DNSSEC / Privacy

CS 161: Computer Security

Prof. Vern Paxson

TAs: Devdatta Akhawe, Mobin Javed & Matthias Vallentin

http://inst.eecs.berkeley.edu/~cs161/

April 5, 2011
Today’s Lecture

• Finish discussion of **DNSSEC**
  – Ensuring that DNS results indeed match those from the corresponding authority

• A look at **privacy**
  – Mechanisms & practices that subvert it
  – Technical measures to obtain it
Securing DNS Lookups

• How can we ensure that when clients look up names with DNS, they can trust the answers they receive?

• Idea #1: do DNS lookups over TLS
  – (assuming either we run DNS over TCP, or we use “Datagram TLS”)
  – Issues?
    • Performance: DNS is very lightweight. TLS is not.
    • Caching: crucial for DNS scaling. But then how do we keep authentication assurances?

• Idea #2: make DNS results like certs
  – I.e., a verifiable signature that guarantees who generated a piece of data; signing happens off-line
Operation of DNSSEC

• DNSSEC = standardized DNS security extensions currently being deployed

1. Suppose we look up mail.google.com
   – We get an answer from google.com nameserver (NS)
   – Plus: signature for answer (in Additional section) purportedly signed by google.com NS

2. Look up public key for google.com NS
   – That answer is signed by .com NS

3. Look up public key for .com NS
   – That answer is signed by root (‘.’) NS

4. Root NS’s public key is wired into our resolver

• All of these keys are cacheable
Ordinary DNS:

Client \( \rightarrow \) mail.google.com? \( \rightarrow \) ns1.google.com

Client \( \rightarrow \) mail.google.com A 1.2.3.4

DNSSEC:

Client \( \rightarrow \) mail.google.com? \( \rightarrow \) ns1.google.com

Client \( \rightarrow \) mail.google.com A 1.2.3.4

Client \( \rightarrow \) google.com? \( \rightarrow \) com’s NS

Client \( \rightarrow \) google.com KEY 0x828C..E

Client \( \rightarrow \) SIG 0x90A4..5
DNS:

Client → mail.google.com? → ns1.google.com

mail.google.com A 1.2.3.4

DNSSEC:

Client → mail.google.com? → ns1.google.com

Client → mail.google.com A 1.2.3.4

 SIG 0x1F92..9

 Client → google.com? → com’s NS

 google.com KEY 0x828C..E

 SIG 0x90A4..5

This key …
DNS:

Client → mail.google.com? → ns1.google.com

mail.google.com A 1.2.3.4

DNSSEC:

Client → mail.google.com? → ns1.google.com

mail.google.com A 1.2.3.4

SIG 0x1F92..9

Client → google.com? → com’s NS

google.com KEY 0x828C..E
SIG 0x90A4..5

… validates this signature
Similarly, the root will return .com’s KEY, which will validate this signature.
Finally, we can validate the .com KEY supposedly returned by the root using our *hardwired* knowledge of the root’s public key.
Upon completing all of these signature validations, we have strong confidence in the chain of signatures, and thus in the correctness of the mail.google.com result.
Issues With DNSSEC?

• Issue #1: Replies are Big
  – E.g., “dig +dnssec berkeley.edu” can return 2100+ B
  – DoS amplification
  – Increased latency on low-capacity links
  – Headaches w/ older libraries that assume replies < 512B

• Issue #2: Partial deployment
  – Suppose .com not signing, though google.com is
  – Major practical concern. What do we do?
  – Can wire additional key into resolver (doesn’t scale)
  – Or: outsource to trusted third party (”lookaside”)  
    • Wire their key into resolver, they sign numerous early adopters
Issues With DNSSEC, con’t

• Issue #3: *Partial deployment*
  – What do you do with unsigned/unvalidated results?
  – If you trust them, *weakens incentive* to upgrade
  – If you don’t trust them, a whole lot of things *break*

