CSci 5271
Introduction to Computer Security
Day 23: Anonymizing the network

Stephen McCamant
University of Minnesota, Computer Science & Engineering

## Outline

Anonymous communications techniques

Announcements intermission

Tor basics

Tor experiences and challenges

## Traffic analysis

- What can you learn from encrypted data? A lot
- Content size, timing
- Who's talking to who
  - → countermeasure: anonymity

## Nymity slider (Goldberg)

- Verinymity
  - Social security number
- Persistent pseudonymity
  - Pen name ("George Eliot"), "moot"
- Linkable anonymity
  - Frequent-shopper card
- Unlinkable anonymity
  - (Idealized) cash payments

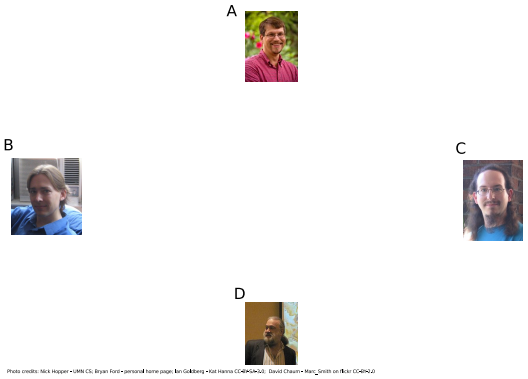
## Nymity ratchet?

- It's easy to add names on top of an anonymous protocol
- The opposite direction is harder
- But, we're stuck with the Internet as is
- So, add anonymity to conceal underlying identities

## Steganography

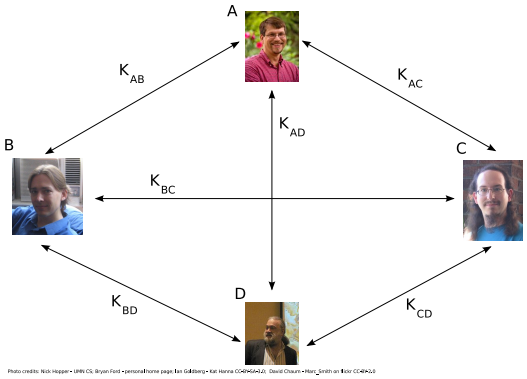
- One approach: hide real content within bland-looking cover traffic
- Classic: hide data in least-significant bits of images
- Easy to fool casual inspection, hard if adversary knows the scheme

# Dining cryptographers



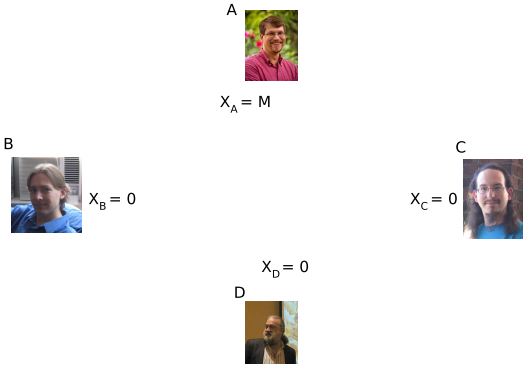



# Dining cryptographers



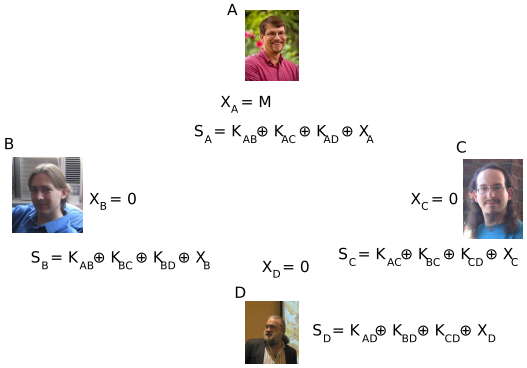



# Dining cryptographers






# Dining cryptographers



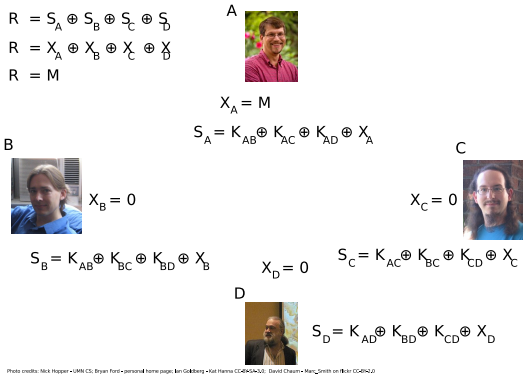



# Dining cryptographers

$R = S_A \oplus S_B \oplus S_C \oplus S_D$

$R = X_A \oplus X_B \oplus X_C \oplus X_D$

$R = M$






# DC-net challenges

- Quadratic key setups and message exchanges per round
- Scheduling who talks when
- One traitor can anonymously sabotage
- Improvements subject of ongoing research

## Mixing/shuffling

- Computer analogue of shaking a ballot box, etc.
- Reorder encrypted messages by a random permutation
- Building block in larger protocols
- Distributed and verifiable variants possible as well

## Anonymous remailers

- Anonymizing intermediaries for email
  - First cuts had single points of failure
- Mix and forward messages after receiving a sufficiently-large batch
- Chain together mixes with multiple layers of encryption
- Fancy systems didn't get critical mass of users

## Outline

Anonymous communications techniques

Announcements intermission

Tor basics

Tor experiences and challenges

## Note to early readers

- This is the section of the slides most likely to change in the final version
- If class has already happened, make sure you have the latest slides for announcements

