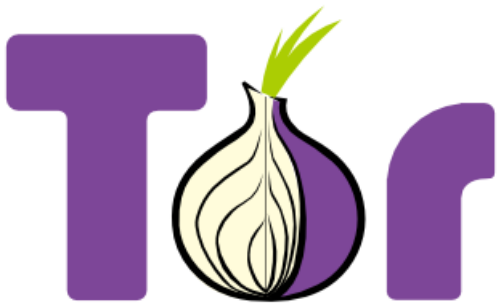
Online Anonymity

Andrew Lewman
andrew@torproject.org

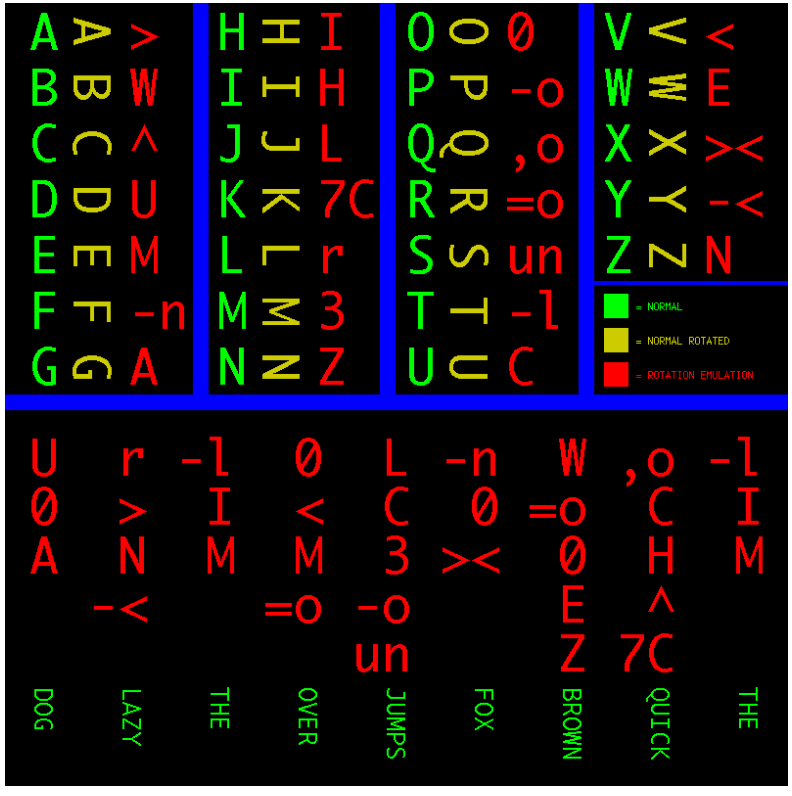
June 8, 2010



What is anonymity?

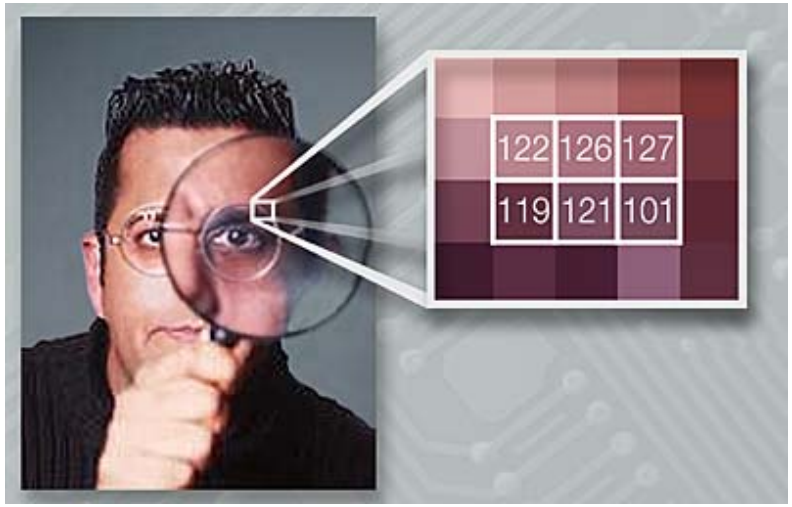
Anonymity isn't cryptography

- Cryptography protects the contents in transit
- You still know who is talking to whom, how often, and how much data is sent.



Anonymity isn't steganography

Attacker can tell Alice is talking to someone, how often, and how much data is sent.



Anonymity isn't just wishful thinking...

- "You can't prove it was me!"

Anonymity isn't just wishful thinking...

- "You can't prove it was me!"
- "Promise you won't look"

Anonymity isn't just wishful thinking...

- "You can't prove it was me!"
- "Promise you won't look"
- "Promise you won't remember"

Anonymity isn't just wishful thinking...

- "You can't prove it was me!"
- "Promise you won't look"
- "Promise you won't remember"
- "Promise you won't tell"

Anonymity isn't just wishful thinking...

- "You can't prove it was me!"
- "Promise you won't look"
- "Promise you won't remember"
- "Promise you won't tell"
- "I didn't write my name on it!"

Anonymity isn't just wishful thinking...

- "You can't prove it was me!"
- "Promise you won't look"
- "Promise you won't remember"
- "Promise you won't tell"
- "I didn't write my name on it!"
- "Isn't the Internet already anonymous?"

..since "weak" isn't anonymity.

- *"You can't prove it was me!"* Proof is a very **strong** word. Statistical analysis allows suspicion to become certainty.

..since "weak" isn't anonymity.

- *"You can't prove it was me!"* Proof is a very **strong** word. Statistical analysis allows suspicion to become certainty.
- *"Promise you won't look/remember/tell"* Will other parties have the abilities and incentives to keep these promises?

..since "weak" isn't anonymity.

- *"You can't prove it was me!"* Proof is a very **strong** word. Statistical analysis allows suspicion to become certainty.
- *"Promise you won't look/remember/tell"* Will other parties have the abilities and incentives to keep these promises?
- *"I didn't write my name on it!"* Not what we're talking about.

..since "weak" isn't anonymity.

- *"You can't prove it was me!"* Proof is a very **strong** word. Statistical analysis allows suspicion to become certainty.
- *"Promise you won't look/remember/tell"* Will other parties have the abilities and incentives to keep these promises?
- *"I didn't write my name on it!"* Not what we're talking about.
- *"Isn't the Internet already anonymous?"* Nope!

Anonymous communication

- People have to hide in a crowd of other people ("anonymity loves company")
- The goal of the system is to make all users look as similar as possible, to give a bigger crowd
- Hide who is communicating with whom
- Layered encryption and random delays hide correlation between input traffic and output traffic

Low versus High-latency anonymous communication systems

- Tor is not the first system; ZKS, mixmaster, single-hop proxies, Crowds, Java Anon Proxy.
- Low-latency systems are vulnerable to end-to-end correlation attacks.
- High-latency systems are more resistant to end-to-end correlation attacks, but by definition, less interactive.

