Anonymity on the Internet

Privacy and Identity

Bart Jacobs and Gergely Alpár
Institute for Computing and Information Sciences – Digital Security
Radboud University
December 15, 2015



Outline

Introduction
- Anonymous communication on the internet
- Mix networks
- Onion routing

Tor

Attacks on Tor
- Correlation
- Further threats

iCIS | Digital Security
Radboud University



Currently we are here...

Introduction
- Anonymous communication on the internet
- Mix networks
- Onion routing

Tor

Attacks on Tor

Why do we need anonymous communication?

- It is required to a free society
  - Privacy and free participation in the society are human rights
  - If communication cannot be anonymous, these rights may be impaired
  - Thus, it is especially important in countries where human rights are weakened
- It used to be possible to talk to someone in private, and we want the same also on the Internet
- Hence, the goal is that:

> *"Parties are free to identify themselves [...] But the use of a public network should not automatically give away the identities and locations of the communicating parties"*
>
> *(Goldschlag, Reed, Syverson, 1996. [2])*

Anonymous communication online

- Anonymous communication enables potentially other applications, such as voting, anonymous participation
  - Thus, we can say, it's an enabler
- Full anonymity is hard to achieve on the Internet
  - Networking is layered, so anonymity is required at all layers
  - Each layer has different functionality, so different techniques needed
- What are the characteristics (+/−) of anonymising different layers?
  - Application: + IP address can be used for optimisation (e.g. packets can be batched before sending); − application specific
  - Transport: + application neutral; − the IP addresses, segment size and timing are visible
  - Network: + high level of anonymity and it can support any protocol; − it requires big (kernel-level) changes in the client

Security goals for anonymous communication

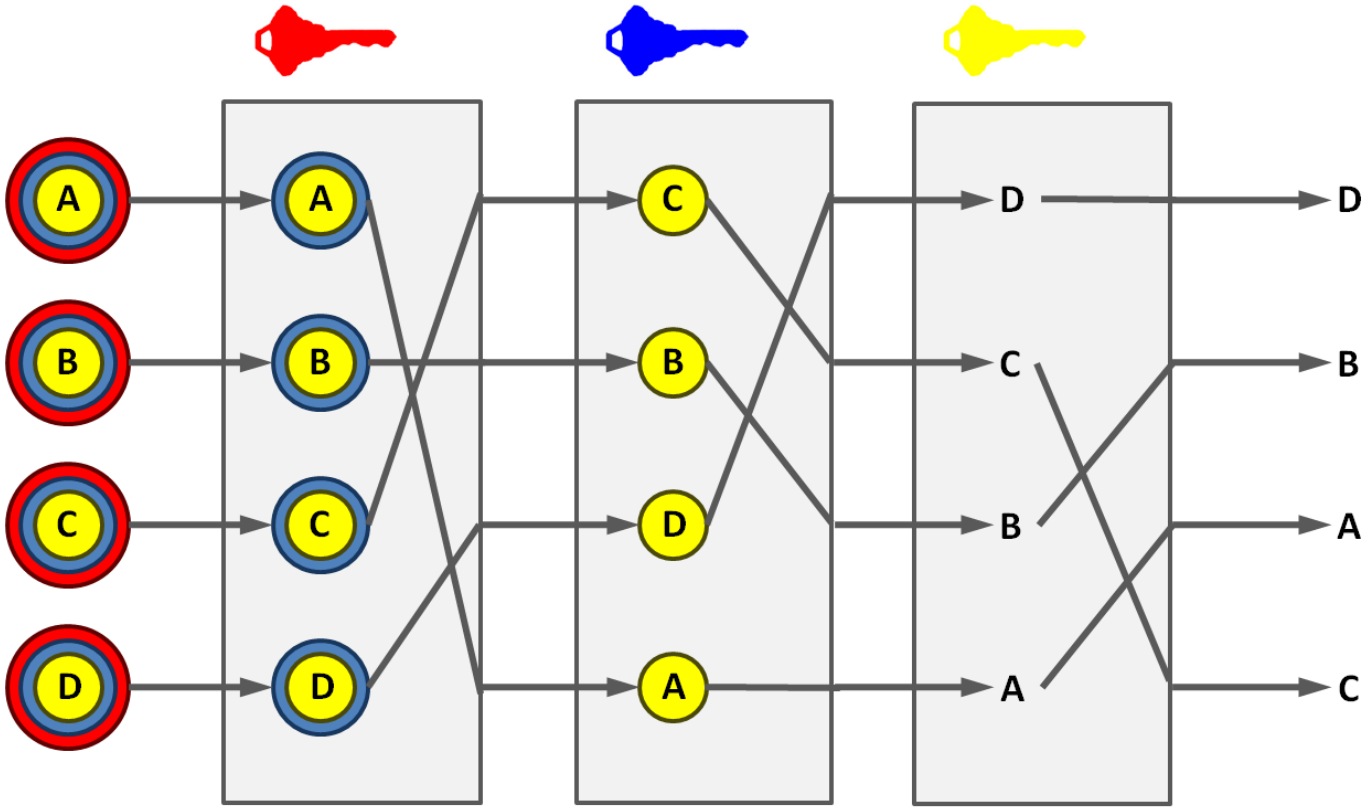
- Anonymity adversary on the internet tries to trace, censor, make associations explicit (who talks to whom?), find out location
  - No confidentiality: the goal is not listening to content; e.g. SSL/TLS

