

Anonymity Theory & Practice

Computer Security

Tom Chothia

Today's Lecture

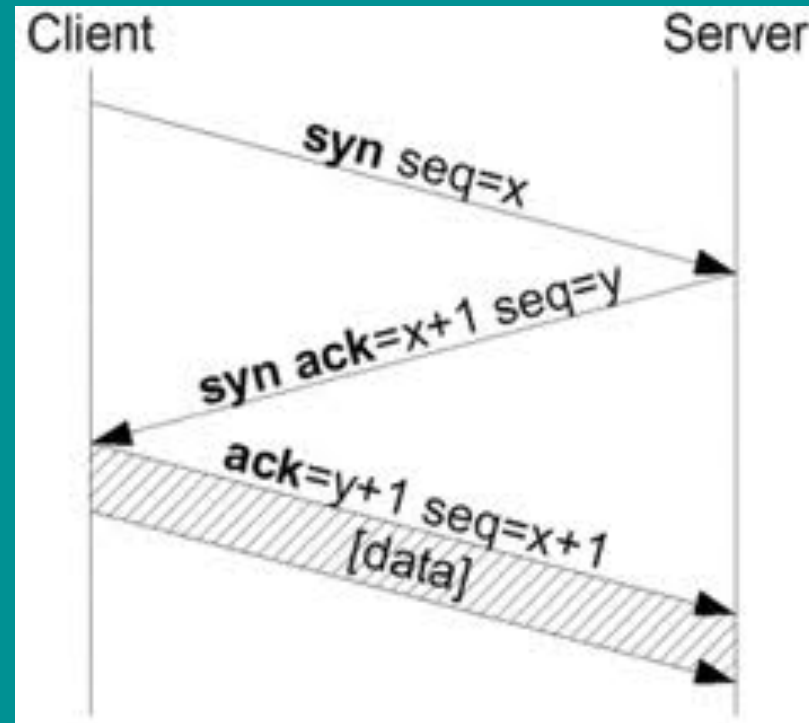
- Definitions of anonymity
- Dining Cryptographers Protocol,
- The Crowds Protocol,
- Mixes,
- Onion routing.
- Tor

IP address

- When you connect to another computer you send it your IP address.
- It is very hard to communicate without revealing an address on which you can receive traffic.
- Recent court cases have decided that your IP address can be used to identify you in court.

The TCP/UDP protocol

- The TCP protocol
 - Requires 2 way communication.
- Uses sequence numbers.



The UDP protocol

- UDP doesn't need 2 way connections.
- But it is often used for protocols that require 2 way communication
- ISPs usually block users from sending packets not from their IP.

Who could fake your IP?

- If you have 2 way communication:
 - Anyone on the route between the hosts:
 - Sys admins at both ends
 - Your ISP
 - Anyone with access to the routers
 - Anyone on your local wi-fi or Ethernet
 - A hacker that has broken into any of the above.

Most Common Assumption:

- *Your IP always explicitly identifies you.*
- E.g. file sharers are sued by firms that
 - Find their IP and a time.
 - Go to the ISP and ask for the address of the user who had that IP at that time.

More Threats to Anonymity

- Your ISP logs all your web activities.
- Cookies in your browser.
 - Ad. companies profile your web surfing.
- Google knows every search you make.

“You have zero privacy
anyway, get over it”