Low-latency systems are generally more attractive to today's user

- Interactive apps: web, instant messaging, VOIP, ssh, X11, cifs/nfs, video streaming (millions of users)
- Multi-hour delays: email, nntp, blog posting? (tens of thousands of users?)

Low-latency systems are generally more attractive to today's user

- Interactive apps: web, instant messaging, VOIP, ssh, X11, cifs/nfs, video streaming (millions of users)
- Multi-hour delays: email, nntp, blog posting? (tens of thousands of users?)
    - And if anonymity loves company...

Who wants anonymity online?

- Normal people
- Law Enforcement
- Human Rights Activists
- Business Execs
- Militaries
- Abuse Victims

What is Tor?

- online anonymity, circumvention software and network
- open source, free software (BSD 3-clause & GPLv2 licenses)

What is Tor?

- online anonymity, circumvention software and network
- open source, free software (BSD 3-clause & GPLv2 licenses)
- active research environment:
  Rice, UMN, NSF, NRL, Drexel, Waterloo, Cambridge UK,
  Bamberg Germany, Boston U, Harvard, MIT, RPI, GaTech

What is Tor?

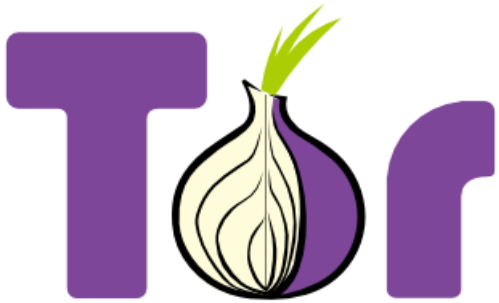
- online anonymity, circumvention software and network
- open source, free software (BSD 3-clause & GPLv2 licenses)
- active research environment:
  Rice, UMN, NSF, NRL, Drexel, Waterloo, Cambridge UK,
  Bamberg Germany, Boston U, Harvard, MIT, RPI, GaTech
- increasingly diverse toolset:
  Tor, Torbutton, Tor Browser Bundle, TorVM, Incognito
  LiveCD, Tor Weather, Tor auto-responder, Secure Updater,
  Orbot, TorFox, Torora, Portable Tor, Tor Check, Arm,
  Nymble, Tor Control, Tor Wall

Who is The Tor Project, Inc?

The 501(c)(3) non-profit organization dedicated to the research and development of tools for online anonymity and privacy

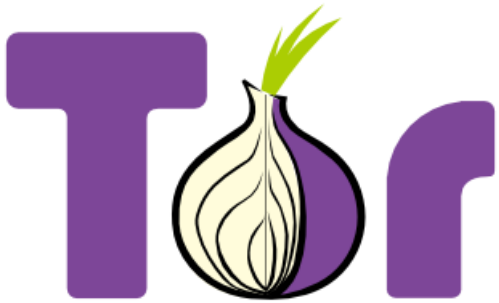
Tor is a low-latency anonymity system

- Based on technology developed in the Onion Routing project



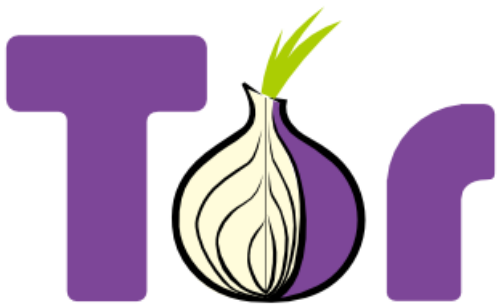
Tor is a low-latency anonymity system

- Based on technology developed in the Onion Routing project
- Privacy by design, not by policy (no data collected)



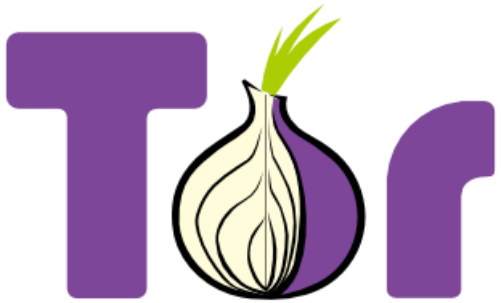
Tor is a low-latency anonymity system

- Based on technology developed in the Onion Routing project
- Privacy by design, not by policy (no data collected)
- Commonly used for web browsing and instant messaging (works for any TCP traffic)



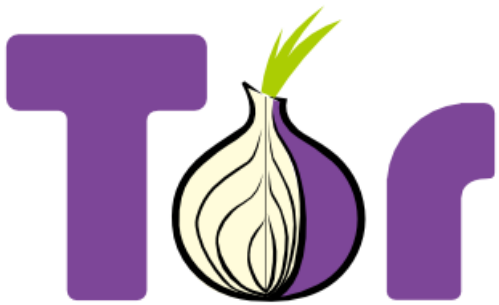
Tor is a low-latency anonymity system

- Based on technology developed in the Onion Routing project
- Privacy by design, not by policy (no data collected)
- Commonly used for web browsing and instant messaging (works for any TCP traffic)
- Originally built as a pure anonymity system (hides who is talking to whom)



Tor is a low-latency anonymity system

- Based on technology developed in the Onion Routing project
- Privacy by design, not by policy (no data collected)
- Commonly used for web browsing and instant messaging (works for any TCP traffic)
- Originally built as a pure anonymity system (hides who is talking to whom)
- Now designed to resist censorship too (hides whether someone is using the system at all)



TorProject.org

Tor code stats

Lines of Code By Language

Language	Code Lines	Comment Lines	Comment Ratio	Blank Lines	Total Lines
C	80,135	16,382	17.0%	9,792	106,309
C++	3,138	1,778	36.2%	669	5,585
shell script	2,033	824	28.8%	506	3,363
XML	1,050	0	0.0%	3	1,053
Autoconf	738	27	3.5%	128	893
Python	673	169	20.1%	130	972
Perl	382	73	16.0%	53	508
Automake	243	45	15.6%	83	371
Make	143	46	24.3%	39	228
HTML	105	16	13.2%	23	144
Ruby	62	35	36.1%	18	115

stats from http://www.ohloh.net/p/tor

How many people use Tor?

No idea. It's an anonymity system.

