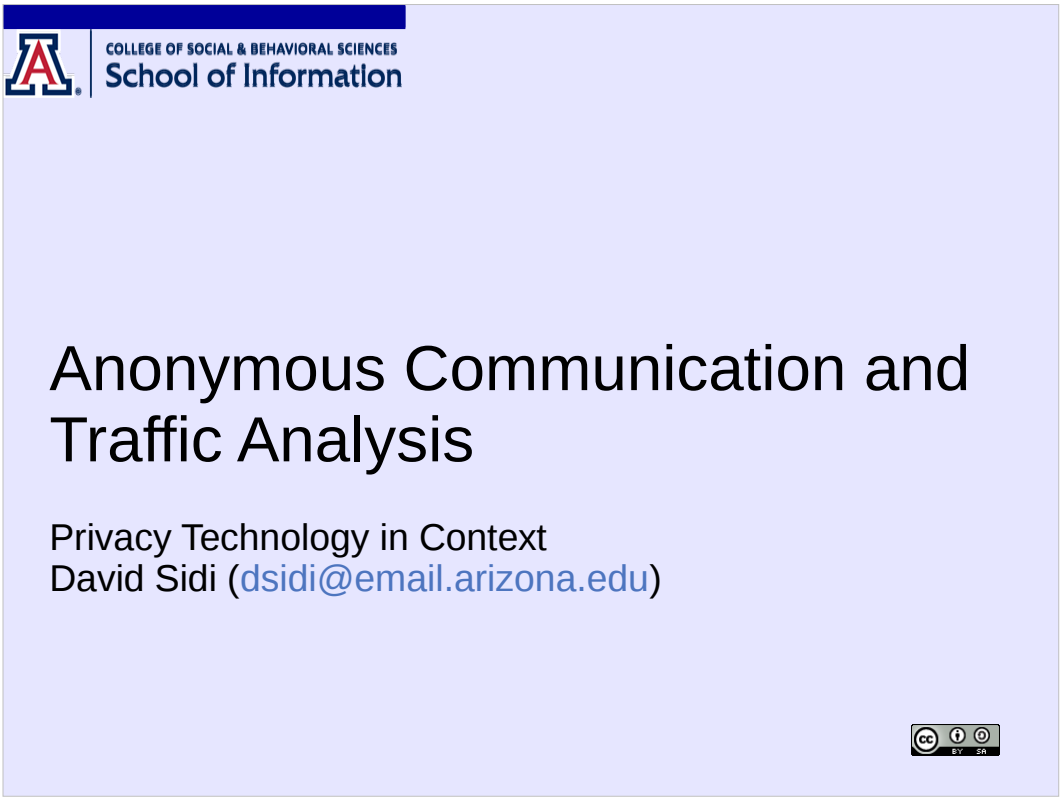


Today: background for discussion of anonymity

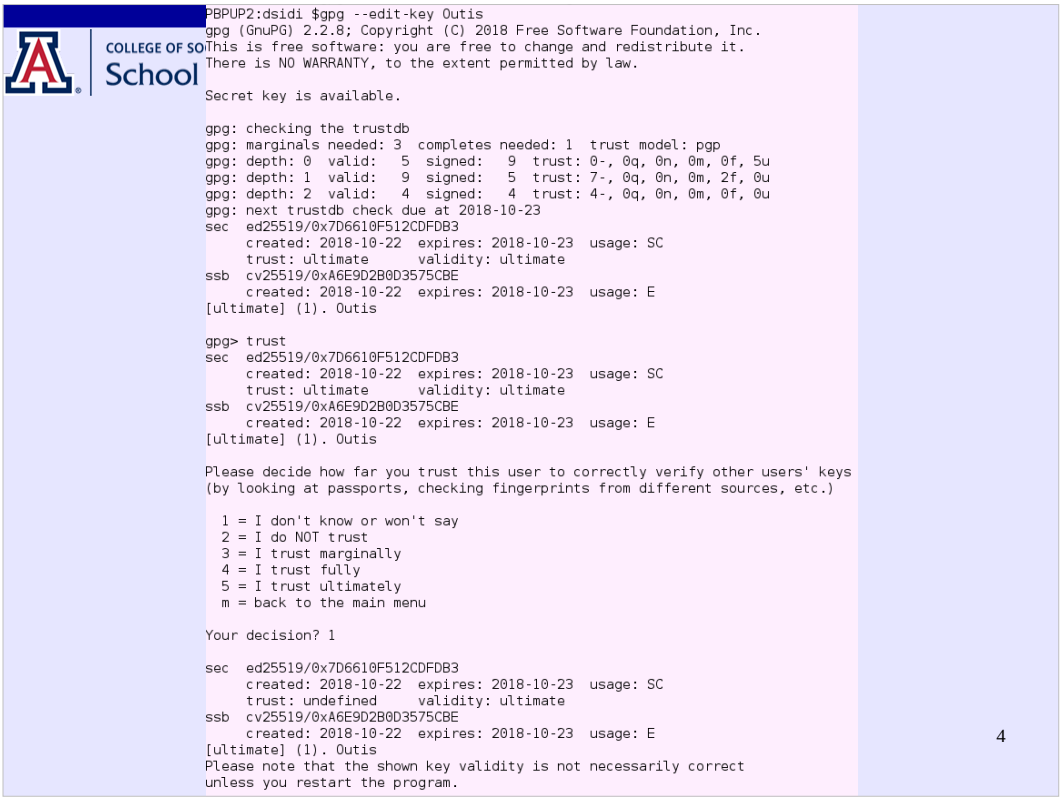
Small mention of interesting things

- open access week
- star wars via telnet
- Assignment 4

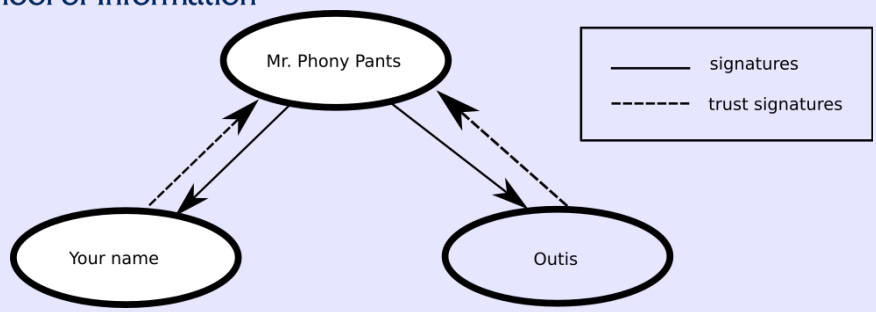
```
PBPUP2:dsidi $gpg --expert --full-gen-key
gpg (GnuPG) 2.2.8; Copyright (C) 2018 Free Software Foundation, Inc.
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, to the extent permitted by law.

Please select what kind of key you want:
   (1) RSA and RSA (default)
   (2) DSA and Elgamal
   (3) DSA (sign only)
   (4) RSA (sign only)
   (7) DSA (set your own capabilities)
   (8) RSA (set your own capabilities)
   (9) ECC and ECC
  (10) ECC (sign only)
  (11) ECC (set your own capabilities)
  (13) Existing key
Your selection? 9
```

```
PBPUP2:dsidi $echo "I am the very model of a modern major general" > foo
PBPUP2:dsidi $gpg -u 0x7D6610F512CDFDB3 --sign ./foo
PBPUP2:dsidi $
```

