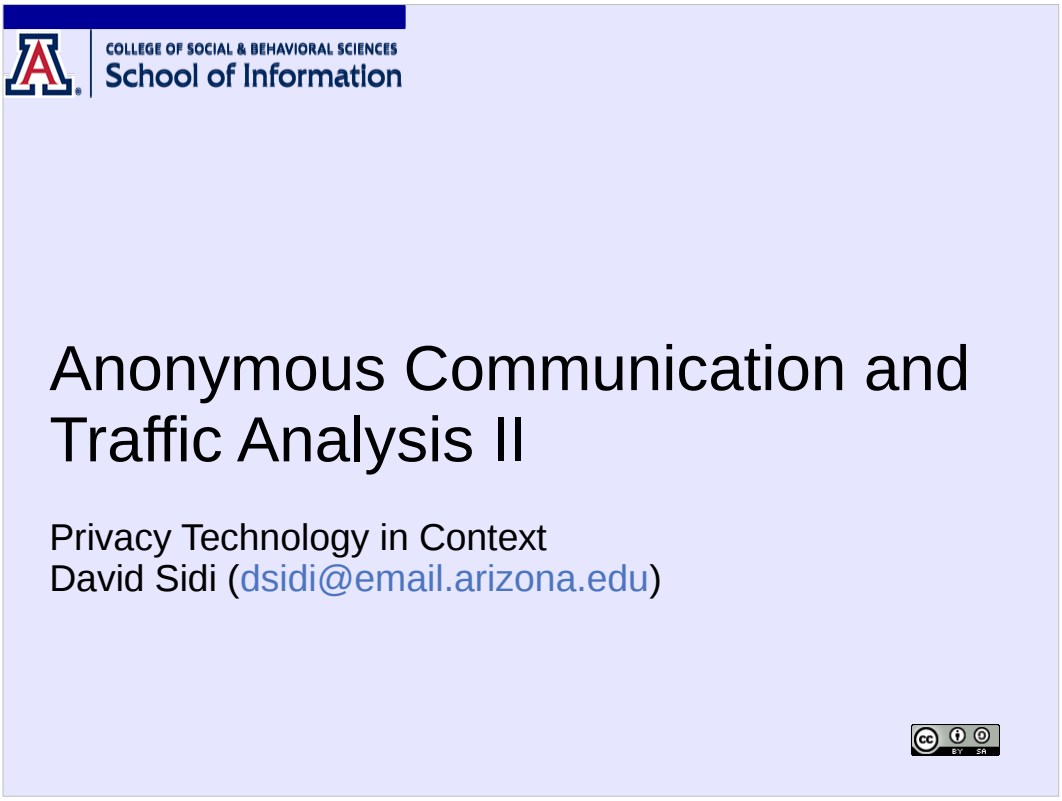
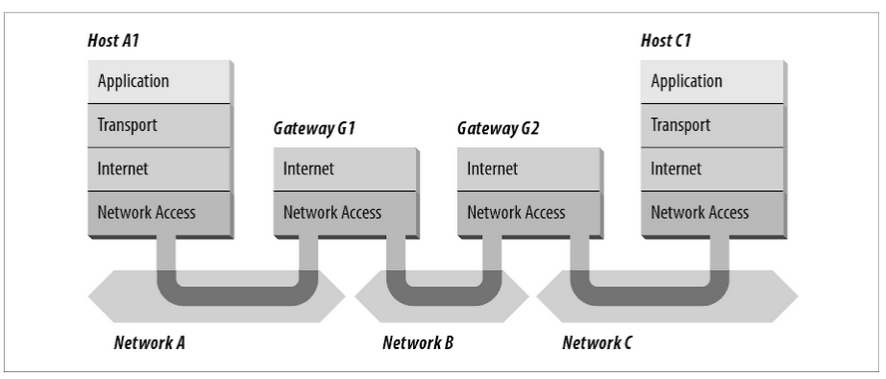


we're going to finish up some background before moving to anonymity, remaining fairly high level. Next time we will discuss further details.

Small mention of interesting things

- Grading for assignment 4
- MIT student work on freehaven led to Tor

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information



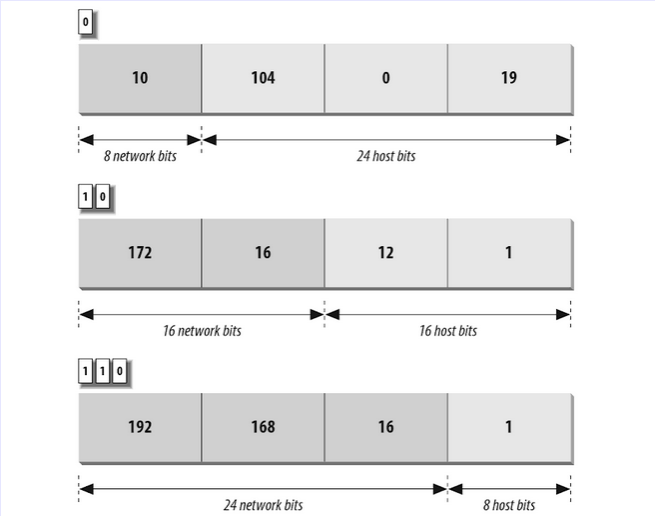
A bunch of physically heterogeneous networks that use TCP/IP form an internet. The Internet is, basically but not exactly, all the TCP/IP internets.

IP is central to the internet.

Hosts go through the whole protocol stack; gateways just go up to IP

Thus packets providing routing info get data from one network to another, then to a host on that network

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information



4

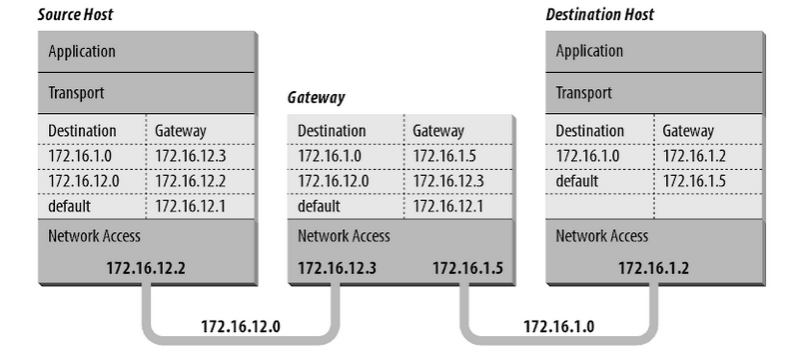
IP defines an addressing scheme, and uses it in IP packets (called datagrams) to move data hop-by-hop to the network that has the destination host on it.

This diagram is a little like what I drew on the board last time, showing the IP addressing scheme.

There is a network part and a host part to the address. The network part takes up less of the address for larger networks, and less of the address for smaller networks. The size of the network part is indicated using subnet masks (see: CIDR).