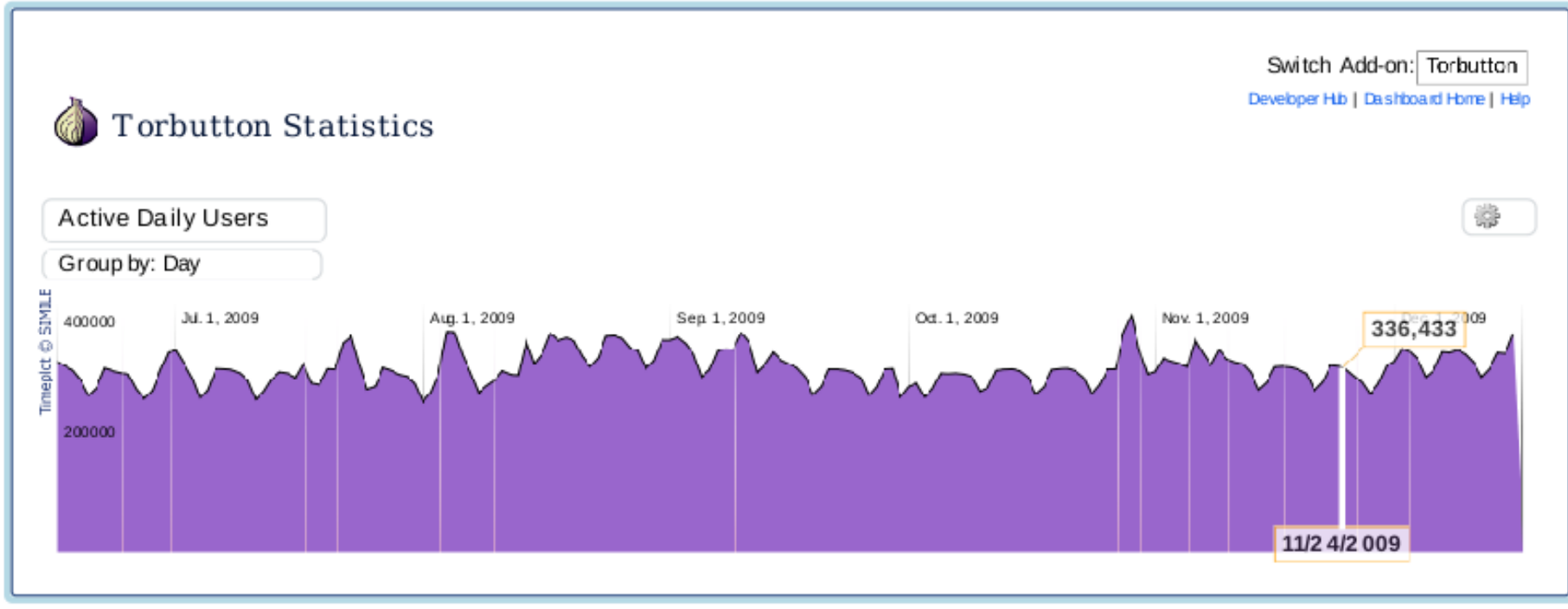
How many people use Tor?

No idea. It's an anonymity system.

http://metrics.torproject.org/ for an idea.

estimated 500,000 daily users

No really, how many people use Tor?



Total Downloads Since Mar. 23, 2006	**3,392,240**	**Active Daily Users** On Wednesday, Dec. 16	**403,079**
Last Day Count Wednesday, Dec. 16	2,720	Change from previous count 365,969 on Dec. 15	+10.14%
Average Daily Downloads	3,765	Average Daily Active Users	298,291
Downloads in the last 7 days	20,508	Average Daily Users this Week +0.63% from last week	360,676