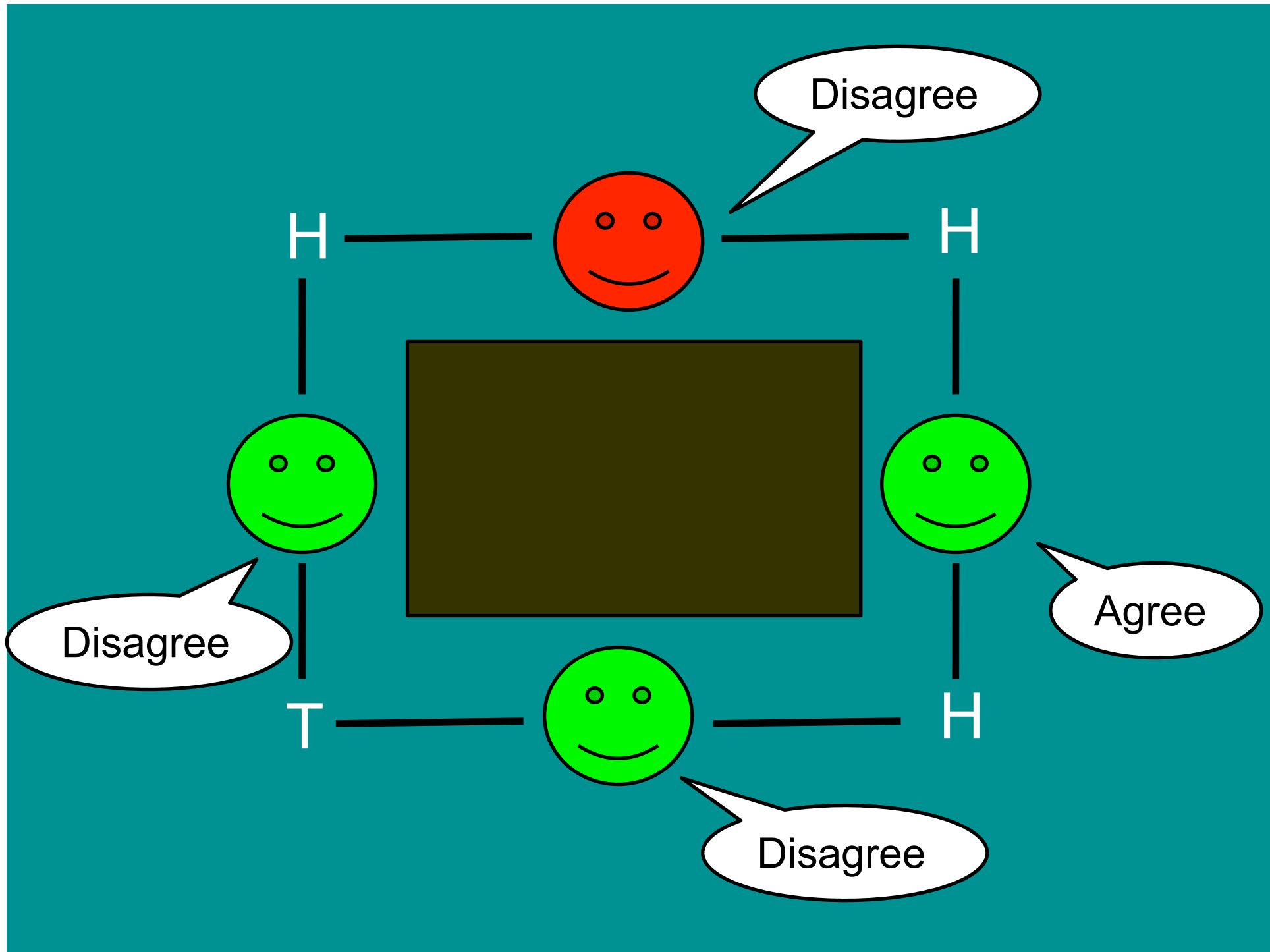
Scott McNealy,
CEO of SUN Microsystems.

“With your permission, you give us more information about you, about your friends, and we can improve the quality of our searches. We don't need you to type at all. We know where you are. We know where you've been. We can more or less know what you're thinking about.”