(In practice, big blocks of contiguous addresses are given to service providers who are best placed to provide routing information. Example: UA has 150.135.0.0; these are then delegated to groups in the university (the iSchool has 150.135.15.0, I think). The same is true of your ISP--- everyone is assigned an IP by their ISP in order to connect to the internet.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

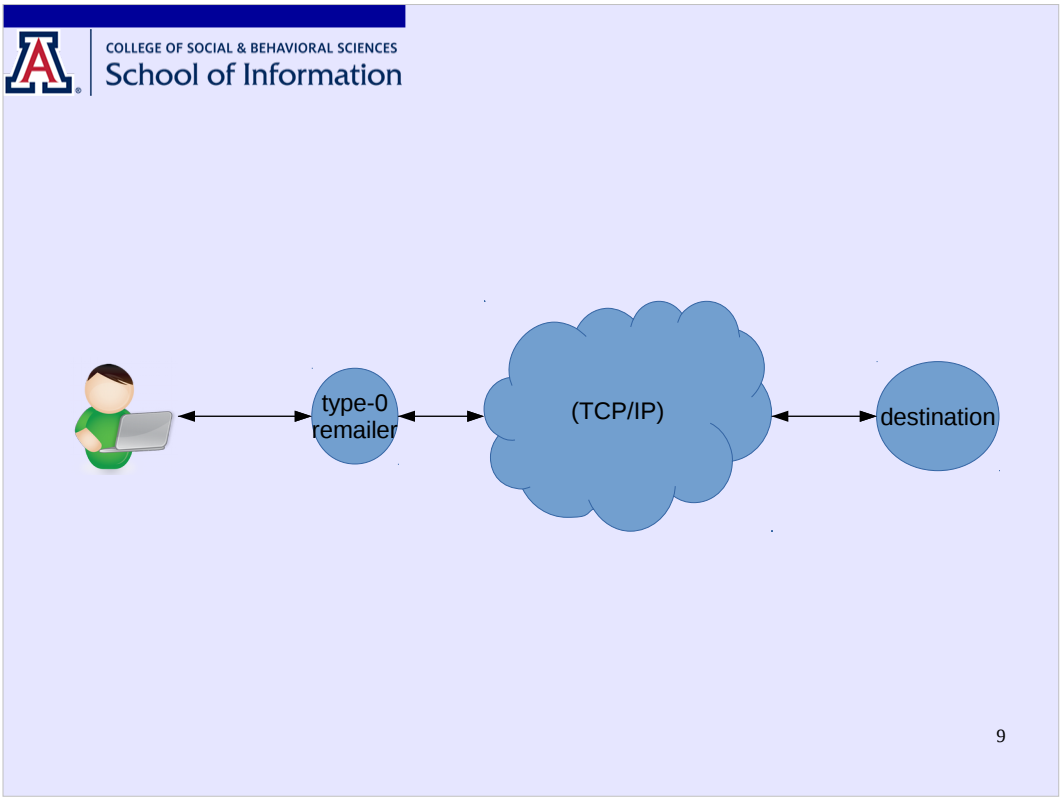


How does the IP addressing scheme support routing?
Essentially, with (deferred) table lookups

here 172.16.12.2 sends to 172.16.1.2 (walk through).

notice the source and gateway determine the next hop with a table lookup. They also have a default route for addresses they can’t look up.

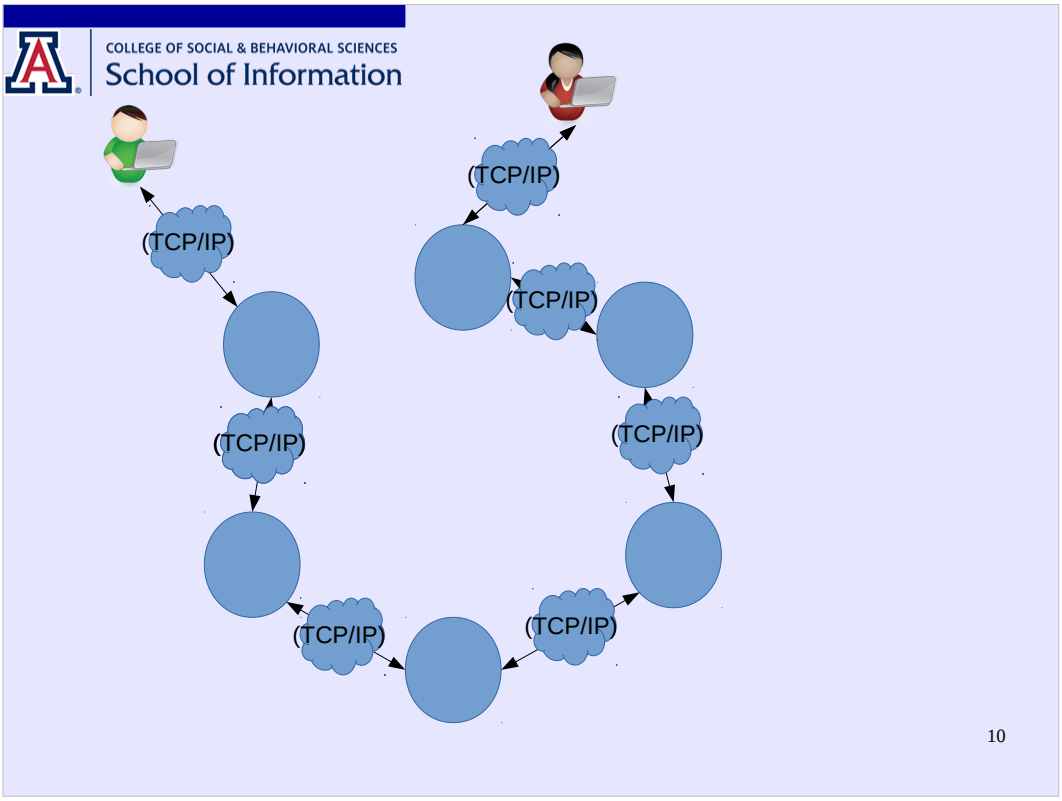
routing information isn’t hidden in IP: its only used bit-by-bit, but its all available in principle. But censorship systems often work by ensuring they are on the route, and filtering. So what to do?



One simple answer is to proxy: encrypted traffic is sent to an intermediate node outside of which the censor is not on the route. The censor just sees your connection to some box (ideally, it doesn't know its a proxy), and the proxy handles forwarding your traffic for you.

A variant of this idea is this simplest (degenerate) form of mixnet, which forwards mail.