Tor hides communication patterns by relaying data through volunteer servers

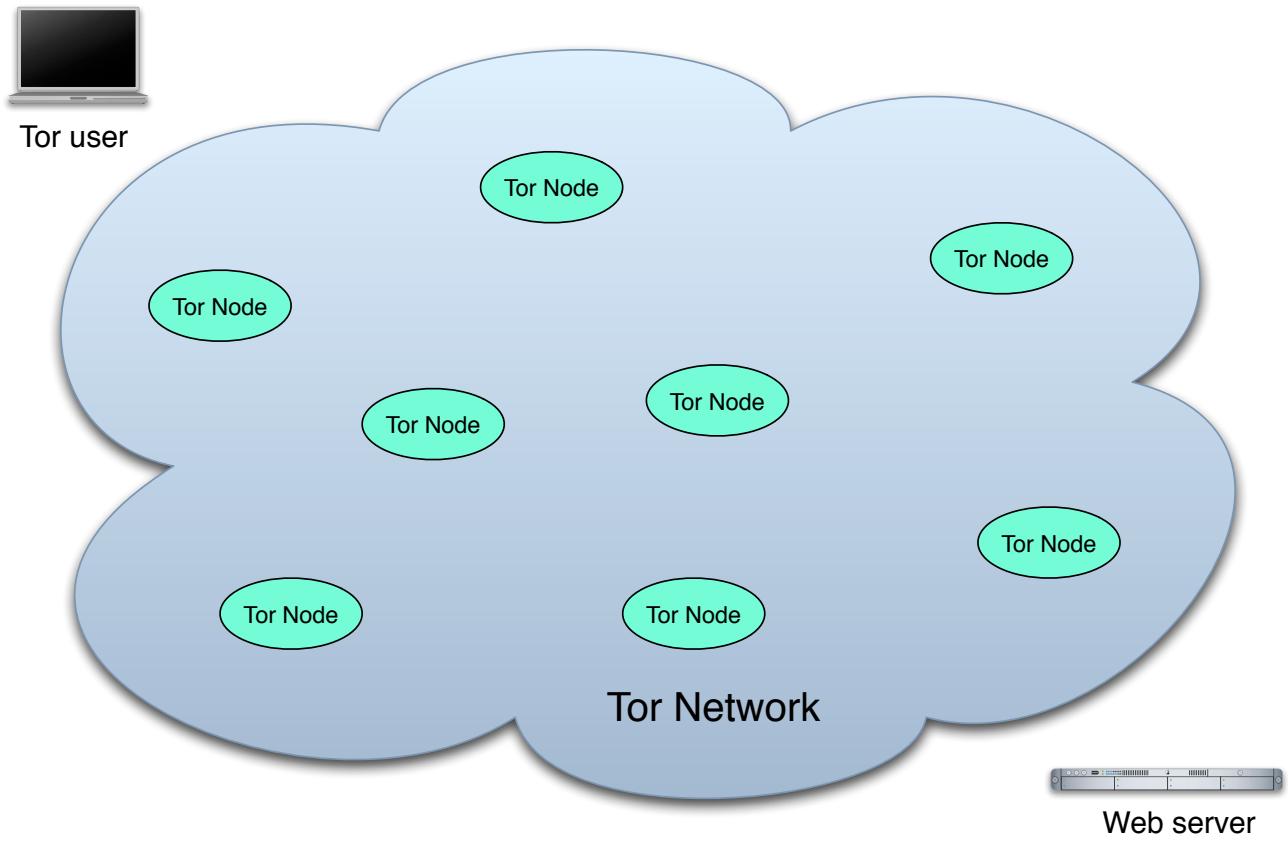


Diagram: Robert Watson

Tor hides communication patterns by relaying data through volunteer servers

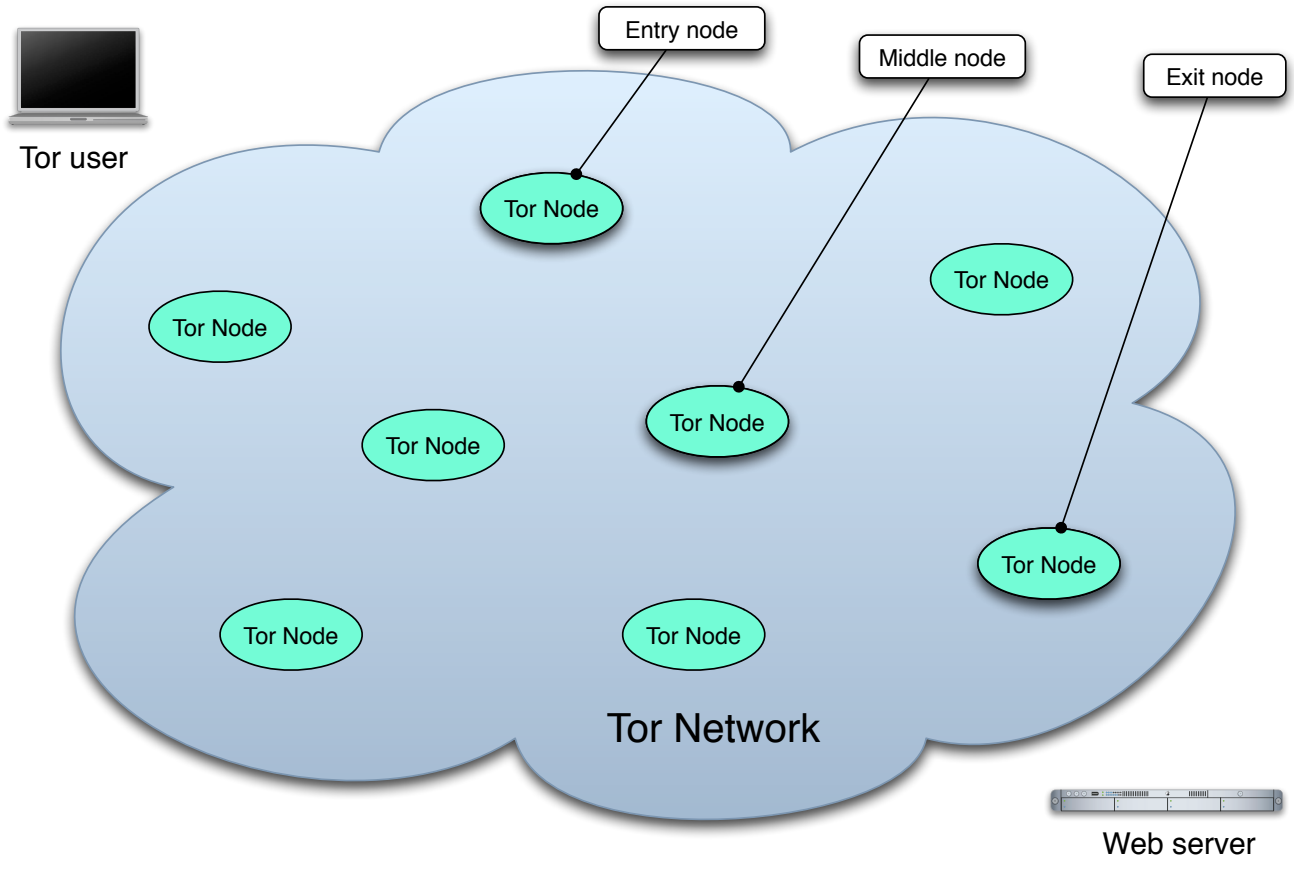