Outis's keygrip



need to save and exit




```
pub   rsa4096/0xE0B4CBFADE84A90F 2018-10-22 [SC] [expires: 2018-10-23]
      Key fingerprint = 6669 5394 AA1C 8483 42D9  6FCE E0B4 CBFA DE84 A90F
uid                   [  full  ] Mr. Phony Pants
sig 3        0xE0B4CBFADE84A90F 2018-10-22  Mr. Phony Pants
sig          1 0x7D6610F512CDFDB3 2018-10-22  Outis
sig          1 0x43BA803E7B77C59C 2018-10-22  Joe Bloggs
sub   rsa4096/0x24B29D8050E407C2 2018-10-22 [E] [expires: 2018-10-23]
sig          0xE0B4CBFADE84A90F 2018-10-22  Mr. Phony Pants
```

```
pub   ed25519/0x7D6610F512CDFDB3 2018-10-22 [SC] [expires: 2018-10-23]
      Key fingerprint = FB35 960C CD67 FA34 D8E4  8E0F 7D66 10F5 12CD FDB3
uid                   [  full  ] Outis
sig 3        0x7D6610F512CDFDB3 2018-10-22  Outis
sig          0xE0B4CBFADE84A90F 2018-10-22  Mr. Phony Pants
sub   cv25519/0xA6E9D2B0D3575CBE 2018-10-22 [E] [expires: 2018-10-23]
sig          0x7D6610F512CDFDB3 2018-10-22  Outis
```

```
pub   rsa4096/0x43BA803E7B77C59C 2018-10-22 [SC] [expires: 2018-10-23]
      Key fingerprint = 4A1C 7CDC A7DA E218 0430  236C 43BA 803E 7B77 C59C
uid                   [ultimate] Joe Bloggs
sig 3        0x43BA803E7B77C59C 2018-10-22  Joe Bloggs
sig          0xE0B4CBFADE84A90F 2018-10-22  Mr. Phony Pants
sub   rsa4096/0x9846E4B70B131BE0 2018-10-22 [E] [expires: 2018-10-23]
sig          0x43BA803E7B77C59C 2018-10-22  Joe Bloggs
```

```
PBPUP2:dsidi $gpg -u "Joe Bloggs" --verify foo.gpg
gpg: Signature made Mon 22 Oct 2018 12:03:58 PM MST
gpg:                using EDDSA key FB35960CCD67FA34D8E48E0F7D6610F512CDFDB3
gpg: Good signature from "Outis" [unknown]
gpg: WARNING: This key is not certified with a trusted signature!
gpg:          There is no indication that the signature belongs to the owner.
Primary key fingerprint: FB35 960C CD67 FA34 D8E4  8E0F 7D66 10F5 12CD FDB3
```

```
PBPUP2:dsidi $gpg -u "Joe Bloggs" --verify foo.gpg
gpg: Signature made Mon 22 Oct 2018 12:03:58 PM MST
gpg:                using EDDSA key FB35960CCD67FA34D8E48E0F7D6610F512CDFDB3
gpg: Good signature from "Outis" [full]
Primary key fingerprint: FB35 960C CD67 FA34 D8E4  8E0F 7D66 10F5 12CD FDB3
```

```
fal8-course:dsidi $gnutls-cli -V --print-cert localhost </dev/null
Processed 151 CA certificate(s).
Resolving 'localhost:443'...
Connecting to '127.0.0.1:443'...
- Certificate type: X.509
- Got a certificate list of 1 certificates.
- Certificate[0] info:
 - X.509 Certificate Information:
        Version: 3
        Serial Number (hex): 01
        Issuer: CN=Mr_Phony_Pants,O=UA Privacy Course,L=Tucson,ST=Arizona,C=US
        Validity:
                Not Before: Mon Aug 20 19:00:37 UTC 2018
                Not After: Sun May 16 19:00:37 UTC 2021
        Subject: CN=Mr. Phony Pants,O=UA Privacy Class,L=Tucson,ST=AZ,C=US
        Subject Public Key Algorithm: RSA
        Algorithm Security Level: Medium (2048 bits)
                Modulus (bits 2048):
                        00:d7:29:2f:68:a2:02:c7:2f:bc:58:68:40:23:a1:ae
                        98:8c:14:4f:25:7a:0c:20:b5:fc:ae:2e:6e:ba:df:be
                        79:de:e0:32:28:d6:29:d0:5e:3b:4b:23:3c:3f:8c:5d
                        e2:12:1a:db:11:60:6f:69:8b:5e:17:38:ff:a9:d1:b9
                        af:72:ed:6a:74:c3:ab:26:eb:59:57:b9:6b:0c:3d:f5
                        cd:59:6f:80:de:69:99:ea:4d:0d:b5:2f:d0:98:79:84
                        76:d8:32:b0:dd:29:39:d1:3a:41:92:72:e5:ef:89:52
                        7b:4b:e9:99:4d:7c:b1:84:74:7f:db:0c:9f:a7:20:40
                        eb:cb:da:ef:98:49:31:eb:0a:d7:88:f8:52:a3:0f:51
                        43:00:74:76:f1:48:b8:f9:ab:74:14:f8:49:e1:84:d1
                        d0:ac:3c:81:b8:57:8d:e5:b9:c9:57:21:cc:33:bf:95
                        f2:11:82:4c:f0:e6:ff:f2:45:06:0b:52:df:e9:b8:ce
                        b4:58:d0:cc:02:49:b9:09:e1:2c:11:66:1e:4a:f8:34
                        d4:bc:19:84:03:28:40:53:5c:80:42:d1:f4:1b:48:25
                        d7:54:c4:3a:3f:64:19:d2:64:be:05:84:21:c1:45:26
                        6f:16:39:8d:51:11:fe:42:89:76:db:98:f6:56:50:e9
                        d7
                Exponent (bits 24):
                        01:00:01
        Extensions:
                Subject Key Identifier (not critical):
                        33e4731d3550969edecacec8dcdedc18768179b9
                Authority Key Identifier (not critical):
                        directoryName: CN=Mr_Phony_Pants,O=UA Privacy Course,L=Tucson,ST=Arizona,C=US
                        serial: 00bc4be139218655bb
                Basic Constraints (not critical):
                        Certificate Authority (CA): FALSE
                Key Usage (not critical):
                        Digital signature.
                        Key encipherment.
                Unknown extension 1.3.6.1.4.1.743 (not critical):
                        ASCII: ..Many cheerful facts
```

743

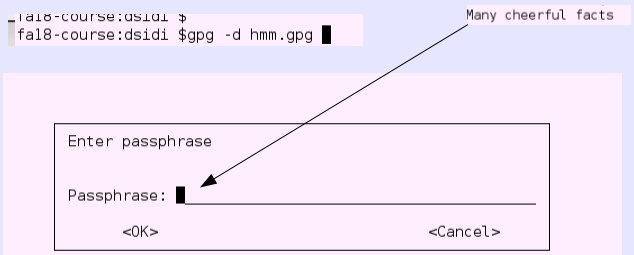
- Format of this page
- Modify this OID
- Create child OID
- Create brother OID
- Find similar OIDs

OID description

	{iso(1) identified-organization(3) dod(6) internet(1) private(4) enterprise(1) 743}	(ASN.1 notation)
OID:	1.3.6.1.4.1.743	(dot notation)
	/ISO/Identified-Organization/6/1/4/1/743	(OID-IRI notation)

Description: Central Intelligence Agency (CIA)

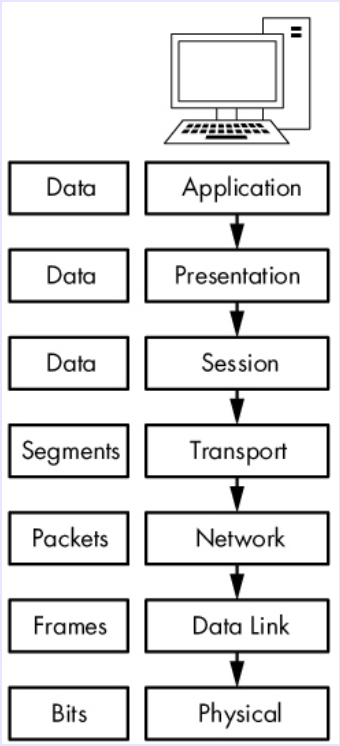
```
fa18-course:dsidi $gnutls-cli -V --print-cert localhost </dev/null
Processed 151 CA certificate(s).
Resolving 'localhost:443'...
Connecting to '127.0.0.1:443'...
- Certificate type: X.509
- Got a certificate list of 1 certificates.
- Certificate[0] info:
 - X.509 Certificate Information:
        Version: 3
        Serial Number (hex): 01
        Issuer: CN=Mr_Phony_Pants,O=UA Privacy Course,L=Tucson,ST=Arizona,C=US
        Validity:
                Not Before: Mon Aug 20 19:00:37 UTC 2018
                Not After: Sun May 16 19:00:37 UTC 2021
        Subject: CN=Mr. Phony Pants,O=UA Privacy Class,L=Tucson,ST=AZ,C=US
        Subject Public Key Algorithm: RSA
        Algorithm Security Level: Medium (2048 bits)
                Modulus (bits 2048):
                        00:d7:29:2f:68:a2:02:c7:2f:bc:58:68:40:23:a1:ae
                        98:8c:14:4f:25:7a:0c:20:b5:fc:ae:2e:6e:ba:df:be
                        79:de:e0:32:28:d6:29:d0:5e:3b:4b:23:3c:3f:8c:5d
                        e2:12:1a:db:11:60:6f:69:8b:5e:17:38:ff:a9:d1:b9
                        af:72:ed:6a:74:c3:ab:26:eb:59:57:b9:6b:0c:3d:f5
                        cd:59:6f:80:de:69:99:ea:4d:0d:b5:2f:d0:98:79:84
                        76:d8:32:b0:dd:29:39:d1:3a:41:92:72:e5:ef:89:52
                        7b:4b:e9:99:4d:7c:b1:84:74:7f:db:0c:9f:a7:20:40
                        eb:cb:da:ef:98:49:31:eb:0a:d7:88:f8:52:a3:0f:51
                        43:00:74:76:f1:48:b8:f9:ab:74:14:f8:49:e1:84:d1
                        d0:ac:3c:81:b8:57:8d:e5:b9:c9:57:21:cc:33:bf:95
                        f2:11:82:4c:f0:e6:ff:f2:45:06:0b:52:df:e9:b8:ce
                        b4:58:d0:cc:02:49:b9:09:e1:2c:11:66:1e:4a:f8:34
                        d4:bc:19:84:03:28:40:53:5c:80:42:d1:f4:1b:48:25
                        d7:54:c4:3a:3f:64:19:d2:64:be:05:84:21:c1:45:26
                        6f:16:39:8d:51:11:fe:42:89:76:db:98:f6:56:50:e9
                        d7
                Exponent (bits 24):
                        01:00:01
        Extensions:
                Subject Key Identifier (not critical):
                        33e4731d3550969edecacec8dcdedc18768179b9
                Authority Key Identifier (not critical):
                        directoryName: CN=Mr_Phony_Pants,O=UA Privacy Course,L=Tucson,ST=Arizona,C=US
                        serial: 00bc4be139218655bb
                Basic Constraints (not critical):
                        Certificate Authority (CA): FALSE
                Key Usage (not critical):
                        Digital signature.
                        Key encipherment.
                Unknown extension 1.3.6.1.4.1.743 (not critical):
                        ASCII: ..Many cheerful facts
```