- Anonymity requirements online:
  - Sender anonymity
  - Unlinkability of communication instances, messages from the same user (or to the same user)
  - Unlinkability of communicating parties
  - Untraceability of particular users

Mixnet basics

- Introduced by David Chaum in 1981; see [1]
- Each participant (client, server) is assumed to have a public key
- Mixnet: short for a mixing network
  - users communicating via (one or) several mix nodes
  - a mix node is a trusted server randomly permuting and relaying users' messages
  - a mix node decrypts and drops randomness from each packet, and (after shuffling them) sends the result further; e.g. to send the message to Bob, mix node 3:
    $K_3(R_3, Bob, K_B(R_B, msg)) \rightarrow K_B(R_B, msg)$
- Untraceability of messages in the network
  - The adversary can't correlate the input and the output
  - Reasons: permutation, encryption and randomness

Mixnet diagram



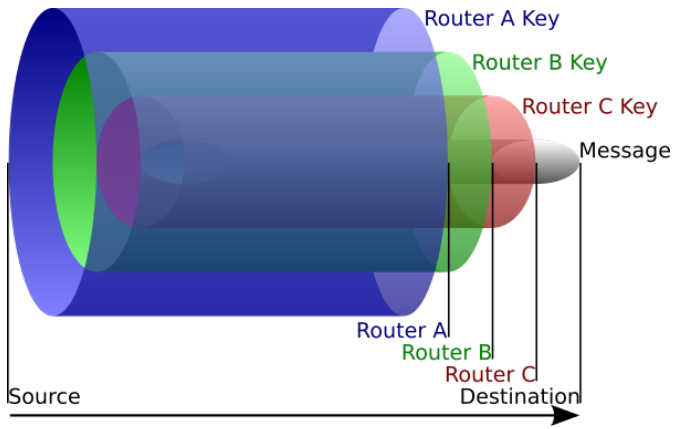
[source: Wikipedia / Mix network]

Mixnet analysis in more detail

- To prevent sender and receiver identification
  - The sender applies layered public-key encryption to each mix nodes (and possibly to the receiver)
- To increase the level of anonymity
  - Nodes decrypt, permute batches and forward to the next node
  - The number of users (anonymity set) should be large: The permutation is done over the messages of many users
- To prevent traffic analysis
  - based on message size, fixed output length (truncating or padding) applied
  - based on timing: random delays applied
- In sum, a mixnet can achieve high-level of anonymity at the cost of large latencies
  - For instance, it may be suitable for e-mail but not for web browsing

Onion routing basics

- Onion routing is introduced in 1996 by Goldschlag, Reed and Syverson; see [2]
- Onion: a message is wrapped in multiple layers of (symmetric-key) encryption
- Onion routers (ORs) relay messages
  - Thus, the adversary is assumed not to be able to observe the whole path!