Diagram: Robert Watson

Tor hides communication patterns by relaying data through volunteer servers

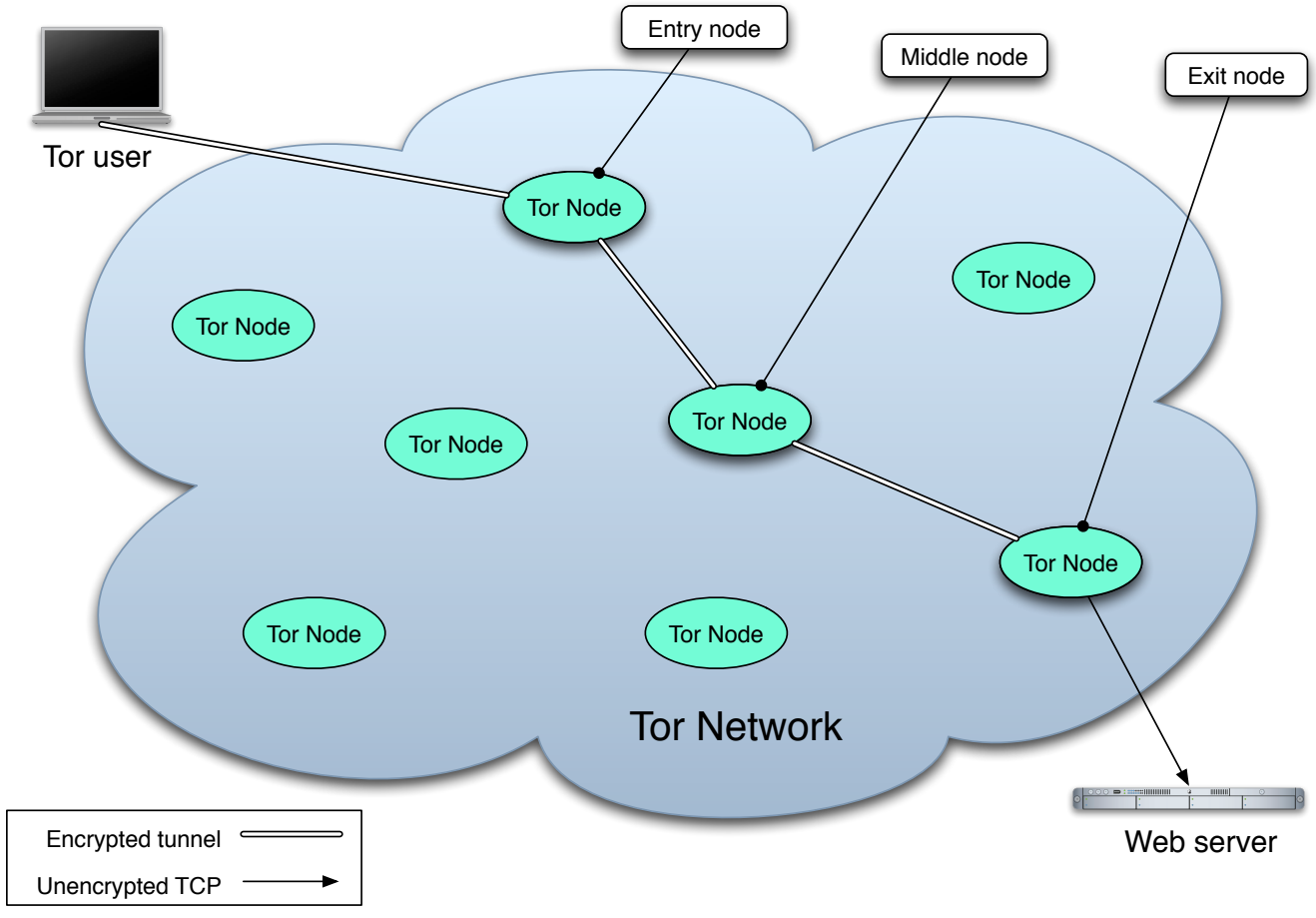
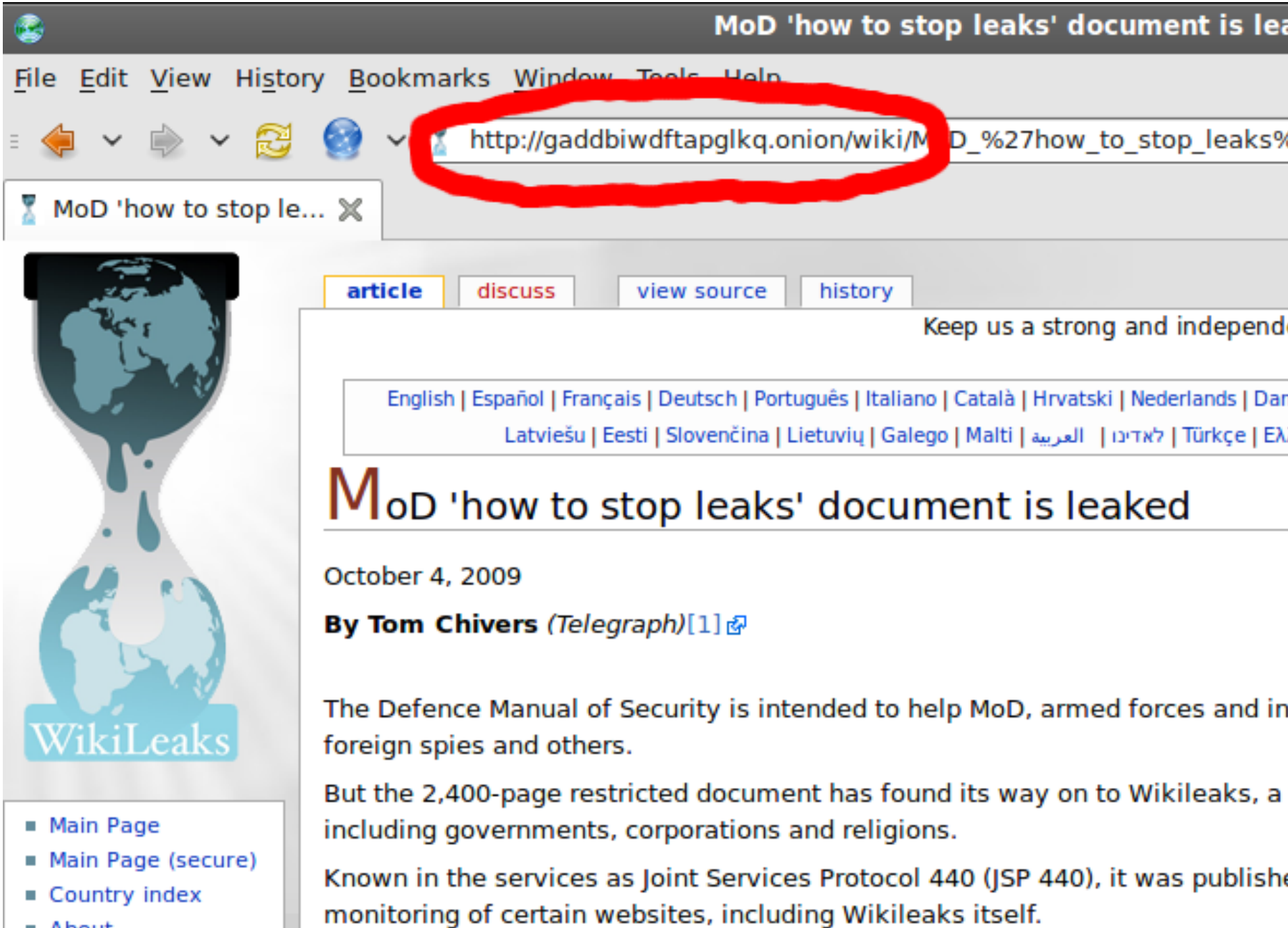
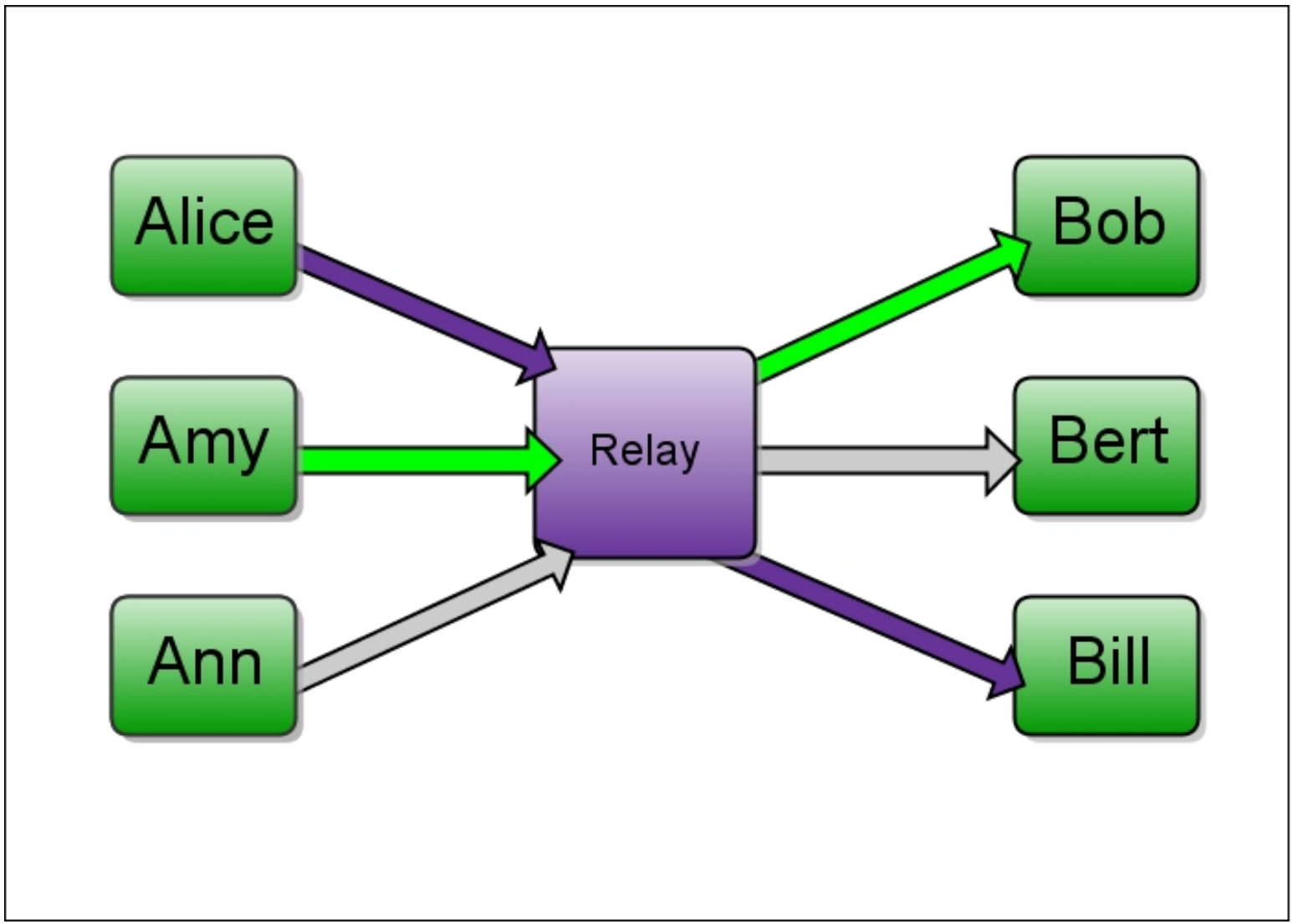