fa18-course:dsidi $
fa18-course:dsidi $gpg -d hmm.gpg
Many cheerful facts
Enter passphrase
Passphrase:
<OK>
<Cancel>

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

The Seven-Layer OSI Reference Model

"the way up is the way down." -Heraclitus

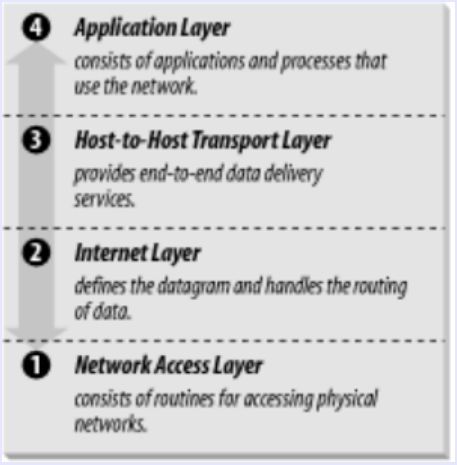


We'll mostly stick with a simplification of this model, the TCP/IP protocol stack (book diagram). A stack is composed of layers, which only interact with their immediate neighbor layers, via encapsulation (book diagram).

The TCP/IP networking model has four layers.

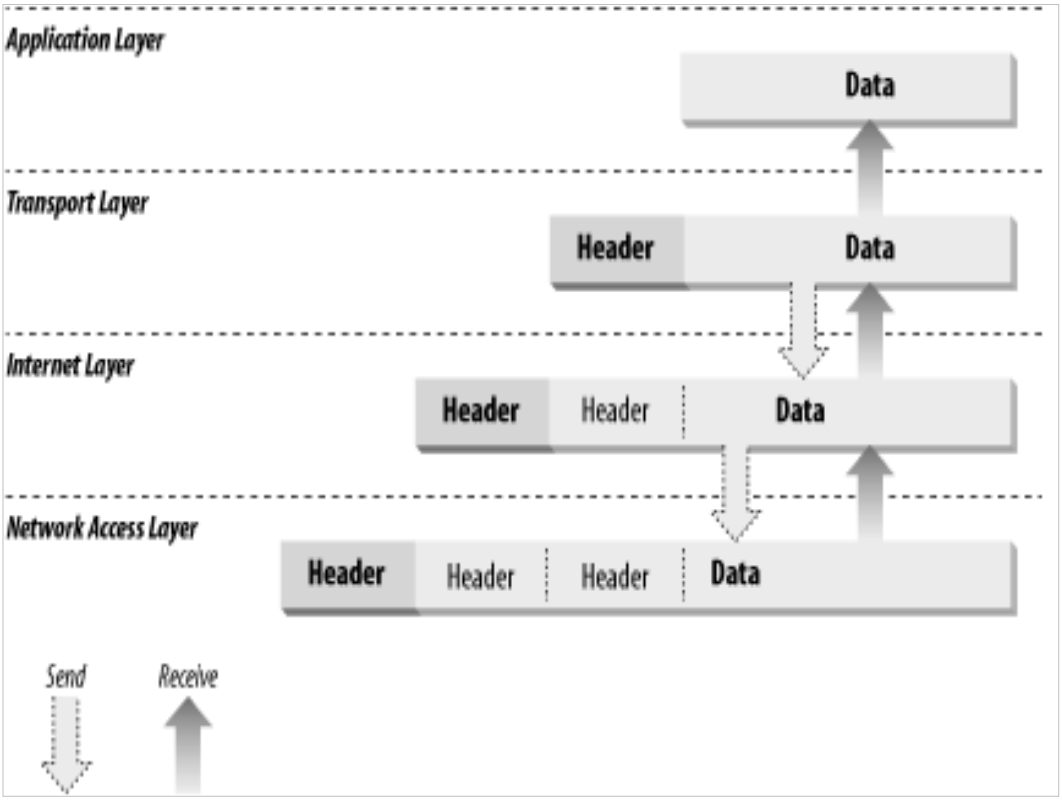
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

TCP/IP Protocol Stack



At its beginnings in the application layer, data is generated carrying information in a form useful to a particular network service. The data *payload* progresses “downward” through the other layers in a fixed order, accumulating control information as it goes in the form of layer-specific headers.

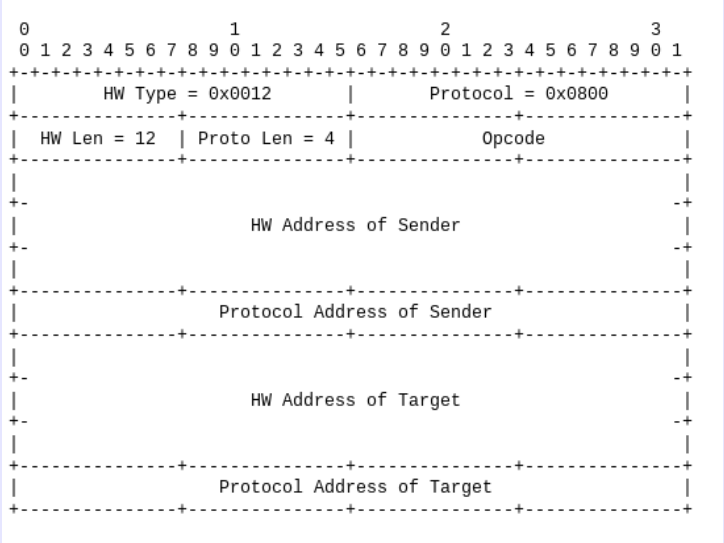
In general there are lots of network services trying to use transport layer protocols like TCP, lots of transport protocols using internet layer protocols like IP, and lots of internet protocols using network access layer protocols like Ethernet. Need multiplexing and demultiplexing



encapsulation.