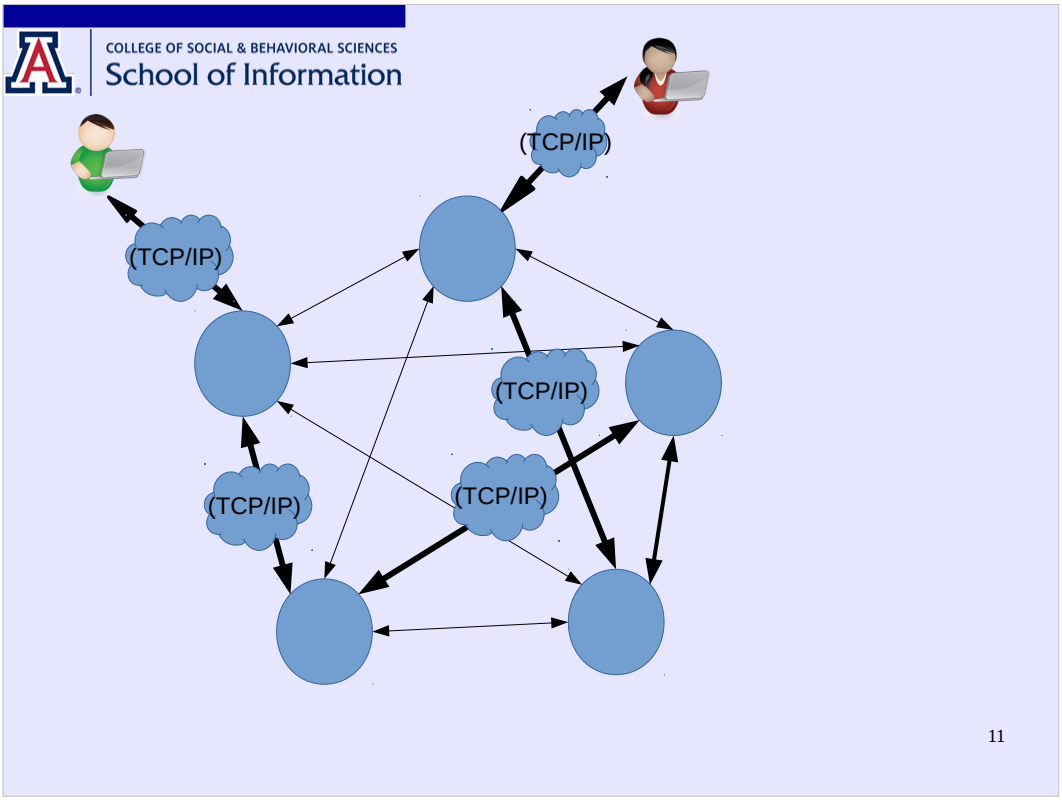
Problems here?



A more sophisticated approach combines several intermediate nodes to ameliorate the problem of centralized trust in one intermediate. Even if some of these intermediates are malicious, the traffic can continue.

Here's the classic topology for mixnets: a cascade. These have a fixed route through all the mixes

Problems here? (Suddenly lots of traffic going to an intermediate node---might it attract attention of censors? Are you hidden as a user of such a system?)



free route topology

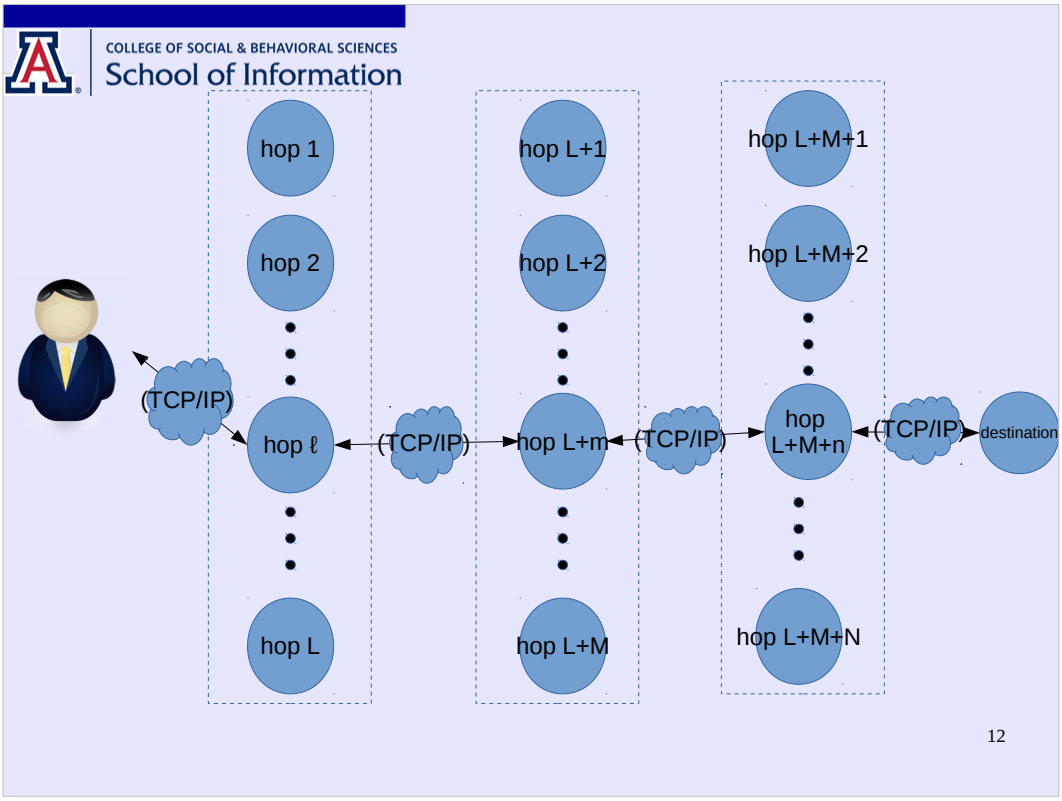
it is well known that this topology is the most heavy metal (see pentagram). (If you are reading this and are confused: it's a joke).

Old-school onion routing is free-route like this.

more scalable, so more anonymous in practice in some ways; but also less anonymous in some ways than mixes.

Problems? are the entry points to the network known?

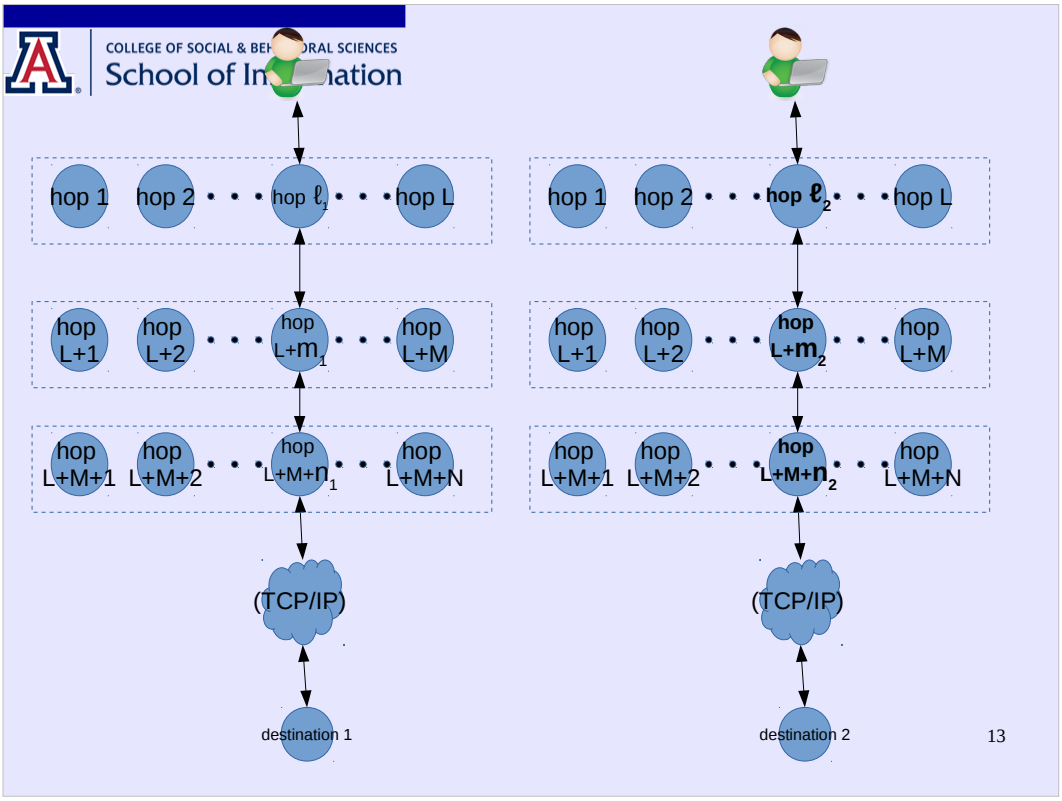
scalability and anonymity are tied together



stratified topology. The choice at each layer is random (see next slide)

