CPSC 467: Cryptography and Computer Security

Instructor: Michael Fischer
Lecture by Ewa Syta

Lecture 25
December 4, 2013

Anonymous Communication

Anonymity[1]

Anonymity is the state of being not identifiable within a set of subjects, the anonymity set.

Anonymity set is the set of all possible participants in a system that could have been the sender or recipient of a particular message.

Sender anonymity is the state of being not identifiable within the sender anonymity set as the sender of a particular message.

Receiver anonymity is the state of being not identifiable within the receiver anonymity set as the receiver of a particular message.

[1] A. Pfitzmann and M. Hansen, *Anonymity, unobservability, and pseudonymity: A consolidated proposal for terminology*, 2008

Goal of anonymity systems[1]

In general, anonymity systems seek to provide *unlinkability* between sent messages and their true recipients (receiver anonymity), and between received messages and their true senders (sender anonymity).

Benefits of anonymous communication[2]

- Investigative journalism
- Whistleblowing
- Law enforcement
- Self-help
- Personal privacy protection
- Avoiding persecution

[2] R. Kling, Y-C. Lee, and A. Teich, *Assessing Anonymous Communication on the Internet: Policy Deliberations*, Journal of Information Society, 1999

Harms of anonymous communication[2]

- Spamming
- Deception
- Hate mail
- Impersonation and misrepresentation
- Online financial fraud
- Other illegal activities

Approaches to anonymous communication

There are two main approach to anonymous communication:

- DC-nets
- Mix networks

Dining cryptographers problem[3]

Three cryptographers are sitting down to dinner at their favorite three-star restaurant. The waiter informs them that the meal has been paid anonymously by one of the cryptographers *or* the National Security Agency. The cryptographers respect each other's right to make an anonymous payment, but they wonder if the NSA paid.

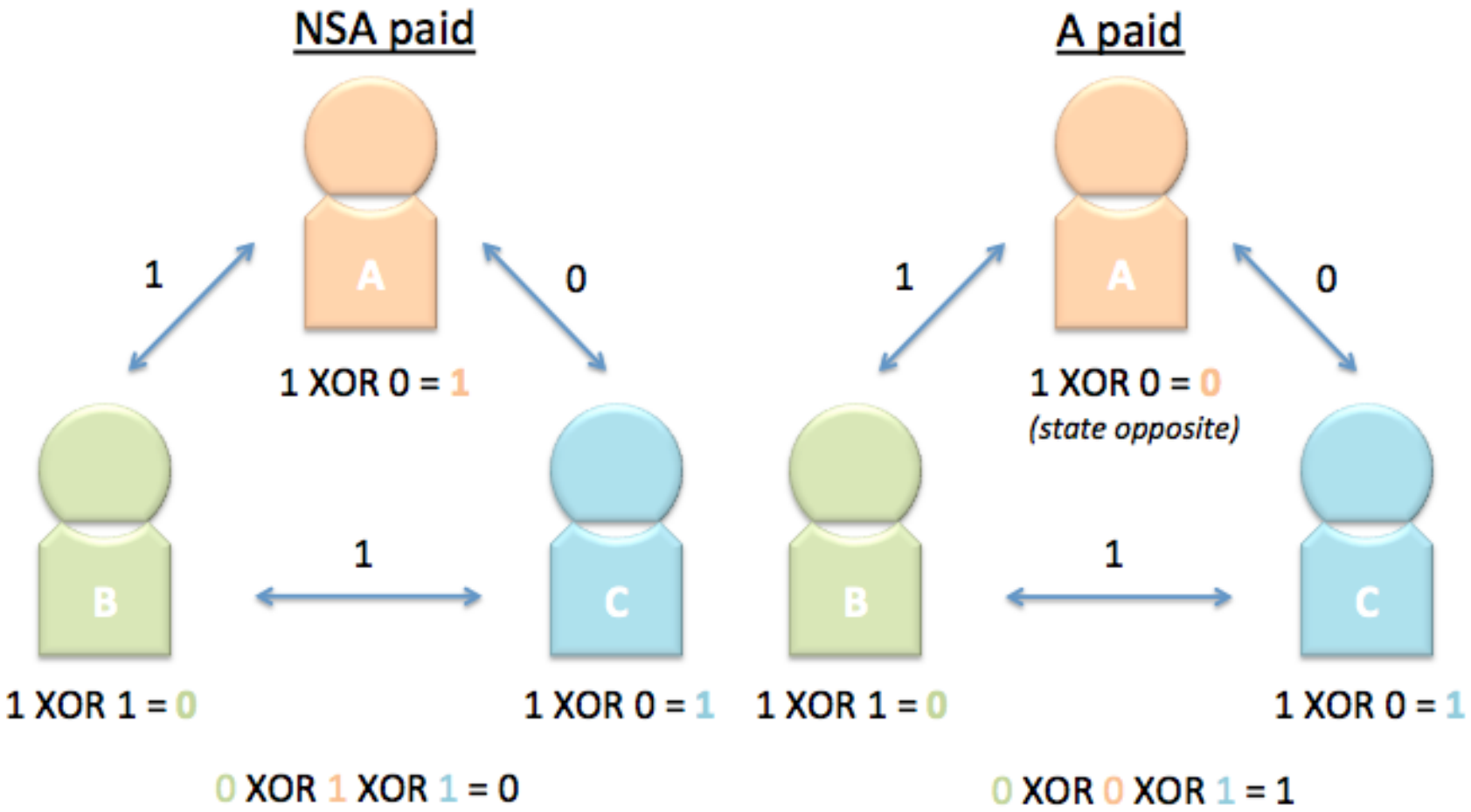
They find out by executing a two-stage protocol.

[3] D. Chaum, *Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms*, Communications of the ACM, 1981

DC-nets

1. Establish a secret.
   - Each cryptographer secretly flips an unbiased coin with the cryptographer on his right.
2. Reveal the secret.
   - Each cryptographer says whether the two coins he sees fell on the same side or on different sides.
   - If one of the cryptographers is the payer, he states the opposite of what he sees.

If the result is 0, then NSA paid. If it is 1, then one of the cryptographer paid. However, cryptographers who did not pay do not know which one did.


NSA paid
A
1
0
1 XOR 0 = 1
B
C
1
1 XOR 1 = 0
1 XOR 0 = 1
0 XOR 1 XOR 1 = 0
A paid
A
1
0
1 XOR 0 = 0
(state opposite)
B
C
1
1 XOR 1 = 0
1 XOR 0 = 1
0 XOR 0 XOR 1 = 1

DC-nets

The protocol is simple, elegant and unconditionally secure (assuming a secure channel) if carried out faithfully.

However, there are two major issues:

- Collisions – if two cryptographers paid for the dinner, their messages will cancel each other out. It means that only one participant can transmit at the same time.
- Disruptions - the last cryptographer can change the final result.

Herbivore[4]

Herbivore is a distributed anonymous communication system, providing private file sharing and messaging over the Internet.

It lets people anonymously publish and retrieve documents, and guarantees that even the most resourceful adversary cannot compromise this anonymity.

Built to be self-organizing, Herbivore relies on neither central servers nor trusted parties, and ultimately provides anonymity by drawing on its community of users.

[4] http://www.cs.cornell.edu/people/egs/herbivore/

Dissent[6]

DISSENT stands for Dining cryptographers Shuffled Send Network. It provides accountability by allowing to discover misbehaving members.

DISSENT is mostly suitable for latency-tolerant applications because its scalability is limited.[5]

[5] The version presented here. There are other scalable versions of DISSENT.

[6] H. Corrigan-Gibbs and B. Ford, Dissent: Accountable Group Anonymity, CCS 2010

Security Goals

Anonymity - the adversary cannot guess the sources of the messages from honest users with probability significantly greater than random guessing.

Integrity - every honest member for whom the protocol completes successfully has the same output and receives the messages of all the other honest members.

Accountability - every honest member for whom the protocol failed obtains proof of some member's misbehavior, and the adversary cannot produce a valid proof of misbehavior by an honest member.

