

Privacy in Asynchronous Messaging

Privacy Seminar (NWI-I00136)

Onno de Gouw

Laura Kolijn

Stefan Popa

Denise Verbakel

02-06-2022

Outline

- ▶ Introduction
 - ▶ Practical examples
 - ▶ What is metadata?
 - ▶ Social graphs
- ▶ Core problem
 - ▶ Legal implications
 - ▶ Societal implications
- ▶ Solving the core problem: 3 attempts
 - ▶ Encryption
 - ▶ The Onion Router
 - ▶ Three PETs
 - ▶ Transmission Protocol
 - ▶ RIPOSTE
 - ▶ DP5: will not explain
- ▶ Conclusion
- ▶ Questions



Introduction

- ▶ What is Asynchronous Messaging?
 - ▶ Does anyone have an idea?

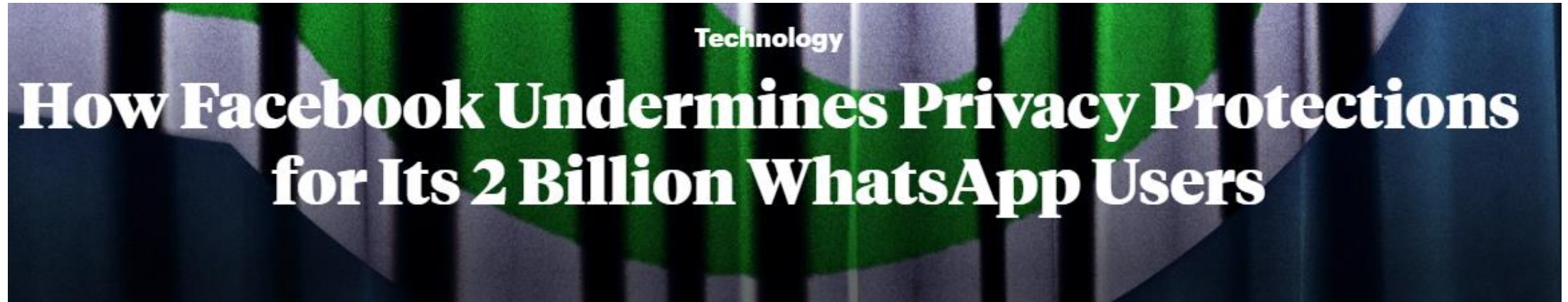


Asynchronous messaging

“Asynchronous Messaging is a communication method where participants on both sides of the conversation have the freedom to start, pause, and resume conversational messaging on their own terms, eliminating the need to wait for a direct live connection.” [1]

- ▶ Question time!
 - ▶ Can you think of possible privacy issues associated with asynchronous messaging?
 - ▶ What could practically be going wrong at e.g. WhatsApp when it comes to privacy?
 - ▶ Hint: Message content end-to-end encrypted, but? Moderation?

WhatsApp extensive monitoring [2]



WhatsApp assures users that no one can see their messages — but the company has an extensive monitoring operation and regularly shares personal information with prosecutors.

- ▶ Investigated by ProPublica, Sept. 2021
- ▶ Investigative Journalism

WhatsApp extensive monitoring [2]

- ▶ ProPublica
 - ▶ Investigation of data, documents, dozens of interviews
- ▶ Monitoring billions of users: Flagged content
 - ▶ Over 1000 contract workers
 - ▶ Content reviewers go through millions of messages
 - ▶ Artificial intelligence systems and account information for message examination
 - ▶ Respond to dozens of lawful requests; sharing metadata
- ▶ Mark Zuckerberg's vision: very secure
 - ▶ WhatsApp's focus on privacy using end-to-end encryption
- ▶ Privacy statement assures full metadata control: "Trust us"
- ▶ Metadata is powerful!



6 / 85

Metadata

- ▶ Metadata is powerful
- ▶ Question time!
 - ▶ What is Metadata actually?



Metadata

- ▶ Data providing information about data
- ▶ Pre-digital analogy: Outside of an envelope; inside protected
- ▶ There are many different types of metadata:
 - ▶ Usage data
 - ▶ Location data
 - ▶ Who are you contacting and for how long?
 - ▶ Social Graph -> Relevant concept w.r.t. privacy!
 - ▶ Etc...



“Metadata absolutely tells you everything about somebody’s life. If you have enough metadata, you don’t really need content.” [2]

- Former NSA General Counsel Stewart Baker

Metadata is crucial information!

► Edward Snowden [3,4]

NSA collecting phone records of millions of Verizon customers daily

Exclusive: Top secret court order requiring Verizon to hand over all call data shows scale of domestic surveillance under Obama

- [Read the Verizon court order in full here](#)
- [Obama administration justifies surveillance](#)

Support the Guardian Available for everyone, funded by readers
Contribute → Subscribe →

Search jobs Sign in Search International edit

The Guardian

News Opinion Sport Culture Lifestyle More ▾

World ► Europe US Americas Asia Australia Middle East Africa Inequality Global development

The NSA files

NSA files decoded/
Edward Snowden's surveillance revelations explained



Under the terms of the order, the numbers of both parties on a call are handed over, as is location data and the time and duration of all calls. Photograph: Matt Rourke/AP

The National Security Agency is currently collecting the telephone records of millions of US customers of Verizon, one of America's largest telecoms providers, under a top secret court order issued in April.

Metadata is crucial information!

- ▶ Edward Snowden [3,4]
- ▶ Leaking NSA top secret documents
- ▶ Mass surveillance
 - ▶ Millions of regular, innocent Americans
- ▶ Building a ‘pattern of life’
 - ▶ Detailed profile of a target and anyone associated with them
- ▶ By gathering metadata
- ▶ Find out individual’s connections and associations



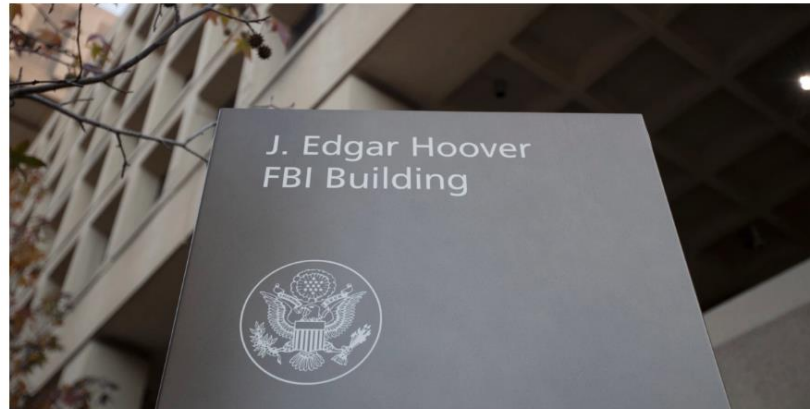
10 / 85

Metadata: FBI & WhatsApp [5,6]

FBI Document Says the Feds Can Get Your WhatsApp Data — in Real Time

A previously unreported FBI document obtained by *Rolling Stone* reveals that “private” messaging apps WhatsApp and iMessage are deeply vulnerable to law-enforcement searches

By **ANDY KROLL**



The J. Edgar Hoover FBI building in Washington Nov. 30, 2017 (AP Photo/Carolyn Kaster)

AP

WASHINGTON — As **Apple** and WhatsApp have built themselves into multibillion-dollar behemoths, they’ve done it while preaching the importance of privacy, especially when it comes to secure messaging.

But in a **previously unreported FBI document** obtained by *Rolling Stone*, the bureau claims that it’s particularly easy to harvest data from **Facebook’s** WhatsApp and Apple’s iMessage services, as long as the **FBI** has a warrant or subpoena. Judging by this document, “the most popular encrypted messaging apps iMessage and WhatsApp are also the most permissive,” according to Mallory Knodel, the chief technology officer at the Center for Democracy and Technology.

UNCLASSIFIED//LAW ENFORCEMENT SENSITIVE

FEDERAL BUREAU OF INVESTIGATION

LAWFUL ACCESS

(U//FOUO) FBI's Ability to Legally Access Secure Messaging App Content and Metadata

(U//LES) As of November 2020, the FBI's ability to legally access secure content on leading messaging applications is depicted below, including details on accessible information based on the applicable legal process. Return data provided by the companies listed below, with the exception of WhatsApp, are actually logs of latent data that are provided to law enforcement in a non-real-time manner and may impact investigations due to delivery delays.

UNCLASSIFIED//LAW ENFORCEMENT SENSITIVE

App	iMessage	Line	Signal	Telegram	Threema	Viber	WeChat	WhatsApp	Wickr
Information Accessed	<ul style="list-style-type: none"> Message Content: Limited Subpoena: can render basic subscriber information 18 U.S.C. §2703(d): can render 25 days of iMessage lookups to and from a target number* Pen Registers: no capability* Search Warrant: can render backups of a target device; if target uses iCloud backup, the encryption keys should also be provided with content return; can also acquire iMessages from iCloud returns if target has enabled Messages in iCloud 	<ul style="list-style-type: none"> Message Content: Limited* Suspect's and/or victim's, registered information (profile image, display name, email address, phone number, LINE ID, date of registration, etc.) Information on usage <p>*Maximum of seven days; worth of specified users' text chats (Only when E2EE has not been elected and applied and only when receiving an effective warrant; however, video, picture, files, location, phone call audio and other such data will not be disclosed)</p>	<ul style="list-style-type: none"> No Message Content Date and time a user registered Last date of a user's connectivity to the service 	<ul style="list-style-type: none"> No Message Content No contact information provided for law enforcement to pursue a court order. As per Telegram's privacy statement, for confirmed terrorist investigations, Telegram may disclose IP address and phone number to relevant authorities 	<ul style="list-style-type: none"> No Message Content Hash of phone number and email address, if provided by user Push Token, if push service is used Public Key Date (no time) of Threema ID creation Date (no time) of last login 	<ul style="list-style-type: none"> No Message Content Provides account (i.e. phone number) registration data and IP address at time of creation Message History: time, date, source number and destination number* 	<ul style="list-style-type: none"> No Message Content Accepts preservation letters and subpoenas, but cannot provide records in China created in China For non-China accounts, they can provide basic information (name, phone number, email, IP address), which is retained for as long as the account is active 	<ul style="list-style-type: none"> Message Content: Limited* Subpoena: can render basic subscriber records Court Orders Subpoena return as well as information like blocked users Search Warrant: Provides address book contacts and WhatsApp users who have the target in their address book contacts Pen Registers: Sent every 15 minutes, provides source and destination for each message <p>*If target is using an iPhone and iCloud backups enabled, iCloud returns may contain WhatsApp data, to include: message content</p>	<ul style="list-style-type: none"> No Message Content Date and time account created Type of device(s) app installed on Date of last use Total number of messages Number of external IDs (email addresses and phone numbers) connected to the account, but not plaintext external IDs themselves Avatar image Limited records of recent changes to account setting such as adding or suspending a device (does not include message content or routing and delivery information) Wickr Version Number*
Legal Process & Additional Details	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>SUBSCRIBER DATA</p> </div> <div style="text-align: center;"> <p>MESSAGE SENDER/RECEIVER DATA</p> </div> <div style="text-align: center;"> <p>DEVICE BACKUP</p> </div> <div style="text-align: center;"> <p>IP ADDRESS</p> </div> <div style="text-align: center;"> <p>ENCRYPTION KEYS</p> </div> <div style="text-align: center;"> <p>DATE/TIME INFORMATION</p> </div> <div style="text-align: center;"> <p>REGISTRATION TIME DATA</p> </div> <div style="text-align: center;"> <p>USER'S CONTACTS</p> </div> </div>								