• Issue #4: Negative results ("no such name")
  – What statement does the nameserver sign?
  – If “gabluph.google.com” doesn’t exist, then have to do dynamic key-signing (expensive) for any bogus request
    • DoS vulnerability
  – Instead, sign (off-line) statements about order of names
    • E.g., sign “gabby.google.com followed by gabrunk.google.com”
    • Thus, can see that gabluph.google.com can’t exist
  – But: now attacker can *enumerate* all names that exist :-(

Issues With DNSSEC, con’t

• Issue #5: *Who do you really trust?*
  – For your laptop, say, who does all the “grunt work” of fetching keys & validating DNSSEC signatures?

• Convenient answer: your laptop’s local resolver
  – … which you acquire via DHCP in your local coffeeshop
  – I.e., exactly the most-feared potentially *untrustworthy* part of the DNS resolution process!

• Alternatives?
  ⇒ Your laptop needs to do all the validation work itself
Privacy
Defining Privacy

• Privacy = right to control who knows certain aspects about you / your communications / your activities
  – Control over disclosure
  – And ideally over subsequent use

• How much of an issue is this?
  E.g., how much information about you do web sites learn as you surf?
Privacy & Web Surfing

• The sites you visit learn:
  – The URLs you’re interested in
    • Google/Bing also learns what you’re searching for
  – Your IP address
    • Thus, your service provider & geo-location
    • Can often link you to other activity including at other sites
  – Your browser’s capabilities, which OS you run, which language you prefer
  – Which URL you looked at that took you there
    • Via “Referer” header
Privacy & Web Surfing, con’t

- Oh and also cookies.
- Cookies = **state** that server tells browser to store locally
  - Name/value pair, plus expiration date
- Browser returns the state any time visiting the same site

- Where’s the harm in that?
  And are these used much anyway?
Let’s remove all of our cookies.
We do a Google search on “private browsing” and click on the top result.
Note that this mode is privacy from your family, not from web sites!
What on earth is Google tracking in this one?

Whoa - we gained 11 cookies!

It sticks around for 6 months.
Hmmm. Mozilla is tracking us too. And for 5 years!
They're even remembering just how we visited them.
And something else (as we’ll see in a bit) until the End Of Time
Without doing anything else, we’ve gained a 12th cookie …
We now do just one more operation, opening the home page of www.nytimes.com
What a lot of yummy cookies!

doubleclick.net - who's that? And how did it get there from visiting www.nytimes.com?
Third-Party Cookies

• How can a web site enable a third party to plant cookies in your browser & later retrieve them?
  – Answer: using a “web bug”
  – Include on the site’s page (for example):
    • `<img src="http://doubleclick.net/ad.gif" width=1 height=1>`

• Why would a site do that?
  – Site has a business relationship w/ DoubleClick*
  – Now DoubleClick sees all of your activity that involves their web sites (each of them includes the web bug)
    • Because your browser dutifully sends them their cookies for any web page that has that web bug
    • Identifier in cookie ties together activity as = YOU

* Owned by Google, by the way
Remember this till-the-End-of-Time cookie?

<table>
<thead>
<tr>
<th>Site</th>
<th>Cookie Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>google.com</td>
<td>NID, PREF, SS</td>
</tr>
<tr>
<td>mozilla.com</td>
<td>s_vi, s_sq, s_cc</td>
</tr>
<tr>
<td>support.mozilla.com</td>
<td>__utmz, __utmc, __utmb, utma, SUMOv1</td>
</tr>
</tbody>
</table>

Name: __utma
Content: 92405663.30107794.1269986049.1269986049.1269986049.1269986049.1
Domain: .support.mozilla.com
Path: /
Send For: Any type of connection
Expires: January 17, 2038 4:00:00 PM
Google Analytics