Mix-networks

A *mix* is a process that accepts encrypted messages as input, groups several messages together into a batch, and then decrypts and forwards some or all of the messages in the batch.

A mix network consists of mix routers. Mix routers forward the message along the specified path. Some mix routers may intentionally delay sending messages, send them in batches of specified size, etc.

The goal is to obscure the associations between incoming and outgoing messages.

Onion routing[7]

Onion routing is based on the idea of mix networks.

A message is iteratively wrapped with layers of encryption to form "an onion" that specifies properties of the connection along the route.

Each onion router along the route uses its public key to decrypt the entire onion. It exposes the identity of the next onion router, cryptographic information used, and the embedded onion. The router pads the onions and forwards to the next hop.

The last node in the chain removes the remaining layer of encryption and forwards the message to its destination.

[7] `http://www.onion-router.net/`

Tor (The Onion Router)[8]

Tor is a low latency anonymity network based on onion routing. It consists of a number of relays constantly encrypting and then randomly bouncing messages.

The random path packets take is extended one hop at a time, and each relay along the way knows only the previous and the next relay. A separate set of encryption keys is negotiated for each hop along the circuit.

Once a circuit has been established, it is used to exchanged different kinds of data. For efficiency, Tor uses the same circuit for connections that happen within the same ten minutes or so.

[8] The Tor Project `https://www.torproject.org/`

Tor

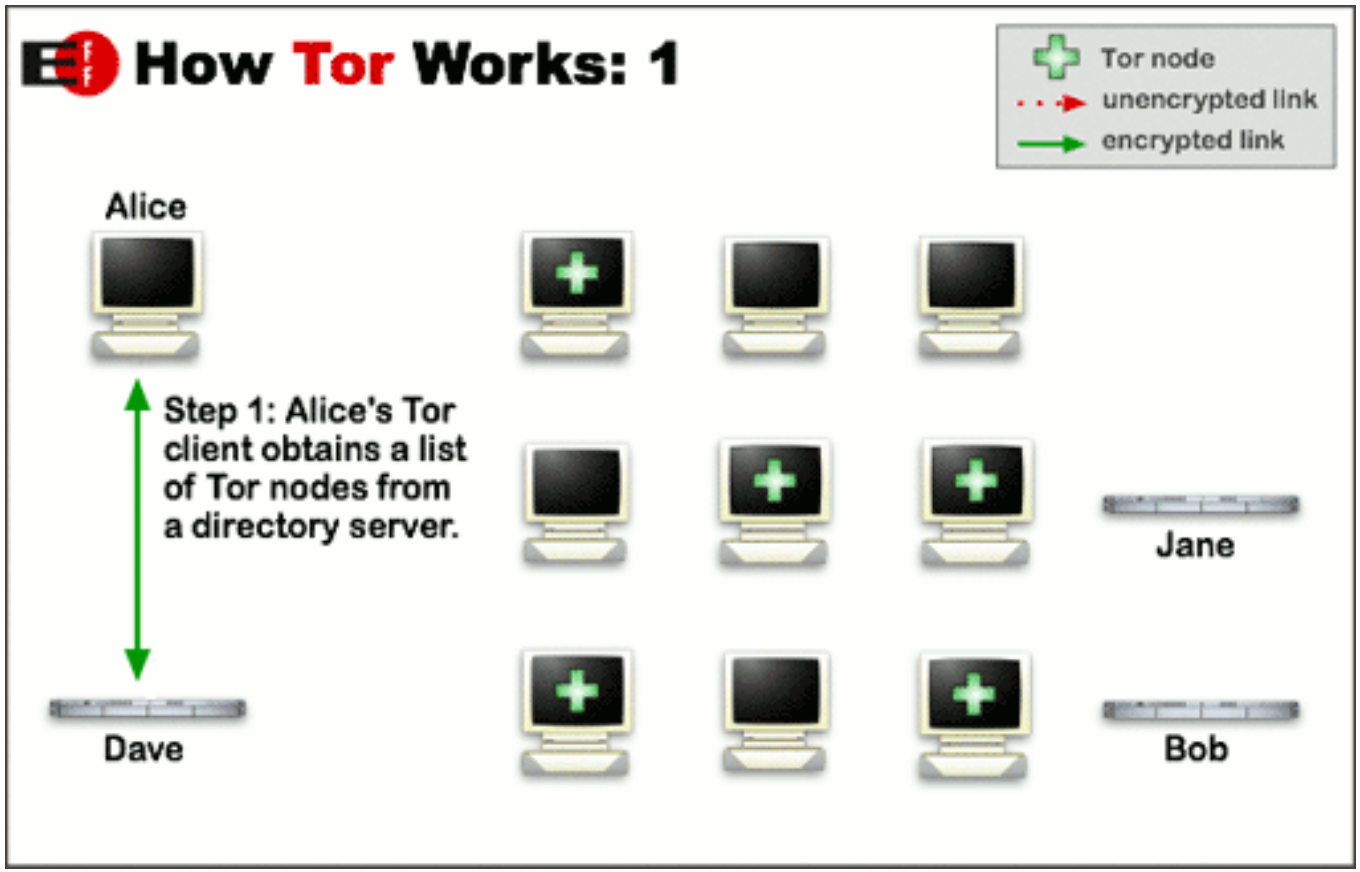
The goal is to protect users' from network surveillance.

No individual relay ever knows the complete path.

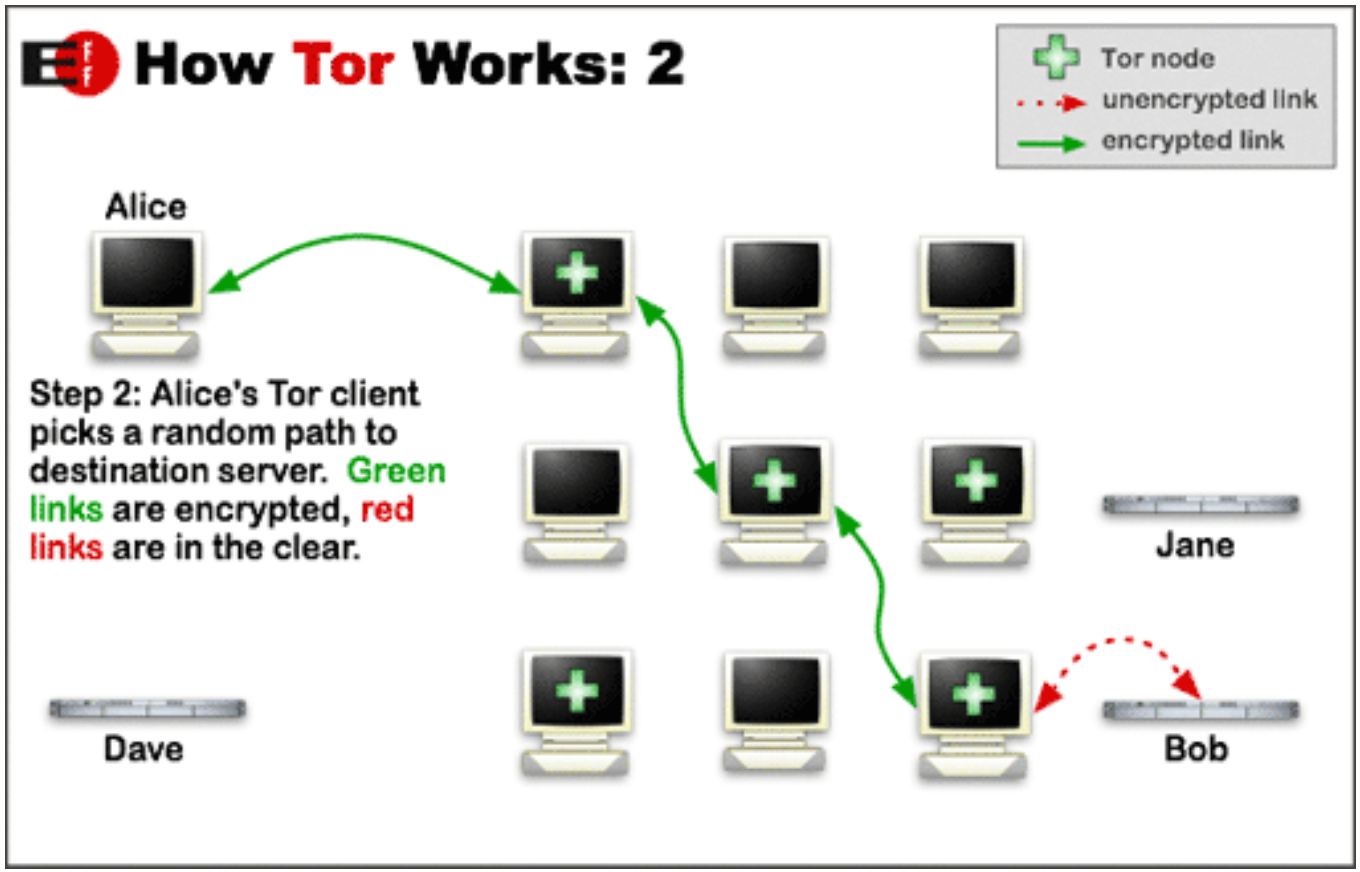
Neither an eavesdropper nor a compromised relay can link the connection's source and destination.