Eric Schmidt
CEO of Google

Dining Cryptographers

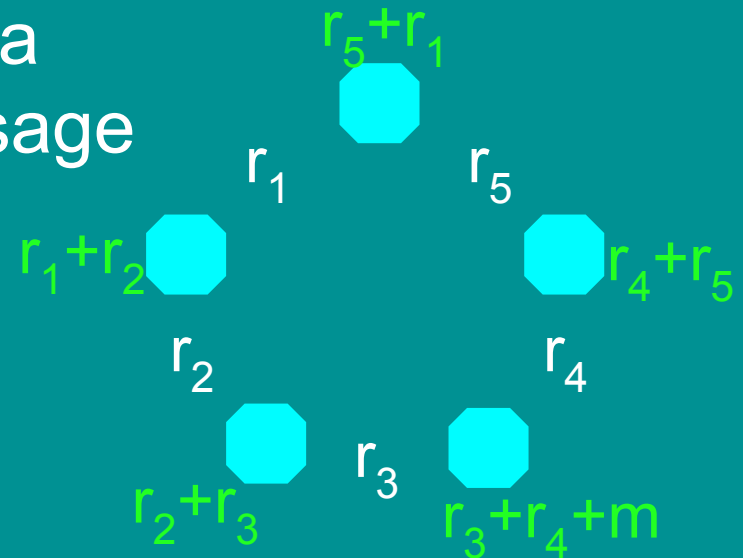
- 4 Cryptographers are sitting round a table, when the waiter tells them someone has paid the bill.
- Can they find out if it was one of them, or someone else, while respecting the payer's anonymity?



Dining Cryptographers

- Nodes form a ring
- Each adjacent pair picks a random number
- Each node broadcasts the sum (xor) of the adjacent numbers
- The user who wants to send a message also adds the message
- The total sum (xor) is:

$$\begin{aligned} & r_1 + r_2 + r_2 + r_3 + r_3 + \\ & r_4 + r_4 + r_5 + r_5 + r_1 + m = m \end{aligned}$$



Dinning Cryptographers

- It's impossible to tell who added m .
- Beyond suspicion even to a global attacker.
- Very inefficient: everyone must send the same amount of data as the real sender.

Some Kinds of Anonymity

- Sender anonymity.
- Receiver anonymity.
- Sender-receiver unlinkability.

From the

- Sender
- Receiver
- Another participant in the protocol.
- Outside observer that can see all/some of the network

Levels of Anonymity

Reiter and Rubin provide the classification:

- ***Beyond suspicion***: the user appears no more likely to have acted than any other.
- ***Probable innocence***: the user appears no more likely to have acted than to not to have.
- ***Possible innocence***: there is a nontrivial probability that it was not the user.