## Outline

Anonymous communications techniques

Announcements intermission

Tor basics

Tor experiences and challenges

## Tor: an overlay network

- Tor (originally from "the onion router")
  - `https://www.torproject.org/`
- An anonymous network built on top of the non-anonymous Internet
- Designed to support a wide variety of anonymity use cases

## Low-latency TCP applications

- Tor works by proxying TCP streams
  - (And DNS lookups)
- Focuses on achieving interactive latency
  - WWW, but potentially also chat, SSH, etc.
  - Anonymity tradeoffs compared to remailers

## Tor Onion routing

- Stream from sender to $\mathrm{D}$ forwarded via $\mathrm{A}$, $\mathrm{B}$, and $\mathrm{C}$
  - One Tor circuit made of four TCP hops
- Encrypt packets (512-byte "cells") as $\mathrm{E}_{\mathrm{A}}(\mathrm{B}, \mathrm{E}_{\mathrm{B}}(\mathrm{C}, \mathrm{E}_{\mathrm{C}}(\mathrm{D}, \mathrm{P})))$
- TLS-like hybrid encryption with "telescoping" path setup

## Client perspective

- Install Tor client running in background
- Configure browser to use Tor as proxy
  - Or complete Tor+Proxy+Browser bundle
- Browse web as normal, but a lot slower
  - Also, sometimes `google.com` is in Swedish

## Entry/guard relays

- "Entry node": first relay on path
- Entry knows the client's identity, so particularly sensitive
  - Many attacks possible if one adversary controls entry and exit
- Choose a small random set of "guards" as only entries to use
  - Rotate slowly or if necessary
- For repeat users, better than random each time

## Exit relays

- Forwards traffic to/from non-Tor destination
- Focal point for anti-abuse policies
  - E.g., no exits will forward for port 25 (email sending)
- Can see plaintext traffic, so danger of sniffing, MITM, etc.

## Centralized directory

- How to find relays in the first place?
- Straightforward current approach: central directory servers
- Relay information includes bandwidth, exit polices, public keys, etc.
- Replicated, but potential bottleneck for scalability and blocking

## Outline

Anonymous communications techniques

Announcements intermission

Tor basics

Tor experiences and challenges

## Anonymity loves company

- Diverse user pool needed for anonymity to be meaningful
  - Hypothetical Department of Defense Anonymity Network
- Tor aims to be helpful to a broad range of (sympathetic sounding) potential users

## Who (arguably) needs Tor?

- Consumers concerned about web tracking
- Businesses doing research on the competition
- Citizens of countries with Internet censorship
- Reporters protecting their sources
- Law enforcement investigating targets

## Tor and the US government

- Onion routing research started with the US Navy
- Academic research still supported by NSF
- Anti-censorship work supported by the State Department
  - Same branch as Voice of America
- But also targeted by the NSA
  - Per Snowden, so far only limited success

## Volunteer relays

- Tor relays are run basically by volunteers
  - Most are idealistic
  - A few have been less-ethical researchers, or GCHQ
- Never enough, or enough bandwidth
- P2P-style mandatory participation?
  - Unworkable/undesirable
- Various other kinds of incentives explored

## Performance

- Increased latency from long paths
- Bandwidth limited by relays
- Currently 1-2 sec for 50KB, 5-10 sec for 1MB
- Historically worse for many periods
  - Flooding (guessed botnet) fall 2013

## Anti-censorship

- As a web proxy, Tor is useful for getting around blocking
- Unless Tor itself is blocked, as it often is
- *Bridges* are special less-public entry points
- Also, protocol obfuscation arms race (currently behind)

## Hidden services

- Tor can be used by servers as well as clients
- Identified by cryptographic key, use special rendezvous protocol
- Servers often present easier attack surface

## Undesirable users

- P2P filesharing
  - Discouraged by Tor developers, to little effect
- Terrorists
  - At least the NSA thinks so
- Illicit e-commerce
  - "Silk Road" and its successors

## Intersection attacks

- Suppose you use Tor to update a pseudonymous blog, reveal you live in Minneapolis
- Comcast can tell who in the city was sending to Tor at the moment you post an entry
  - Anonymity set of 1000 → reasonable protection
- But if you keep posting, adversary can keep narrowing down the set

## Exit sniffing

- Easy mistake to make: log in to an HTTP web site over Tor
- A malicious exit node could now steal your password
- Another reason to always use HTTPS for logins

## Browser bundle JS attack

- Tor's Browser Bundle disables many features try to stop tracking
- But, JavaScript defaults to on
  - Usability for non-expert users
  - Fingerprinting via NoScript settings
- Was incompatible with Firefox auto-updating
- Many Tor users de-anonymized in August 2013 by JS vulnerability patched in June

## Traffic confirmation attacks

- If the same entity controls both guard and exit on a circuit, many attacks can link the two connections
  - "Traffic confirmation attack"
  - Can't directly compare payload data, since it is encrypted
- Standard approach: insert and observe delays
- Protocol bug until last year: covert channel in hidden service lookup

## Hidden service traffic conf.

- Bug allowed signal to guard when user looked up a hidden service
  - Non-statistical traffic confirmation
- For 5 months in 2014, 115 guard nodes (about 6%) participated in this attack
  - Apparently researchers at CMU's SEI/CERT
- Beyond "research," they also gave/sold info. to the FBI
  - Apparently used in Silk Road 2.0 prosecution, etc.

## Next time

- How usability affects security