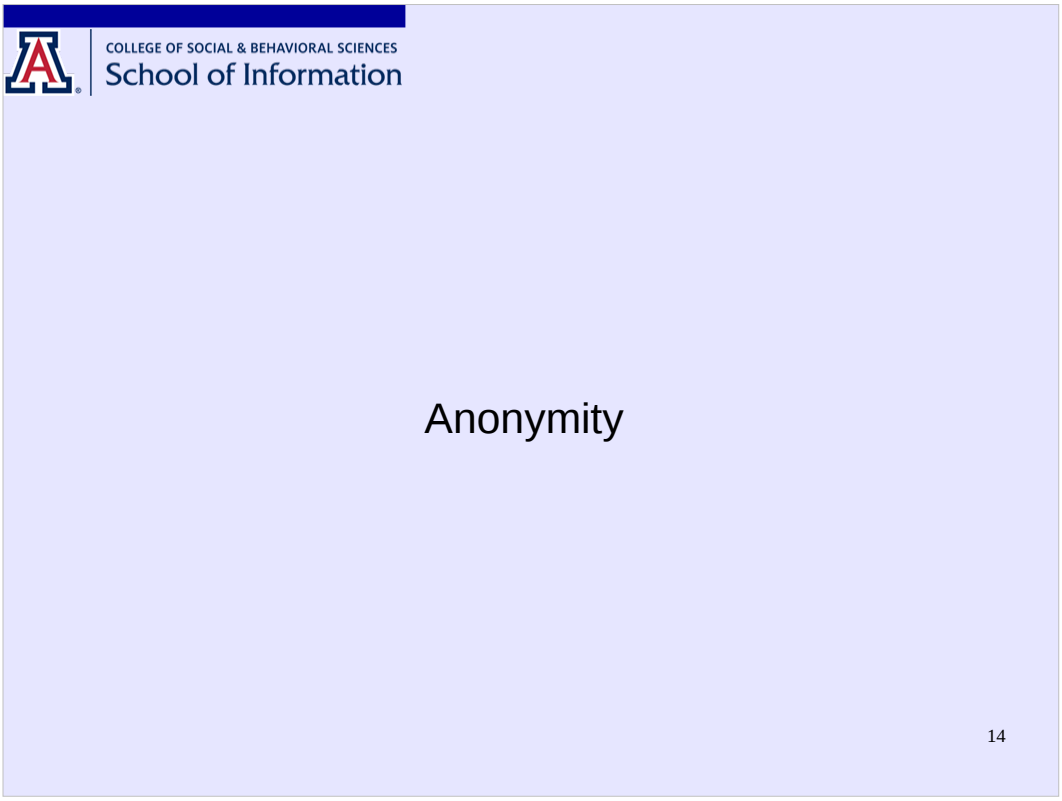
balances scalability with anonymity. Tor does this with helper nodes (“guards”), intermediate nodes (“relays”), and exit nodes.

To address public knowability of entry nodes, Tor uses bridges. An elaboration of this to avoid recognition of tor traffic by traffic analysis is pluggable transports.



random choice at each layer.

we have talked



so far I've relied on an intuitive understanding of anonymity, but when we're building technologies for anonymity, it's better to be precise about what is protected and what is not.

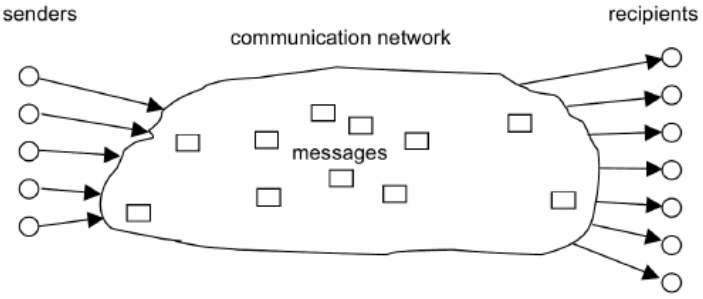
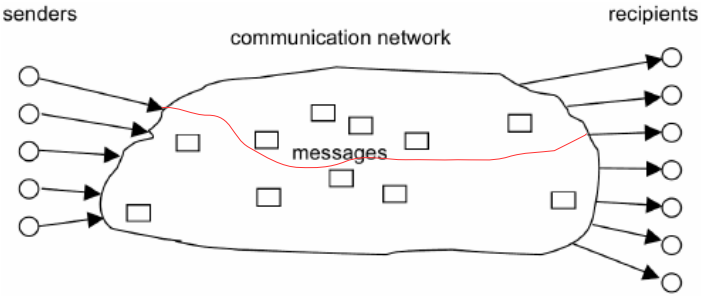


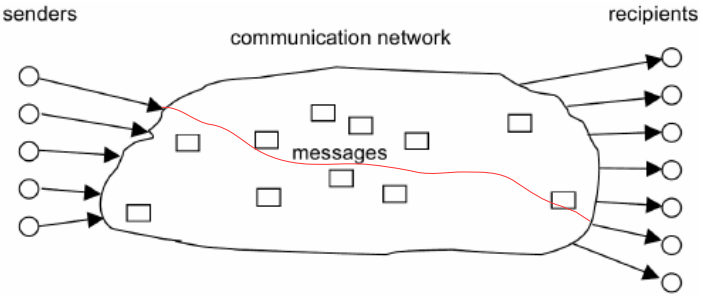
Image credit (before modification):
Christina Pöpper
Ruhr-University Bochum

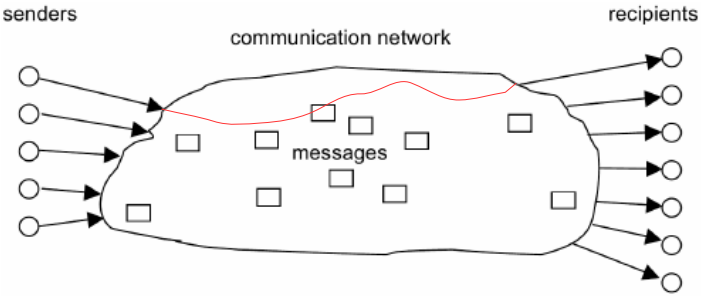
senders communicating messages over a channel with recipient.