(U) Prepared by Science and Technology Branch and Operational Technology Division

(U//LES) Apple provided logs only identify if a lookup occurred. Apple returns include a disclaimer that a log entry between parties does not indicate a conversation took place. These query logs have also contained errors.

(U) LAW ENFORCEMENT SENSITIVE: The information marked (U//LES) in this document is the property of FBI and may be distributed within the Federal Government (and its contractors), US intelligence, law enforcement, public safety or protection officials and individuals with a need to know. Distribution beyond these entities without FBI authorization is prohibited. Precautions should be taken to ensure this information is stored and/or destroyed in a manner that precludes unauthorized access. Information bearing the LES caveat may not be used in legal proceedings without first receiving authorization from the originating agency. Recipients are prohibited from subsequently posting the information, marked LES on a website or an unclassified network.

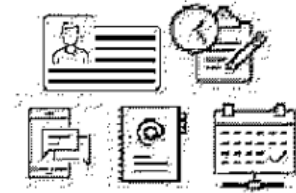
UNCLASSIFIED//LAW ENFORCEMENT SENSITIVE

7 January 2021

Metadata: FBI & WhatsApp [5,6]

- ▶ Rolling Stone, American magazine
- ▶ Obtained unreported FBI document
- ▶ Popular message apps deeply vulnerable to law enforcement searches
 - ▶ Easy for FBI to harvest data; warrant or subpoena
- ▶ Chat apps claim privacy and transparency
- ▶ WhatsApp offers the most real-time information
 - ▶ With a search warrant: all address book contacts
 - ▶ Serious consequences for e.g. journalists working with confidential source or facing governmental threats

WhatsApp



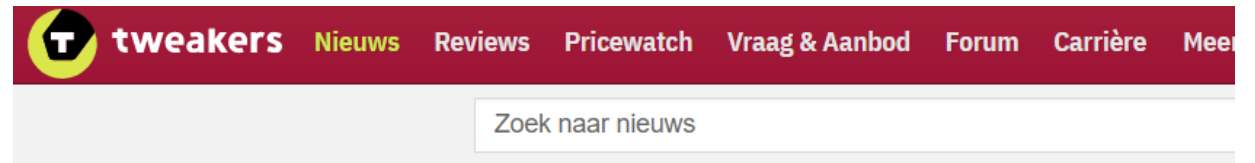
- **Message Content: Limited***
- **Subpoena:** can render basic subscriber records
- **Court Order: Subpoena:** return as well as information like blocked users
- **Search Warrant:** Provides address book contacts and WhatsApp users who have the target in their address book contacts
- **Pen Register:** Sent every 15 minutes, provides source and destination for each message

*If target is using an iPhone and iCloud backups enabled, iCloud returns may contain WhatsApp data, to include message content.

12 / 85

Metadata: Belgium Government [7]

- ▶ Obligatory for chat apps to store metadata by law
- ▶ In particular:
 - ▶ Identification data
 - ▶ Traffic data
 - ▶ Location data
 - ▶ Who contacts who
- ▶ Replacement of data retention law
 - ▶ Privacy concerns



Belgische overheid verplicht chatapps zoals WhatsApp metadata op te slaan

De Belgische overheid gaat communicatiediensten zoals WhatsApp, Facebook Messenger en Telegram verplichten om metadata op te slaan. De diensten moeten bijhouden wie met wie in contact stond en waar dat gebeurde. Wetgeving hiervoor moet eind dit jaar ingaan.

De verplichting om metadata te bewaren staat [volgens De Standaard](#) in een wetsontwerp van de Belgische overheid. Dat ontwerp is al goedgekeurd door de federale regering en de planning is dat de wet in het najaar ingevoerd kan worden. Ook andere Belgische kranten, zoals [De Morgen](#) en [De Tijd](#), schrijven daarover.

In het wetsontwerp zou benadrukt worden dat versleuteling van communicatie is toegestaan, maar dat dit niet mag verhinderen dat chatdiensten de identificatie-, verkeer- en locatiegegevens bewaren. De inhoud van het communicatieverkeer hoeft niet bewaard te worden; het gaat alleen om metadata. Ook telecomproviders moeten dergelijke metadata opslaan.


Metadata: Social Engineering

- ▶ Metadata can be used against you [8]
 - ▶ Social Engineering Attacks – Trick users
- ▶ Legitimacy of popups and people

DARKReading  The Edge DR Tech Sections Events

Tech Insight: How Attackers Use Your Metadata Against You

Using easily accessible data about your files, bad guys can wreak havoc on your sensitive information

 Dark Reading Staff
Dark Reading

February 13, 2009

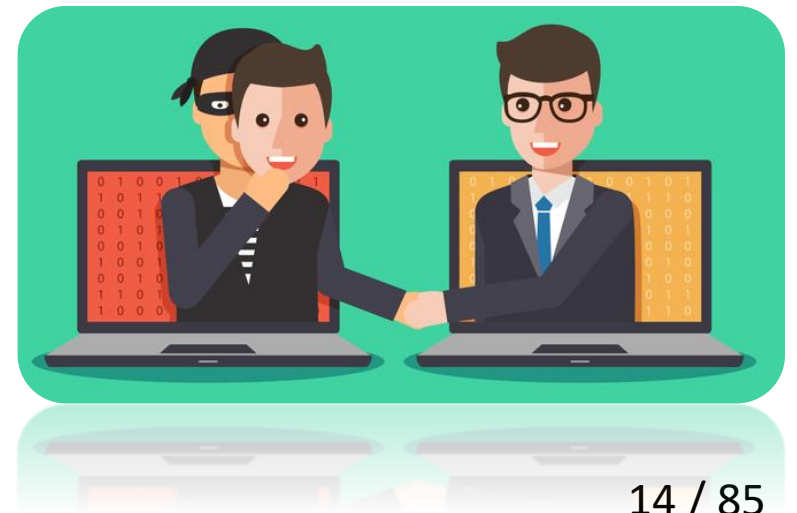


A Special Analysis For Dark Reading

First of Two Articles

To steal your identity, a cybercriminal doesn't have to have direct access to your bank account or other personal information. Often, he collects information *about* you from a variety of seemingly innocuous sources, then uses that data to map out a strategy to crack your online defenses and drain your accounts.

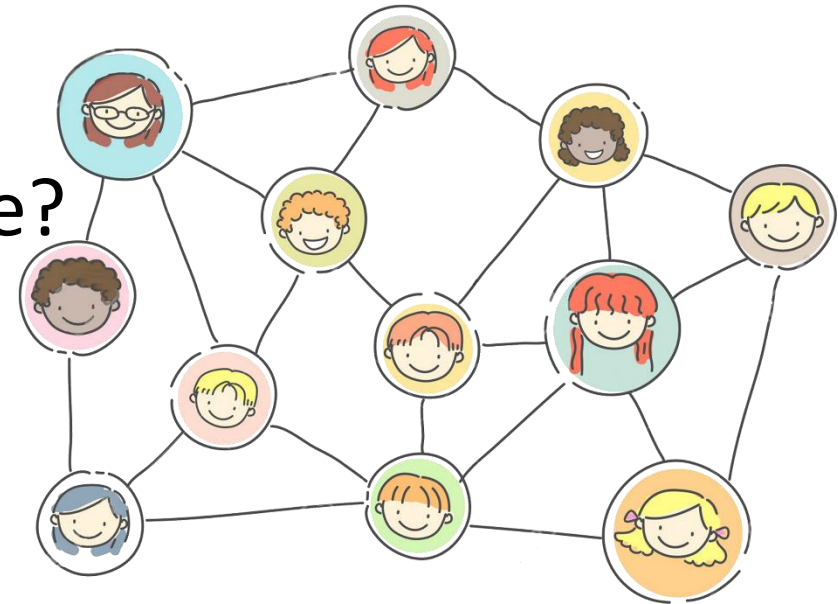
Such methods are well-known to security professionals. But what those same professionals often overlook is this approach also can be used to crack the defenses of sensitive business files, as well. Rather than trying to gain access to your data, itself, the bad guys are analyzing the so-called harmless information about your files – collectively known as metadata – and using it to develop attacks that can drain your business of its most sensitive information.