However, Tor cannot (and does not attempt to) protect the traffic entering and exiting the network.

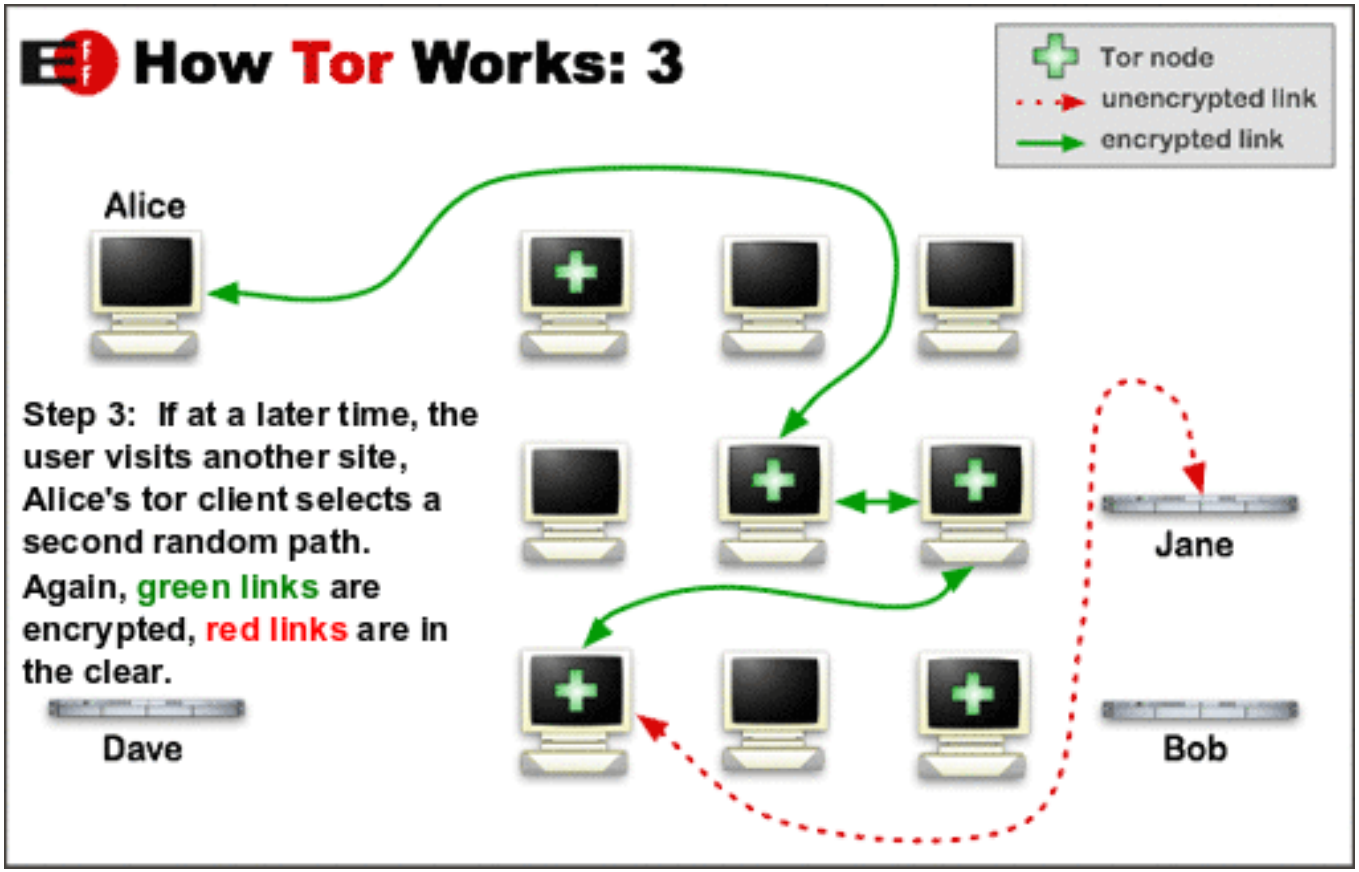
How Tor works?



How Tor works?



How Tor works?



Attacks on Anonymity

Limitations of existing schemes

There's a whole array of attacks on anonymity ranging from exploiting flaws of the anonymous system, leveraging network level information, to exploiting vulnerabilities of the client-side machines.

Method	Weakness
Mix Nets, Tor	Traffic analysis attacks
Group and Ring Signatures	Traffic analysis attacks
Voting Protocols	Short, fixed-length messages
DC Nets	Anonymous DoS attacks

Understanding the Adversary

Many attacks assume a powerful adversary who can see large chunks of the network at the same time or can obtain large amounts of traffic data.

Is an existence of a global adversary a reasonable assumption?

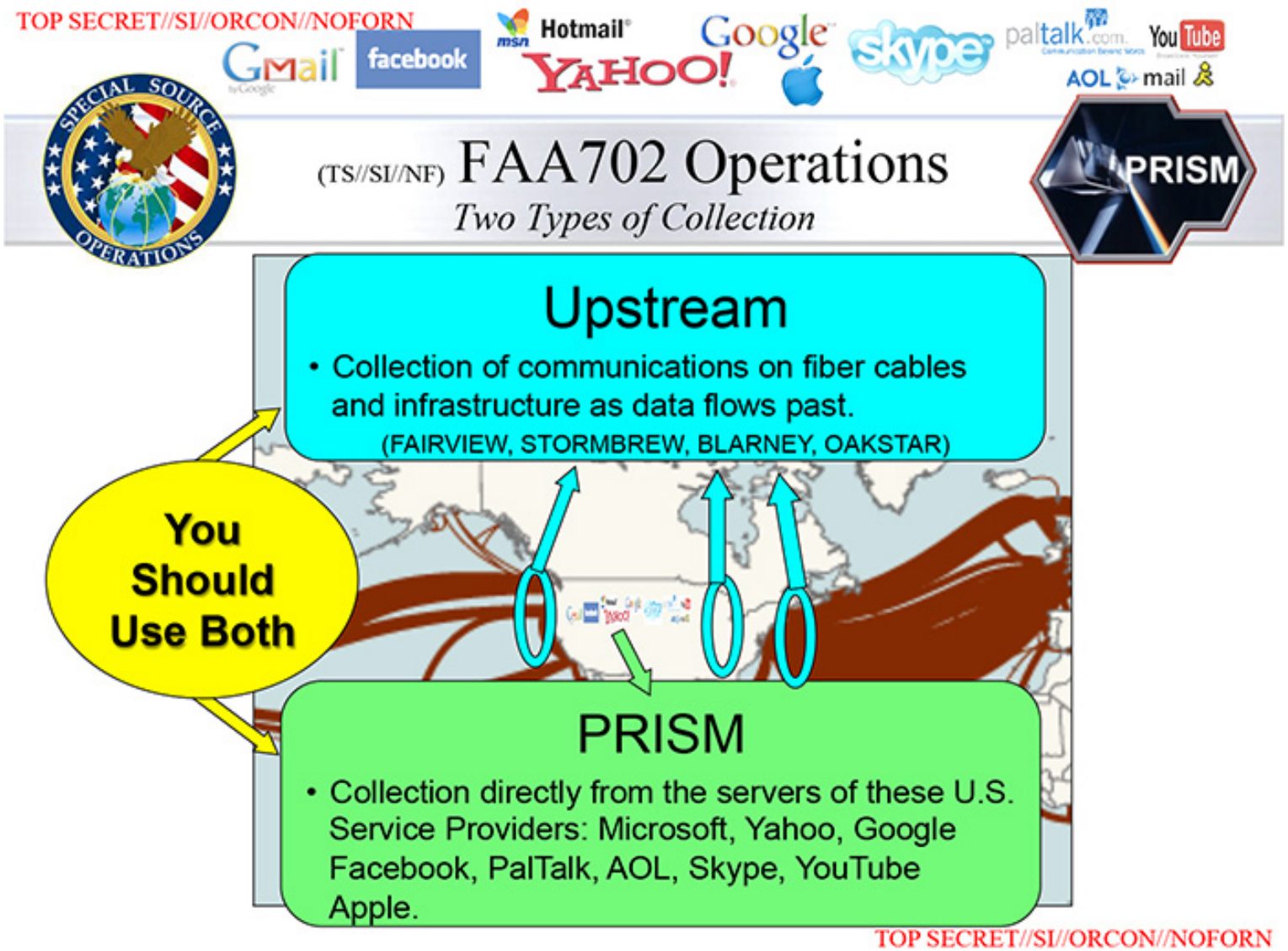
PRISM[9]