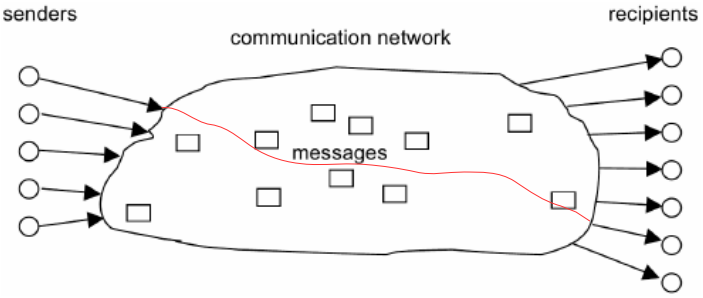
An adversary (or attacker) tries to reduce the anonymity of some or all of the parties to the communication

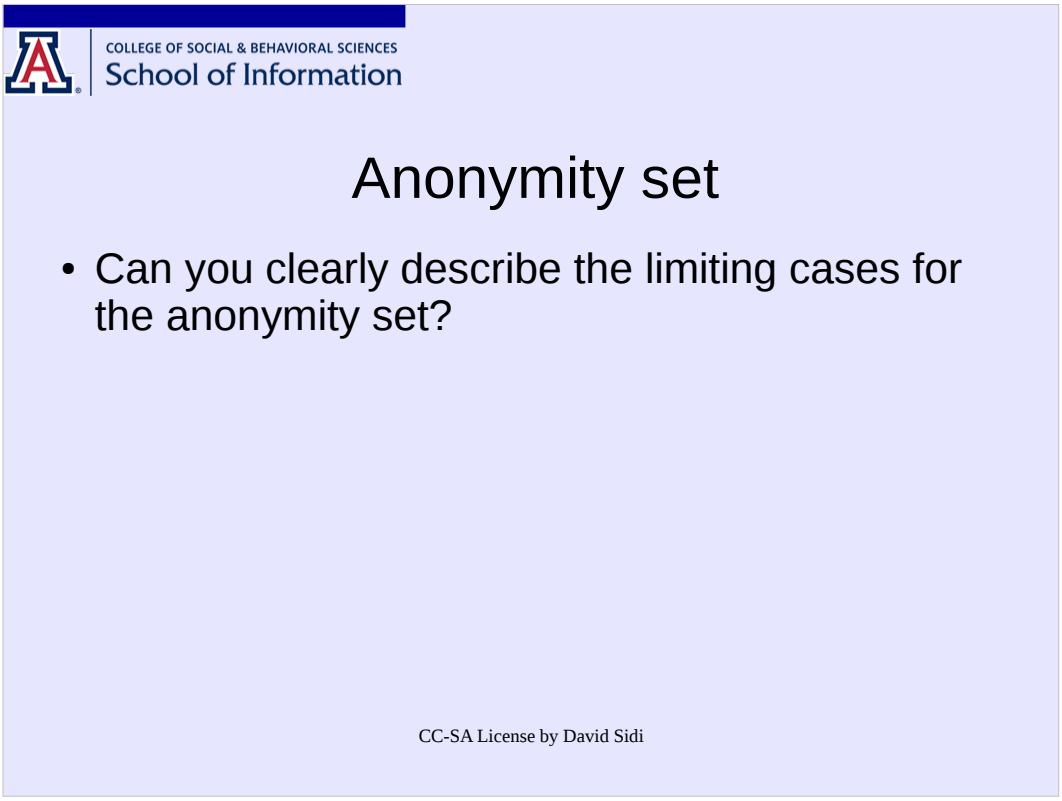
Anonymity is a property of a channel.


senders
communication network
recipients
messages


senders
communication network
recipients
messages


senders
communication network
recipients
messages

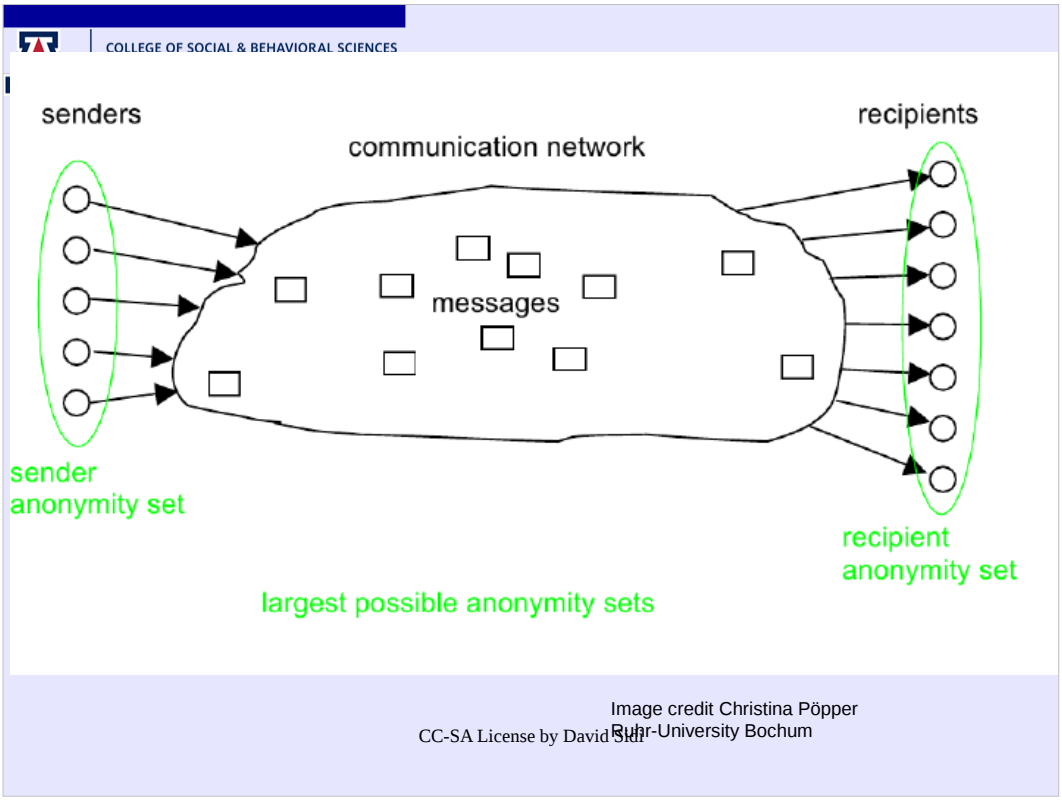




1. singleton. The subject has no anonymity; he is perfectly identifiable relative to the set of subjects.

2. universe. The subject has perfect anonymity in the set of users of the system.

Notice what this means (Berthold's metric)

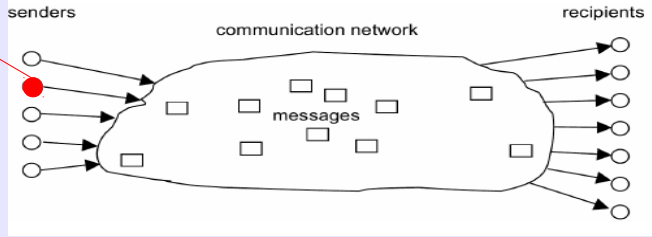


- What is a sender? i.e., how do we get the set of all senders? (Think about the definitions)
  - something that sends messages over the network to recipients (implements protocols, etc.). People, personal computers, cameras, phones, etc.?
- if that were all, all senders would be the same! But the anonymity set is intended to be useful, not trivial, in its separation of senders that cannot be distinguished from those that can be
  - senders should have attributes to distinguish them

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Senders and recipients are thought of as rows in a database table