14 / 85

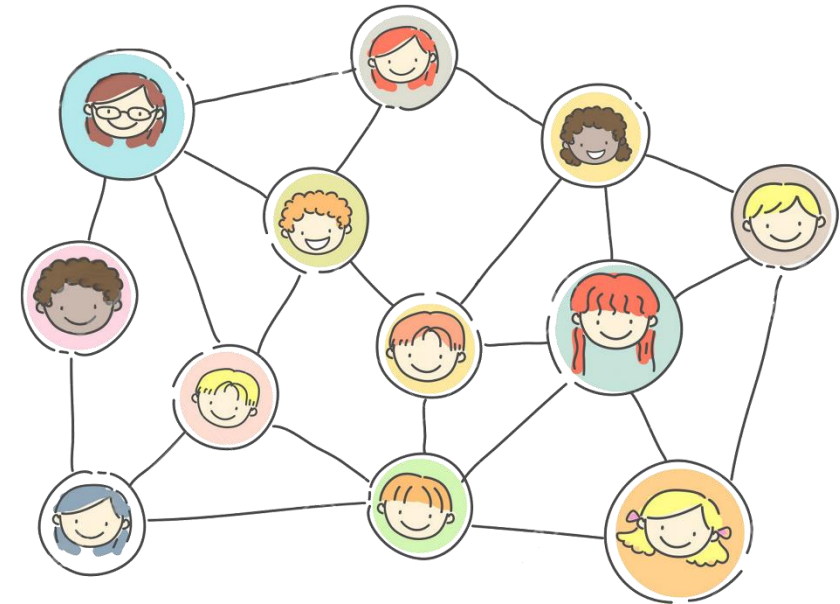
Another Relevant Concept

- ▶ We have seen practical cases regarding metadata
- ▶ Metadata can be used against you
- ▶ Who payed attention? Are you still awake?
- ▶ What relevant concept with regard to privacy did I mention a few slides back?



Social Graphs

- ▶ Can be created from metadata
- ▶ What is a social graph?
 - ▶ “A representation of the interconnection of relationships in an online social network.” [9]
- ▶ Linkability
- ▶ What can it be used for?
 - ▶ Advantages and disadvantages



Social Graphs Use Cases

- ▶ Advantages:
 - ▶ Law enforcement
 - ▶ Forensic research
 - ▶ Behavioural studies
 - ▶ Spreading of viruses



Social Graphs Use Cases

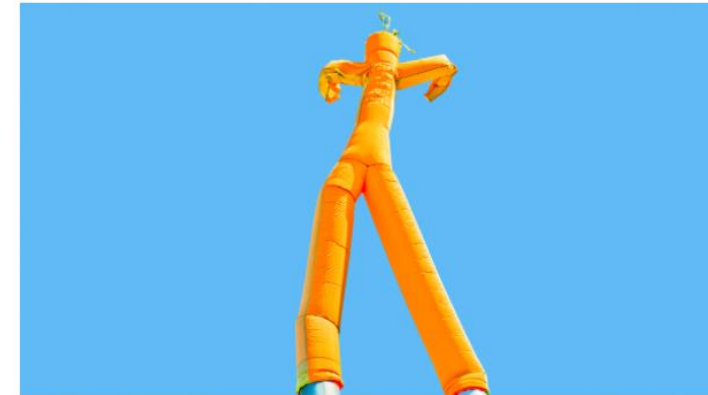
- ▶ Disadvantages:
 - ▶ Criminal activities and stalking
 - ▶ Commercial use for businesses
 - ▶ Prejudices/profiling
 - ▶ Refuse of hiring
 - ▶ Unknown connection to illegal activities

Reference article: [10]

The Rise of Social Graphs for Businesses

by Sangeet Paul Choudary

February 02, 2015

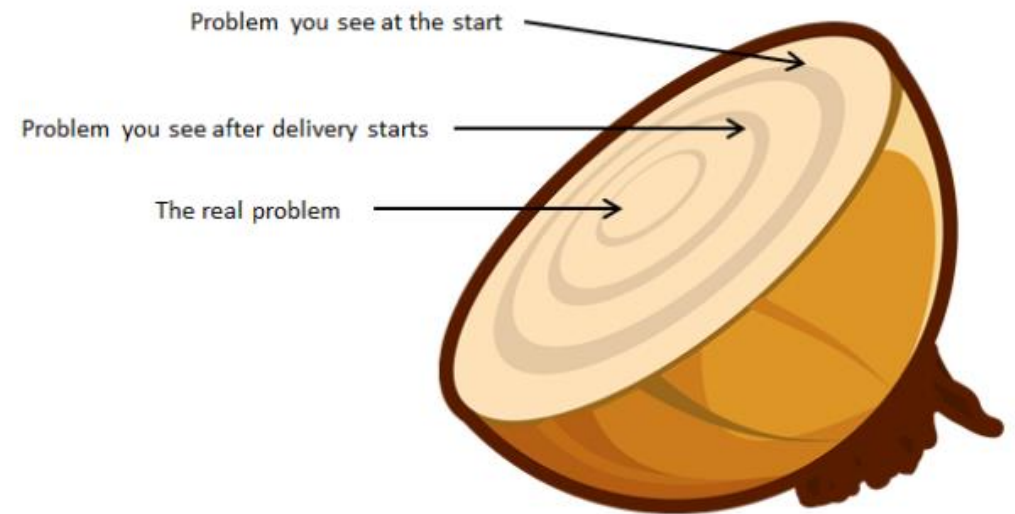


If you're a savvy social media user then you've already figured out that the knowledge a tool like Facebook is able to gather about your social connections is not only valuable to you. For you, Facebook's ability to depict your network of friends and the varying strengths of those relationships supports all your mutual information sharing. For others — third parties — this "social graph" makes it possible to make personalized recommendations to you, and everyone else. For example,

18 / 85

Core problem

- ▶ Content of communication
- ▶ When, where and/or with who
 - ▶ Metadata
 - ▶ Behavioral data



"We kill people based on metadata"

- ▶ Johns Hopkins University's Foreign Affairs Symposium
- ▶ Mass surveillance programs of NSA
- ▣ Michael Hayden
 - ▣ Former director of NSA and CIA



"We kill people based on metadata"

Reference
video: [11]



21 / 85

"We kill people based on metadata"

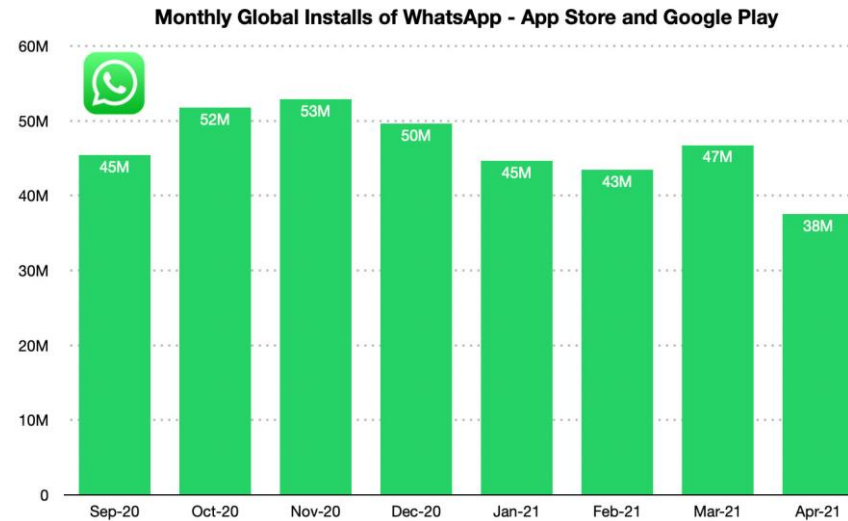
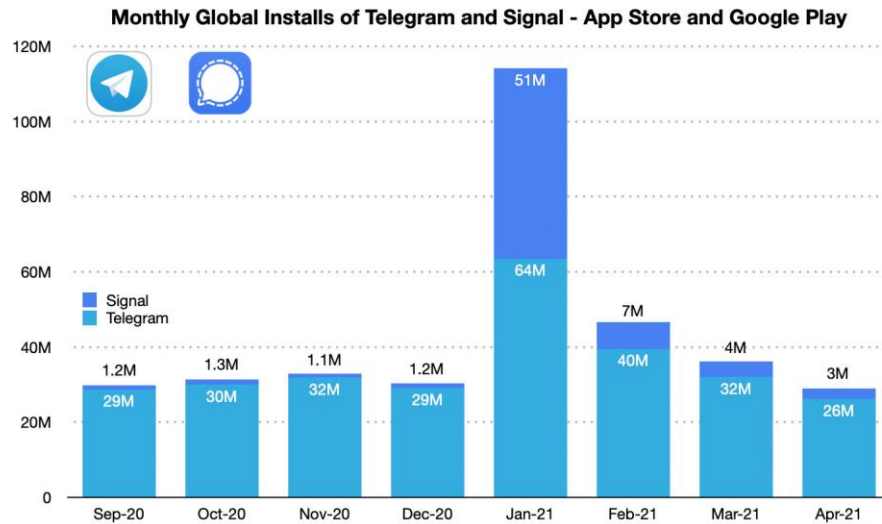
Reference
video: [11]



22 / 85

Societal Perspective

- ▶ "Only" metadata
- ▶ Challenge to move
 - ▶ Signal and Telegram <-> WhatsApp [12]



Societal Perspective – Whistle-blowers

► What is a whistle-blower?

“One who ‘blows the whistle’ on a person or activity.” [13]

“To bring an activity to a sharp conclusion, as if by the blast of a whistle; now usually by informing on (a person) or exposing (an irregularity or crime).” [13]



Societal Perspective – Whistle-blowers

- ▶ Speaking truth to power [14]
 - ▶ Congressmen Ted Lieu and Don Beyer

Know your options.

 <p>Encrypted message to the press</p>	<p>Chat apps that employ end-to-end encryption are a safe bet, like WhatsApp, Signal, or Telegram.</p>
 <p>Snail mail to the press</p>	<p>Mail without a return address is a low tech--but secure--way to communicate. To intercept mail in transit, authorities would require a warrant under Fourth Amendment protections.</p>



Agency Inspector General

Each federal agency houses an independent Inspector General that stands ready to receive information about fraud, waste, and abuse. Visit <http://bit.ly/2kJ415p> for more information about each Agency's IG.

For more on federal employee protections, visit <http://bit.ly/2IMsL2j>

Societal Perspective – Whistle-blowers

- ▶ Natalie Edwards [15]
- ▶ Financial Crimes Enforcement Network (FinCEN)
- ▶ 50,000 documents of which 2,000 SARs
- ▶ Presidential elections of 2016
- ▶ BuzzFeed - The money trail [16]



Societal Perspective – Whistle-blowers

18. Based upon my training, experience, my conversations with other law enforcement agents with training and experience in cyber technology, and my conversations with law enforcement agents who have reviewed records received in response to a judicially-authorized pen register and trap and trace order for the EDWARDS Cellphone (the "EDWARDS Pen"), I have learned, among other things, that:

Reference
article: [17]

a. The EDWARDS Cellphone utilized a mobile messaging service that utilizes end-to-end encryption (the "Encrypted Application"), that is, a method of secure communication that prevents third-parties from accessing data, including the companies that host the end-to-end encrypted services, and law enforcement.

b. On or about August 1, 2018, within approximately six hours of the EDWARDS Pen becoming operative—and the day after the July 2018 Article was published—the EDWARDS Cellphone exchanged approximately 70 messages via the Encrypted Application with the Reporter-1 Cellphone during an approximately 20-minute time span between 12:33 a.m. and 12:54 a.m.