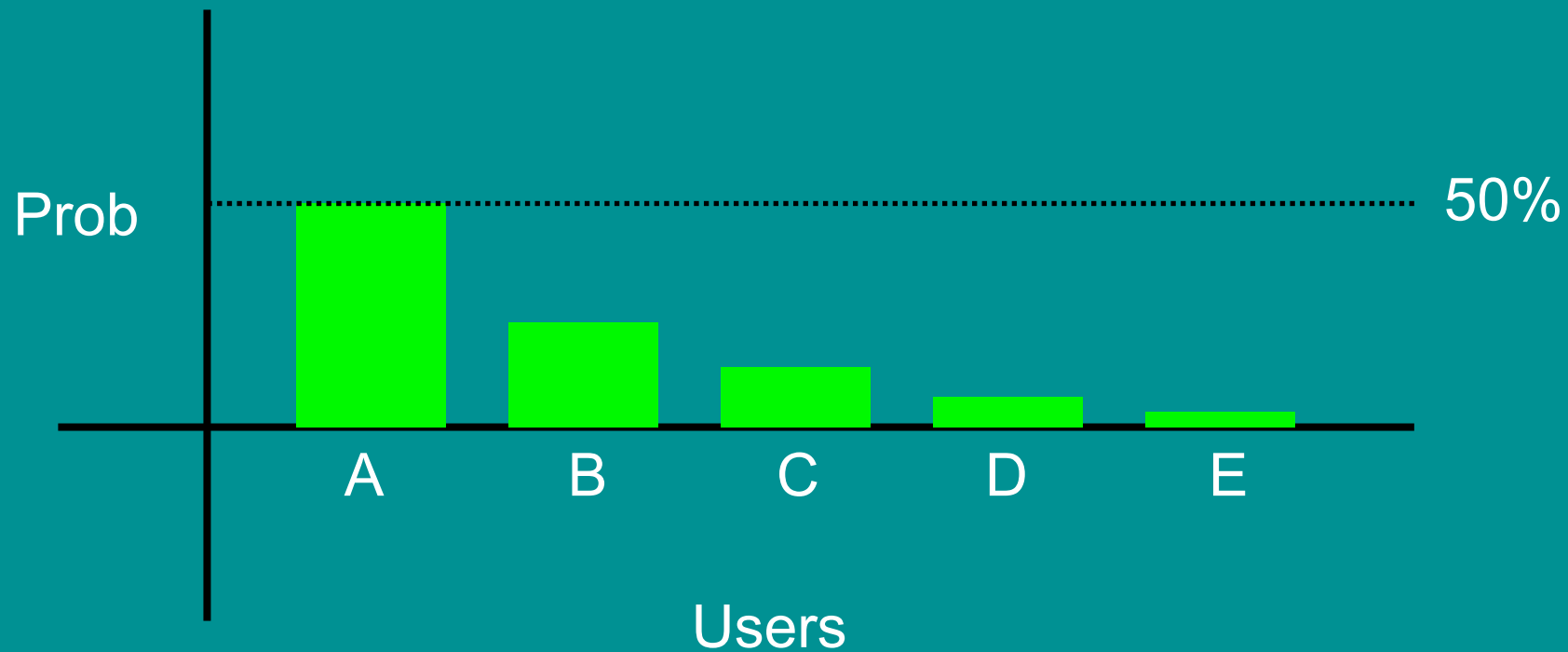
Beyond suspicion

- All users are Beyond suspicion:



Probable Innocence

- All users are Probably Innocence



Possible Innocence

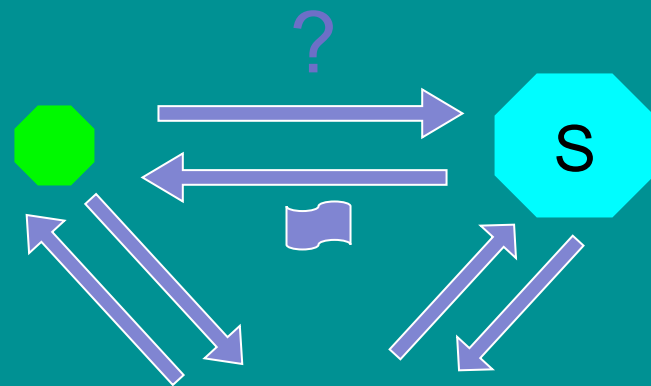
- There a small change it is not you.



Definitions

- These definitions do not take into account how likely each principal is to be guilty to start off with.
- Or quantify what the attacker learns from a run of the protocol.
- This is currently a hot research topic, so far there are lots of complicated definitions but no clear winner.

Proxy Example: The Anonymizer



The Anonymizer

An Internet connection reveals your IP number.

The Anonymizer promises “Anonymity”

Connection made via The Anonymizer.

The server see only the Anonymizer.

Proxy Example: The Anonymizer

The sender is *Beyond Suspicion* to the server.

The server knows The Anonymizer is being used.

If there is enough other traffic, you are *Probably Innocent* to a global observer.

The global observer knows you are using the “The Anonymizer”

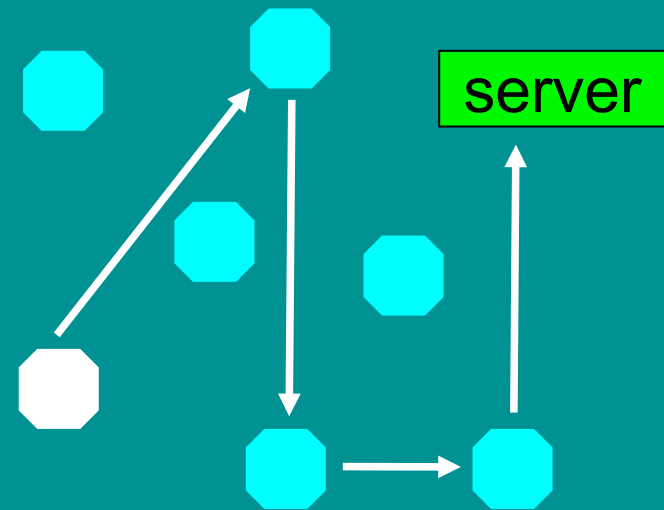
There is no anonymity to “The Anonymizer”