[source: Wikipedia / Onion routing]

Onion routing in operation

- A user (or a proxy) determines the routing
  - selects randomly a circuit, i.e. a sequence of nodes that will route the traffic
  - establishes a symmetric key with each OR in the route (possibly for several communications)
- An OR, upon receiving an onion, decrypts one of the layers and forwards the message to the next node
  - Thus, an OR knows only the previous and the next nodes
- Typically, three-node onion routing is applied
  - Main example: Tor

Mixnet vs onion routing

	Mixnet	**Onion routing**
Encryption	public-key	symmetric-key
# intermediary nodes	1 (then trusted!)	3
Tasks of node	decrypt, batch, permute, (delay), forward	decrypt, forward
Resistance	end-to-end adversary	local adversary

Currently we are here...

Introduction

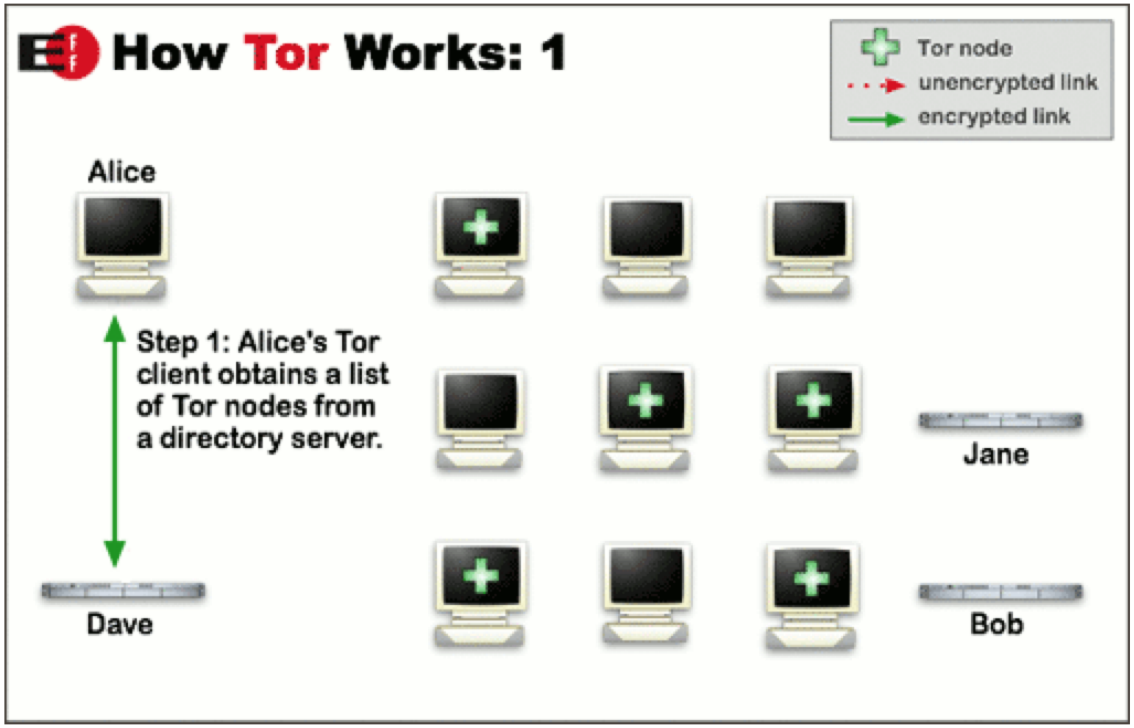
Tor

Attacks on Tor

Tor basics

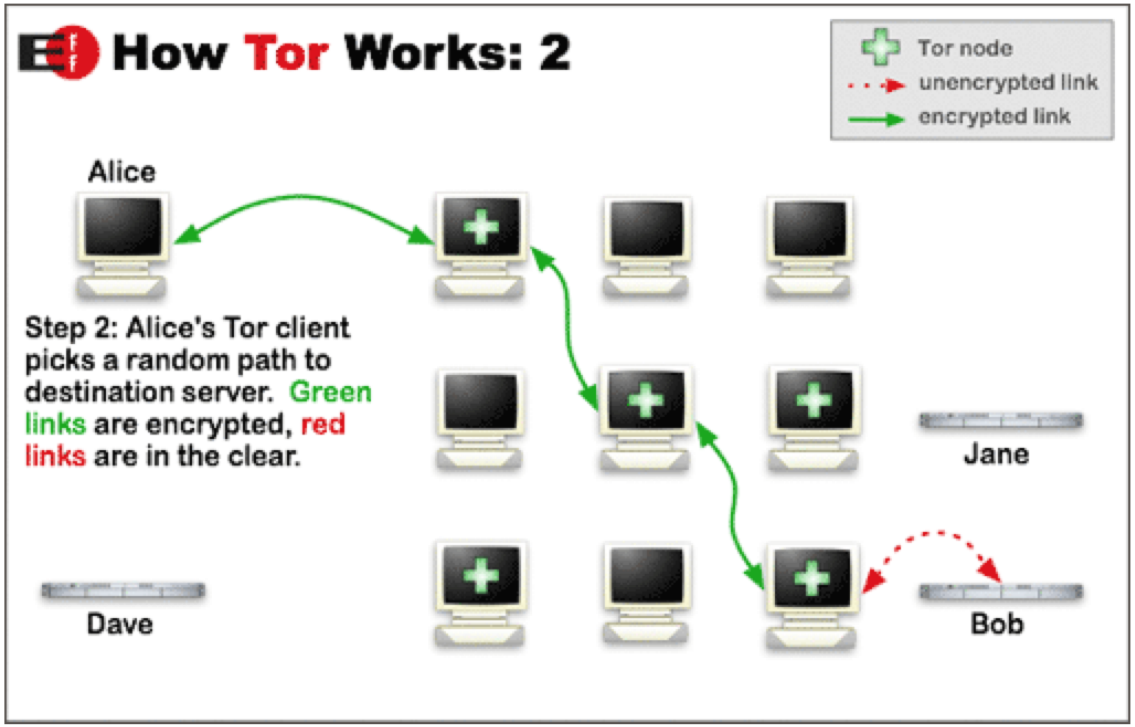
- Onion Routing
  - based on the article "Hiding routing information" from 1996 [2]
  - a project of the US Naval Research Laboratory
  - stopped in 2005
- Tor: The second generation onion routing [4] (torproject.org)
  - transport-layer anonymity network
  - free software
  - goal: low latency, protection against traffic analysis
  - Tor Browser is an additional component that helps protect the user's privacy by correct browser settings
- "Tor aims to solve three privacy problems"
  - "prevents websites and other services from learning your location"
  - "prevents people watching your traffic locally"
  - "routes your connection through more than one Tor relay so no single relay can learn what you're up to"