Societal Perspective – Whistle-blowers

c. Between on or about July 31, 2018 and August 2, 2018, the EDWARDS Cellphone and the personal cellphone of CC-1 exchanged dozens of messages via the Encrypted Application.

d. On or about August 2, 2018, approximately one week prior to the publication of the First August 2018 Article, the EDWARDS Cellphone exchanged approximately 541 messages with the Reporter-1 Cellphone via the Encrypted Application.

e. On or about August 10, 2018, the day of the publication of the First August 2018 Article, the EDWARDS Cellphone and the Reporter-1 Cellphone exchanged approximately 11 messages via the Encrypted Application.

f. On or about August 15, 2018, the EDWARDS Cellphone exchanged approximately 14 messages with the Reporter-1 Cellphone via the Encrypted Application.

Reference
articles: [17]

28 / 85

Legal Perspective – History

- ▶ Data Retention Directive [18]
- ▶ Declared invalid in 2014
- ▶ Why?
 - ▶ Violation of fundamental rights
- ▶ Possible new legislation [19]



Press and Information

Court of Justice of the European Union
PRESS RELEASE No 54/14
Luxembourg, 8 April 2014

Judgment in Joined Cases C-293/12 and C-594/12
Digital Rights Ireland and Seitlinger and Others

The Court of Justice declares the Data Retention Directive to be invalid

It entails a wide-ranging and particularly serious interference with the fundamental rights to respect for private life and to the protection of personal data, without that interference being limited to what is strictly necessary

Legal Perspective – GDPR

- ▶ Metadata not mentioned
- ▶ Article 4(i)
 - ▶ *“‘personal data’ means **any** information relating to an identified or identifiable natural person (‘data subject’)” [20]*
- ▶ e-Privacy Regulation¹
 - ▶ Lex specialis GDPR

Legal Perspective – Human Rights

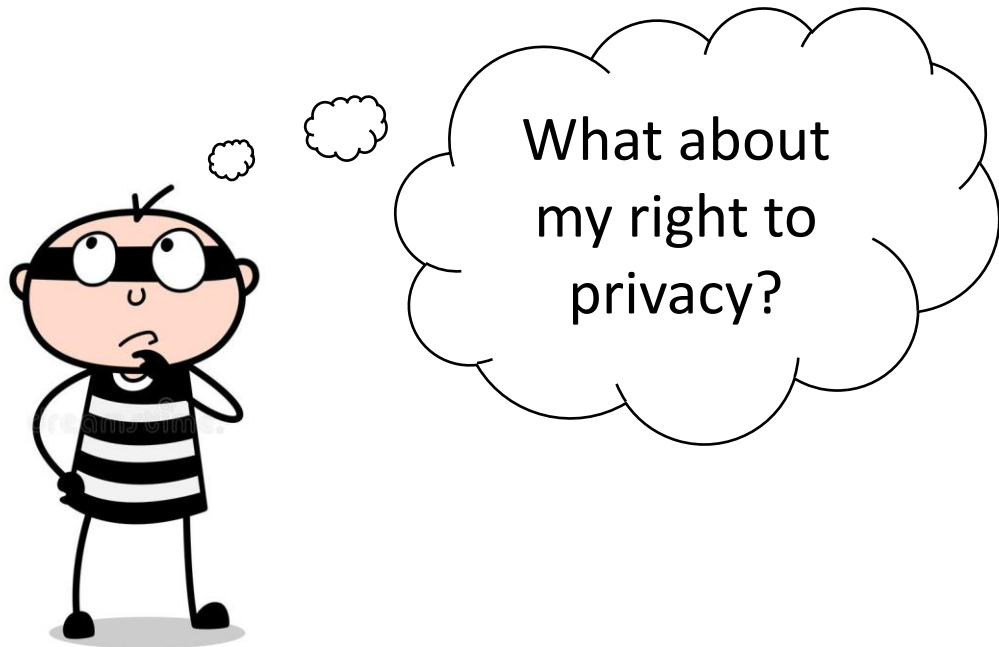
- ▶ Human Rights Council
- ▶ The right to privacy in the digital age [21]
 - ▶ Right to privacy extends to metadata

Legal Perspective – Human Rights

6. The protection of the right to privacy is broad, extending not only to the substantive information contained in communications but equally to metadata as, when analysed and aggregated, such data “may give an insight into an individual’s behaviour, social relationship, private preference and identity that go beyond even that conveyed by accessing the content of a communication” (see *A/HRC/27/37*, para. 19). The protection of the right to privacy is not limited to private, secluded spaces, such as the home of a person, but extends to public spaces and information that is publicly available (see *CCPR/C/COL/CO/7*, para. 32). For example, the right to privacy comes into play when a Government is monitoring a public space, such as a marketplace or a train station, thereby observing individuals. Similarly, when information that is publicly available about an individual on social media is collected and analysed, it also implicates the right to privacy.⁷ The public sharing of information does not render its substance unprotected.⁸

Ethical discussion

► Privacy vs. Safety

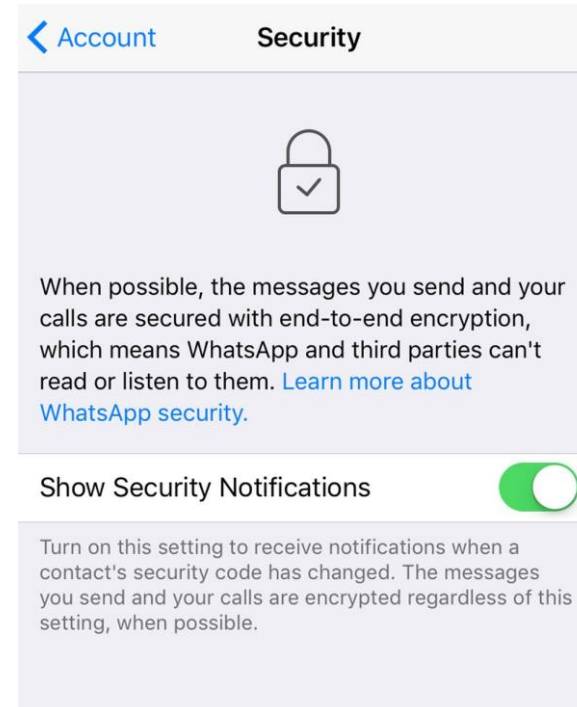


Attempts at Solutions

- ▶ Attempt 1: Encryption
- ▶ Attempt 2: The Onion Router
- ▶ Attempt 3: PETs
 - ▶ Transmission Protocol Using Public Bulletin Board [22]
 - ▶ RIPOSTE Protocol [23]
 - ▶ DP5 Protocol [24]

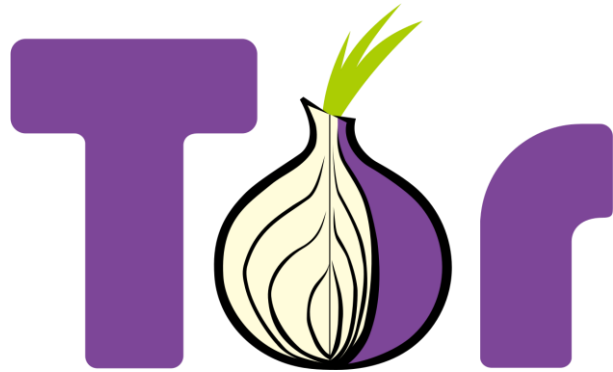
Attempt 1: Encryption

- ▶ End-to-end encryption
- ▶ Secures content



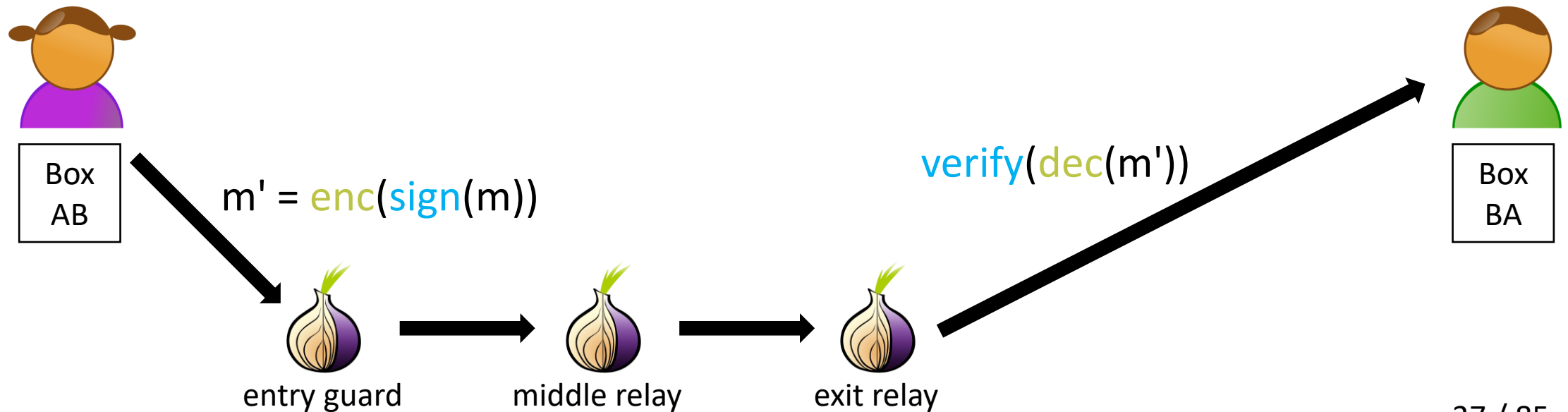
Reference picture: [25]

Attempt 2: The Onion Router [26,27]