Proxy Example: The Anonymizer

- From the small print:
- ... we disclose personal information only in the good faith belief that we are required to do so by law, or that doing so is *reasonably necessary* ...
- ... Note to European Customers: The information you provide us *will* be transferred outside the European Economic Area

Crowds

- A crowd is a group of n nodes
- The *initiator* selects randomly a node (called *forwarder*) and forwards the request to it
- A forwarder:
 - With prob. $1-p_f$ selects randomly a new node and forwards the request to him
 - With prob. p_f sends the request to the server

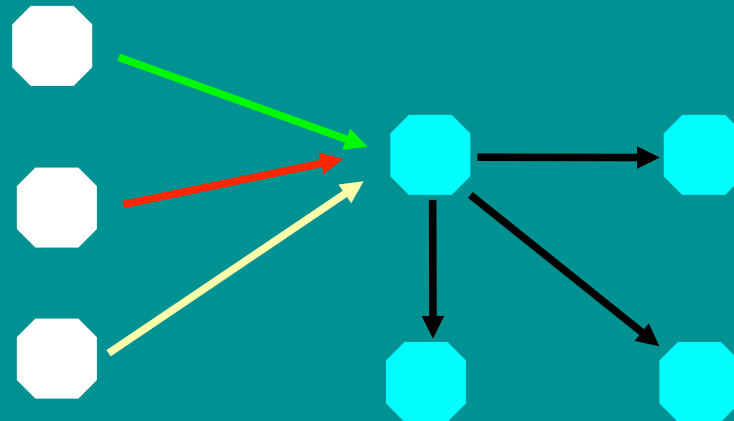


Crowds

- The sender is beyond suspicion to the server.
- Some of the nodes could be corrupted.
- The initiator could forward the message to a corrupted node.
- The sender has probable innocence to other nodes.

MIXes

- MIXes are proxies that forward messages between them
- A user contacts a MIX to send a message
- The MIX waits until it has received a number of messages, then forwards them in different order



MIXes

- It is difficult to trace the route of each message.
- Provides beyond suspicion S-R unlinkability even to a global attacker.
- Messages have to be delayed (can be solved with dummy traffic).
- More complicated when sending series of packets

Some types of Mix

- Time delay mixes
 - Fires after a set time
- Threshold mixes
 - Fires when a set number of messages have arrived.
- Pool mixes
 - Like Threshold, but some messages are randomly held back to the next round.

Onion Routing

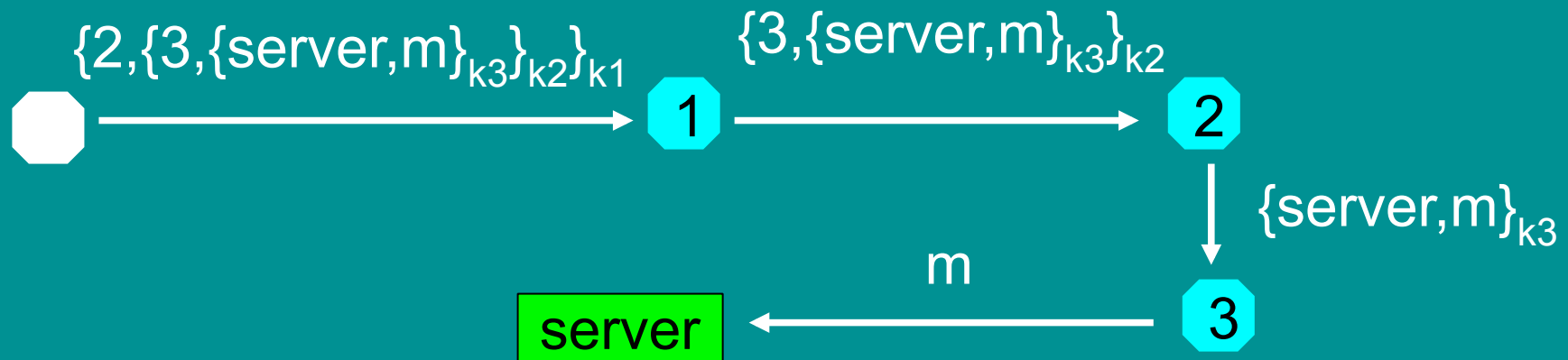
- Onion Routing ensures that your message really is routed via the proxies you want.
- Presented in the research paper:

"Tor: The Second-Generation Onion Router"

By Roger Dingledine, Nick Mathewson,
Paul Syverson

Onion Routing

- Each node publishes a public key.
- The initiator selects the whole route and encrypts the message with all keys in reverse order.
- Each node unwraps a layer and forwards the message to the next one.



Onion Routing

- Each node only learns the IP of the node before it and the node after it.
- End-users can run their own node
 - Better anonymity
- or use an existing one
 - Easier to use
 - User's identity is revealed to the node

Tor

- Tor implements this protocol.
- Several hundred volunteer nodes.
- Firefox plug-in.
- Big supporters include:
 - Electronic Frontier Foundation
 - US Naval Research Laboratory

Tor

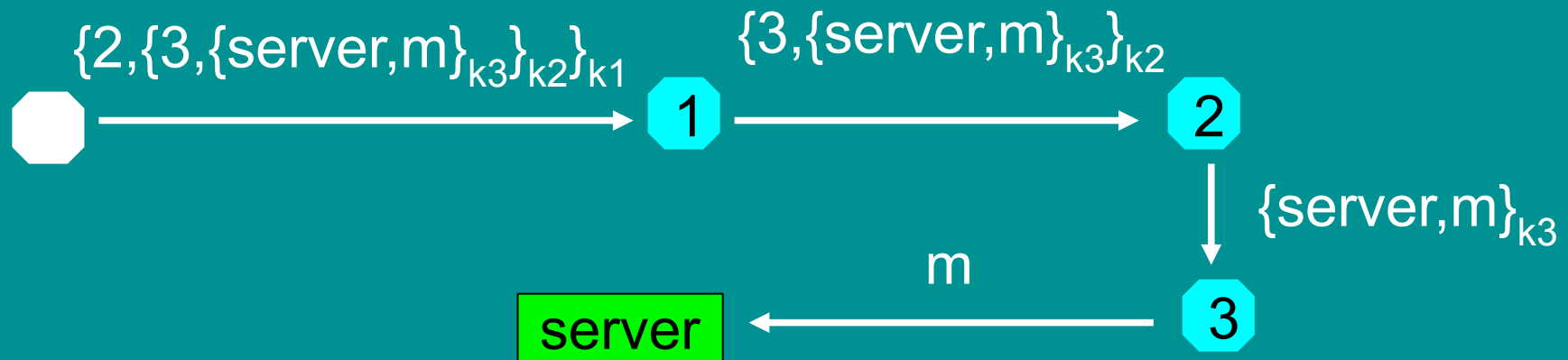
- When you start the Vidalia proxy:
 - It picks routes via a number of proxies,
 - Runs a SOCKS proxy on port 9050,
 - Data sent across the SOCKS proxy is
 - Encoded using onion routing
 - Then sent out across the network
 - The end node listens for replies and sends them back along the path.

Problems with Tor

- Tor provides a very high degree of anonymity.
- No anonymity from an attacker that monitors the whole network.
- Some protocols broadcast their IP address
- If all the nodes on the path work together they can break your anonymity.

Problems with Tor

- You reveal your IP to the first node.
- The final node knows the server and the message



Correlation attacks.