On the way out (sending) each layer in the stack takes the header+payload from the one above it, and adds its own header. On the way in (receiving) it strips off a layer of headers and interprets them to get the data closer to its destination

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information



Network access layer example. IP -> Ethernet is done with ARP

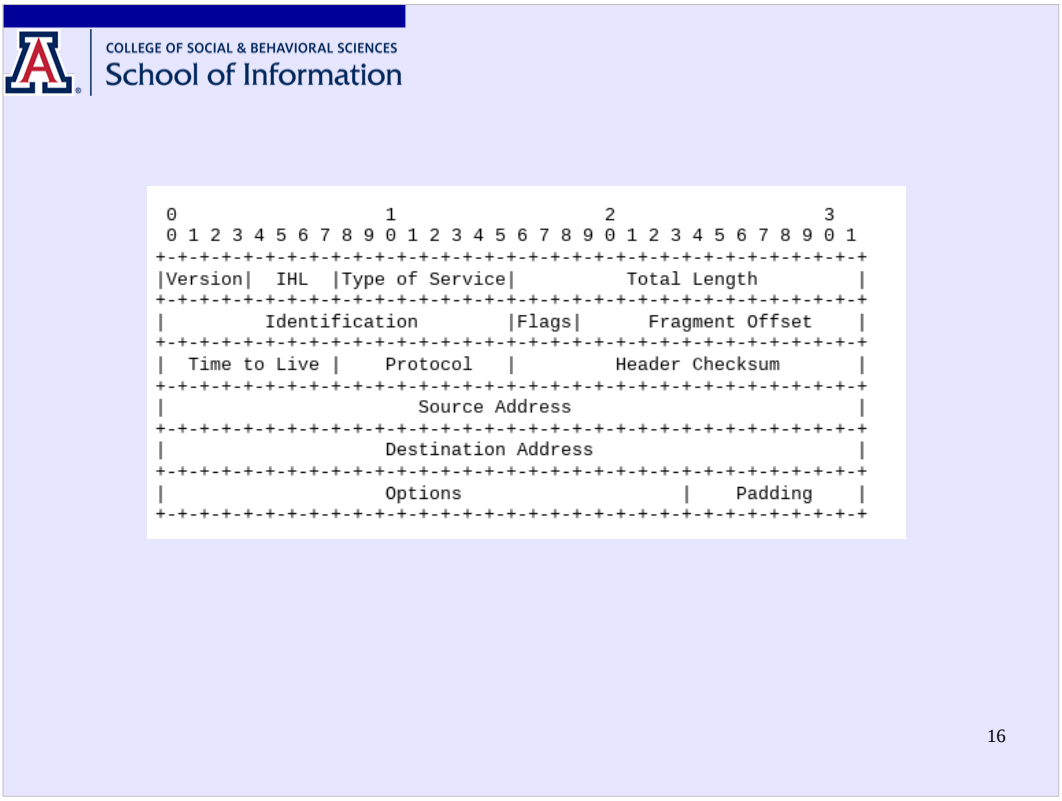
From RFC 4338, Sect. 7: ‘ARP Packet Format’

HW Type value is for Fiber Channel ARP. Ethernet is 0x0001, IEEE 802.11 is 0x0006

Protocol: what kind of message from layer above is being sent? 2048 is IPv4

HW address on ethernet is a MAC address

Physical networks do not understand IP addresses---they have their own addressing scheme.



Internet layer example: IP Datagram header

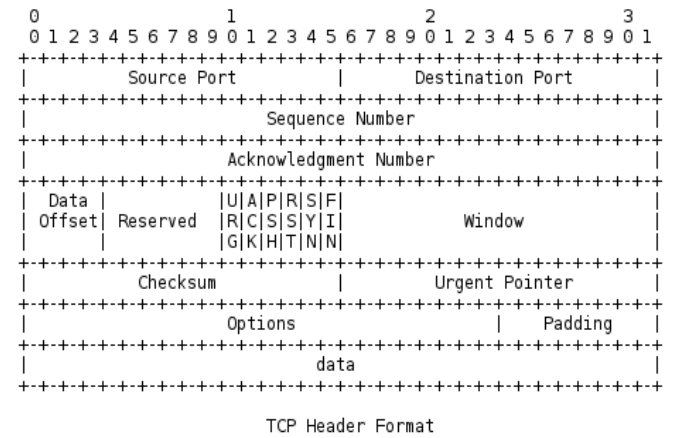
TCP/IP solves the routing problem with an addressing scheme: IP. Every host and gateway on the internet has a globally unique IP address.

Important idea: in general, nobody knows the full route.

(more on IP addresses: address classes, CIDR)

4 bytes. First nibble encodes address classes: prefix length gives network part of the address. Last 1, 2, 3 bytes are for hosts in class C, B, A networks, respectively. CIDR supercedes address classes.

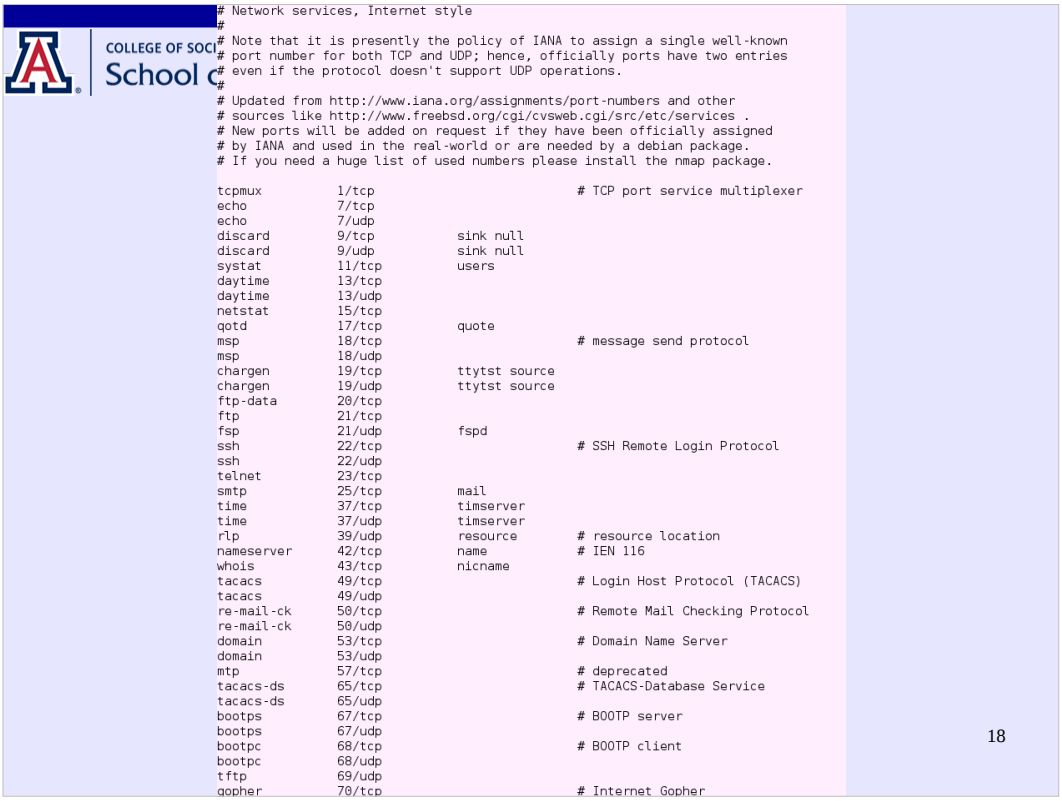
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information



TCP Header Format

Transport layer example: TCP Segment Header

Addressing multiplexing problem: message must not just reach a host, but a process running on the host. Ports are codes for application layer services...(next)



```
# Network services, Internet style
#
# Note that it is presently the policy of IANA to assign a single well-known
# port number for both TCP and UDP; hence, officially ports have two entries
# even if the protocol doesn't support UDP operations.
#
# Updated from http://www.iana.org/assignments/port-numbers and other
# sources like http://www.freebsd.org/cgi/cvsweb.cgi/src/etc/services .
# New ports will be added on request if they have been officially assigned
# by IANA and used in the real-world or are needed by a debian package.
# If you need a huge list of used numbers please install the nmap package.

tcpmux          1/tcp                           # TCP port service multiplexer
echo            7/tcp
echo            7/udp
discard         9/tcp           sink null
discard         9/udp           sink null
systat          11/tcp          users
daytime         13/tcp
daytime         13/udp
netstat         15/tcp
qotd            17/tcp          quote
msp             18/tcp                          # message send protocol
msp             18/udp
chargen         19/tcp          ttytst source
chargen         19/udp          ttytst source
ftp-data        20/tcp
ftp             21/tcp
fsp             21/udp          fspd
ssh             22/tcp                          # SSH Remote Login Protocol
ssh             22/udp
telnet          23/tcp
smtp            25/tcp          mail
time            37/tcp          timserver
time            37/udp          timserver
rlp             39/udp          resource        # resource location
nameserver      42/tcp          name            # IEN 116
whois           43/tcp          nicname
tacacs          49/tcp                          # Login Host Protocol (TACACS)
tacacs          49/udp
re-mail-ck      50/tcp                          # Remote Mail Checking Protocol
re-mail-ck      50/udp
domain          53/tcp                          # Domain Name Server
domain          53/udp
mtp             57/tcp                          # deprecated
tacacs-ds       65/tcp                          # TACACS-Database Service
tacacs-ds       65/udp
bootps          67/tcp                          # BOOTP server
bootps          67/udp
bootpc          68/tcp                          # BOOTP client
bootpc          68/udp
tftp            69/udp
gopher          70/tcp                          # Internet Gopher
```