	MAC	Browser_fingerprint	IP	Sites_visited
SNDER_1	00:a0:ef:eb:5v:ff	af7f098c39728f8cb676e3df82ced01a149ee3aa92af2b88c20c4948a5fad5fd		torproject.org, ischool.arizona.edu, maps.google.com...
SNDER_2	00:c0:ff:dd:ff:ef	a5fad5fdd01a149eeaf7f098c39728f8cb676e3df82ce3aa92af2b88c20c4948		nytimes.com, purple.com



- Behind the anonymity set in Torra's discussion (which is typical) is the database: things that can be identified ("subjects," "persons," etc...) are records in a database---descriptions of people via a set of attribute values

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

CC-SA License by David Sidi

- Suppose I include in a record of UA students a person's weight and height as 150 lbs, 5'3". Is the person anonymous? (think: anonymity set)

- Now suppose further that I do so for a database of male UA basketball players. Obviously, the player is not (as) anonymous.
- Where might you find combinations of attributes that identify people using computer networks?
  - IP, Ethernet, Tor user (Harvard bomb threat example), DNS, third-party tracking, browser fingeprinting
  - Lets look at DNS for a second

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

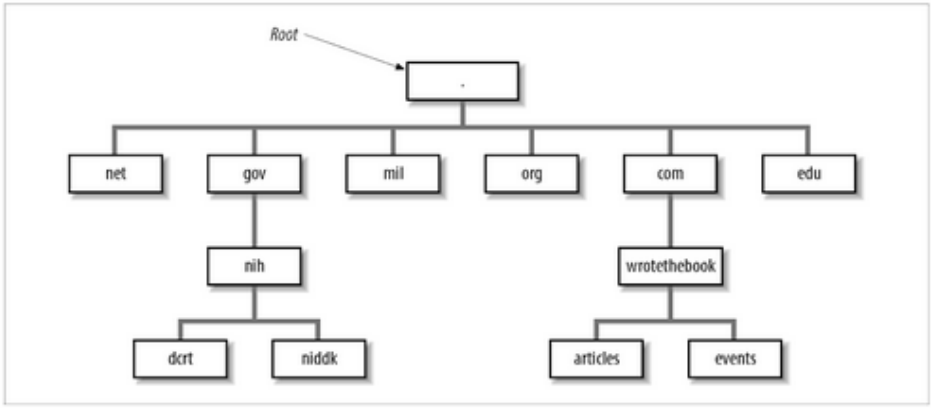


Figure 3-1. Domain hierarchy

- DNS puts a human-friendly layer on top of destination IP addresses, by associating structured strings of words (and abbreviations) with IPs.
- In the beginning was the host table. And it was good (for some things, but not for scalability)
- Distributed, hierarchical
  - again with the layers: root servers have information about TLD servers beneath them
  - Registering a domain name involves telling the TLD servers about your server
  - you can then do subdomains at will
- Forwarding DNS server
- Recursive DNS server (or resolver)
- (Root nameserver)
- (Top Level Domain nameserver)
- Authoritative nameserver

- quick task: get the DNS for arizona.edu using the server

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

DNS

```
$dig arizona.edu

; <<>> DiG 9.9.5-9+deb8u14-Debian <<>> arizona.edu
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 14058
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;arizona.edu.                   IN     A

;; ANSWER SECTION:
arizona.edu.            6813IN     A      128.196.128.233

;; Query time: 32 msec
;; SERVER: 208.67.222.222#53(208.67.222.222)
;; WHEN: Mon Oct 23 12:43:03 MST 2017
;; MSG SIZE  rcvd: 56
```

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

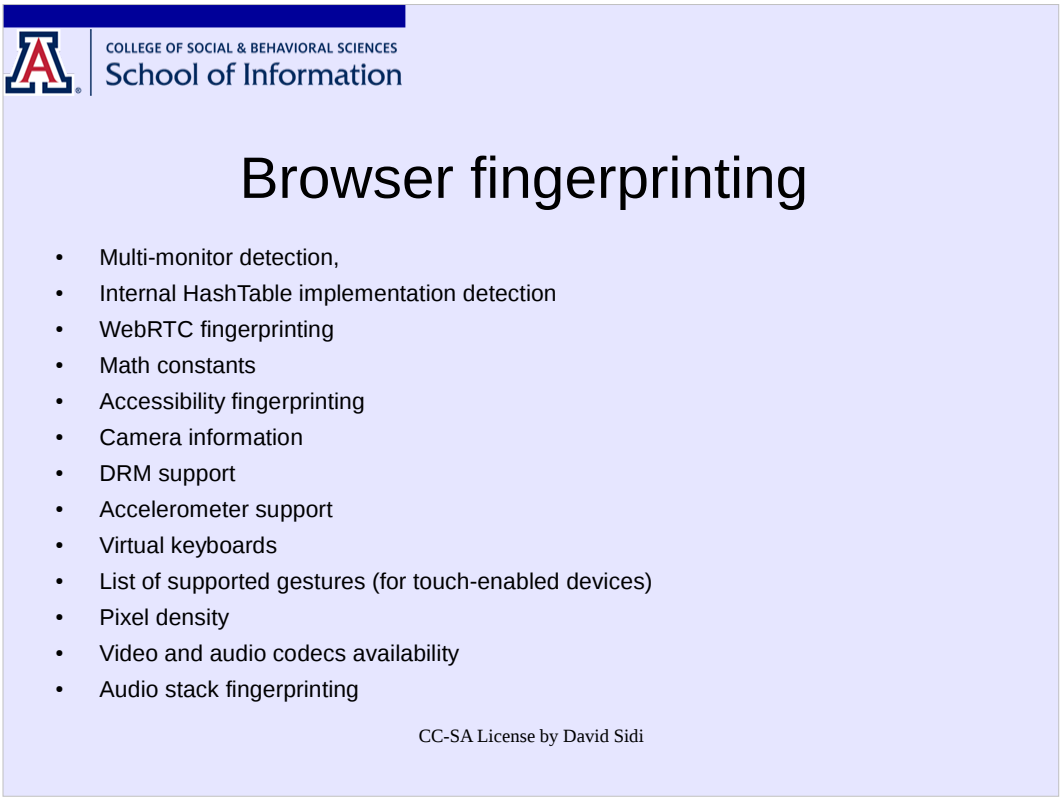
DNS

- Suppose I use a VPN to tunnel my traffic to a server I control, but you serve my DNS requests. What can you learn about me?

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Browser fingerprinting

- UserAgent
- Language
- Color Depth
- Screen Resolution
- Timezone
- Has session storage or not
- Has local storage or not
- Has indexed DB
- Has IE specific 'AddBehavior'
- Has open DB
- CPU class
- Platform
- DoNotTrack or not
- Full list of installed fonts (maintaining their order, which increases the entropy), implemented with Flash.
- A list of installed fonts, detected with JS/CSS (side-channel technique) - can detect up to 500 installed fonts without flash
- Canvas fingerprinting
- WebGL fingerprintingPlugins (IE included)
- Is AdBlock installed or not
- Has the user tampered with its languages 1
- Has the user tampered with its screen resolution 1
- Has the user tampered with its OS 1
- Has the user tampered with its browser 1
- Touch screen detection and capabilities
- Pixel Ratio
- System's total number of logical processors available to the user agent.



note: this has some uncertainty associated with it.
comparison on panopticlick.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Discussion: “Identifiability”

- Why might it be too simple to say that for a sender S, every other potentially-different sender is either completely indistinguishable from S or not?
- 2 minutes alone, 2 minutes with a partner, then we’ll talk as a class

Question

- What is problematic about this definition of anonymity? “Anonymity is thus defined as the state of being not identifiable within a set of subjects, the anonymity set.” (Danezis and Diaz 3)

- Later, they say “A subject carries on the transaction anonymously if he cannot be distinguished **(by an adversary)** from other subjects. **This definition of anonymity captures the probabilistic information often obtained by adversaries trying to identify anonymous subjects**.