Tor operation overview



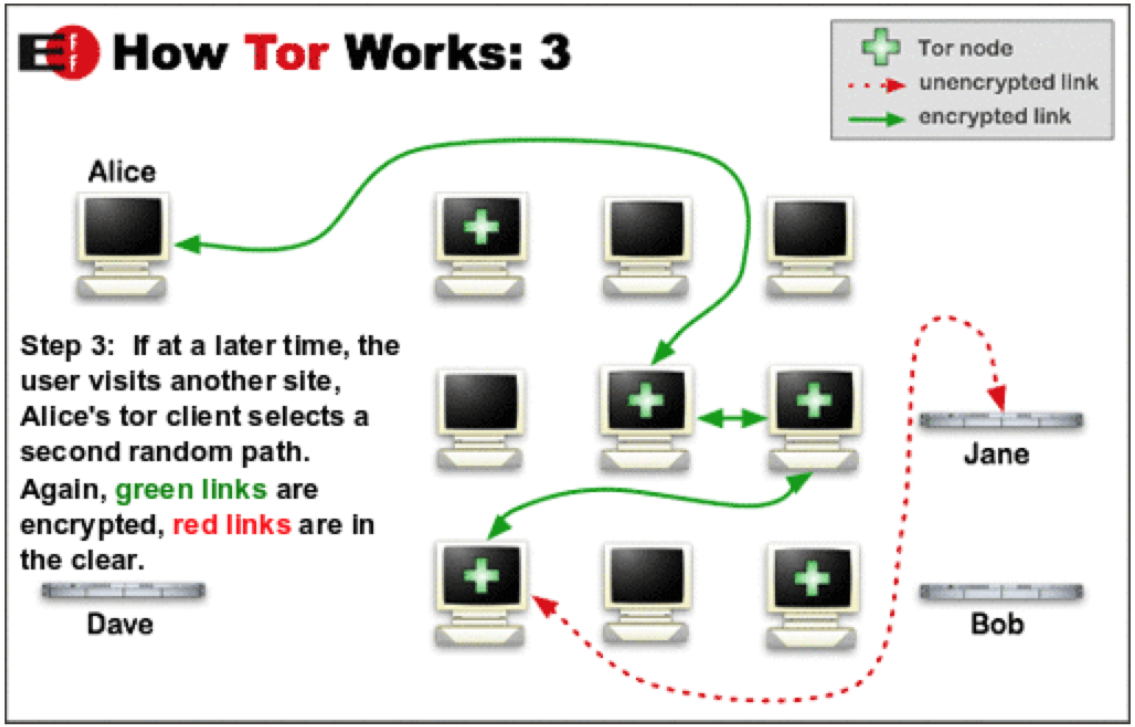
[source: Tor / Overview]

Tor operation overview



[source: Tor / Overview]

Tor operation overview



[source: Tor / Overview]

Currently we are here...

Introduction

Tor

Attacks on Tor
- Correlation
- Further threats

Correlation attacks

- Recap: Layered circuits
  - route: client → entry (guard) relay → middle relay → exit relay → destination
  - a relay knows the identity of only the previous and the next hops
  - the payload is encrypted, so e.g. the middle relay can't see: what is the hop after the next relay?, what is the content of the packet?, ...
  - no participant knows the whole route (except for the client)
- End-to-end timing analysis
  - as Tor happens within the TCP layer, IP headers are in the clear
  - observing traffic between client → entry relay and between exit relay → server, an adversary can correlate packet sizes and packet timing [3]
  - Result: adversary knows the client and the destination. Note: timing analysis works even if the communication is encrypted.

Asymmetric traffic analysis (RAPTOR, 2015) [5]

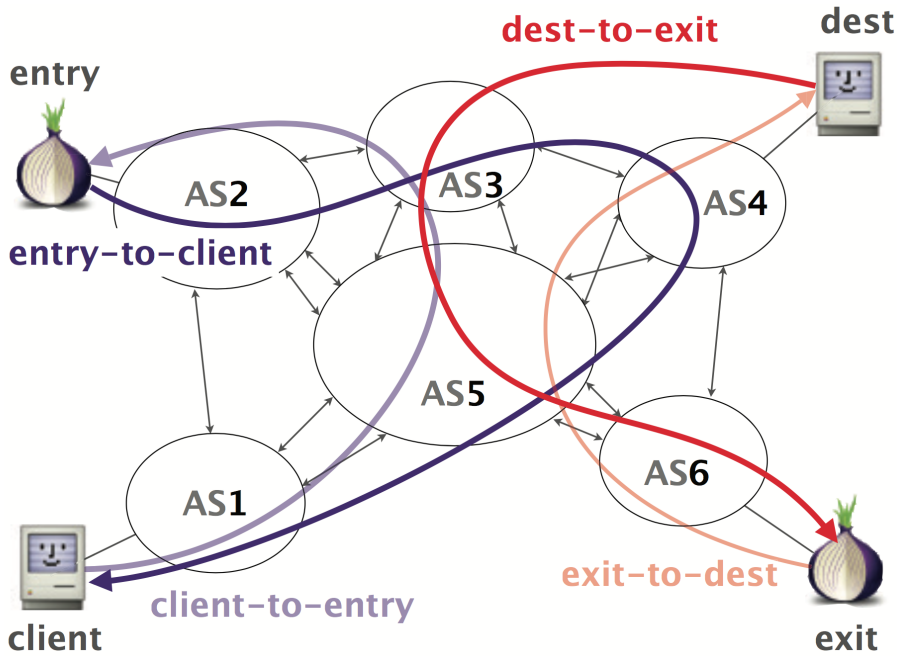
TLS is implemented within TCP; thus, header is not encrypted