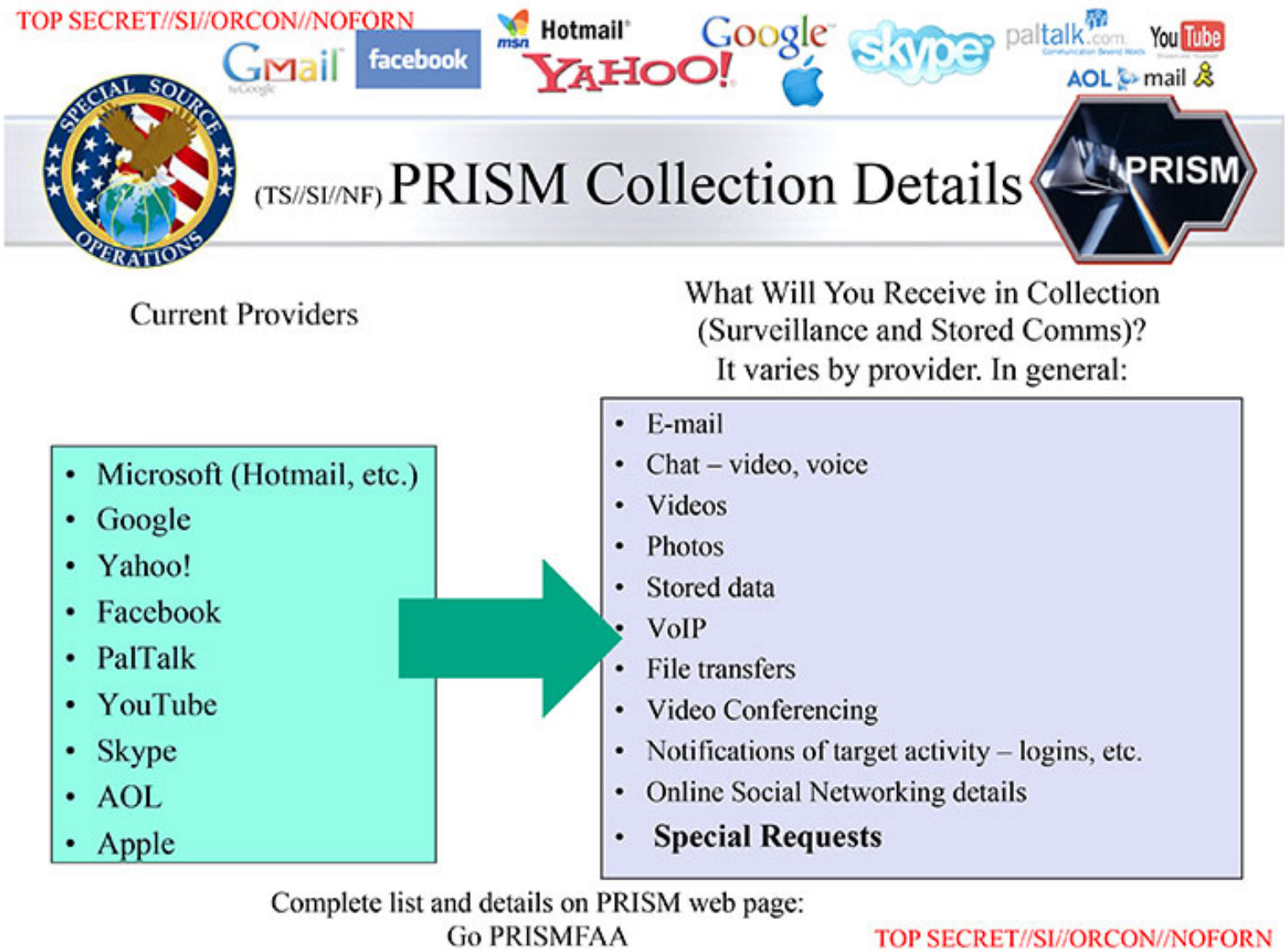
Perhaps the most infamous but not the most comprehensive NSA mass surveillance program established in 2007.

Geared towards collecting the Internet data stored by nine major companies: Facebook, Google, Yahoo, Microsoft, PalTalk, Skype, YouTube, Apple, and AOL.

The program operates through a secretive judiciary body called the Foreign Intelligence Surveillance Court.