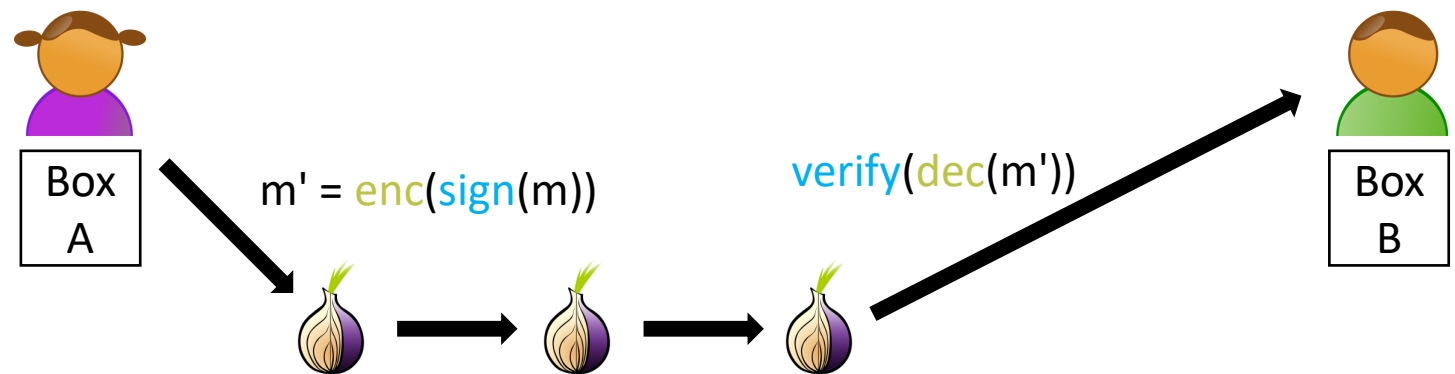
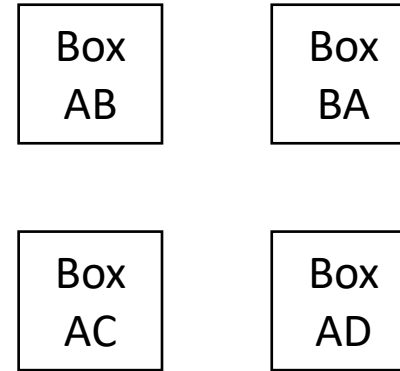
Tor is a network solution for anonymous communication on the internet based on obfuscation of this communication via different and variable routes

The Onion Router – Solution [22]



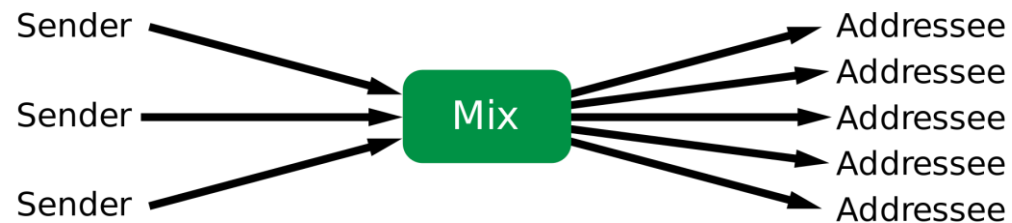
The Onion Router – Problems?

- ▶ Two problems remain
 - ▶ Access in succession
 - ▶ Habit in checking email



Intermezzo – Can we do better?

- ▶ Mix networks [28]?



- ▶ Dissent [29,30] or Vuvuzela [31] protocols?

But these solutions are synchronous!

Attempt 3: Three PETs

- ▶ PET 1: Transmission Protocol Using Public Bulletin Board
 - ▶ October 2015, see [22]
- ▶ PET 2: RIPOSTE Protocol
 - ▶ March 2015, see [23]
- ▶ PET 3: DP5 Protocol
 - ▶ June 2015, see [24]

PET 1: Transmission Protocol

"A secure and privacy friendly asynchronous unidirectional message transmission protocol using a public bulletin board that makes individual send or receive events unlinkable to one another."

"An asynchronous unidirectional private point-to-point message transmission protocol allows one user to send messages asynchronously to another user in private."

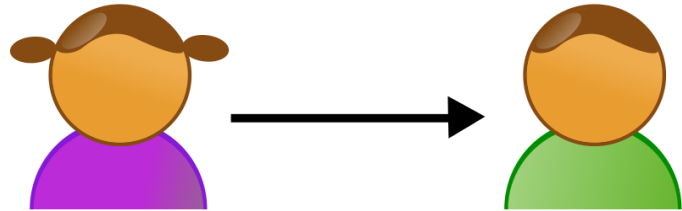
Requirements

- ▶ Correctness
- ▶ Confidentiality
- ▶ Integrity
- ▶ Availability
- ▶ Unlinkability of events
- ▶ Unlinkability of relationships
- ▶ Forward security
- ▶ Authenticity