- Any web site can (anonymously) register with Google to instrument their site for analytics
  - Gather information about who visits, what they do when they visit
- To do so, site adds a small Javascript snippet that loads http://www.google-analytics.com/ga.js
  - You can see sites that do this because they introduce a "__utma" cookie
- Code ships off to Google information associated with your visit to the web site
  - Shipped by fetching a GIF w/ values encoded in URL
  - Web site can use it to analyze their ad “campaigns”
  - Not a small amount of info …
Values Reported via Google Analytics

<table>
<thead>
<tr>
<th>Affiliation</th>
<th>Host Name</th>
<th>Screen Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Billing City</td>
<td>Java-enabled</td>
<td>Shipping Cost</td>
</tr>
<tr>
<td>Billing Country</td>
<td>Language Encoding</td>
<td>Special Event</td>
</tr>
<tr>
<td>Billing Region</td>
<td>Order ID</td>
<td>Start Campaign Sess.</td>
</tr>
<tr>
<td>Browser Lang.</td>
<td>Page Title</td>
<td>Tax</td>
</tr>
<tr>
<td>Complete URL</td>
<td>Product Code</td>
<td>Tracking Code Version</td>
</tr>
<tr>
<td>Cookie Values</td>
<td>Product Name</td>
<td>Unique GIF ID</td>
</tr>
<tr>
<td>Current Page</td>
<td>Profile Number</td>
<td>Unit Price</td>
</tr>
<tr>
<td>Event Tracking</td>
<td>Repeat Campaign Visit</td>
<td>User Defined Var</td>
</tr>
<tr>
<td>Flash Version</td>
<td>Quantity</td>
<td>Variations on an Item</td>
</tr>
<tr>
<td>Grand Total</td>
<td>Screen Color Depth</td>
<td></td>
</tr>
</tbody>
</table>
Still More Tracking Techniques …

• Any scenario where browsers execute programs that manage persistent state can support tracking by cookies
  – Such as …. *Flash* ?
My browser had Flash cookies from 67 sites!

Sure, this is where you’d think to look to analyze what Flash cookies are stored on your machine.

Some Flash cookies “respawn” regular browser cookies that you previously deleted!

My browser had Flash cookies from 67 sites!
Still More Tracking Techniques …

• Any scenario where browsers execute programs that manage persistent state can support tracking by cookies
  – Such as …. Flash ?

• Surely though something as innocuous as cut-and-paste is safe though, right?
  (demo)
Leverage the benefit of copy/paste

Make link-backs to your content effortless for readers and gain new insight into user engagement with Tynt Publisher Tools.

FREE! Get Started

Keep Your Users

*tynt keywords*
Find out what outbound keywords are causing your users to leave.

*tynt seo (formerly insight)*
Generate more search engines visible links back to your content.

Measure Social Impact

*tynt social*
Understand what social channels are most effective.

tynt® labs / api
See how Tynt is showcasing the most engaging content on the web. ...more →
I Can Stalk U
Raising awareness about inadvertent information sharing

Who have we stalked recently?

ICanStalkU was able to stalk RangelLifeEnt at 51 Great Jones St New York NY
1 minute ago • Map Location • View Tweet • View Picture • Reply to RangelLifeEnt

ICanStalkU was able to stalk Inicklsson at http://maps.google.com/?q=57.134444444,12.714166667
2 minutes ago • Map Location • View Tweet • View Picture • Reply to Inicklsson

ICanStalkU was able to stalk Welerson13 at http://maps.google.com/?q=-15.738055556,-47.898611111
2 minutes ago • Map Location • View Tweet • View Picture • Reply to Welerson13

ICanStalkU was able to stalk BritBangert at 920 Hawley St Taylorville IL
1 minute ago • Map Location • View Tweet • View Picture • Reply to BritBangert

ICanStalkU was able to stalk jiggy_Owla at http://maps.google.com/?q=13.7830055879,100.518500685
4 minutes ago • Map Location • View Tweet • View Picture • Reply to jiggy_Owla

ICanStalkU was able to stalk gcolony at http://maps.google.com/?q=37.785166667,-122.40416667
4 minutes ago • Map Location • View Tweet • View Picture • Reply to gcolony

Links
- Mayhemic Labs
- PaulDotCom
- SANS ISC
- Electronic Frontier Foundation
- Center for Democracy & Technology

How did you find me?

