Hey there! You are using WhatsApp

Enumerating Three Billion Accounts for Security and Privacy

Gabriel Gegenhuber, Philipp Frenzel, Maximilian Günther, Johanna Ullrich, Aljosha Judmayer

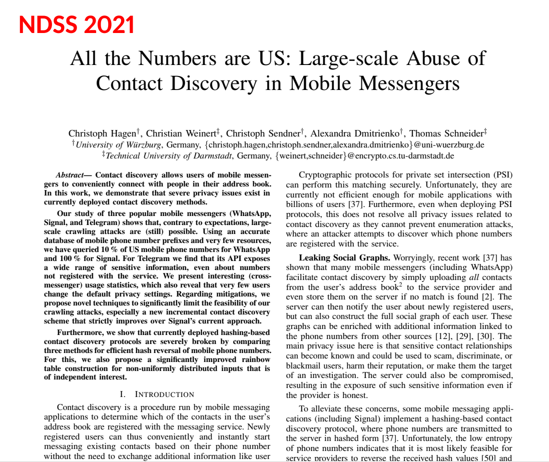
WhatsApp: World’s Largest Messenger

- WhatsApp is incredibly popular
    - Used for personal and business communication
    - Official statement: more than ~~2 billion~~ 3 billion users
    - Almost public utility?
    - Certainly an interesting research target :)

Contact Discovery in Instant Messengers

- Instant messengers **need a way of discovering other peers**
  - Often telephone number is used as identifier
  - Users need to be able to legitimately discover other accounts
  - **Inherently vulnerable to enumeration attacks**
    - Attacker can learn whether target uses a service and potentially retrieve other metadata (e.g., profile picture)
    - Attacker can generate and **probe predictable IDs** (i.e., phone numbers)

Previous Work



NDSS 2021

All the Numbers are US: Large-scale Abuse of Contact Discovery in Mobile Messengers

Christoph Hagen[†], Christian Weinert[‡], Christoph Sendner[†], Alexandra Dmitrienko[†], Thomas Schneider[‡]
[†]*University of Würzburg*, Germany, {christoph.hagen,christoph.sendner,alexandra.dmitrienko}@uni-wuerzburg.de
[‡]*Technical University of Darmstadt*, Germany, {weinert,schneider}@encrypto.cs.tu-darmstadt.de

***Abstract*— Contact discovery allows users of mobile messengers to conveniently connect with people in their address book. In this work, we demonstrate that severe privacy issues exist in currently deployed contact discovery methods.**

Our study of three popular mobile messengers (WhatsApp, Signal, and Telegram) shows that, contrary to expectations, large-scale crawling attacks are (still) possible. Using an accurate database of mobile phone number prefixes and very few resources, we have queried 10 % of US mobile phone numbers for WhatsApp and 100 % for Signal. For Telegram we find that its API exposes a wide range of sensitive information, even about numbers not registered with the service. We present interesting (cross-messenger) usage statistics, which also reveal that very few users change the default privacy settings. Regarding mitigations, we propose novel techniques to significantly limit the feasibility of our crawling attacks, especially a new incremental contact discovery scheme that strictly improves over Signal's current approach.

Furthermore, we show that currently deployed hashing-based contact discovery protocols are severely broken by comparing three methods for efficient hash reversal of mobile phone numbers. For this, we also propose a significantly improved rainbow table construction for non-uniformly distributed inputs that is of independent interest.

I. Introduction

Contact discovery is a procedure run by mobile messaging applications to determine which of the contacts in the user's address book are registered with the messaging service. Newly registered users can thus conveniently and instantly start messaging existing contacts based on their phone number without the need to exchange additional information like user

Cryptographic protocols for private set intersection (PSI) can perform this matching securely. Unfortunately, they are currently not efficient enough for mobile applications with billions of users [37]. Furthermore, even when deploying PSI protocols, this does not resolve all privacy issues related to contact discovery as they cannot prevent enumeration attacks, where an attacker attempts to discover which phone numbers are registered with the service.