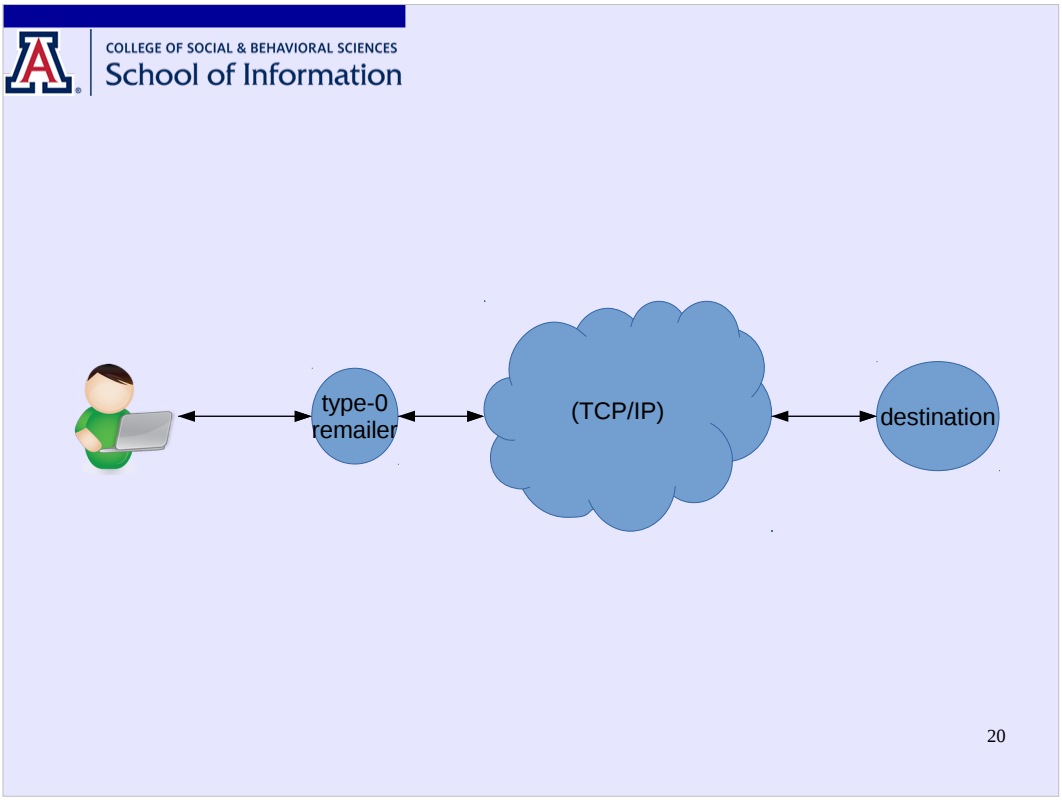

this is /etc/services, showing port numbers of well-known services.

Chaum 1981:

process each item of mail before it is delivered. A participant prepares a message M for delivery to a participant at address A by sealing it with the addressee's public key K_a, appending the address A, and then sealing the result with the mix's public key K_1. The left-hand side of the following expression denotes this item which is input to the mix:

$$K_1(R_1, K_a(R_0, M), A) \rightarrow K_a(R_0, M), A.$$

How does the message get to A?
One answer: overlay network



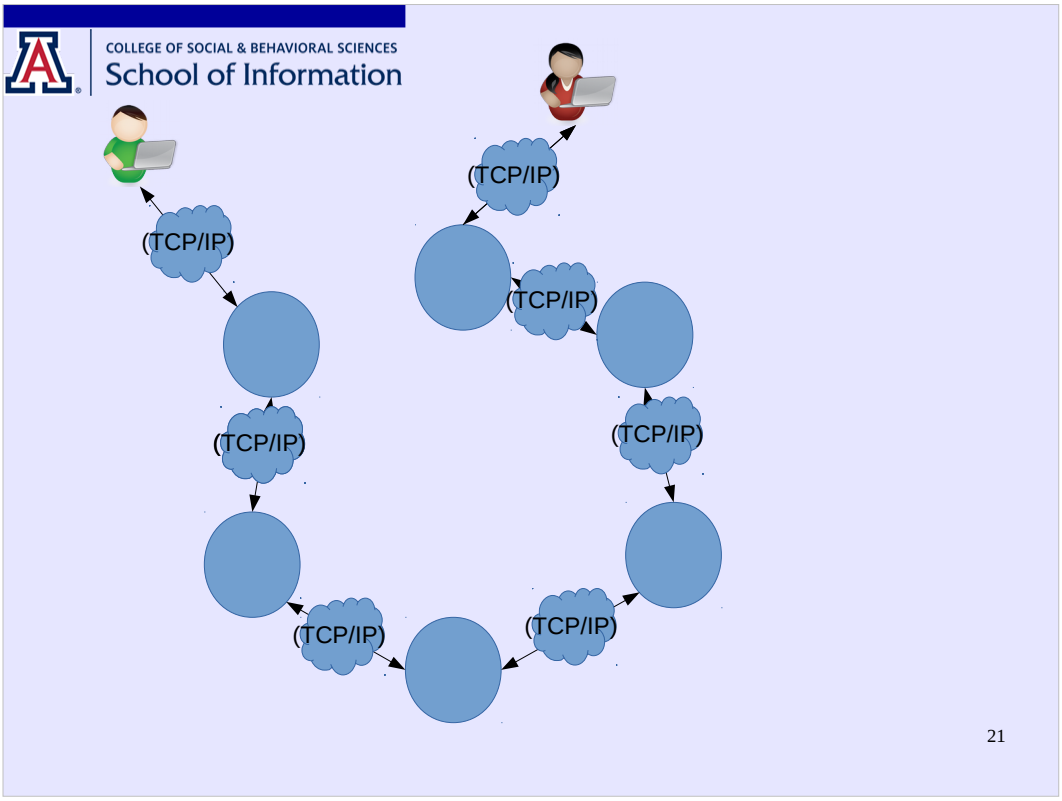
The particular kind of application layer network services we are going to be interested in are those that perform routing *on top of* TCP/IP, to provide some additional anonymity-related properties. These are called *overlay networks*

we said IP solves the routing problem across heterogeneous physical networks

in doing so, it reveals the final destination to every intermediate hop

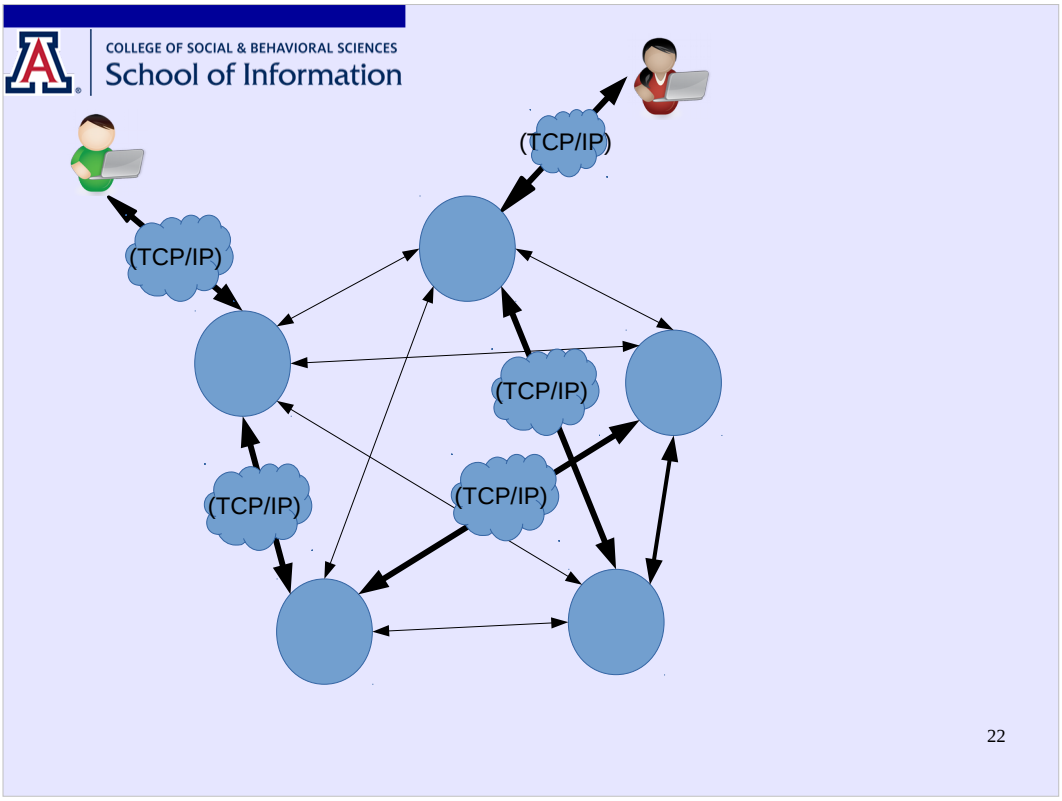
Not so great for anonymity, but you can implement anonymous routing protocols using TCP/IP