Diagram: Robert Watson

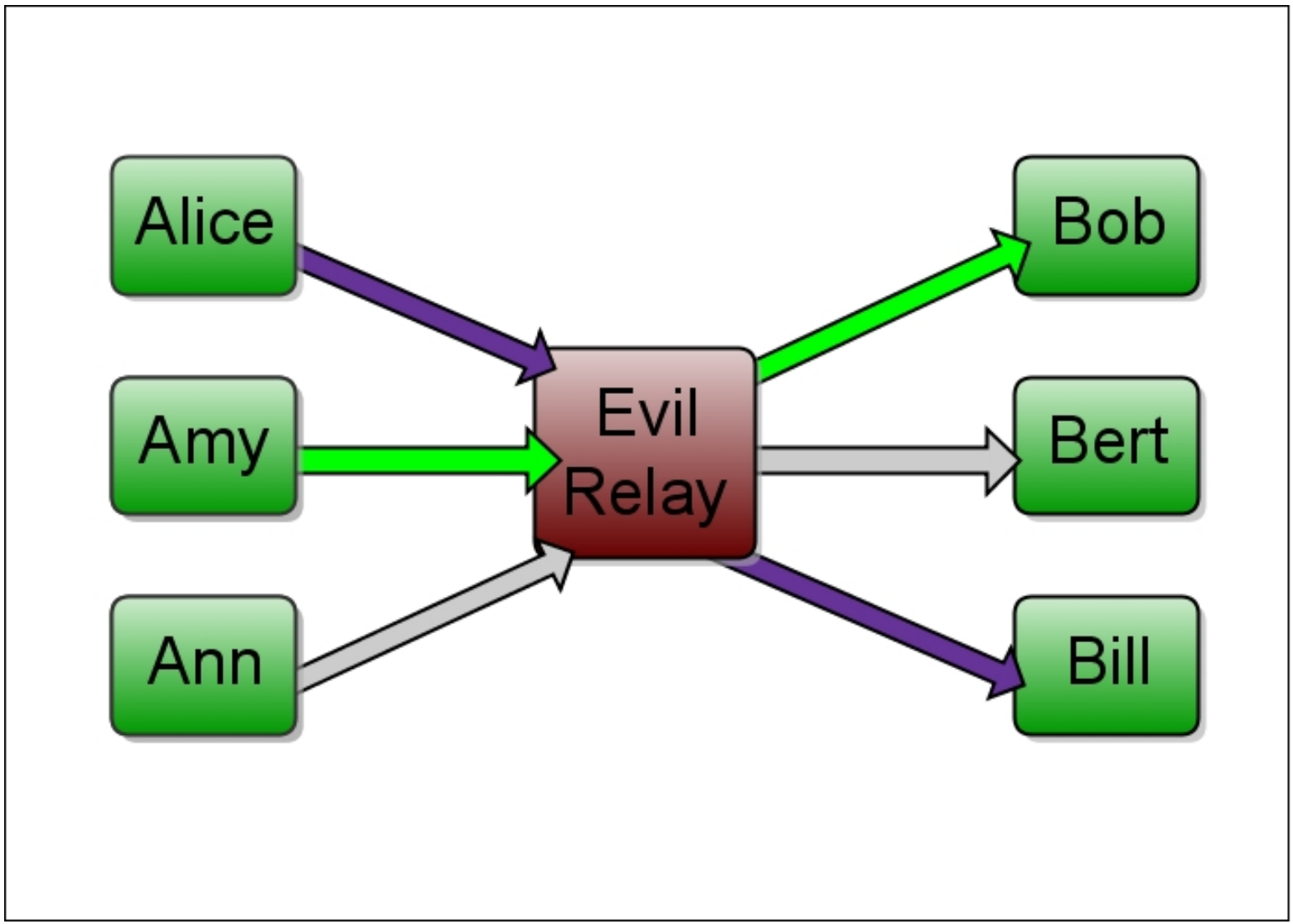
Tor hidden services allow censorship resistant services



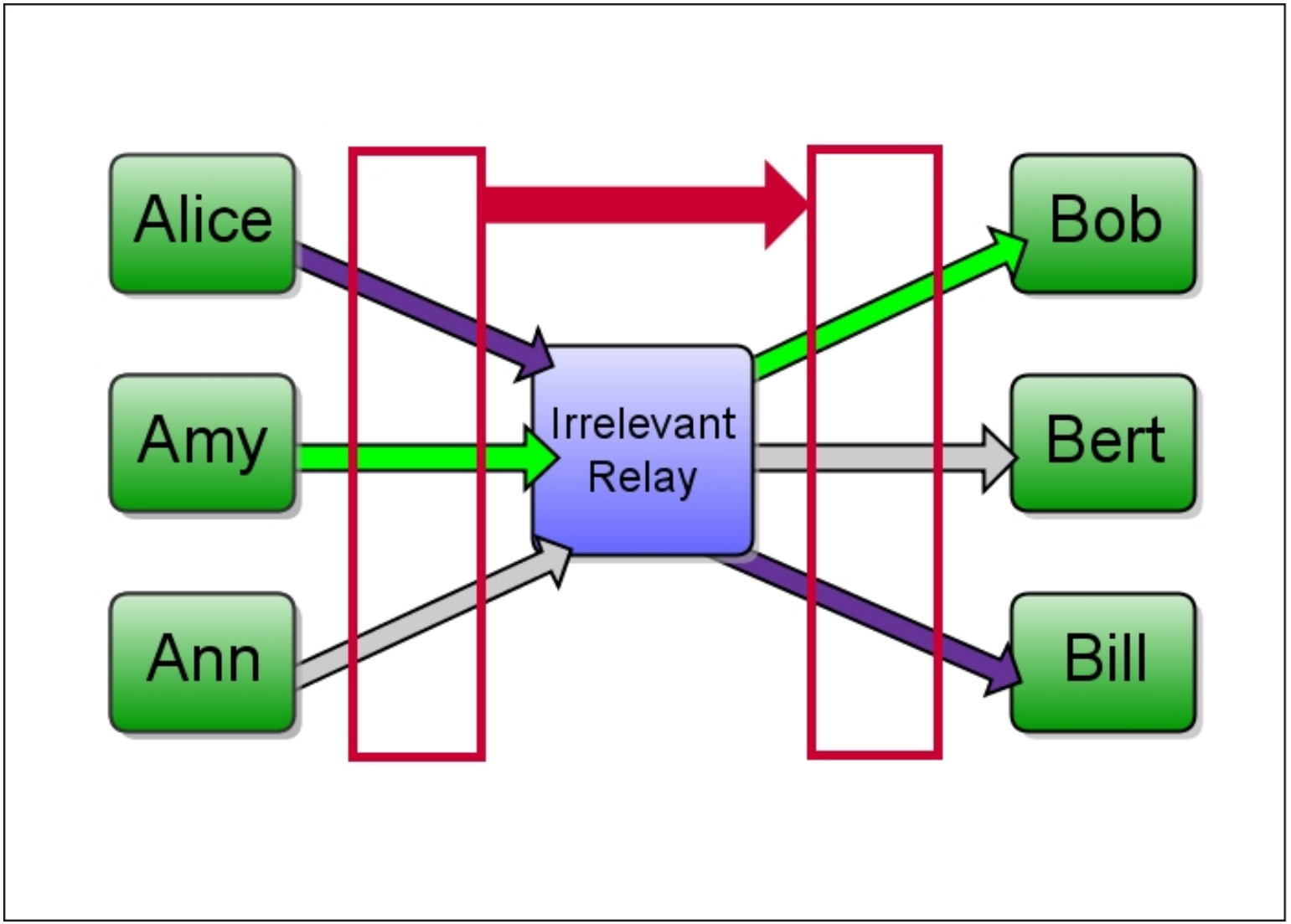
How is Tor different from other systems?