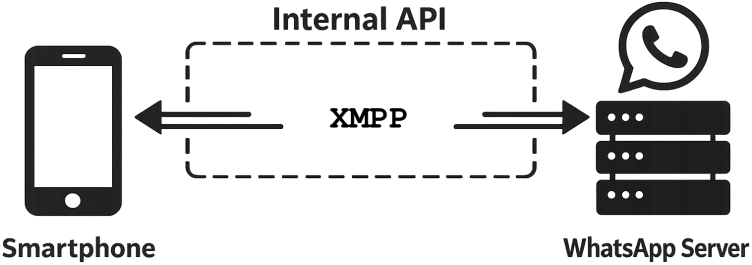
Leaking Social Graphs. Worryingly, recent work [37] has shown that many mobile messengers (including WhatsApp) facilitate contact discovery by simply uploading *all* contacts from the user's address book[2] to the service provider and even store them on the server if no match is found [2]. The server can then notify the user about newly registered users, but can also construct the full social graph of each user. These graphs can be enriched with additional information linked to the phone numbers from other sources [12], [29], [30]. The main privacy issue here is that sensitive contact relationships can become known and could be used to scam, discriminate, or blackmail users, harm their reputation, or make them the target of an investigation. The server could also be compromised, resulting in the exposure of such sensitive information even if the provider is honest.

To alleviate these concerns, some mobile messaging applications (including Signal) implement a hashing-based contact discovery protocol, where phone numbers are transmitted to the server in hashed form [37]. Unfortunately, the low entropy of phone numbers indicates that it is most likely feasible for service providers to reverse the received hash values [50] and

- Investigating account enumeration at WhatsApp
- Via **app-automation**, Combined multiple (emulated) devices over longer time
- **Rate-limit 2019/09: 60,000 accounts / day**
- Overall queried about 50 M accounts

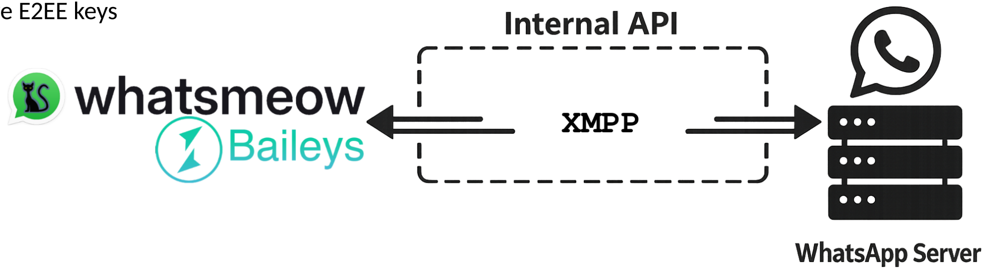
WhatsApp: Client to Server Communication (Official Client)

- Enumeration vector = telephone book sync
- Method restricted to UI automation
  - **Little/No control over individual API queries**

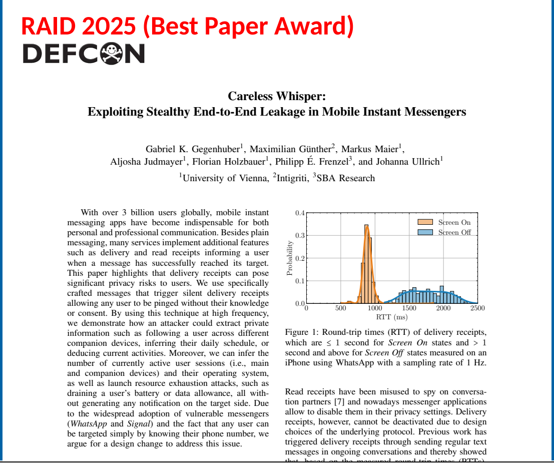


WhatsApp: Client to Server Communication (Reverse Engineered OSS)

- **Full control over API queries** (respecting protocol and allowed syntax)
- Potentially, **many different enumeration vectors**
  - Telephone book sync
  - Retrieve profile picture
  - Retrieve E2EE keys



(Our) Previous Work and Responsible Disclosures



RAID 2025 (Best Paper Award)