Here's the simplest (degenerate) form of mixnet



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

route unpredictability is not a thing; all clients use the same route. Mix strategies are what give mixnets their security properties



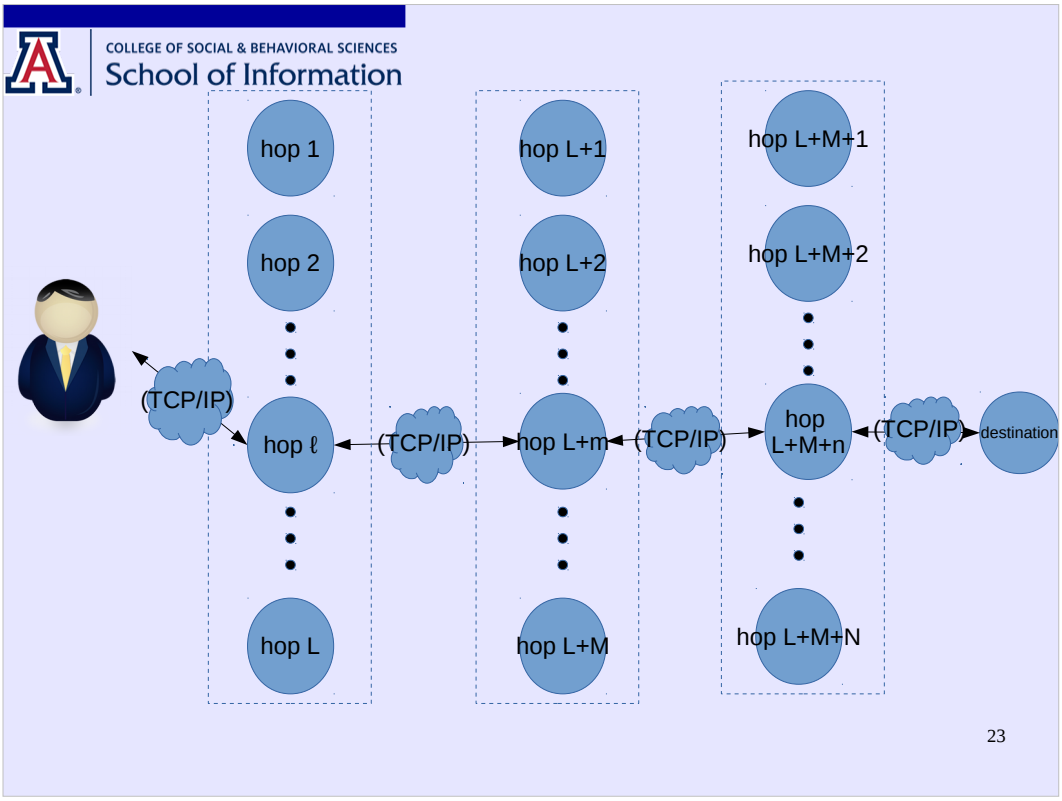
free route topology

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.

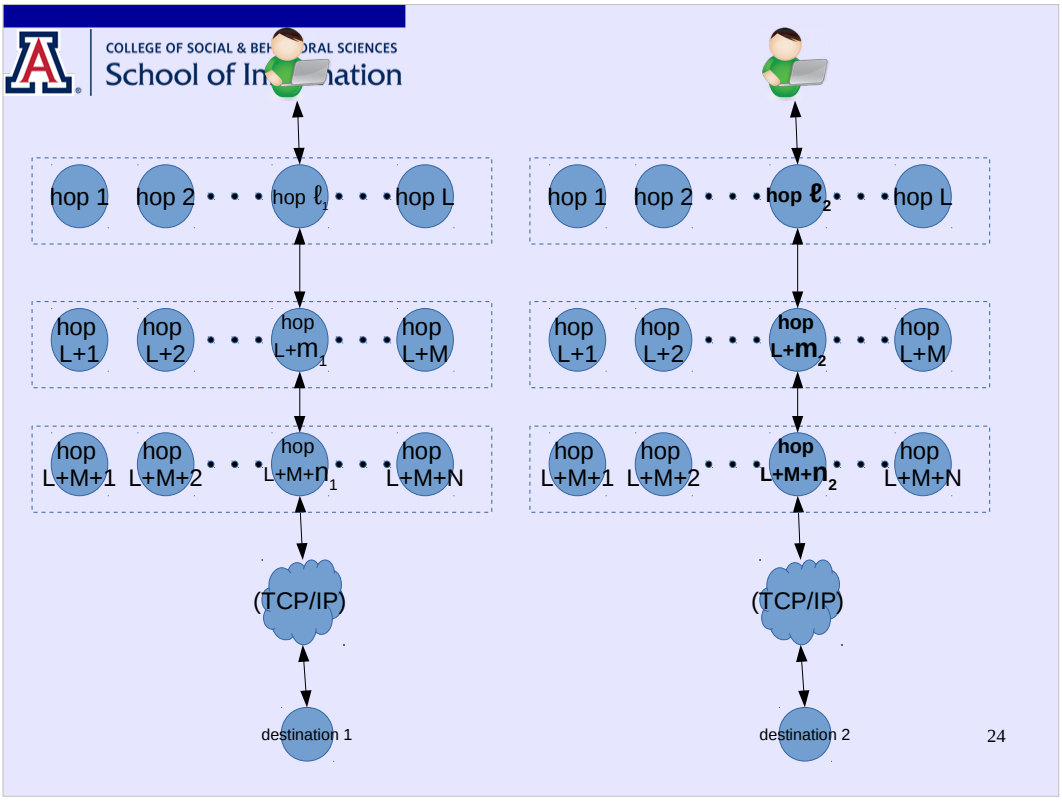
more scalable, so more anonymous in practice in some ways; but also less anonymous in some ways than mixes.