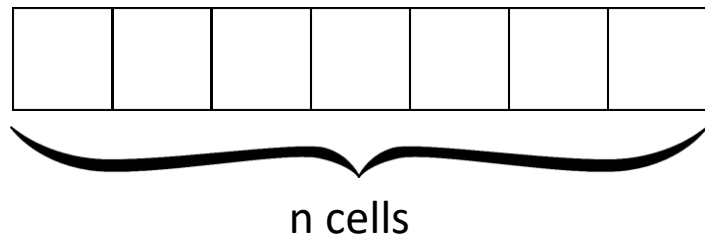


Ingredients

- ▶ (Unidirectional) protocol between e.g., Alice and Bob

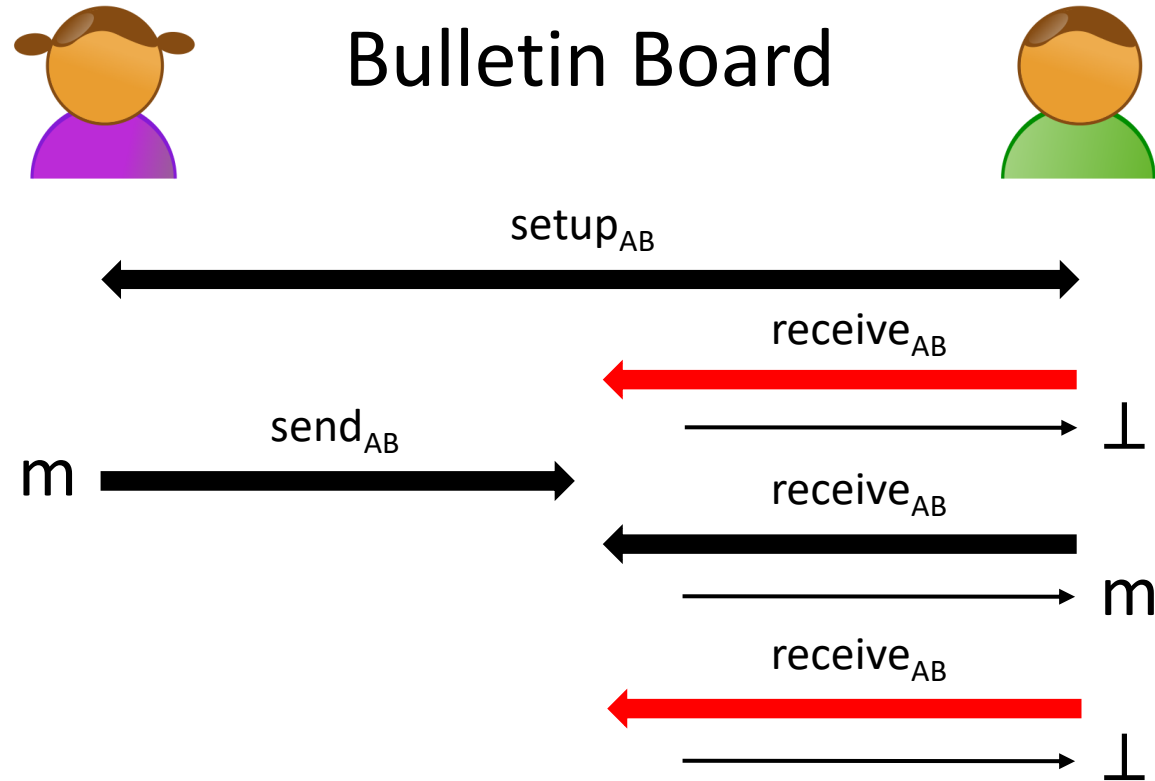


- ▶ Bulletin board

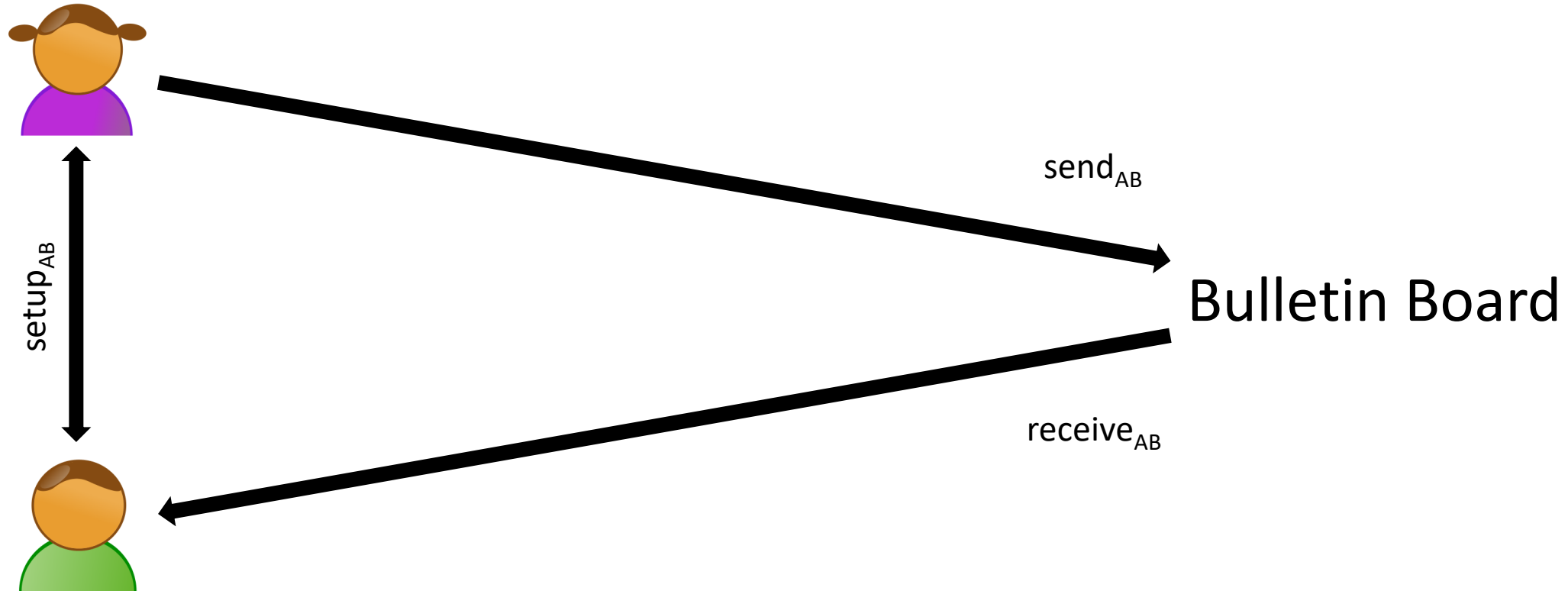


Simple Protocol

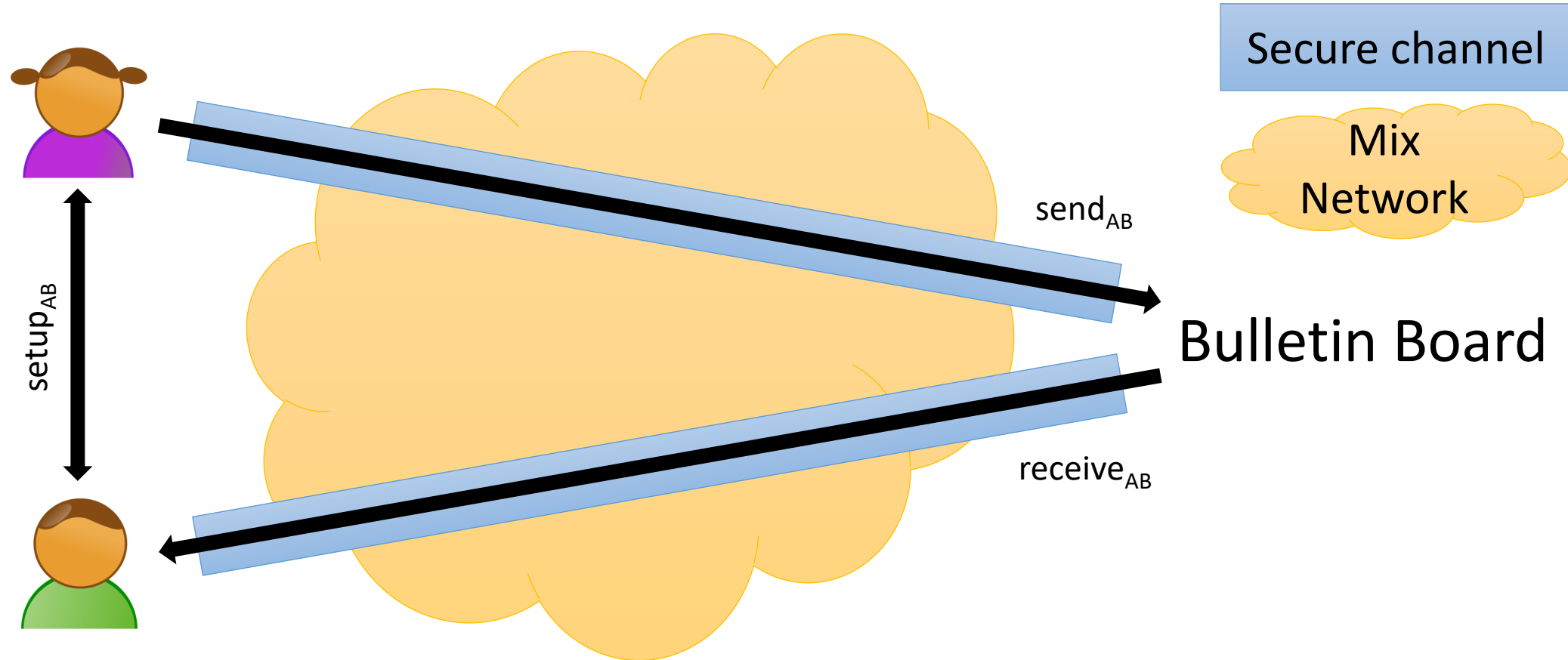
- ▶ setup_{AB}
- ▶ send_{AB}
- ▶ receive_{AB}



Simple Protocol?



High-Level Protocol



Bulletin Board

- ▶ n cells
- ▶ Randomly indexed: $B[i]$
- ▶ Value-tag pairs $\langle v, t \rangle$
- ▶ Hash $H(\cdot)$ with $t = H(b)$
- ▶ Two operations:
 - ▶ $\text{add}(i, v, t)$
 - ▶ $\text{get}(i, b)$

1	6	0	3	4	2	5
---	---	---	---	---	---	---

↓ $\text{add}(2, v, t)$

1	6	0	3	4	$\langle v, t \rangle$	5
---	---	---	---	---	------------------------	---

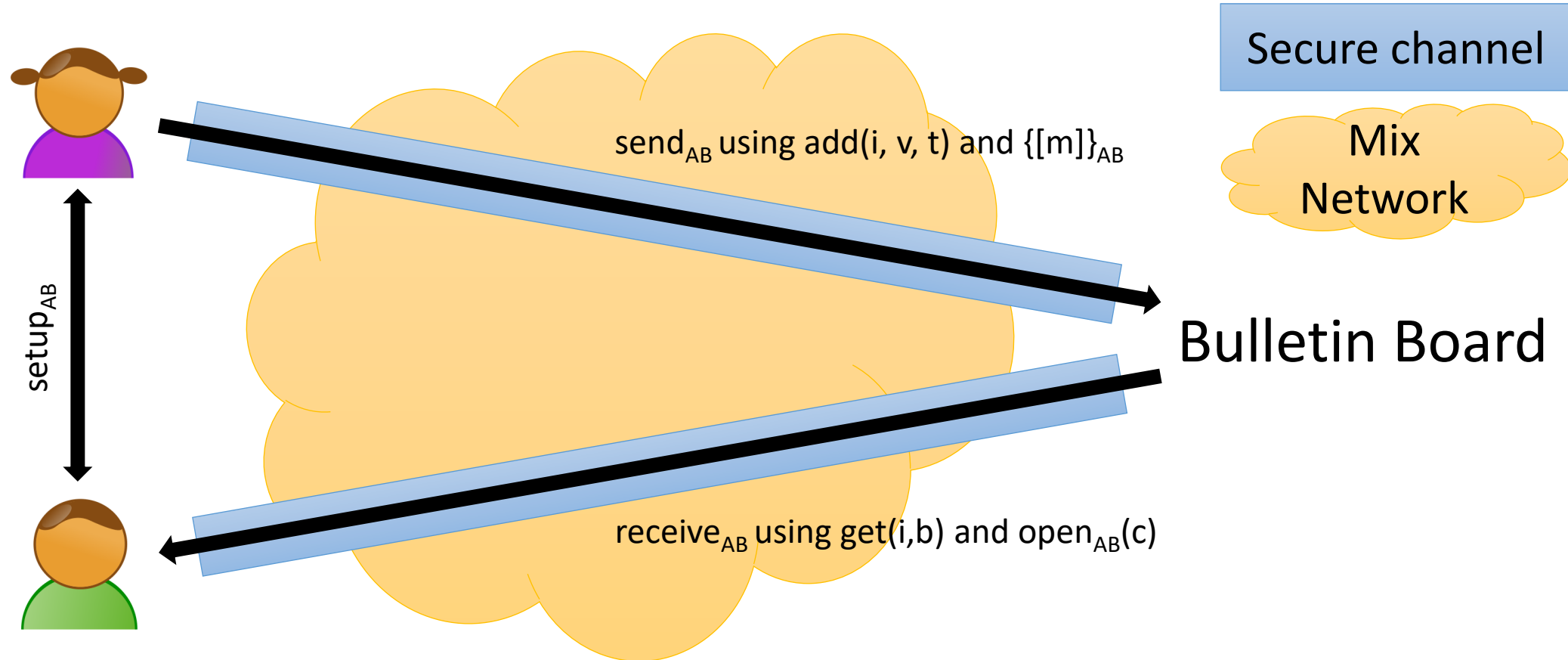
↓ $\text{get}(2, b)$

1	6	0	3	4	2	5
---	---	---	---	---	---	---

Secure Channel

- ▶ Authenticated encryption scheme
- ▶ Authenticated encryption: $c = \{[m]\}_{AB}$
 - ▶ Using K_{AB}
 - ▶ Assumed to leak no info on $\{[m']\}_{A'B'}$
- ▶ Decryption: $\text{open}_{AB}(c)$
 - ▶ Returns m or \perp

Overview Protocol



(Semi) Formal Notation

*send*_{AB}(*m*):

$idx' \in_R \{0, \dots, n-1\}$

$tag' \in_R T$

$u := \{[m || idx' || tag']\}_{AB}$

$add(idx_{AB}, u, H(tag_{AB}))$

$(idx_{AB}, tag_{AB}) := (idx', tag')$

$K_{AB} := KDF(K_{AB})$

*receive*_{AB}():

$u := get(idx_{AB}, tag_{AB})$

if $u \neq \perp \wedge$

$(m || idx' || tag') := open_{AB}(u)$

then $(idx_{AB}, tag_{AB}) := (idx', tag')$

$K_{AB} := KDF(K_{AB})$

return *m*

else return \perp

Requirements Check

- ▶ Requirements are met
- ▶ Unlinkability
 - ▶ Remember Tor example?
 - ▶ Mailboxes randomly allocated
 - ▶ Privacy
- ▶ Physical challenges
 - ▶ Synchronization
 - ▶ A lot of users?



Stronger Adversary?

- ▶ Availability
 - ▶ DoS attacks
 - ▶ Adversary fills every cell
- ▶ Unlinkability!
 - ▶ Block every user except Alice and Bob
 - ▶ Access cell?

Alternative to WhatsApp?

