Anonymity / Sneakiness

### CS 161: Computer Security Prof. Vern Paxson

# TAs: Devdatta Akhawe, Mobin Javed & Matthias Vallentin

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

April 7, 2011

### **Today's Lecture**

- A look at technical means for one form of anonymity: hiding one's IP address –"Onion routing"
- A look sneakiness
  - –Ways of communicating or computing by cheating

### Gaining Privacy Through Technical Means

- How can we surf the web truly anonymously?
- Step #1: remove browser leaks
  - Delete cookies (oops also Flash cookies!)
  - Turn off Javascript (so Google Analytics doesn't track you)
- Step #2: how do we hide our IP address?
- One approach: trusted third party
  - E.g.



### Gaining Privacy Through Technical Means

- How can we surf the web truly anonymously?
- Step #1: remove browser leaks
  - Delete cookies (oops also "Flash cookies"!)
  - Turn off Javascript (so Google Analytics doesn't track you)
- Step #2: how do we hide our IP address?
- One approach: trusted third party
  - E.g. hidemyass.com
    - You set up an encrypted VPN to their site
    - All of your traffic goes via them

Alice wants to send a message M to Bob ...

... but ensuring that Eve can't determine that she's indeed communicating with Bob.

# HMA accepts messages encrypted for it. Extracts destination and forwards.

### Gaining Privacy Through Technical Means

- How can we surf the web truly anonymously?
- Step #1: remove browser leaks
  - Delete cookies (oops also "Flash cookies"!)
  - Turn off Javascript (so Google Analytics doesn't track you)
- Step #2: how do we hide our IP address?
- One approach: trusted third party
  - E.g. hidemyass.com
    - You set up an encrypted VPN to their site
    - All of your traffic goes via them
  - Issues?
    - Performance
    - (\$80-\$200/year)
    - "*rubber hose cryptanalysis*" (cf. anon.penet.fi & Scientologists)

Alice wants to send a message M to Bob ...

... but ensuring that Eve can't determine that she's indeed communicating with Bob ...

... and that HMA can't determine it, either.



## **Onion Routing**

- This approach generalizes to an arbitrary number of intermediaries ("mixes")
- As long as any of the mixes is honest, no one can link Alice with Bob



## **Onion Routing Issues/Attacks?**

- Performance: message bounces around a lot
- Key management: the usual headaches
- Attack: rubber-hose cryptanalysis of mix operators
  - Defense: use mix servers in different countries
    - Though this makes performance worse :-(
- Attack: adversary operates all of the mixes
  - Defense: have lots of mix servers (Tor today: ~2,000)
- Attack: adversary observes when Alice sends and when Bob receives, links the two together
  - A "confirmation" attack
  - Defenses: pad messages, introduce significant delays
    - Tor does the former, but notes that it's not enough for defense

## **Onion Routing Attacks, con't**

- Issue: leakage
- Suppose all of your HTTP/HTTPS traffic goes through Tor, but the rest of your traffic doesn't
  - Because you don't want it to suffer performance hit
- How might the operator of sensitive.com deanonymize your web session to their server?
- Answer: they inspect the logs of their DNS server to see who looked up sensitive.com just before your connection to their web server arrived
- Hard, general problem: anonymity often at risk when adversary can correlate separate sources of information

### **Sneakiness**

### Steganography

 Transmitting hidden messages using a known communication channel

– No one knows the message is even there

 Same notion applies to hiding extra hidden data inside known storage

- Again, no one knows the data is there

- Goal: Sneak communication past a reference monitor ("warden")
- Does not imply confidentiality
  - If message is discovered, it's revealed
  - (Though you could decide to also encrypt it)

### Steganography, con't

- Examples?
  - Zillions: tattooed heads of slaves, least-significant bits of image pixels, extra tags in HTML documents, ...
  - All that's necessary is agreement between writer of message & reader of message ...
  - ... and some extra capacity
- Security?
  - Brittle: relies on *security-by-obscurity*
  - If well designed, and warden can only watch, then can be difficult to detect
  - If however warden can modify communication (e.g., recode images, canonicalize HTML, shave slave heads) then warden can disrupt/discover

## **Covert Channels**

- Communication between two parties that uses a hidden (secret) channel
- Goal: evade reference monitor inspection entirely

– Warden doesn't even realize communication is possible

- Again, main requirement is agreement between sender and receiver (established in advance)
- Example: suppose (unprivileged) process A wants to send 128 bits of secret data to (unprivileged) process B ...
  - But can't use pipes, sockets, signals, or shared memory; and can only read files, can't write them

### **Covert Channels, con't**

- Method #1: A syslog's data, B reads via /var/log/...
- Method #2: select 128 files in advance. A opens for read only those corresponding to 1-bit's in secret.

– **B** recovers bit values by inspecting access times on files

- Method #3: divide A's running time up into 128 slots. A either runs CPU-bound - or idle - in a slot depending on corresponding bit in the secret. B monitors A's CPU usage.
- Method #4: Suppose A can run 128 times. Each time it either exits after 2 seconds (0 bit) or after 30 seconds (1 bit).
- Method #5: ...
  - There are zillions of Method #5's!