[9] *The definitive guide to NSA spy programs*, Joe Kloc, August 14, 2013

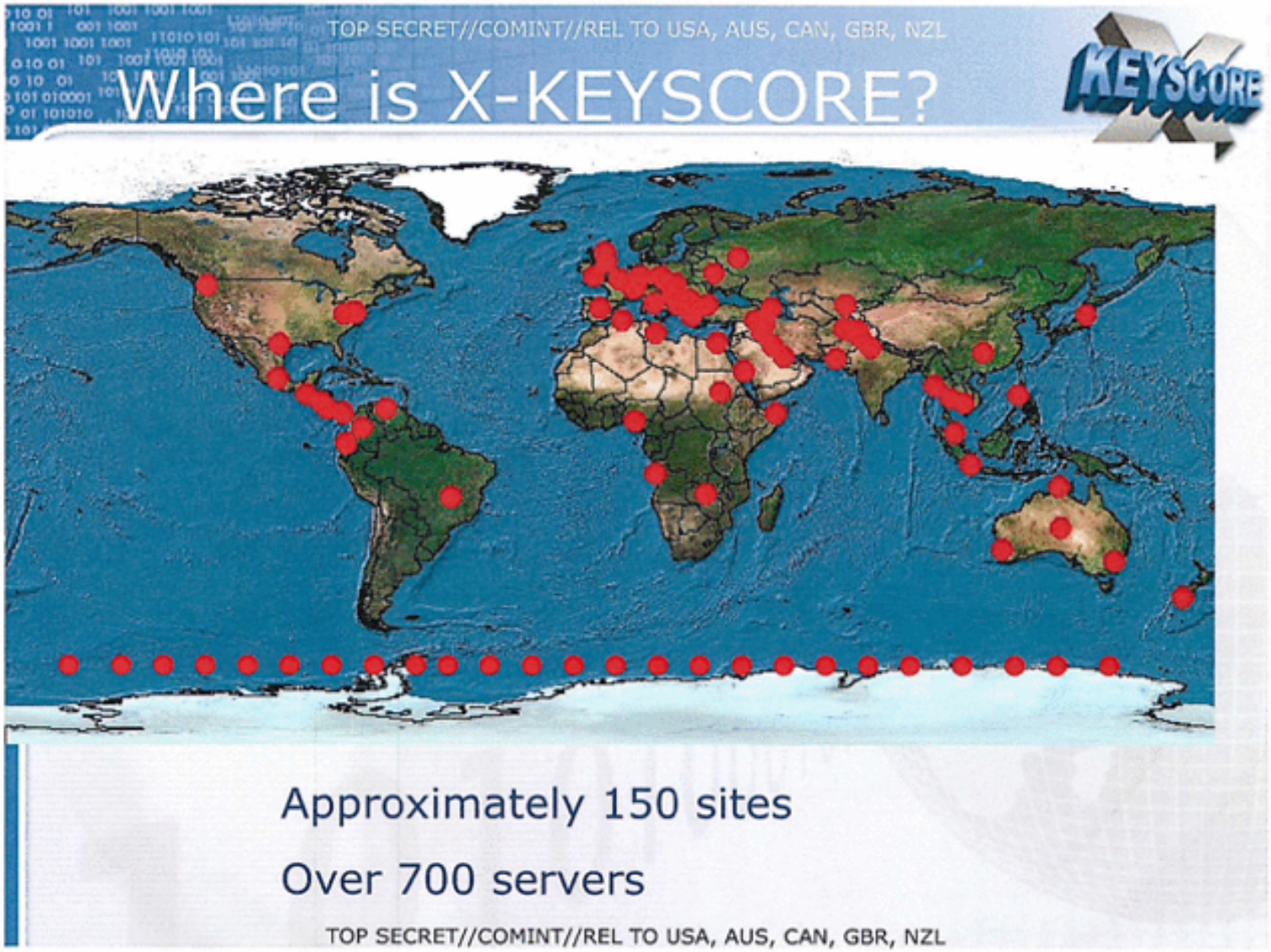

TOP SECRET//SI//ORCON//NOFORN
Gmail
facebook
Hotmail
msn
YAHOO!
Google
skype
paltalk.com
YouTube
AOL mail
SPECIAL SOURCE OPERATIONS
(TS//SI//NF) FAA702 Operations
Two Types of Collection
PRISM
Upstream
• Collection of communications on fiber cables and infrastructure as data flows past.
(FAIRVIEW, STORMBREW, BLARNEY, OAKSTAR)
You Should Use Both
PRISM
• Collection directly from the servers of these U.S. Service Providers: Microsoft, Yahoo, Google Facebook, PalTalk, AOL, Skype, YouTube Apple.
TOP SECRET//SI//ORCON//NOFORN

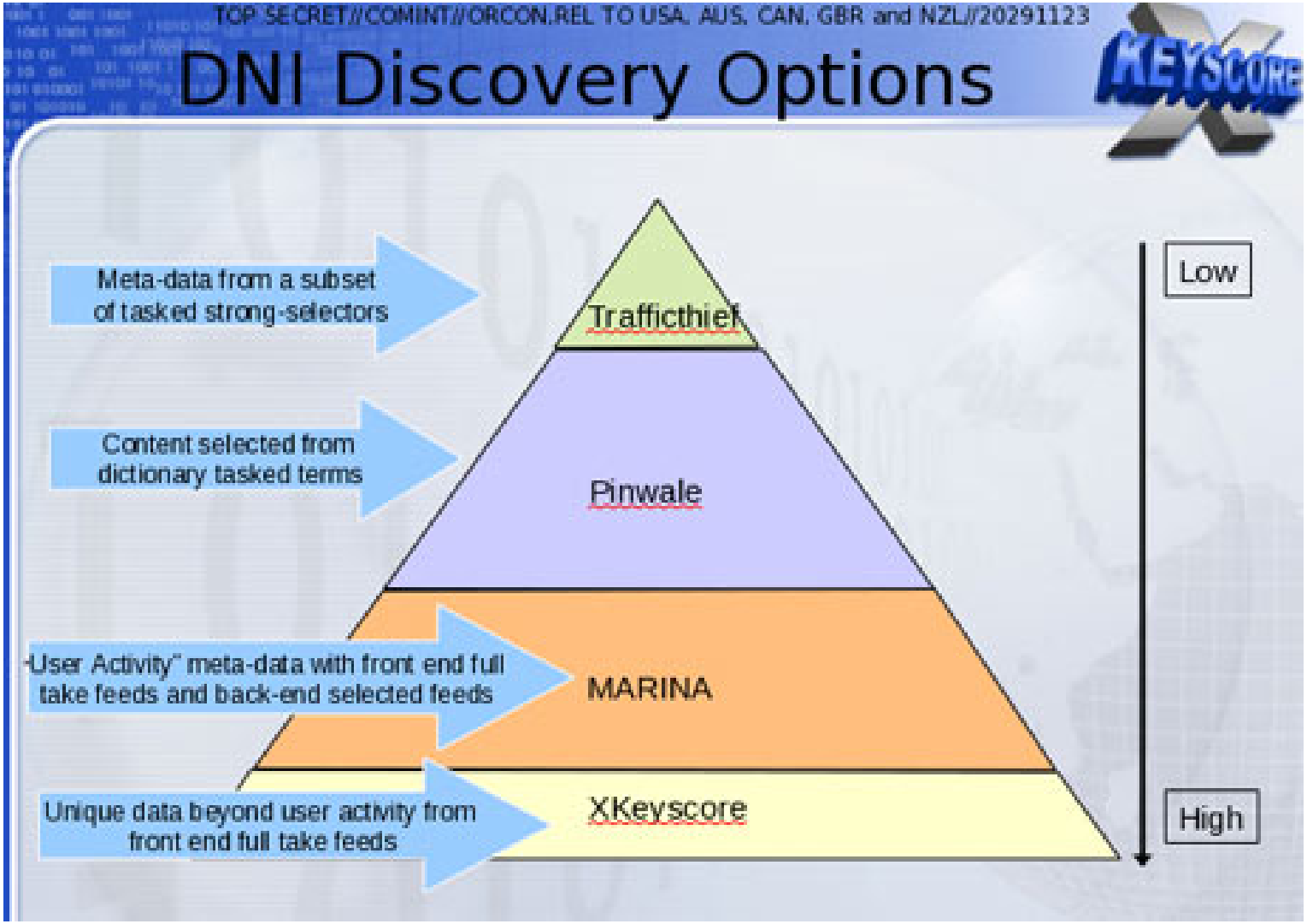

TOP SECRET//SI//ORCON//NOFORN
Gmail
facebook
Hotmail
YAHOO!
Google
skype
paltalk.com
YouTube
AOL mail
SPECIAL SOURCE OPERATIONS
(TS//SI//NF) PRISM Collection Details
PRISM
Current Providers
What Will You Receive in Collection (Surveillance and Stored Comms)? It varies by provider. In general:
• Microsoft (Hotmail, etc.)
• Google
• Yahoo!
• Facebook
• PalTalk
• YouTube
• Skype
• AOL
• Apple
• E-mail
• Chat – video, voice
• Videos
• Photos
• Stored data
• VoIP
• File transfers
• Video Conferencing
• Notifications of target activity – logins, etc.
• Online Social Networking details
• Special Requests
Complete list and details on PRISM web page:
Go PRISMFAA
TOP SECRET//SI//ORCON//NOFORN

XKEYSCORE[9]

NSA's widest-reaching program. It collects nearly everything a typical user does on the Internet, including the content of emails and chats, visible in real time.

It builds a searchable database of both metadata and communications content collected from around the world.


TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL
Where is X-KEYSCORE?
KEYSCORE
Approximately 150 sites
Over 700 servers
TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL


TOP SECRET//COMINT//ORCON,REL TO USA, AUS, CAN, GBR and NZL//20291123
DNI Discovery Options
XKEYSCORE
Meta-data from a subset of tasked strong-selectors
Content selected from dictionary tasked terms
"User Activity" meta-data with front end full take feeds and back-end selected feeds
Unique data beyond user activity from front end full take feeds
Trafficthief
Pinwale
MARINA
XKeyscore
Low
High

Practical Evidence of Traffic Hijacking[11] [12]

Researchers from network intelligence firm Renesys observed numerous live Man-In-the-Middle (MITM) hijacks so far this year in events lasting from minutes to days, by attackers working from various countries.