- ▶ Bidirectional use
- ▶ Online/offline messages



PET 2: Riposte

“Riposte is a new system for anonymous broadcast messaging. Riposte is the first such system, [...] that simultaneously protects against traffic-analysis attacks, prevents anonymous denial-of-service by malicious clients, and scales to million-user anonymity sets.”

Why Riposte?

- ▶ Low-latency anonymizing proxies are vulnerable to traffic analysis attacks
- ▶ Other anonymous messaging systems are not scalable

Properties

- ▶ Protects against traffic analysis attacks
- ▶ Prevents malicious clients from anonymously executing denial-of-service attacks
- ▶ Scales to anonymity set sizes of millions of users

Ingredients

- ▶ Database
- ▶ Write requests, split into shares
- ▶ Time epochs
- ▶ Coalition of (**non-colluding**) servers
- ▶ Two variants: Large network size vs. reliability

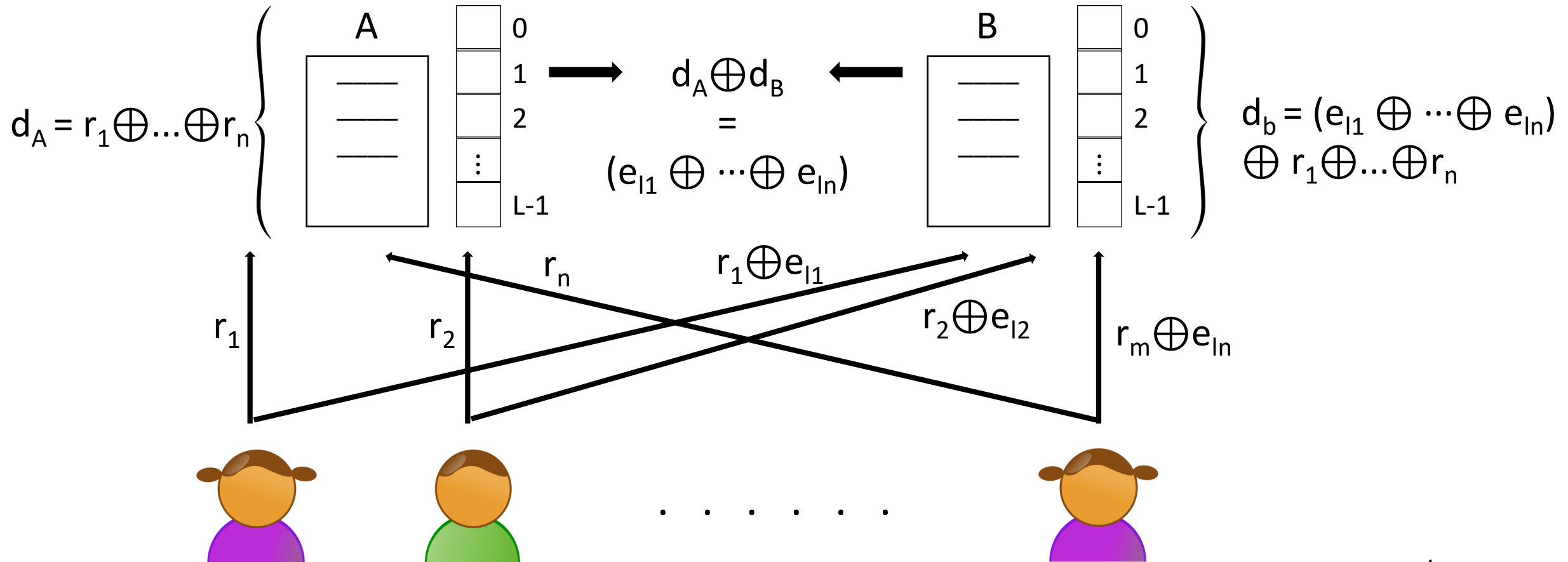


Security Goals

- ▶ Write-private - (s, t) -**Write Privacy**
- ▶ Disruption resistant – $\leq n$ compromised rows



Toy Protocol 1/2



Toy Protocol 2/2

n Clients:

$$\{r_1, \dots, r_n\} \in_R \{0, 1\}^L$$

$$\{e_{l1}, \dots, e_{ln}\}, e[l]=1$$

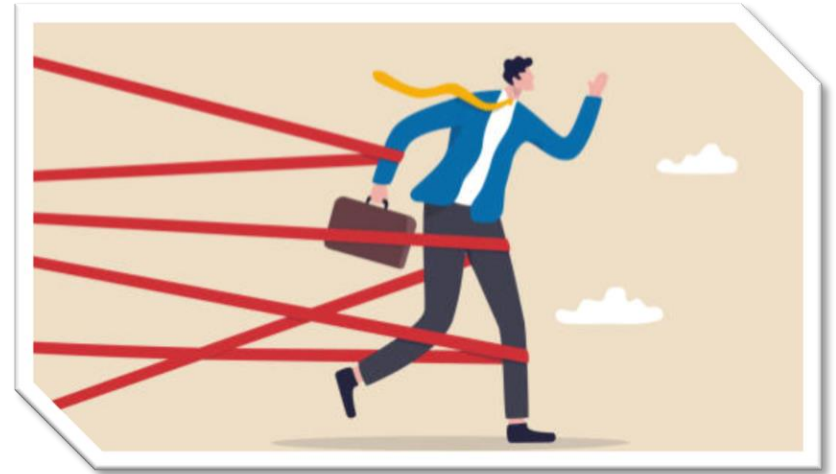
2 Servers:

$$d_A = r_1 \oplus \dots \oplus r_n$$

$$d_B = (e_{l1} \oplus \dots \oplus e_{ln}) \oplus (r_1 \oplus \dots \oplus r_n)$$

Limitations

- ▶ Bandwidth efficiency
- ▶ Disruption resistance



Intermezzo 1 – Collisions and Forward Secrecy 1/2

- ▶ Clients write at random locations in a database
- ▶ Chance of messages being overwritten
- ▶ *k-way* collisions, client *i* writes in each database cell
(m_i, m_i^2, \dots, m_i^k)
- ▶ By increasing number of clients *k*, we reduce the table size

Intermezzo 1 – Collisions and Forward Secrecy 2/2

- ▶ Adversary can compromise all servers
- ▶ n write requests, before epoch end
- ▶ Unlinkability
- ▶ Servers are honest at the moment requests are sent

Distributed Point Functions 1/4

- ▶ Converse of private information retrieval, e.g., write action
- ▶ Core building block is a distributed point function
- ▶ $P_{l,m}: \mathbb{Z}_L \rightarrow \mathbb{F}$, such that $P_{l,m}(l) = m$ and $P_{l,m}(l') = 0, l \neq l'$
- ▶ $P_{3,1} = (0, 0, 0, 1, 0)$, for $l \in (0, 1, 2, 3, 4)$

Distributed Point Functions 2/4

- ▶ Second step: (s, t) -distributed point function
 - ▶ $\text{Gen}(l, m) \rightarrow (k_0, \dots, k_{s-1})$
 - ▶ $\text{Eval}(k, l') \rightarrow m'$

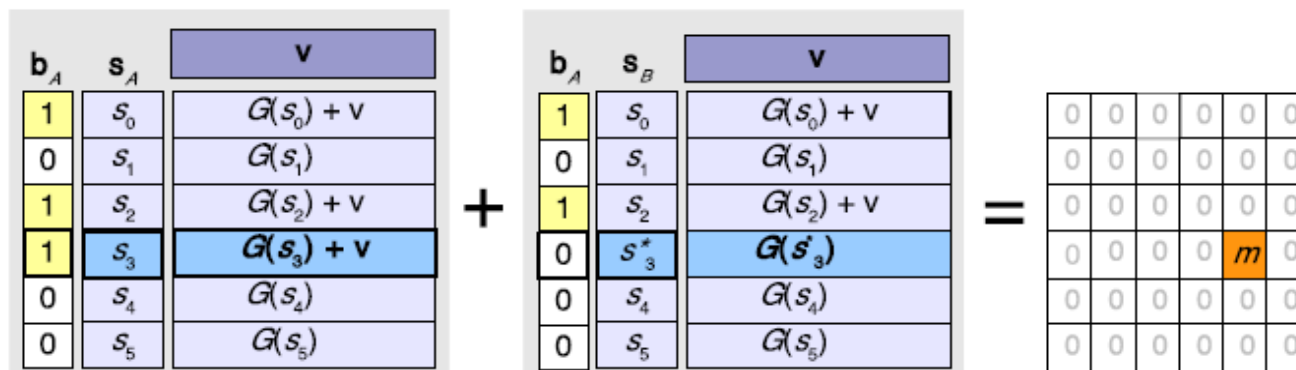
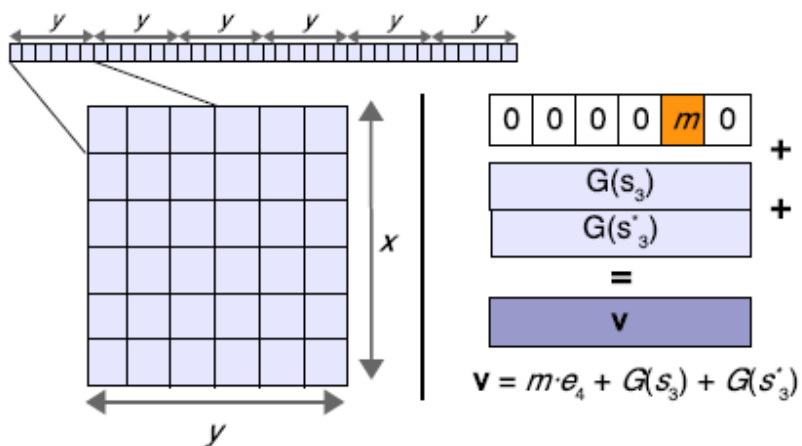
Distributed Point Functions 3/4

- ▶ Toy Construction: $(s, s-1)$ -distributed point function
 - ▶ $\text{Gen}(l, m) \rightarrow (k_0, \dots, k_{s-1})$
 - ▶ k_0, \dots, k_{s-2} generated randomly
 - ▶ $k_{s-1} = m \times e_l - \sum_{i=0}^{s-2} k_i$
 - ▶ $\text{Eval}(k, l') \rightarrow m'$