How is Tor different from other systems?



How is Tor different from other systems?



https://torproject.org/volunteer

Limitations of anonymous communication

- There is something for everyone to like, and something for everyone to dislike, going on with online anonymity systems
- Bad people do use them to do bad things (for many different definitions of bad people)
- It is impossible to block bad uses, even if we could come up with a common definition of bad content
- The systems are not perfect, so it is possible some people will be caught

United States Constitution: 1st Amendment

> *Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.*

McIntyre v. Ohio Elections Commission

> *Protections for anonymous speech are vital to democratic discourse. Allowing dissenters to shield their identities frees them to express critical, minority views . . . Anonymity is a shield from the tyranny of the majority . . . It thus exemplifies the purpose behind the Bill of Rights, and of the First Amendment in particular*

Twitter in Iran: Good.

Iran Protests: Twitter, the Medium of the Movement

By **LEV GROSSMAN** Wednesday, Jun. 17, 2009

Related

Photos

Behind the Scenes with Mousavi

Stories

- In Iran, Rival Regime Factions Play a High-Stakes Game of Chicken
- Latest Tweets on Fallout from Iran's

MORE

Share The U.S. State Department doesn't usually take an interest in the maintenance schedules of dotcom start-ups. But over the weekend, officials there reached out to Twitter and asked them to delay a network upgrade that was scheduled for Monday night. The reason? To protect the interests of

From http://www.time.com/time/world/article/0,8599,1905125,00.html

Twitter in USA: Bad.

FBI Raids Queens Home in G20 Protest Twitter Crackdown

AP Photo/Matt Rourke

That's right, a Twitter crackdown. A lawyer for Jackson Heights social worker Elliot Madison, 41, says that the feds searched his client's house for 16 hours on Thursday after Madison was arrested on September 24th at a Pittsburgh hotel room with another man. What were they up to? Sitting at laptops sending Twitter messages advising G20 demonstrators about riot police activity in the streets. And yet *real* Twitter threats like Lindsay Lohan and Courtney Love remain at large.

Madison, a self-described anarchist, was in Pittsburgh volunteering for the Tin Can Comms Collective, a group that uses Twitter to send mass text messages during protests describing events observed on the streets or over police scanners; stuff like "SWAT teams rolling down 5th Ave." Tin Can was active during the St. Paul RNC protests, and the authorities are now on to them. Madison was charged with hindering apprehension or prosecution, criminal use of a communication facility and possession of instruments of crime; he's currently out on bail.