Traffic has been improperly redirected to routers at Belarusian or Icelandic service providers. The hacks exploit implicit trust placed in BGP (border gateway protocol).[10]

Huge chunks of Internet traffic belonging to financial institutions, government agencies, and network service providers have repeatedly been diverted under unexplained circumstances.

---

[10] Protocol to exchange routing and reachability information between autonomous systems (AS)

[11] *The New Threat: Targeted Internet Traffic Misdirection*, Jim Cowie, Renesys, November 19, 2013

[12] *Repeated attacks hijack huge chunks of Internet traffic, researchers warn*, Dan Goodin, November 20, 2013


Traceroute Path 1: from Guadalajara, Mexico to Washington, D.C. via Belarus
7. Moscow, Russia
8. Minsk, Belarus
6. London, UK
9. Frankfurt, Germany
4. Ashburn, VA
10. New York, NY
5. Washington, D.C.
11. Washington, D.C. END
3. Laredo, TX
2. Monterrey, Mexico
START 1. Guadalajara, Mexico
renesys
Source: Renesys Path Measurements

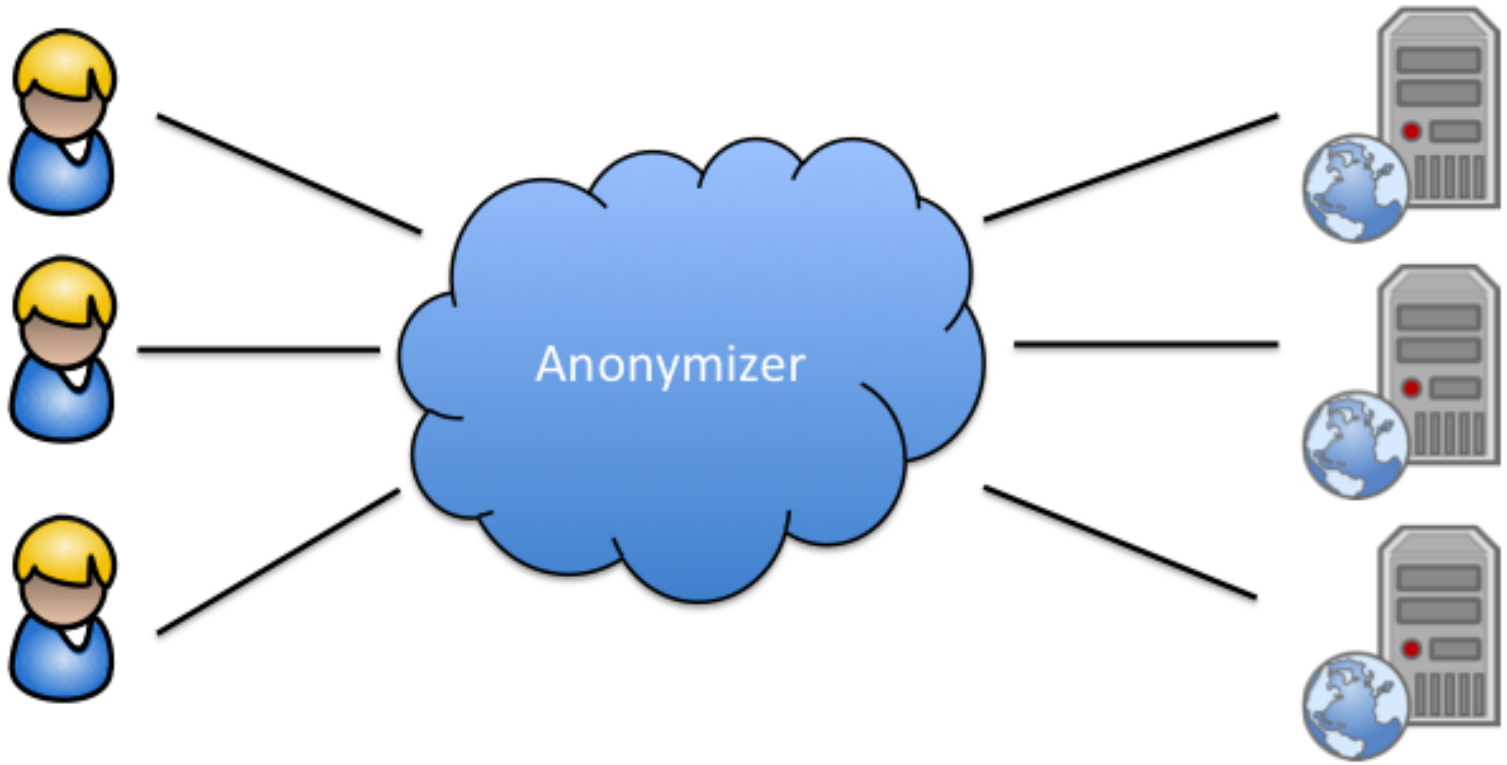

Traceroute Path 2: from Denver, CO to Denver, CO via *Iceland*
6. Reykjavik, Iceland
5. London, UK
7. Montreal, Canada
2. Chicago, IL
8. Chicago, IL
START/END
1. Denver, CO
12. Denver, CO
4. New York, NY
9. New York, NY
11. Kansas City, MO
3. Ashburn, VA
10. Dallas, TX
renesys
Source: *Renesys Path Measurements*

Traffic Analysis Attack

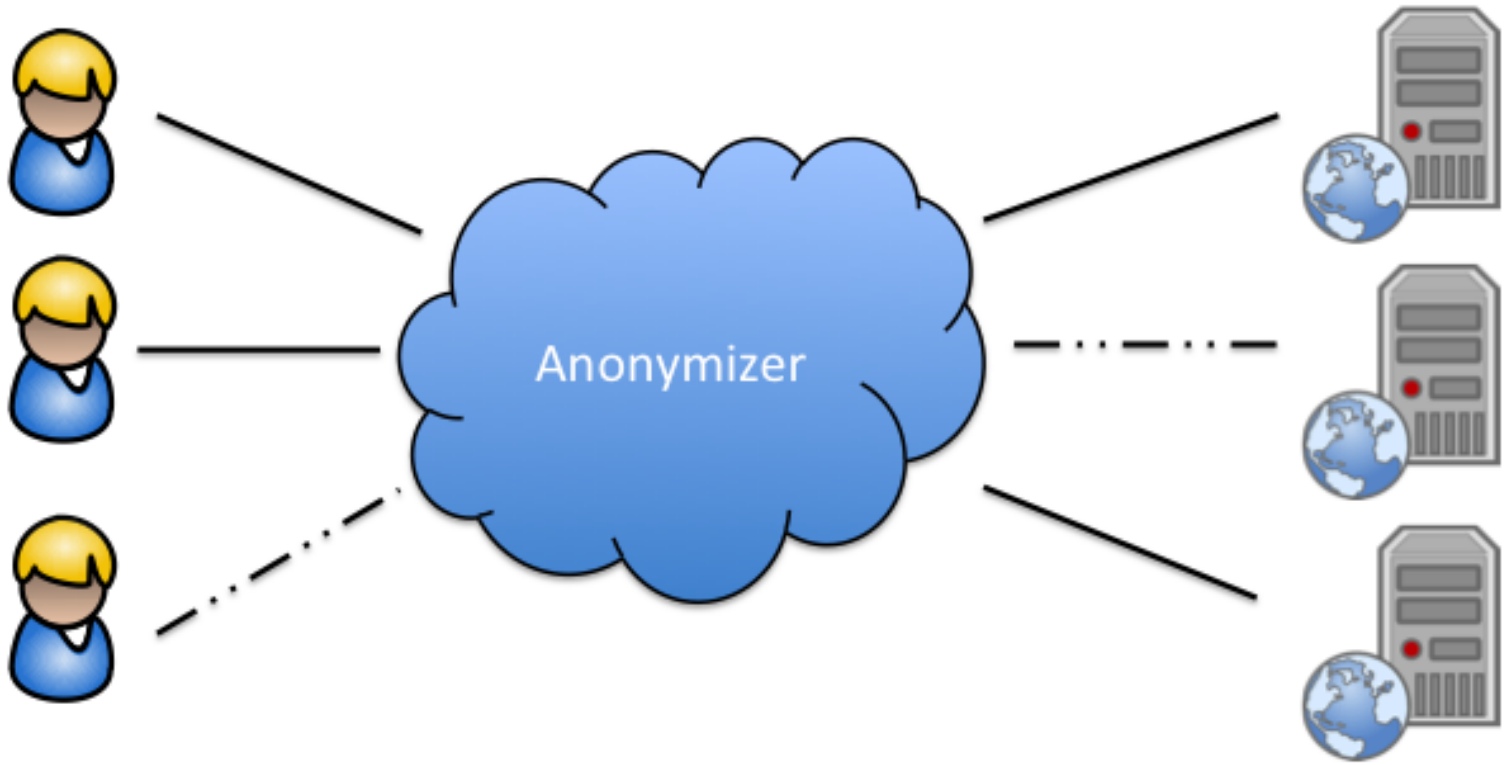
Traffic analysis is the process of analyzing patterns in communication in order to learn some information, frequently the sender's identity.