scalability and anonymity are tied together

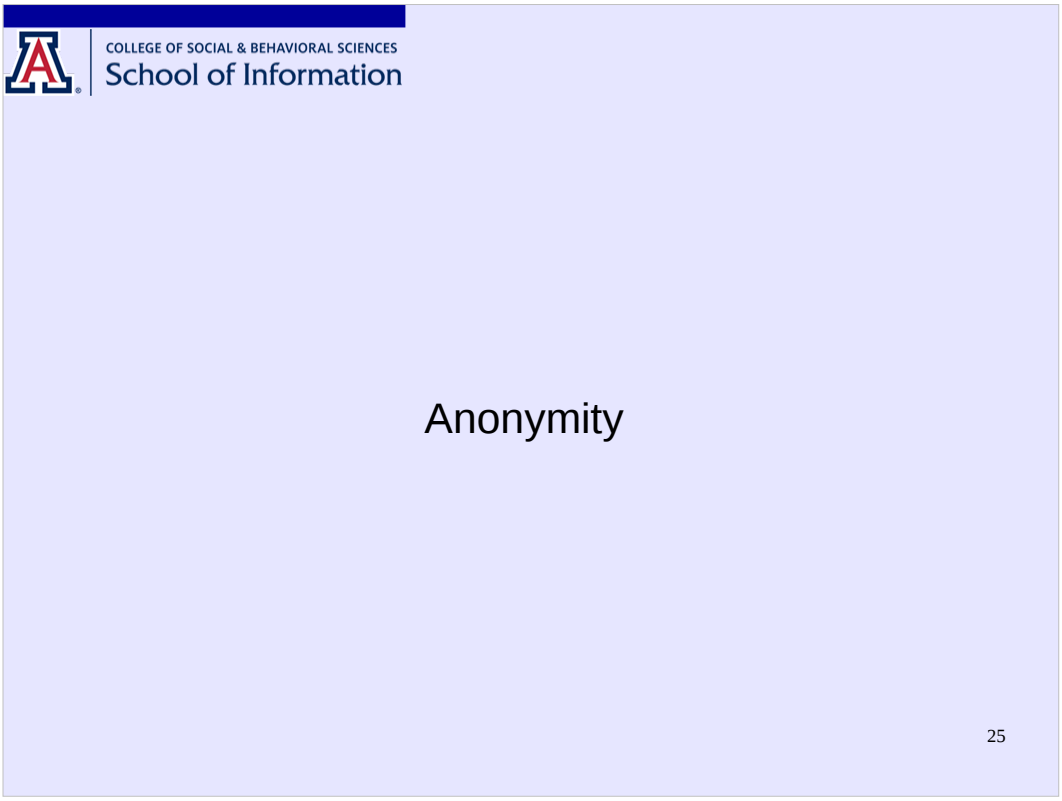


stratified topology

balances scalability with anonymity. Tor does this with helper nodes (“guards”), intermediate nodes (“relays”), and exit nodes. (Other stuff too--e.g., bridges). The choice at each layer is random (see next slide)



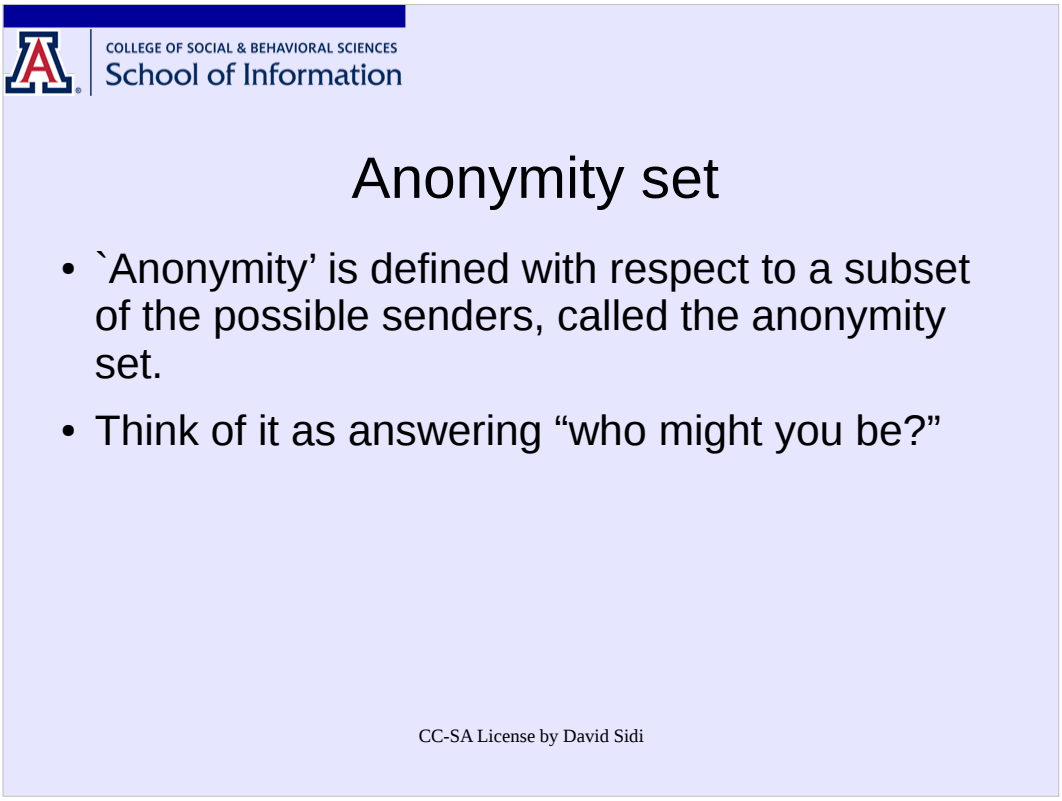
random choice at each layer



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.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Terminology Review



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 subjects

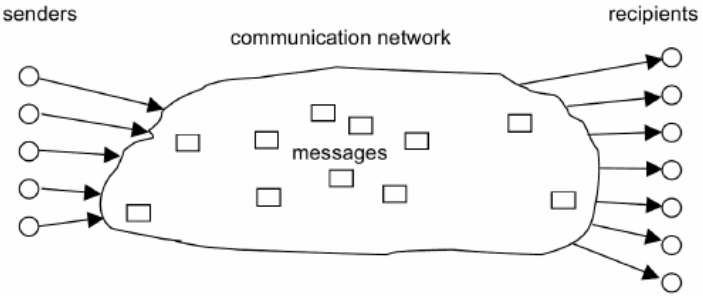
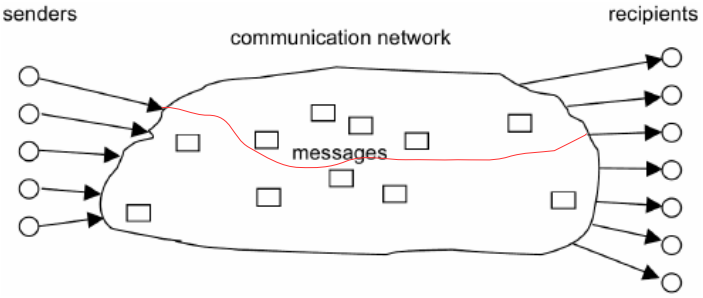
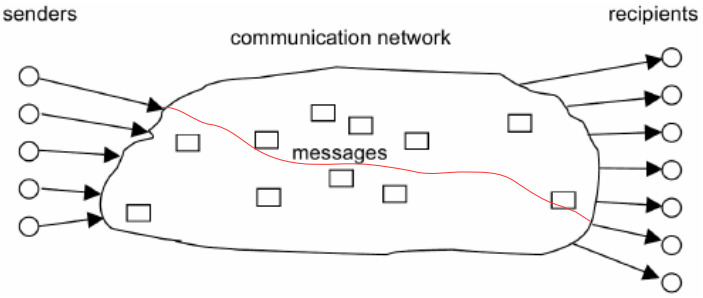


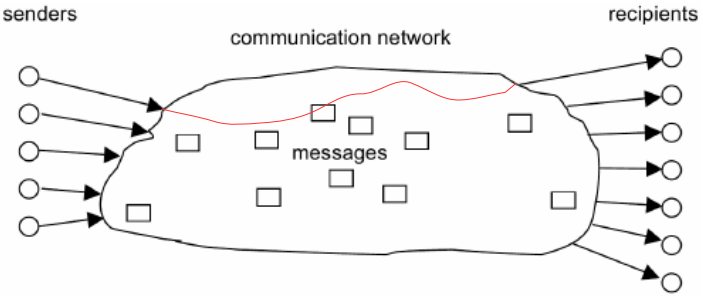
Image credit (before modification):
Christina Pöpper
Ruhr-University Bochum

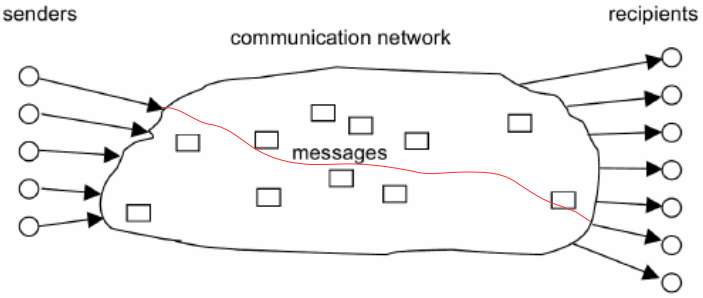
senders communicating messages with recipient.