Question

- What is problematic about this definition of anonymity? “Anonymity is thus defined as the state of being not identifiable within a set of subjects, the anonymity set.” (Danezis and Diaz 3)
- We need to say something about an adversary and an attack model

- Later, they say “A subject carries on the transaction anonymously if he cannot be distinguished **(by an adversary)** from other subjects. **This definition of anonymity captures the probabilistic information often obtained by adversaries trying to identify anonymous subjects**.

- **Definition 1.2** From an adversary's perspective, anonymity of a subject s means that the adversary cannot achieve a certain *level of identification* for the subject s within the anonymity set. (Torra)

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

- Torra's terminology is confusing. His own idea of `n-Confusion' is in the background here :-).
- Simplifying, the point is: for each row in the database and some auxiliary information, we have a distribution over the subjects.
- The closer to uniform this distribution is, the "less identifiable" an entity is within the anonymity set, and the better the anonymization

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity metrics

- Reviewing the discussion from a recent breakout session in a workshop on Privacy as Engineering Practice, Deirdre Mulligan said that there is a need for more formal measures of privacy (including anonymity)
- privacy loss in terms of information flow analysis, measures that take into account inference and not only disclosure, etc.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity metrics

- Degree of anonymity is a distribution 1 - p, where p is the probability assigned to the senders in the anonymity set (Reiter and Rubin 1998)
  - I talk about senders just for convenience, this applies more generally
- Think “worst case” -- who’s got the highest probability of being identified, and how high is that probability?
- This doesn’t account for how evenly distributed the probability is over the anonymity set, in the sense that it just depends on the greatest probability

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity metrics

- Suppose a user u has 0.1 degree of anonymity. Consider two scenarios s1 and s2
  - s1: 2 users u and v, with v also 0.1 degree anonymity
  - s2: 1000 users, all users distinct from u with the same degree anonymity (which is less than 0.001)
- With degree of anonymity as measure, both have equal anonymity

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Degree of anonymity

- One way to account for evenness of distribution is with Information Theory
- “Effective size”: Roughly, how many bits does the attacker need to identify a member of the anonymity set?

Degree of anonymity

- "Effective size": Roughly, how many bits does the attacker need to identify a member of the anonymity set?
- Less roughly, use the Shannon entropy. For Ψ the set of users, $\mathcal{U}$ the posterior of a user being the sender given a message,

Definition 2. *We define the effective size* $\mathcal{S}$ *of an* r *anonymity probability distribution* $\mathcal{U}$ *to be equal to the entropy of the distribution. In other words*

$$\mathcal{S} = -\sum_{u \in \Psi} p_u \log_2(p_u)$$

where $p_u = \mathcal{U}(u, r)$.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Degree of anonymity

- Less roughly, use the Shannon entropy.
  - This gives an expected value
- Danezis and Diaz also mention defining degree of anonymity as $\log_2(N)$, for N the number of users
- Can also use min entropy for worst case
- They don't mention it, but all these are related

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Degree of anonymity

- $H_0 = \log|X|$
- $H_1 = -\sum_{i=1}^{n} p_i \log(p_i)$
- $H_\infty = -\log(\max_i p_i)$

{Sender, Receiver, Relationship} Anonymity

- Sender/receiver anonymity: For a given message m, what is the probability that m came from sender/receiver A?
  - Note this depends on who the adversary is: the recipient? A global passive adversary? An active adversary? ...
- Relationship anonymity: For a given message m, what is the probability that m came from sender A and went to destination B?
- Relationship* anonymity: What is the probability that sender A is communicating with destination B?
  - a persistent relationship, not a single message or request exchange
- Question: how can you have sender anonymity but not relationship* anonymity? Hint: a universal generalization implies all instantiations.

“Suppose the network provides perfect sender anonymity, i.e., any message exiting the network is equally likely to have originated from any active sender. By observing these messages, however, the attacker can easily infer that all of them have the same destination. For every active sender, the attacker can thus determine with 100% certainty that this sender is communicating with the website, completely breaking relationship anonymity.”

Shmatikov and Wang, ‘Measuring Relationship Anonymity in Mix Networks’

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity networks

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity networks address the traffic analysis problem

- Chaum: “Keeping confidential who converses with whom, and when they converse”
- Contrast with secrecy of message content

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Anonymity networks can involve trusted or semi-trusted relays

- Trusted parties are not adversaries: they can break anonymity
- Semi-trusted parties don’t all collude

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Trusted relays

- Example: Nym servers
  - a server keeps a dictionary between real and pseudonymous emails
  - request comes to the remailer, which forwards it, gets the response, and returns it to the user
  - Example: anon.penet.fi
- Other Examples: Anonymous proxies (startpage.com), VPNs

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Trusted relays

- Problem: messages are all linked
  - Stylometric attacks: the frequency of function words in the English language can be used in the long term to identify users (Rao & Rohatgi (2000), “Can Pseudonymity Really Guarantee Privacy?”)
  - Correspondent sets of each nym
- Anonymity is compromised if one node is compromised. (“Single point of failure.”)
  - lots of incentive to coerce
  - or if the node is not honest
- Fails bitwise indistinguishability: sometimes traffic analysis can deanonymize
  - http proxy example
  - timing correlation

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Semi-trusted relays

- Strengths
  - Compromise of more than one is needed, so more coercion resistant than trusted-relay approaches
  - “any single mix is able to provide the secrecy of the correspondence between the input and the outputs of the entire cascade” (Chaum)
- Weaknesses
  - Tagging attacks violate unlinkability (blind signing attack)
  - replay attacks
  - slow (public-key cryptography)

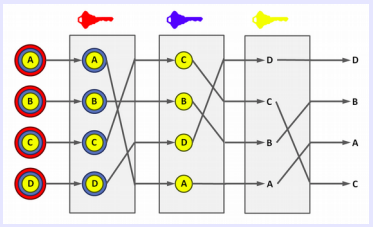
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Semi-trusted relays

- What are the problems with a mixnet with only one node? (Chaum)

Mixnets are anonymity systems with semi-trusted relays

- Routing protocol with a cascade of cryptographic relays called ‘mixes’
- Mixes only know their neighbors
- User-specifiable routing (Chaum’s “new kind of mix”)



wikipedia.org

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Mix-nets are anonymity systems with semi-trusted relays

- Suppose we are at a mix A1, which receives message m.
- m is split into a fixed number blocks, ℓ

$$A_1\colon [K_{A_1}(R_{A_1}, A_2)], [R_{A_1}^{-1}(K_{A_2}(R_{A_2}, A_3))], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)] \rightarrow.$$

$$A_2\colon [K_{A_2}(R_{A_2}, A_3)], [R_{A_2}^{-1}(K_{A_3}(R_{A_3}, A_4))], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)], [R_{A_1}(J_{A_1})] \rightarrow,$$

$$A\colon [M_1], [M_2], \ldots, [M_{l-n}],$$
$$[R_{A_n}(R_{A_{n-1}} \cdots R_{A_1}(J_{A_1}) \cdots)], \ldots, [R_{A_n}(J_{A_n})].$$