An adversary can be passive and simply monitor the frequency and timing of communication, or be active and affect the communication by performing DoS attacks on chosen nodes, overloading the network on certain links, etc.

Traffic Analysis



Traffic Analysis



Aqua[13]

Alternative high-bandwidth anonymity system that resists traffic analysis.

Architecture similar to Tor: the core consists of Aqua servers that clients connect to.

Traffic entering and exiting an Aqua network is made indistinguishable through a combination of chaffing and delayed flow start up. This results in a payload-independent, uniform traffic rate between Aqua servers as well as clients entering and exiting the Aqua network.

In addition, Aqua makes traffic from a set of k-clients look indistinguishable and consequently the clients themselves.

[13] *Towards Efficient Traffic-analysis Resistant Anonymity Networks* S. Le Blond et al., ACM SIGCOMM 2013

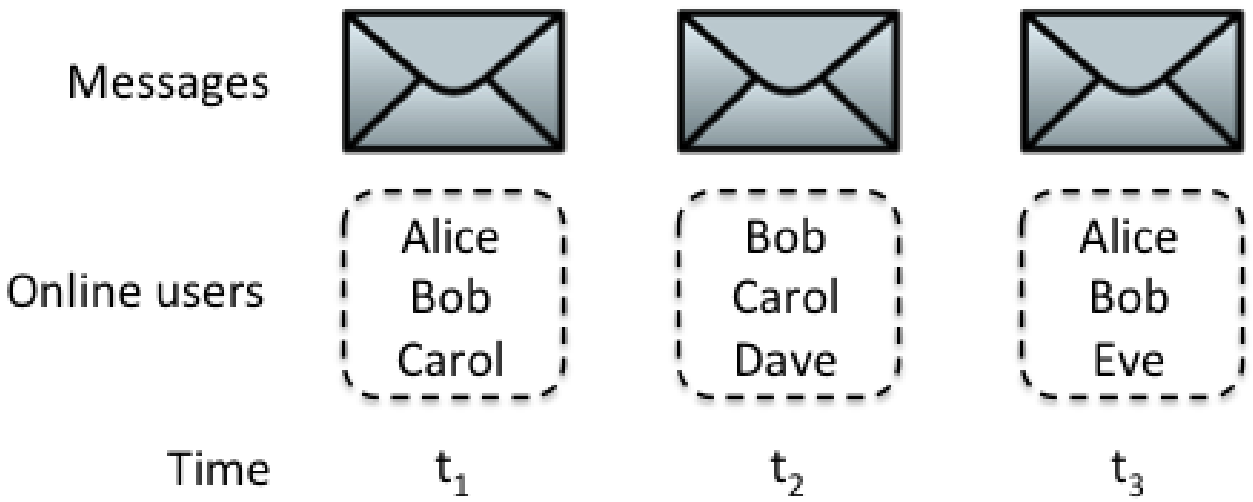
Intersection Attack

An intersection attack is a type of a traffic analysis attack that focuses on the relation between online availability of users and certain actions.

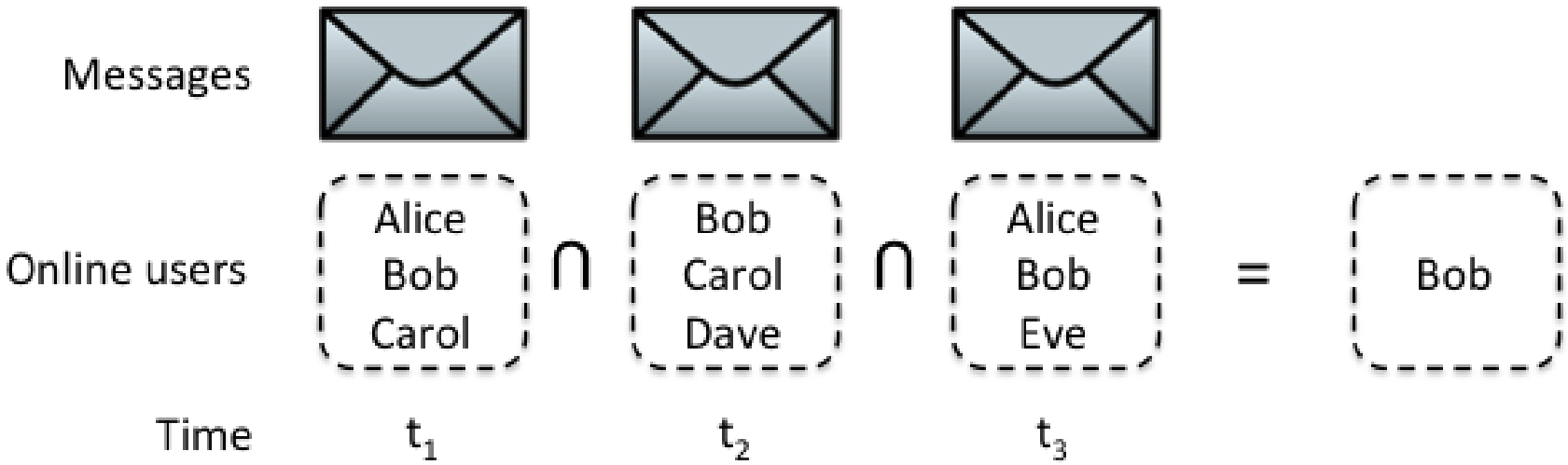
Assume that a series of anonymous blog posts consistently appears every Monday around 10am. Then, the set of users who have been online *each* time a post appears contains the blog owner, as opposed to the much biggest set of users who have been online when some posts appeared.

To do so, an adversary can take snapshots of users available during specific times and *intersect* them.

Intersection Attack



Intersection Attack



Petraeus-Broadwell Debacle[14]

Jill Kelly complains about receiving a series of anonymous stalking emails which seem to involve Gen. Petraeus

Mrs. Broadwell was identified as the sender of the emails by comparing times when emails were sent, locations where the emails were sent from, and the lists of all people who were at those locations

The investigation reveals an extramarital affair between Gen. Petraeus and Mrs. Broadwell

[14] *Online PrivacySurveillance and Security Lessons From the Petraeus Scandal*, C. Soghoian, November 13, 2012

Buddies[15]

Continuously simulate a global adversary to ensure minimum levels of anonymity.