from http://gothamist.com/2009/10/05/fbi_raids_queens_home_in_g20_protes.php

Internet architecture allows surveillance

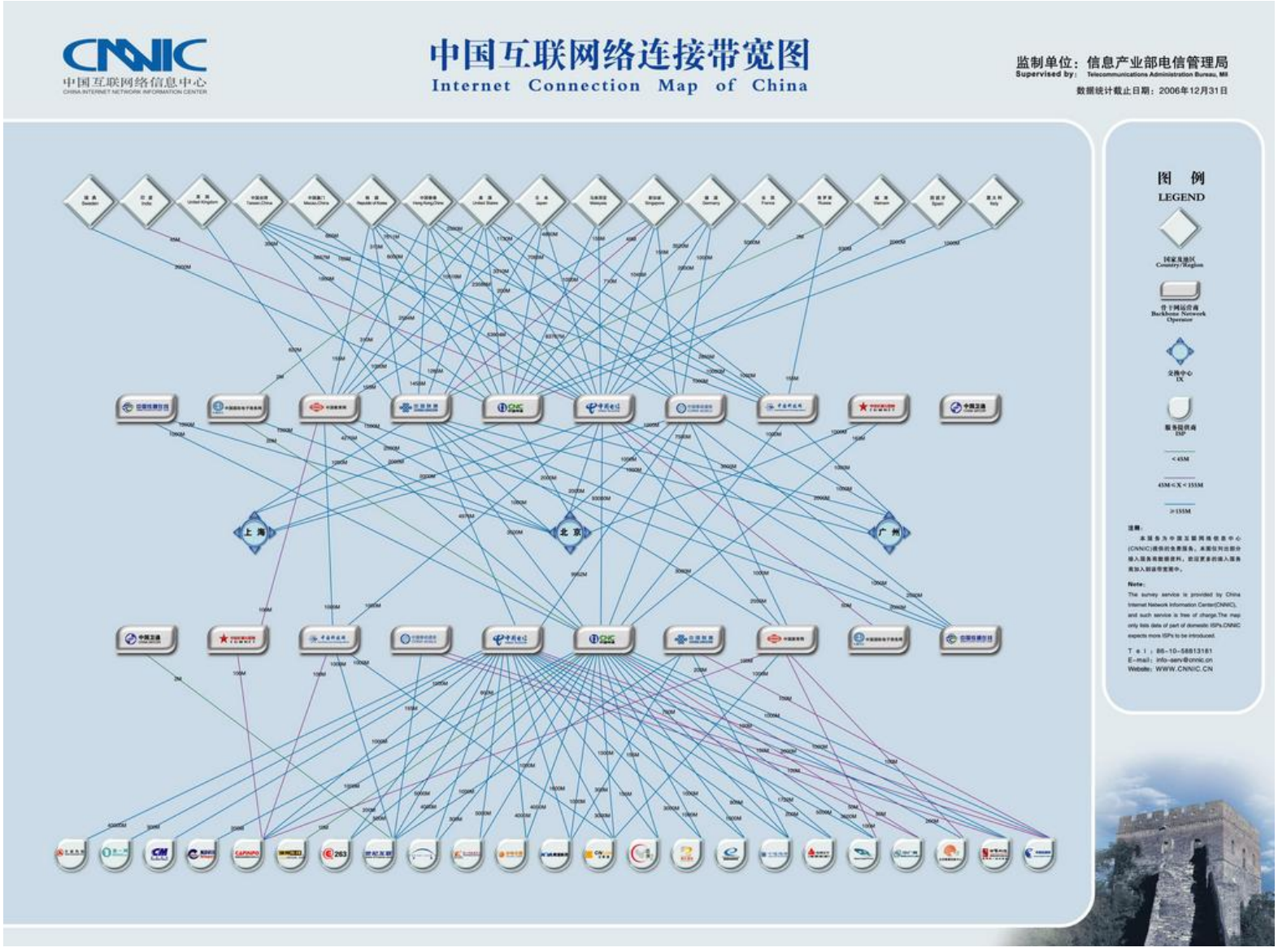


Diagram: China Internet Network Information Center

Internet surveillance is pervasive

- Conventional surveillance methods had to be targeted
- Internet censorship is capable of monitoring everyone, all of the time
- Governments are increasing monitoring: SORM (Russia), Golden Shield (China), Data Retention Directive (EU), and Interception Modernisation Programme (UK), Warrantless Wiretapping (USA)
- 1 in 7 East German citizens worked for the Stasi. Today we can achieve the same results for a fraction of the cost.

Traffic data surveillance

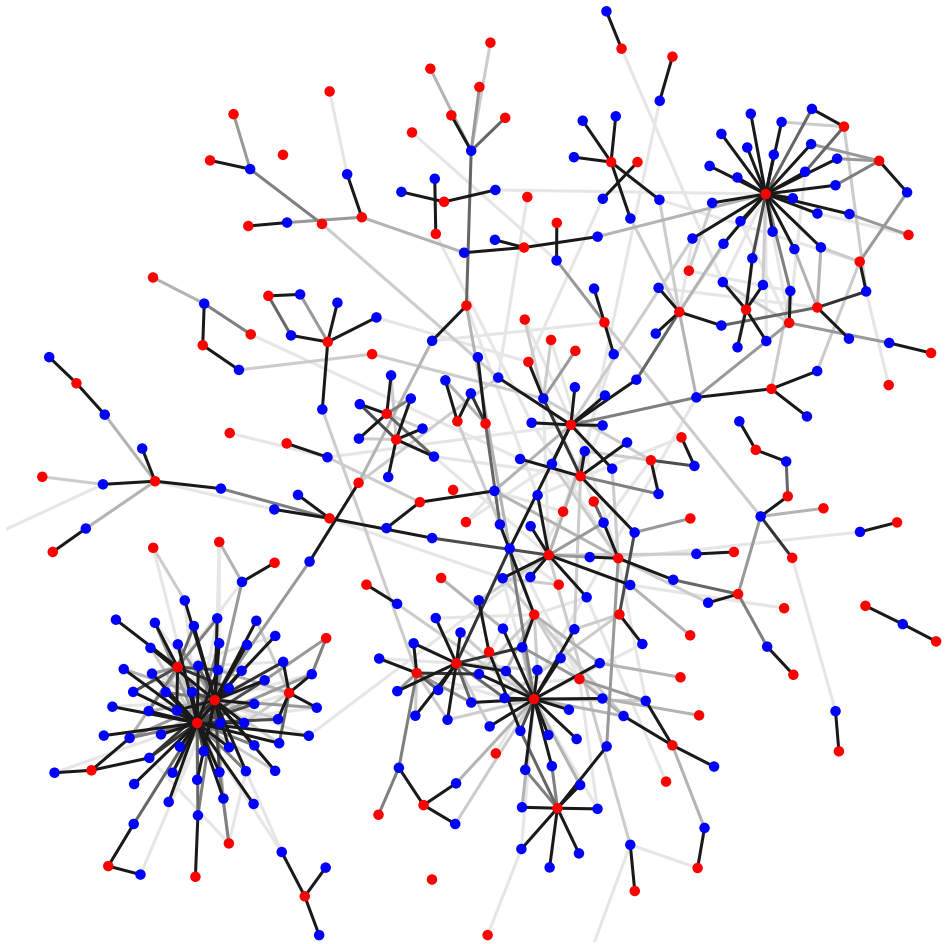
- Traffic data (who talks to whom, how often and for how long) is the core of intelligence capabilities
- This information is cheaper to record and store, compared to full content
- Because it can be easily processed by computer, data mining techniques can be used to understand social structures

> *No government of any colour is to be trusted with such a roadmap to our souls*

— Sir Ken Macdonald, former director of public prosecutions, on the UK Interception Modernisation Program

Importantly, information on social networks can be derived

- Communities
- People



From "The Economics of Mass Surveillance" by George Danezis and Bettina Wittneben

The Transparent Society

- David Brin proposed that in a world of pervasive surveillance, balance could be maintained by allowing everyone to monitor everyone else
- Bruce Schneier retorted that surveillance amplifies existing powers
- Many countries restrict anonymous speech (e.g. Germany and China)
- It is easy for those in power to call on the weak to link their names to opinions

Photo: Manos Simonides

One Version of the Future

> *I'd like to change the design of the Internet by introducing regulation–Internet passports, Internet police and international agreement–about following Internet standards. And if some countries don't agree with or don't pay attention to the agreement, just cut them off.*

— Eugene Kaspersky, Co-Founder & CEO of Kaspersky Labs

Another version of the future

> *Universal identification is impossible. Even attribution – knowing who is responsible for particular Internet packets – is impossible. Attempting to build such a system is futile, and will only give criminals and hackers new ways to hide.*

— Bruce Schneier, Security Technologist & Chief Security Technology Officer of BT

from http://www.schneier.com/blog/archives/2010/02/anonymity$_a nd_{t3}$.html

Conclusions

- The Internet and centralisation can both improve and harm freedom of speech
- Slogans regarding the borderless nature and inherent freedoms of the Internet are frequently wrong
- Technical details matter: how a system is implemented can make a radical difference
- Technologies are tools, they can be used for good and bad
- However, policies must be changed too and pressure is needed on legislators

Credits

- Thank you to Steven J. Murdoch, `http://www.cl.cam.ac.uk/users/sjm217/`, for the research and basis for this presentation.
- who uses tor? `http://www.flickr.com/photos/mattw/2336507468/siz`, Matt Westervelt, CC-BY-SA.
- 500k, `http://www.flickr.com/photos/lukaskracic/334850378/sizes/l/`, Luka Skracic, used with permission.
- Photographer and Diagram credits as listed throughout the presentation.