Distributed Point Functions 4/4

- ▶ Generalization:
 - ▶ s servers, $\leq t$ collude
 - ▶ \mathbb{F}^L database state
 - ▶ Client uses (s, t) -distributed point function to generate s DPF keys
 - ▶ Client sends one key per server
 - ▶ Server adds this key to its database state

Two Server Scheme 1/2



Two Server Scheme 2/2

▶ Step 1:

- ▶ PRG $G: \mathbb{S} \rightarrow \mathbb{F}^y$
- ▶ $\text{Gen}(l, m) \rightarrow (k_A, k_B)$
- ▶ $l = l_x y + l_y$, with $l_x \in \mathbb{Z}_x$, $l_y \in \mathbb{Z}_y$ and $xy \geq L$
- ▶ $\mathbf{b}_A \in_R \{0, 1\}^x$, $\mathbf{s}_A \in_R \mathbb{S}^x$ and $s_{l_x}^* \in_R \mathbb{S}^x$
 - ▶ $\mathbf{b}_A = (b_0, \dots, b_{l_x}, \dots, b_{x-1})$
 - ▶ $\mathbf{b}_B = (b_0, \dots, \overline{b_{l_x}}, \dots, b_{x-1})$
 - ▶ $\mathbf{s}_A = (s_0, \dots, s_{l_x}, \dots, s_{x-1})$
 - ▶ $\mathbf{s}_B = (s_0, \dots, s_{l_x}^*, \dots, s_{x-1})$
- ▶ $\mathbf{v} = m \times e_l + G(s_{l_x}) + G(s_{l_x}^*)$
- ▶ $k_A = (\mathbf{b}_A, \mathbf{s}_A, \mathbf{v})$ and $k_B = (\mathbf{b}_B, \mathbf{s}_B, \mathbf{v})$

▶ Step 2:

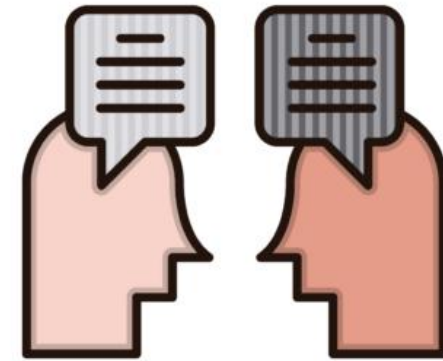
- ▶ $\text{Eval}(k, l') \rightarrow m'$
 - ▶ $k = (\mathbf{b}, \mathbf{s}, \mathbf{v})$
 - ▶ $l' = l'_x y + l'_y$, with $l'_x \in \mathbb{Z}_x$, $l'_y \in \mathbb{Z}_y$ and $xy \geq L$
- ▶ $\mathbf{g} = G(\mathbf{s}[l'_x])$
- ▶ $\mathbf{m}' = (\mathbf{g}[l'_y] + \mathbf{b}[l'_x] \mathbf{v}[l'_y])$

Disruptors Prevention

- ▶ Previous protocols only addressed the bandwidth efficiency
- ▶ Malicious servers could malformed keys, and thus gain additional information about users
- ▶ Option 1: Third non-colluding party
- ▶ Option 2: Expensive zero-knowledge proofs

Discussion

- ▶ Novel applications of PIRs and secure multiparty computation techniques
- ▶ Practical protocol, capable of handling big anonymity sets
- ▶ Protects whistle-blowers
- ▶ Is it enough?



Conclusion

- ▶ Social and legal perspectives
- ▶ Different attempts at anonymizing asynchronous messaging
- ▶ Open questions remain



Questions?

73 / 85

References

- ▶ [1] TTEC. What is Asynchronous Messaging? 2022.
url: <https://www.ttec.com/glossary/asynchronous-messaging>.
- ▶ [2] Nicolas Ortega. How Facebook Undermines Privacy Protections for Its 2 Billion WhatsApp Users. 2021. url:
<https://www.propublica.org/article/how-facebook-undermines-privacy-protections-for-its-2-billion-whatsapp-users>.
- ▶ [3] Glenn Greenwald. NSA collecting phone records of millions of Verizon customers daily. June 2013.
url: <https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order>.

References

- ▶ [4] The Guardian. The NSA files. 2022. url: <https://www.theguardian.com/us-news/the-nsa-files>.
- ▶ [5] Andy Kroll. FBI Document Says the Feds Can Get Your WhatsApp Data — in Real Time. November 2021. url: <https://www.rollingstone.com/politics/politics-features/whatsapp-message-facebook-apple-fbi-privacy-1261816/>.
- ▶ [6] PROPERTY OF THE PEOPLE. Jan. 2021 FBI Infographic re Lawful Access to Secure Messaging Apps Data. November 2021. url: <https://propertyofthepeople.org/document-detail/?doc-id=21114562>.

References

- ▶ [7] Julian Huijbregts. Belgische overheid verplicht chatapps zoals WhatsApp metadata op te slaan. June 2021. url: <https://tweakers.net/nieuws/183702/belgische-overheid-verplicht-chatapps-zoals-whatsapp-metadata-op-te-slaan.html>
- ▶ [8] Dark Reading Staff. Tech Insight: How Attackers Use Your Metadata Against You. February 2009. url: <https://www.darkreading.com/vulnerabilities-threats/-em-tech-insight-em-how-attackers-use-your-metadata-against-you>

References

- ▶ [9] Oxford Reference. Social graph.
url: https://www.oxfordreference.com/view/10.1093/acref/9780195392883.001.0001/m_en_us14436
- ▶ [10] Sangeet Paul Choudary. The Rise of Social Graphs for Businesses. February 2015. url: <https://hbr.org/2015/02/the-rise-of-social-graphs-for-businesses>
- ▶ [11] Johns Hopkins University. The Johns Hopkins Foreign Affairs Symposium Presents: The Price of Privacy: Re-Evaluating the NSA, April 2014. url: <https://www.youtube.com/watch?v=kV2HDM86Xgl>. A Debate.

References

- ▶ [12] Arooj Ahmed. Signal and Telegram have witnessed a rise of 1200% in usage before the implementation of the controversial WhatsApp privacy policy. May 2021. url: <https://www.digitalinformationworld.com/2021/05/signal-and-telegram-have-witnessed-rise.html>
- ▶ [13] Oxford English Dictionary. whistle, n. url: <https://www.oed.com/view/Entry/22854>
- ▶ [14] Press Release. REPRESENTATIVES LIEU AND BEYER RELEASE RESOURCE GUIDE FOR FEDERAL EMPLOYEES TO SHARE KEY INFORMATION WITH PUBLIC. February 2017. url: <https://lieu.house.gov/media-center/press-releases/representatives-lieu-and-beyer-release-resource-guide-federal-employees>

References

- ▶ [15] Sarah Ellison. How May Edwards became the forgotten whistleblower. July 2021. url: https://www.washingtonpost.com/lifestyle/media/may-edwards-treasury-buzzfeed-fincen-whistleblower/2021/07/07/28f5c1f8-da05-11eb-8fb8-aea56b785b00_story.html.
- ▶ [16] BuzzFeed News. The Money Trail. url: <https://www.buzzfeednews.com/collection/themoneytrail>.
- ▶ [17] Martin Shelton. Read This Before You Whistleblow With an App. October 2019. url: <https://onezero.medium.com/read-this-before-you-whistleblow-with-an-app-f193114e7596>. Downloaded File

References

- ▶ [18] Court of Justice of the European Union. The Court of Justice declares the Data Retention Directive to be invalid, April 2014. url: https://curia.europa.eu/jcms/jcms/Jo2_7052/en/?annee=2014. Judgment of the Court of Justice in Joined Cases C-293/12, C-594/12. Digital Rights Ireland. No 54/2014
- ▶ [19] Statewatch. EU: Data retention strikes back? Options for mass telecoms surveillance under discussion again. December 2021. url: <https://www.statewatch.org/news/2021/december/eu-data-retention-strikes-back-options-for-mass-telecoms-surveillance-under-discussion-again/>

References

- ▶ [20] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal, L 119:1–88, April 2016.
- ▶ [21] United Nations High Commissioner for Human Rights. Report on the right to privacy in the digital age. August 2018. url: <https://www.ohchr.org/en/calls-for-input/reports/2018/report-right-privacy-digital-age>. Downloaded File.

References

- ▶ [22] Jaap-Henk Hoepman. Privately (and unlinkably) exchanging messages using a public bulletin board. In Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society. ACM, October 2015. doi: 10.1145/2808138.2808142.
- ▶ [23] Henry Corrigan-Gibbs, Dan Boneh, and David Mazières. Riposte: an anonymous messaging system handling millions of users, March 2015. arXiv: 1503.06115 [cs.CR].
- ▶ [24] Nikita Borisov, George Danezis, and Ian Goldberg. DP5: a private presence service. Proceedings on Privacy Enhancing Technologies, 2015(2):4–24, June 2015. doi: 10.1515/popets-2015-0008.

References

- ▶ [25] BBC. Why you probably shouldn't worry about the WhatsApp security 'bug'. January 2017. url: <https://www.bbc.com/news/newsbeat-38608762>
- ▶ [26] Myra Security GmbH. What is the Tor network? 2022. url: <https://www.myrasecurity.com/en/tor-network/>
- ▶ [27] Inc The Tor Project. Tor — History. url: <https://www.torproject.org/about/history/>.
- ▶ [28] Wikipedia. Mix network. May 2022. url: https://en.wikipedia.org/wiki/Mix_network

References

- ▶ [29] Henry Corrigan-Gibbs and Bryan Ford. Dissent: accountable anonymous group messaging. In Ehab Al-Shaer, Angelos D. Keromytis, and Vitaly Shmatikov, editors, Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS 2010, Chicago, Illinois, USA, October 4-8, 2010, pages 340–350. ACM, 2010.
- ▶ [30] David Isaac Wolinsky, Henry Corrigan-Gibbs, Bryan Ford, and Aaron Johnson. Dissent in numbers: Making strong anonymity scale. In Chandu Thekkath and Amin Vahdat, editors, 10th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2012, Hollywood, CA, USA, October 8-10, 2012, pages 179–182. USENIX Association, 2012.

References

- ▶ [31] Jelle van den Hooff, David Lazar, Matei Zaharia, and Nickolai Zeldovich. Vuvuzela: Scalable private messaging resistant to traffic analysis. In HotPETs 2015, Philadelphia, PA, USA, July 2 2015. (full version to appear).