### **Covert Channels, con't**

- Defenses?
- As with steganography, #1 challenge is identifying the mechanisms
- Some mechanisms can be very hard to completely remove
  - E.g., duration of program execution
- Fundamental issue is the covert channel's capacity (same for steganography)

- Bits (or bit-rate) that adversary can obtain using it

- Crucial for defenders to consider their threat model
- Usual assumption is that Attacker Wins (can't effectively stop communication, esp. if *low rate*)

### Side Channels

- Inferring information meant to be hidden / private by exploiting how system is structured
  - Note: unlike for steganography & covert channels, here we do *not* assume a cooperating sender / receiver
- Can be difficult to recognize because often system builders "abstract away" seemingly irrelevant elements of system structure
- Side channels can arise from physical structure ...



### Side Channels

- Inferring information meant to be hidden / private by exploiting how system is structured
  - Note: unlike for steganography & covert channels, here we do not assume a cooperating sender / receiver
- Can be difficult to recognize because often system builders "abstract away" seemingly irrelevant elements of system structure
- Side channel can arise from physical structure ...
  - ... or higher-layer abstractions

```
/* Returns true if the password from the
 * user, 'p', matches the correct master
 * password. */
                                 Attacker knows code,
bool check password(char *p)
                                but not this value
{
     static char *master_pw = "T0p$eCRET";
     int i;
    for(i=0; p[i] && master_pw[i]; ++i)
          if(p[i] != master pw[i])
               return FALSE;
    /* Ensure both strings are same len. */
     return p[i] == master pw[i];
```

### Inferring Password via Side Channel

 Suppose the attacker's code can call check\_password many times (but not millions)

- But attacker can't breakpoint or inspect the code

- How could the attacker infer the master password using side channel information?
- Consider layout of **p** in memory:

```
...
if(check_password(p))
BINGO();
```

wildGUe\$s

Spread p across different memory pages:



If master password doesn't start with 'w', then loop exits on first iteration (i=0):

If it *does* start with 'w', then loop proceeds to next iteration, generating a page fault that the caller can observe

#### T0p\$eCRET ?

Ajunk	No page fault
Bjunk	No page fault
•••	
Tjunk	Page fault!



```
bool check_password2(char *p)
{
     static char *master pw = "T0p$eCRET";
     int i;
     bool is_correct = TRUE;
    for(i=0; p[i] && master pw[i]; ++i)
          if(p[i] != master_pw[i])
               is correct = FALSE;
     if(p[i] != master_pw[i])
          is correct = FALSE;
     return is correct;
}
              Note: still leaks length of master password
```

### Side Channels in Web Surfing

• Suppose Alice is surfing the web and all of her traffic is encrypted ...

– ... and running through an anonymizer like HMA

- Eve can observe the presence of Alice's packets & their size, but can't read their contents or ultimate destination
- How can Eve deduce that Alice is visiting FoxNews (say)?



Done

Page Info – http://www.foxnews.com/			
	General Media Feeds Permissions Security	$\sim$	
Address		Size	₽.
http://www.fo	xnews.com/ucat/images/255017_laundry90.jpg	10.9 KB	
http://www.fo	xnews.com/i/90x70_us.jpg	9.76 KB	
http://www.fo	xnews.com/i/90x70_world.jpg	7.77 KB	
http://www.fo	xnews.com/i/90x70_politics.jpg	6.2 KB	
http://a57.fox	news.com/static/managed/img/Entertainment/2010/90/70/What-Makes-a-Bombshell.jpg	8.54 KB	
http://video.fo	xnews.com/thumbnails/042310/90/70/ASL-033110HEALTHFNEFACEVEINS-1FEFRC0A_FNC_042310	7.88 KB	
http://a57.fox	news.com/static/managed/img/Leisure/2009/90/70/vw400.jpg	7.51 KB	n
http://a57.fox	news.com/static/managed/img/Scitech/90/70/Asphalt%20Volcanoes%20in%20Pacific.jpg	6.76 KB	
http://a57.fox	news.com/static/managed/img/Opinion/90/70/Hendricks_ChristinaR307.jpg	7.72 KB	
http://www.fo	xnews.com/i/new/fncshed-bg.gif	0.46 KB	
http://www.fo	xnews.com/images/374022/1_51_90_oreilly_new.jpg	3.81 KB	
http://www.fo	xnews.com/images/604051/0_51_90_042310_han_newt.jpg	16.53 KB	
http://www.fo	xnews.com/images/604009/0_51_90_042310_greta_palin.jpg	6.56 KB	
http://www.fo	xnews.com/images/545380/0_51_90_baier_new.jpg	4.03 KB	
http://www.fo	xnews.com/images/604066/0_51_90_beck_regulations.jpg	3.6 KB	<b></b>
http://www.fo	xnews.com/images/604065/0_51_042310_90_yw_porn.jpg	14.5 KB	Ŧ
Location:	http://www.foxnews.com/i/new/right-head_bg.gif	$\smile$	

Eve "fingerprints" web sites based on the specific sizes of the items used to build them. Looks for groups of ciphertext that total the same sizes.