- TCP segment header: sender and receiver IPs, seq. nr., ACK, ...
- TCP sequence number is the byte-stream number of the first byte of the segment
- TCP acknowledgements are cumulative: ACK number is the expected sequence number of other's segment
- The adversary can relate the lengths of segments and the acknowledgement numbers in TCP headers. Four cases:

	client – entry	exit – destination
case 1	$\overrightarrow{\text{data}}$	$\overrightarrow{\text{data}}$
case 2	$\overrightarrow{\text{data}}$	$\overleftarrow{\text{ACK}}$
case 3	$\overleftarrow{\text{ACK}}$	$\overrightarrow{\text{data}}$
case 4	$\overleftarrow{\text{ACK}}$	$\overleftarrow{\text{ACK}}$

- In sum: The adversary can de-anonymise users if she can observe any direction of traffic!

Asymmetric traffic analysis (RAPTOR, 2015) [5]



The arrows show the routes of the (unencrypted!) TCP segments

Which autonomous systems (ASs) can de-anonymise if

(1) only client-to-entry and exit-to-destination can be observed?

(2) both forward and backward traffic can be observed?

Summary and further info about Tor security

- Tor is resilient to local adversaries (no end-to-end observation), but Tor is vulnerable to end-to-end adversaries: users de-anonymised
  - A new attack: asymmetric traffic analysis is 95% accurate!
- Powerful adversaries can control whole ASs (forcing to be involved)
  - BGP's (border gateway protocol) weaknesses can be exploited
  - Snowden documents show: NSA, GCHQ leverage them
- Another goal of active adversaries: block Tor users
  - This is especially important for the authorities of some countries, such as North Korea, China
  - Censorship circumvention is a major objective of Tor.
- Censors also try to find and block Tor services
  - Bridges (secret entry relays) and hidden services
  - A new direction in Tor: Reputation system for bridges with anonymous credentials (to filter censors acting as bridges)
- Full anonymity is hard, but Tor is successful and used in practice.

Currently we are here...

Introduction

Tor

Attacks on Tor



References: Papers

(1) Chaum, David L. "Untraceable electronic mail, return addresses, and digital pseudonyms." *Communications of the ACM* 24.2 (1981).

(2) Goldschlag, D., Reed, M., and Syverson, P. "Hiding routing information." *Information Hiding*. Springer Berlin Heidelberg, 1996.

(3) Syverson, P., Tsudik, G., Reed, M., and Landwehr, C. "Towards an analysis of onion routing security." *Designing Privacy Enhancing Technologies*. Springer Berlin Heidelberg, 2001.

(4) Dingledine, R., Mathewson, N., and Syverson, P. Tor: "The second-generation onion router." Naval Research Lab Washington DC, 2004.

(5) Sun, Y., Edmundson, A., Vanbever, L., Li, O., Rexford, J., and Chiang, M. "RAPTOR: Routing Attacks on Privacy in Tor." In *Proc. of the 24th conference on USENIX Security Symposium*, Washington, DC, US. August 2015.

References: links

Besides academic paper, you can find lots of information on the web. In particular, if you want to find out further details about Tor, you can study

- its website (https://www.torproject.org),
- its open-source implementation (https://gitweb.torproject.org/tor.git?a=tree;hb=HEAD) or
- this comprehensive summary presentation: https://ritter.vg/blog-all_about_tor.html.

Finally, you can find the NSA presentations about Tor and possible attacks on the bottom of this page:

www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded

- Tor: The king of high-secure, low-latency anonymity
- Peeling back the layers of Tor with EgotisticalGiraffe
- Tor Stinks