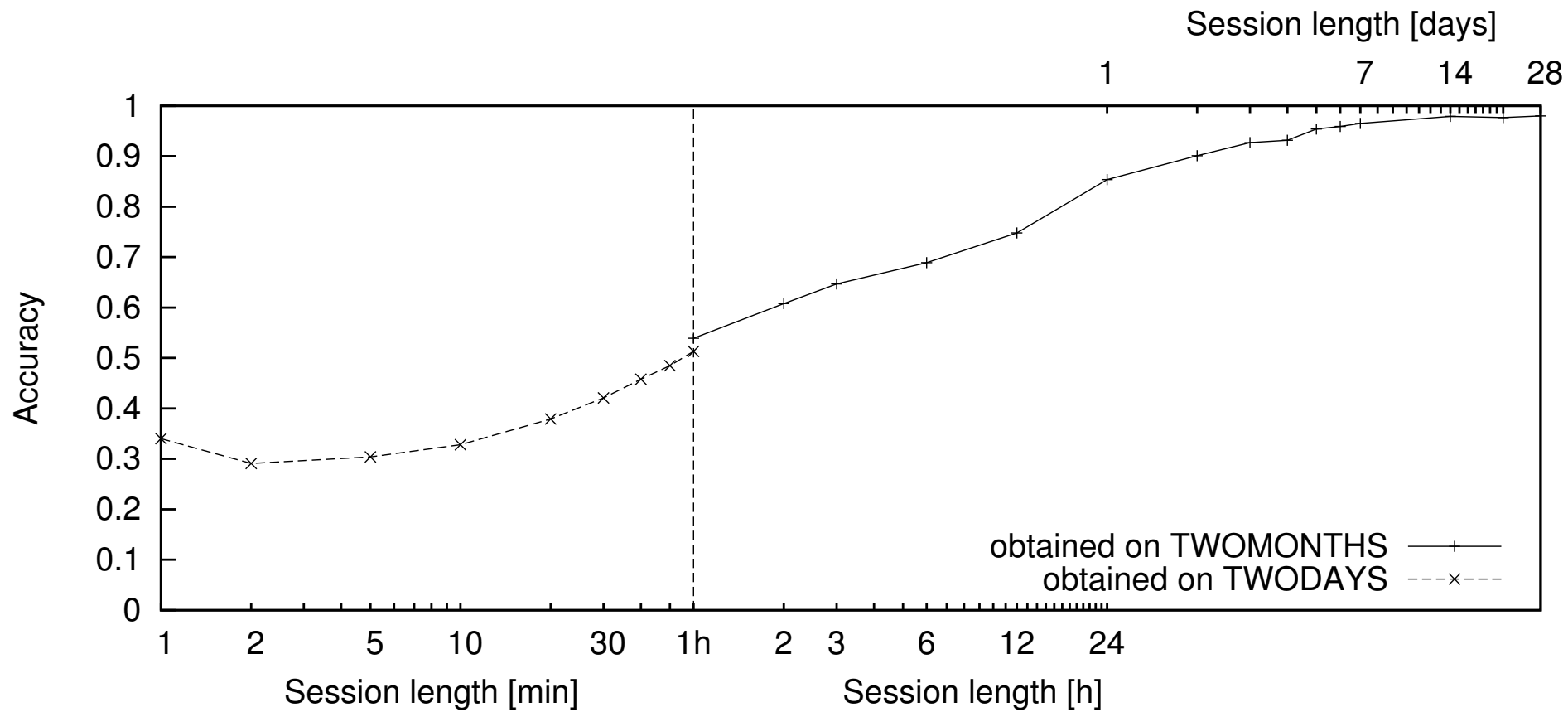


Backup

Result: session linkage relies on most popular hostnames only



Result: linking activities works also with shorter sessions



References

Casey, E.: Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet, p. 15, 2011

Eckersley, P.: How unique is your web browser? in: Privacy Enhancing Technologies. Springer Berlin Heidelberg, 2010.

Herrmann, D., Banse, C. & Federrath, H.: Behavior-based Tracking: Exploiting Characteristic Patterns in DNS Traffic. Computers & Security, Volume 39, Part A, pp. 17-33, Elsevier, November 2013.

Herrmann, D., Wendolsky, W. & Federrath, H.: Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naïve-Bayes Classifier. in: CCSW '09: Proceedings of the 2009 ACM workshop on Cloud computing security. ACM, New York, NY, 2009.

Kohno, T., Broido, A., & Claffy, K. C: Remote physical device fingerprinting. *Dependable and Secure Computing, IEEE Transactions on*, 2(2), 2005.

Krishnan, S. & Monroe, F.: DNS prefetching and its privacy implications: When good things go bad. Proceedings of the 3rd USENIX conference on Large-scale exploits and emergent threats: botnets, spyware, worms, and more. USENIX Association, 2010.

Mowery, K., & Shacham, H.: Pixel perfect: Fingerprinting canvas in HTML5. *Proceedings of W2SP*, 2012.

Lacy, E. A.: Radar Signature Analysis. in: February 1967 Electronics World, 1967.

Srihari, S. N., Cha, S. H., Arora, H. & Lee, S.: Individuality of handwriting: a validation study. in: Proceedings of Sixth International Conference on Document Analysis and Recognition, IEEE, 2001.

Talbot, K. I., Duley, P. R. & Hyatt, M. H.: Specific Emitter Identification and Verification. in: Technology Review Journal, Spring/Summer 2003.