- The data coming into the network is encrypted,
- But the size and timing of the packets depends on the webpage,
- Murdoch & Danezis showed that if you run the first and last node on a path you could statistically match the streams. See paper:

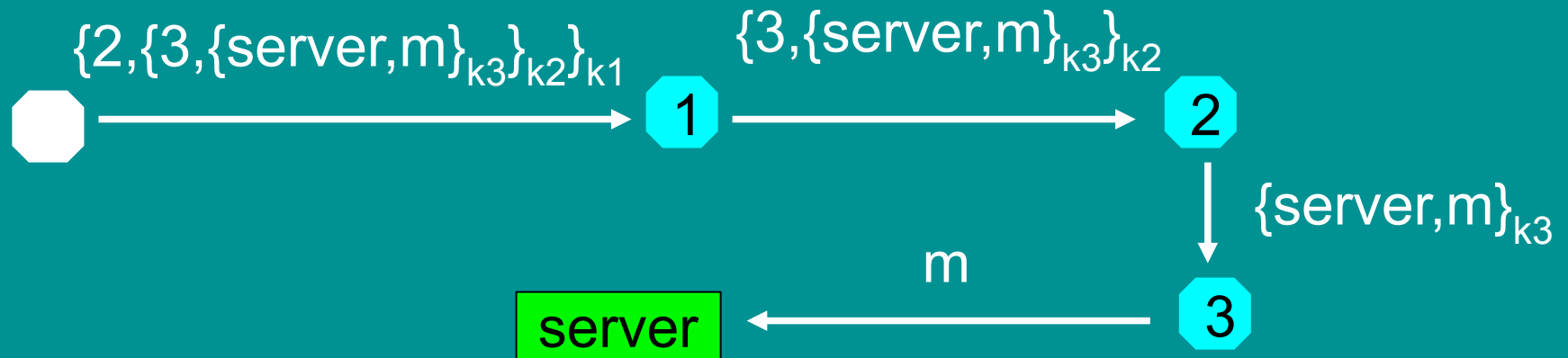
Low-Cost Traffic Analysis of Tor

Poorly Configured Browsers Aren't Anonymous with Tor

- Browser could give out cookies that identify you.
- Tor is TCP only.
 - server could use Flash to make the client connect to the server using UDP.
 - server then learns users real IP address.
- Need to turn off cookies and flash, JavaScript etc.

Question

- Does Tor keep your data secure?



Groups that did not realise this:

- Embassies of Japan, Russia, Kazakhstan, Uzbekistan, Tajikistan, India, Iran, Mongolia.,.
- UK Visa Application Centre in Nepal
- The office of the Dalai Lama
- Several Hong Kong Human Rights Groups
- Over 1,000 businesses ...

Monitoring Tor's Output

- In August 2007 Dan Egerstad (Security Researcher) ran five Tor nodes.
- He monitored all unencrypted traffic
- "It's [mostly] just porn,... It's kind of sad."
- But also unencrypted e-mail traffic, including user names and passwords

Monitoring Tor's Output

- He then tried to contact the embassies involved via the e-mail address.
- Got no or little response
- Finally he posted the login names and passwords on his website.

Some BAD passwords:

- Iranian embassies used their host country or cities name as their password.
- Hong Kong Liberal Party used "123456" and "12345678".
- An Indian embassy used "1234".
- An India Ministry of Defence account used "password+1"
- The Mongolian embassy in the U.S. used "temp"

Reactions

- India, Iran and Uzbekistan were "friendly" and acted quickly to fix the problem,
- China filed a criminal complaint over the posting,
- U.S. authorities had his website taken down,
- and the Swedish authorities arrested him.

The real story:

Further research found Tor nodes that:

- only accept unencrypted traffic for : DNS, POP3, IMAP, MSN Messenger, VNC and IRC.
- only accept HTTP packets bound for MySpace and Google (N.B. not HTTPS).
- A Tor Node that replaces any SSL certificate with a self signed certificate.

Conclusion

- Lots of different kinds and levels of anonymity.
- Key anonymity technologies include:
 - Proxy
 - Onion routing (e.g. Tor)
 - Mix networks