### Side Channels in Web Surfing

• Suppose Alice is surfing the web and all of her traffic is encrypted

– ... and running through an anonymizer like HMA

- Eve can observe the presence of Alice's packets & their size, but can't read their contents or ultimate destination
- How can Eve deduce that Alice is visiting FoxNews (say)?
- What about inferring what terms Alice is searching on?

8 × s	Q		si Q	
ns RSS	southwest ai Suggest	ns PS	sierra at ta Sugges	- 1
S	sfgate	p3 103.	singapore airlines	
S	skype		sierra trading post	F
S	safeway		sidereel	
S	sears		simon monjack	
S	super bowl 2010		silverlight	
S	san jose mercury news	;	sirius	
S	sports authority		silver legacy reno	
S	starbucks		sidestep	
S	speed test		six flags	



8	side	Q	
	sidereel	Suggest	
<b>53 N.S.</b>	sidestep		
	sidebar oak	land	
	sidekick		
	sidestep.com flights		
	sideways		
	sideboard danville		-
	side effects	of h1n1 v	
	side effects	of predni	
_	sideshow c	ollectibles	



8	side ch	Q)
	side chairs Sugg	est
ps KS	side chaining	
	side chain compress	sion
	side chain	
	side channel attack	
	side charging ar-15	
	sidechaining in logi	c
	side chairs contemp	o
	side charging upper	
	side chignon	



#### 102 chars.

#### 136 chars.

8	•	d	Q
ans	RSS	dictionary	Suggest
103	TLD.	dmv california	1
		delta airlines	
		disneyland	
		dominos	
		disney channe	el 🚺
		de young museum	
		doppelganger	· .
		daylight savin	gs time
	_	direct tv	



125 chars.

#### 101 chars.





#### 107 chars.

#### 102 chars.

8	f	Q
ns RS	facebook	Suggest
.ps 105	facebook login	
	fandango	
	firefox	
	food network	
	fedex	
	fafsa	
	fox news	
	frys	
	forever 21	

### Exploiting Side Channels For Stealth Scanning

- Can attacker using system A scan the server of victim V to see what services V runs ...
- ... without V being able to learn A's IP address?
- Seems impossible: how can A receive the results of probes A sends to V, unless probes include A's IP address for V's replies?

### **IP Header Side Channel**





-



**V**ictim

























### **UI Side Channel Snooping**

 Scenario: Ann the Attacker works in a building across the street from Victor the Victim. Late one night Ann can see Victor hard at work in his office, but can't see his CRT display, just the glow of it on his face.



 How might Ann snoop on what Victor's display is showing?



### CRT display is made up of an array of phosphor pixels









(a) Emission decay of a single pixel ( $f_p = 36$  MHz)



So if Ann can synchronize a high-precision clock with when the beam starts up here ...



Then by looking for changes in light level (flicker) matched with high-precision timing, she can tell whether say *this* pixel is on or off ...





Photomultiplier + high-precision timing + deconvolution to remove noise

# CAN YOU **READ THIS?** This image was captured with the help of a light sensor from the high-frequency fluctuations in the light emitted by a cathode-ray tube computer monitor which I picked up as a diffuse reflection from a nearby wall.

Markus Kuhn, University of Cambridge, Computer Laboratory, 2001

## **UI Side Channel Snooping**

- Victor switches to an LCD display. Any other ways Ann can still steal his display contents or his keystrokes?
- Cables from computer to screen & keyboard act as crude antennas!
  - Broadcast weak RF signals corresponding to data streams (as does a CRT's operation - "Tempest")
  - Even induce faint voltage fluctuations in power lines

### Stealing keystrokes through electric lines Relatively simple equipment can tap power lines to intercept what is being typed on nearby keyboards.









### • | 0 | 00111000 | 0 | 1 | = letter 'a'



Copyright 2009 Inverse Path Ltd.

## **UI Side Channel Snooping**

- Victor switches to an LCD display. Any other ways Ann can still steal his display contents or his keystrokes?
- Cables from computer to screen & keyboard act as crude antennas!
  - Broadcast weak RF signals corresponding to data streams
  - Even induce faint voltage fluctuations in power lines
- Keystrokes create sound
  - Audio components unique per key
  - Timing reflects key sequencing / touch typing patterns
    - If language known, can employ spell-checking to clean up errors
  - Can listen w/ any convenient microphone (e.g, telephone!)
  - Can "listen" from a distance using laser + telescope!





Figure 6. Reflections in two other tea pots, taken from a distance of 5m. The 18pt font is readable from the reflection in the left picture, and almost readable in the right picture.



Figure 7. Reflection of a Word document with small 12pt font size in a tea pot, taken from a distance of 5m. The 12pt font is readable from the reflection.