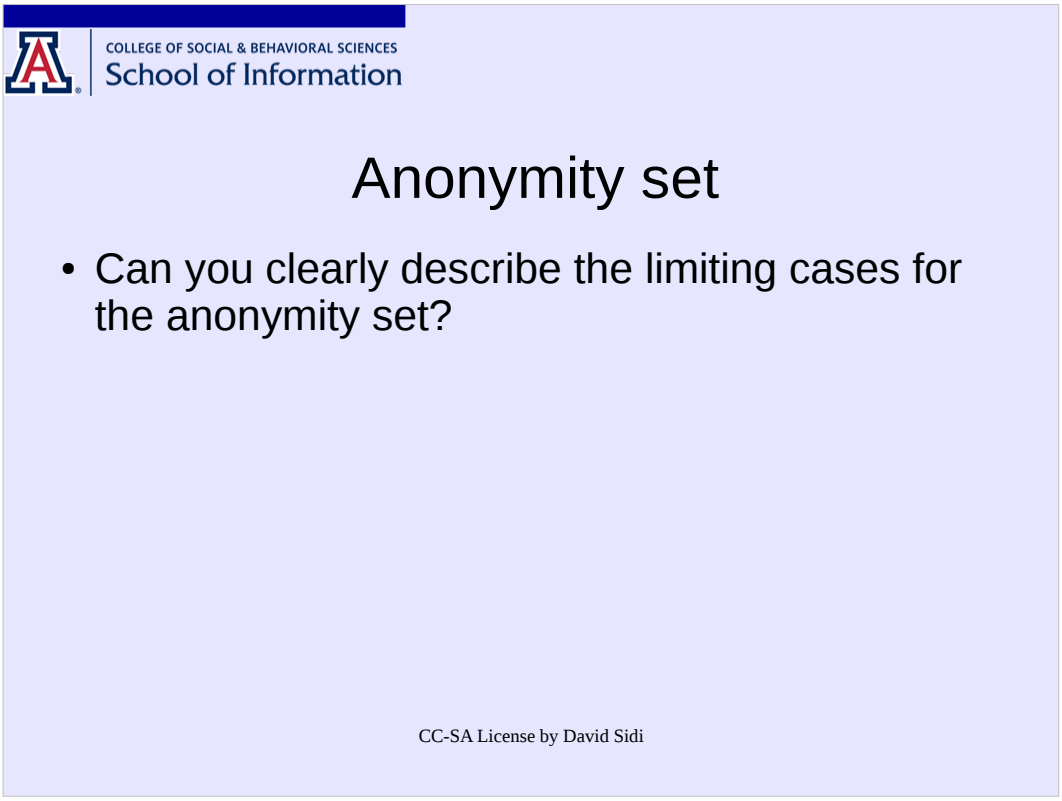
An adversary tries to reduce the anonymity of some or all of the parties to the communication


senders
communication network
recipients
messages


senders
communication network
recipients
messages


senders
communication network
recipients
messages

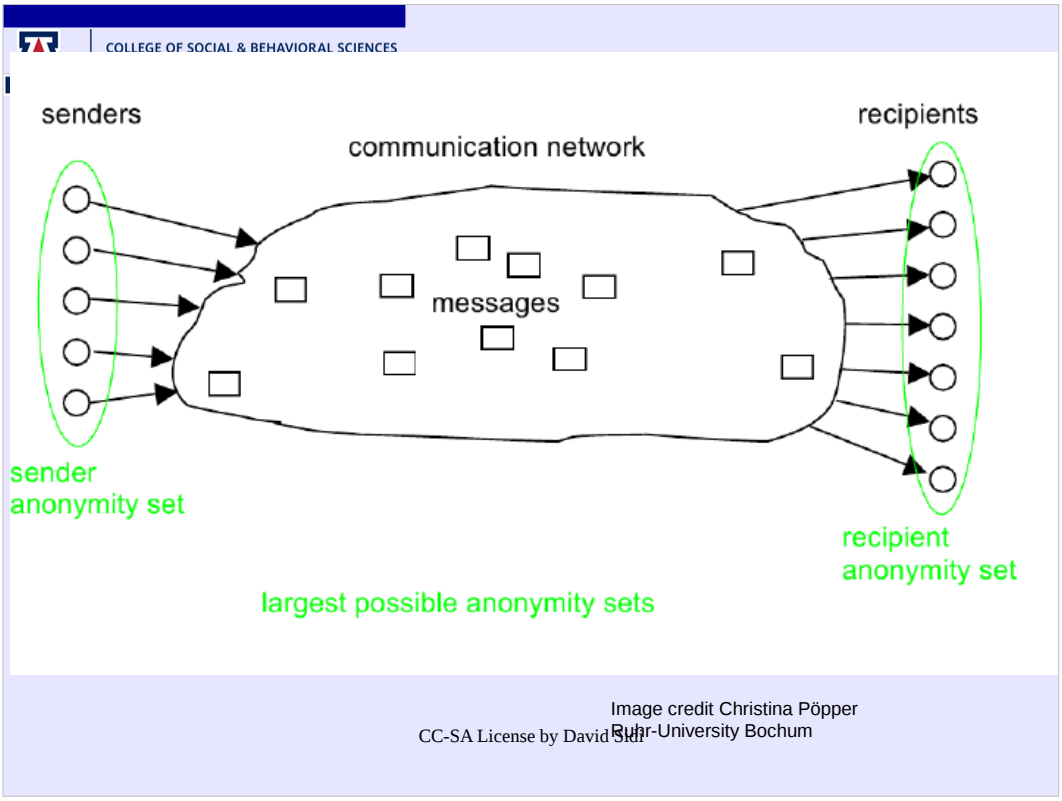




1. singleton. The subject has no anonymity; he is perfectly identifiable relative to the set of subjects.

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

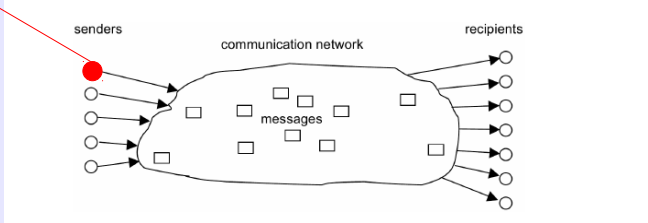
Notice what this means (Berthold's metric)



- 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

	MAC	Browser_fingerprint	IP	Sites_visited
SNDER_1	00:a0:ef:eb:5v:ff	af7f098c39728f8cb676e3df82ced01a149ee3aa92af2b88c20c4948a5fad5fd	172.16.1.5	torproject.org, ischool.arizona.edu, maps.google.com...
SNDER_2	00:c0:ff:dd:ff:ef	a5fad5fdd01a149eeaf7f098c39728f8cb676e3df82ce3aa92af2b88c20c4948	172.16.12.4	nytimes.com, purple.com



- one way to think of senders is as rows in a database: things that can be identified by descriptions in terms of 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. Is the player 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 fingerprinting

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

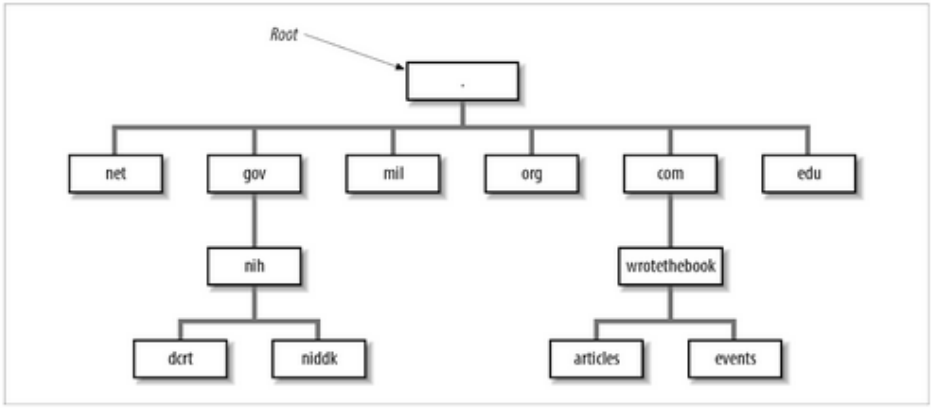


Figure 3-1. Domain hierarchy

- 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 beneaththem
  - Registering a domain name involves telling the TLD servers about you
  - you can then do subdomains at will
- Forwarding DNS server
- Recursive DNS server (or resolver)
- (Root nameserver)
- (Top Level Domain nameserver)
- Authoritative nameserver

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. What can you learn about me from my DNS requests?

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.

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

Browser fingerprinting

- Multi-monitor detection,
- Internal HashTable implementation detection
- WebRTC fingerprinting
- Math constants
- Accessibility fingerprinting
- Camera information
- DRM support
- Accelerometer support
- Virtual keyboards
- List of supported gestures (for touch-enabled devices)
- Pixel density
- Video and audio codecs availability
- Audio stack fingerprinting

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

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

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)

COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Information

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**.”

- **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)