Did you know that a lot of smart phones encode the location of where pictures are taken? Anyone who has a copy can access this information.

read more

Help me fix this!

Disabling Geo-Tagging on your phone is easy.
My baby girl.... http://t.co/5qLfLV6

2 minutes ago via Twitter for Android

BritBangert
Brittany Bangert
Privacy - What’s the Big Deal?

• Cookies etc. form the core of how Internet advertising works today
  – Without them, arguably you’d have to pay for content up front a lot more
    • (and payment would mean you’d lose anonymity anyway)
  – A “better ad experience” is not necessarily bad
    • Ads that reflect your interests; not seeing repeated ads

• But: ease of gathering so much data so easily ⇒ concern of losing control how it’s used
  – Mission creep …
    • Consider how ordering a pizza in the near future might work ([http://www.aclu.org/ordering-pizza](http://www.aclu.org/ordering-pizza))
  – Content shared with friends doesn’t just stay with friends …
More Employers Screening Candidates via Social Networking Sites

Five tips for creating a positive online image

Rosemary Haefner, Vice President of Human Resources at CareerBuilder

Gone are the days when all job seekers had to worry about were their résumés and cover letters. Today, those documents remain a staple of the job-search process, but they are joined by a growing phenomenon: social networking.

Forty-five percent of employers reported in a June 2009 CareerBuilder survey that they use social networking sites to screen potential employees, compared to only 22 percent of employers last year. Eleven percent of employers plan to start using social networking sites for the screening process. More than 2,600 hiring managers participated in the survey.
Why employers disregard candidates after screening online

Thirty-five percent of employers reported they have found content on social networking sites that caused them not to hire the candidate, including:

- Candidate posted provocative or inappropriate photographs or information -- 53 percent
- Candidate posted content about them drinking or using drugs -- 44 percent
- Candidate bad-mouthed their previous employer, co-workers or clients -- 35 percent
- Candidate showed poor communication skills -- 29 percent
- Candidate made discriminatory comments -- 26 percent
- Candidate lied about qualifications -- 24 percent
- Candidate shared confidential information from previous employer -- 20 percent
How To Gain Better Privacy?

• Force of law
  – Example #1: web site privacy policies
    • US sites that violate them commit false advertising
    • But: policy might be “Yep, we sell everything about you, Ha Ha!”
7. Collection of Viewing Information. You acknowledge that you are aware of and consent to the collection of your viewing information during your use of the Software and/or Content. Viewing information may include, without limitation, the time spent viewing specific pages, the order in which pages are viewed, the time of day pages are accessed, IP address and user ID. This viewing information may be linked to personally identifiable information, such as name or address and shared with third parties.
How To Gain Better Privacy?

- Force of law
  - Example #1: web site privacy policies
    - US sites that violate them commit false advertising
    - But: policy might be “Yep, we sell everything about you, Ha Ha!”
  - Example #2: SB 1386
    - Requires an agency, person or business that conducts business in California and owns or licenses computerized 'personal information' to disclose any breach of security (to any resident whose unencrypted data is believed to have been disclosed)
    - Quite effective at getting sites to pay attention to securing personal information
May 8, 2009 1:53 PM PDT

UC Berkeley computers hacked, 160,000 at risk

by Michelle Meyers

This post was updated at 2:16 p.m. PDT with comment from an outside database security software vendor.

Hackers broke into the University of California at Berkeley’s health services center computer and potentially stole the personal information of more than 160,000 students, alumni, and others, the university announced Friday.

At particular risk of identity theft are some 97,000 individuals whose Social Security numbers were accessed in the breach, but it's still unclear whether hackers were able to match up those SSNs with individual names, Shelton Waqqener, UCB's chief technology officer, said in a press conference Friday afternoon.