Mix-nets are anonymity systems with semi-trusted relays

- The first block is like a header: it contains the key R_{A1} and address A2 for the next hop. This is stripped off of the message, and a padding ("junk") block is added to the end

$$A_1\colon [K_{A_1}(R_{A_1}, A_2)], [R_{A_1}^{-1}(K_{A_2}(R_{A_2}, A_3))], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)] \rightarrow.$$

$$A_2\colon [K_{A_2}(R_{A_2}, A_3)], [R_{A_2}^{-1}(K_{A_3}(R_{A_3}, A_4))], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)], [R_{A_1}(J_{A_1})] \rightarrow,$$

$$A\colon [M_1], [M_2], \ldots, [M_{l-n}],$$
$$[R_{A_n}(R_{A_{n-1}} \cdots R_{A_1}(J_{A_1}) \cdots)], \ldots, [R_{A_n}(J_{A_n})].$$

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Mix-nets are anonymity systems with semi-trusted relays

- The rest of the blocks are, first, the header blocks for all remaining routers in the cascade, and next, the message. All of these are encoded using .

$$A_1\colon [K_{A_1}(R_{A_1}, A_2)], [R_{A_1}^{-1}(K_{A_2}(R_{A_2}, A_3))], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)] \rightarrow.$$

$$A_2\colon [K_{A_2}(R_{A_2}, A_3)], [R_{A_2}^{-1}(K_{A_3}(R_{A_3}, A_4))], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)], [R_{A_1}(J_{A_1})] \rightarrow,$$

$$A\colon [M_1], [M_2], \ldots, [M_{l-n}],$$
$$[R_{A_n}(R_{A_{n-1}}, \cdots R_{A_1}(J_{A_1}) \cdots)], \ldots, [R_{A_n}(J_{A_n})].$$

Mix-nets are anonymity systems with semi-trusted relays

- A1 uses the R_{A1} it now has to decode the (ℓ-1) blocks after the header in the original message: these are the first part of the message sent out from A1, they contain the headers for A2, the encoded headers for A3,...An, and then the encoded message
- The blocks are passed to the next node, which could be another mix

$$A_1: [K_{A_1}(R_{A_1}, A_2)], [R_{A_1}^{-1}(K_{A_2}(R_{A_2}, A_3))], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)] \rightarrow.$$

$$A: [M_1], [M_2], \ldots, [M_{l-n}],$$
$$[R_{A_n}(R_{A_{n-1}} \cdots R_{A_1}(J_{A_1}) \cdots)], \ldots, [R_{A_n}(J_{A_n})].$$

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Mix-nets are anonymity systems with semi-trusted relays

- Mixes only know their neighbors. (Question: Why?)
- All nodes have a public key
- Weaknesses
  - active attacks: tagging attacks (blind signing attack), replay attacks
  - slow (public-key cryptography, latency in anonymous remailers).

$$A_1\colon [K_{A_1}(R_{A_1}, A_2)], [R_{A_1}^{-1}(K_{A_2}(R_{A_2}, A_3))], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots], \ldots,$$
$$[R_{A_1}^{-1}(R_{A_2}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)] \rightarrow.$$

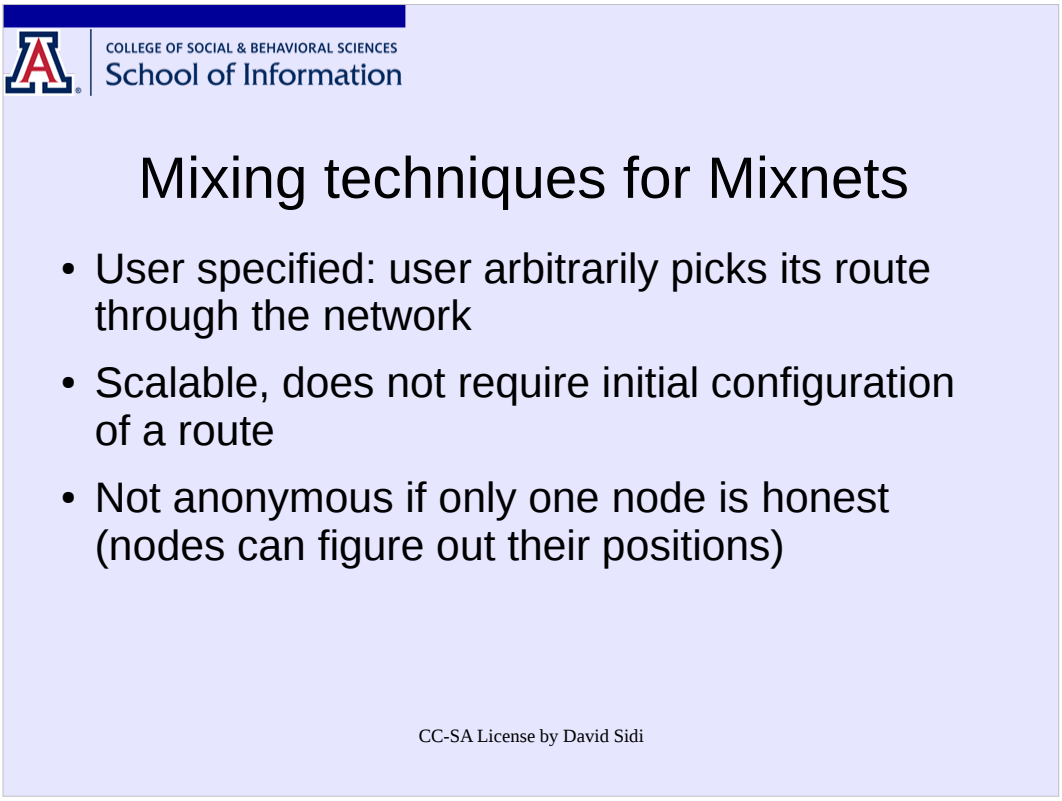
$$A_2\colon [K_{A_2}(R_{A_2}, A_3)], [R_{A_2}^{-1}(K_{A_3}(R_{A_3}, A_4))], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_{n-1}}^{-1}(K_{A_n}(R_{A_n}, A)) \cdots)],$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_1) \cdots)], \ldots,$$
$$[R_{A_2}^{-1}(R_{A_3}^{-1} \cdots R_{A_n}^{-1}(M_{l-n}) \cdots)], [R_{A_1}(J_{A_1})] \rightarrow,$$

$$A\colon [M_1], [M_2], \ldots, [M_{l-n}],$$
$$[R_{A_n}(R_{A_{n-1}} \cdots R_{A_1}(J_{A_1}) \cdots)], \ldots, [R_{A_n}(J_{A_n})].$$

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Mixing techniques for Mixnets

- Cascading: All nodes are always used, in the same order
- Scalability is a problem, requires setting up a fixed route with all nodes
- Only requires one honest node to preserve anonymity



relays can determine their position in the chain, which can be used to deanonymize with enough collusion