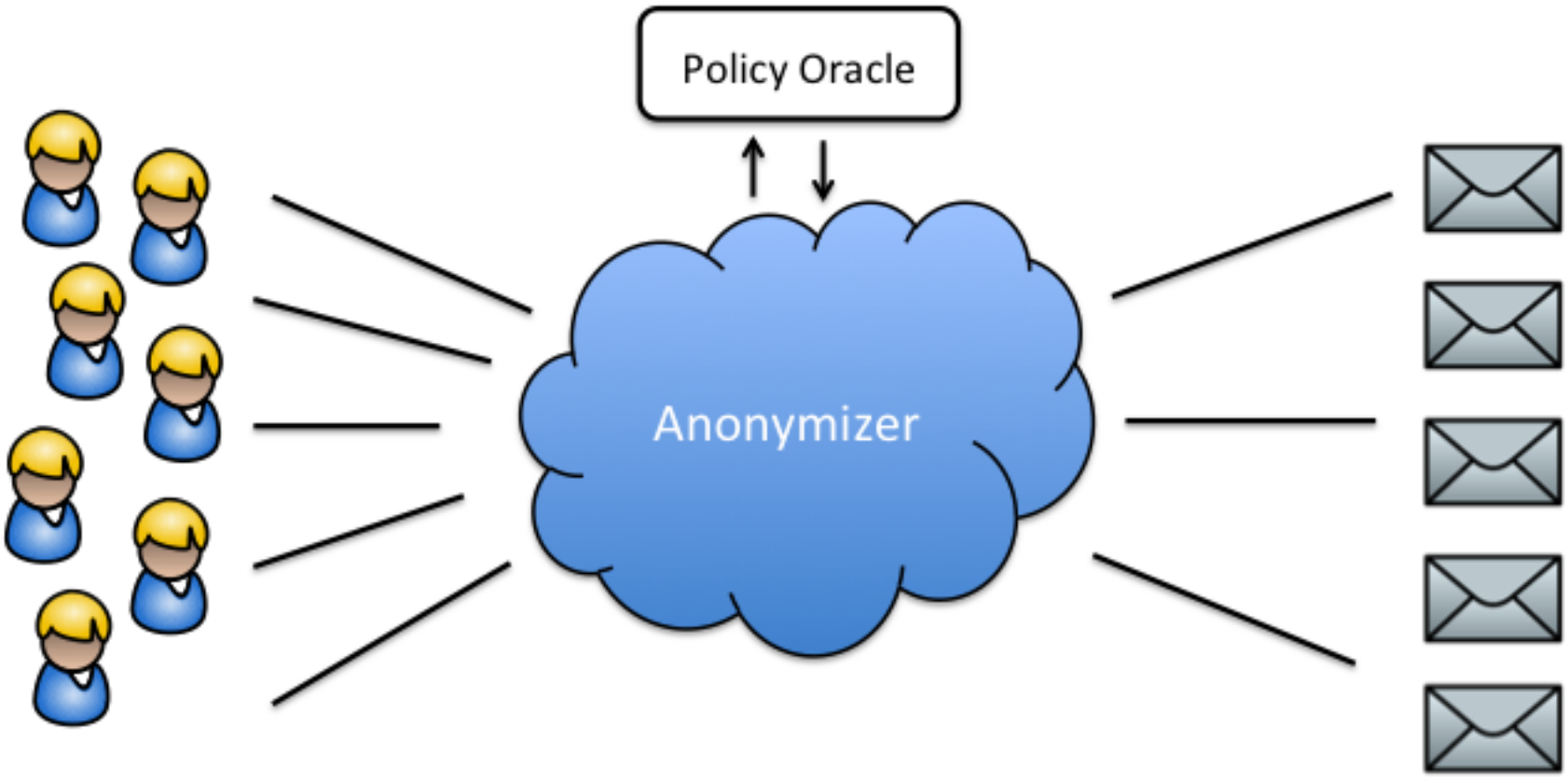
Allow to post messages only when safe ($|$ anonymity set $| > k$).

Two metrics to measure anonymity:

1. Possinymity - plausible deniability
2. Indinymity - indistinguishability within a buddy set

[15] *Hang With Your Buddies to Resist Intersection Attacks*, D. Wolinsky, E. Syta, and B. Ford, , CCS 2013

Buddies



Fingerprinting and Staining Attacks

Fingerprinting attack - an attack that uses specific hardware and/or software specifications of a client's machine to create a unique label (a fingerprint), than can be later used to identify that client.

Staining attack - an attack that relies on the ability to insert an identifiable "stain" (e.g., a cookie) into a client's machine, in order to later on use it to identify the client.

NSA about Tor[16] [17]

"Tor stinks."

"Very Secure."

"Still the King of high secure, low latency Internet Anonymity. There are no contenders for the throne in waiting."

"We will never be able to de-anonymize all Tor users all the time. With manual analysis we can de-anonymize a very small fraction of Tor users, however, no success de-anonymizing a user in response to a TOPI request/on demand."

[16] *Tor: 'The king of high-secure, low-latency anonymity'*, The Guardian. October 4, 2013

[17] *'Peeling back the layers of Tor with EgotisticalGiraffe'*, The Guardian. October 4, 2013

Peeling back the layers of Tor with EGOTISTICALGIRAFFE[17]

Based on what we know, NSA cannot de-anonymize users by breaking the core Tor protocol.

Instead, they target software vulnerabilities in Firefox that the Tor Browser Bundle is built upon.

NSA actively searches for and exploits vulnerabilities "around" Tor.

QUANTUM[18]

The NSA's program to insert packets from the Internet backbone.

Quantum servers are at key places on the Internet backbone ensuring that they can react faster than other servers. This permits "man-on-the-side" attacks that inject arbitrary packets as a fake response to a legitimate request.

The leaked documents included a top-secret NSA diagram that shows a Quantum server impersonating Google.

This same technique is used by the Chinese government to block its citizens from censored Internet content.

[18] *How the NSA Attacks Tor/Firefox Users With QUANTUM and FOXACID*, Bruce Schneier, October 7, 2013

FOXACID[18]

NSA codename for an Internet-enabled system capable of attacking target computers in a variety of ways.

FoxAcid servers are publicly accessible using normal-looking domain names. When a browser visits a FoxAcid server using a special url, called a FoxAcid tag, the server attempts to infect that browser, and then the computer, in an effort to take control of it.

Each target receives its own FoxAcid tag which permits to identify them and use customized exploits.

When FoxAcid servers handle callbacks from infected machines, they are called FRUGALSHOT. A machine may call back to the NSA for more instructions, to upload data from the target computer, etc.

NSA's 6 Step Attack on Tor Users[19]

1. Scan Internet traffic.
   Use "upstream" data collection programs such as Stormbrew, Fairview, Oakstar, and Blarney to tap into the fiberoptic backbone of the Internet.
2. Mark Tor requests using fingerprinting techniques.
   Use XKeyscore to do so.
3. Sift out marked traffic.
   All Tor users look alike so it is easy to tell them apart from non-Tor users.

[19] *How the NSA identifies Tor users in 6 easy steps* , Joe Kloc, October 08, 2013

NSA's 6 Step Attack on Tor Users

4. Send users to NSA servers.
   Use Quantum to re-direct targets to FoxAcid servers which pretend to be the legitimate server that the Tor user is trying to access.
5. Attack users' computers.
   Use FoxAcid servers to deliver exploits.
6. Identify Tor users.
   After obtaining access to the target computer, it is easy to identify the user based on email accounts accessed from the same computer, stored information, etc.

Freedom Hosting[20] & Silk Road

Freedom Hosting, a provider of anonymous hosting, operated as a Tor hidden service. In August, all sites hosted by Freedom Hosting began serving an error message with hidden code embedded into the page. It turned out to be an exploit leveraging a security hole in Firefox to identify Tor users by reporting back the user's IP to a server in Northern Virginia. Later, FBI admitted it was behind this mass malware attack.

Silk Road, a black market website was also operated as a Tor hidden service. FBI shut it down in October with the help from other agencies.

[20] *FBI Admits It Controlled Tor Servers Behind Mass Malware Attack*, K. Poulsen, September 13, 2013

Additional Resources

- The Tor Project, https://www.torproject.org/
- How Tor works? by Artist Molly Crabapple and Writer John Leavitt https://www.eff.org/whatistor
- The definitive guide to NSA spy programs

  http://www.dailydot.com/politics/nsa-spy-prgrams-prism-fairview-blarney/