DEFCON

Careless Whisper: Exploiting Stealthy End-to-End Leakage in Mobile Instant Messengers

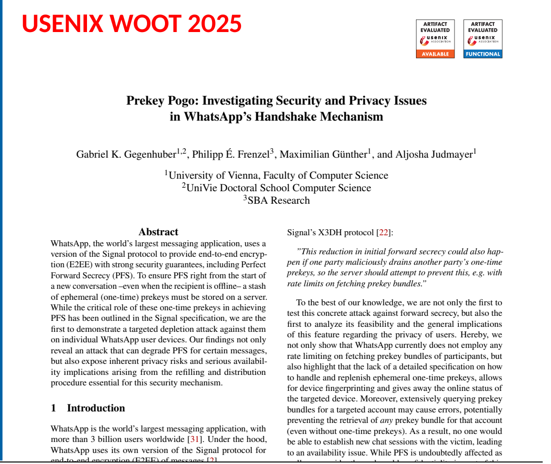
Gabriel K. Gegenhuber[1], Maximilian Günther[2], Markus Maier[1], Aljosha Judmayer[1], Florian Holzbauer[1], Philipp É. Frenzel[3], and Johanna Ullrich[1]

[1]University of Vienna, [2]Intigriti, [3]SBA Research

With over 3 billion users globally, mobile instant messaging apps have become indispensable for both personal and professional communication. Besides plain messaging, many services implement additional features such as delivery and read receipts informing a user when a message has successfully reached its target. This paper highlights that delivery receipts can pose significant privacy risks to users. We use specifically crafted messages that trigger silent delivery receipts allowing any user to be pinged without their knowledge or consent. By using this technique at high frequency, we demonstrate how an attacker could extract private information such as following a user across different companion devices, inferring their daily schedule, or deducing current activities. Moreover, we can infer the number of currently active user sessions (i.e., main and companion devices) and their operating system, as well as launch resource exhaustion attacks, such as draining a user's battery or data allowance, all without generating any notification on the target side. Due to the widespread adoption of vulnerable messengers (*WhatsApp* and *Signal*) and the fact that any user can be targeted simply by knowing their phone number, we argue for a design change to address this issue.

Figure 1: Round-trip times (RTT) of delivery receipts, which are ≤ 1 second for *Screen On* states and > 1 second and above for *Screen Off* states measured on an iPhone using WhatsApp with a sampling rate of 1 Hz.

Read receipts have been misused to spy on conversation partners [7] and nowadays messenger applications allow to disable them in their privacy settings. Delivery receipts, however, cannot be deactivated due to design choices of the underlying protocol. Previous work has triggered delivery receipts through sending regular text messages in ongoing conversations and thereby showed that, based on the measured round-trip times (RTTs)



USENIX WOOT 2025

ARTIFACT EVALUATED usenix ASSOCIATION AVAILABLE

ARTIFACT EVALUATED usenix ASSOCIATION FUNCTIONAL

Prekey Pogo: Investigating Security and Privacy Issues in WhatsApp's Handshake Mechanism

Gabriel K. Gegenhuber[1,2], Philipp É. Frenzel[3], Maximilian Günther[1], and Aljosha Judmayer[1]

[1]University of Vienna, Faculty of Computer Science
[2]UniVie Doctoral School Computer Science
[3]SBA Research

Abstract

WhatsApp, the world's largest messaging application, uses a version of the Signal protocol to provide end-to-end encryption (E2EE) with strong security guarantees, including Perfect Forward Secrecy (PFS). To ensure PFS right from the start of a new conversation –even when the recipient is offline– a stash of ephemeral (one-time) prekeys must be stored on a server. While the critical role of these one-time prekeys in achieving PFS has been outlined in the Signal specification, we are the first to demonstrate a targeted depletion attack against them on individual WhatsApp user devices. Our findings not only reveal an attack that can degrade PFS for certain messages, but also expose inherent privacy risks and serious availability implications arising from the refilling and distribution procedure essential for this security mechanism.

1 Introduction

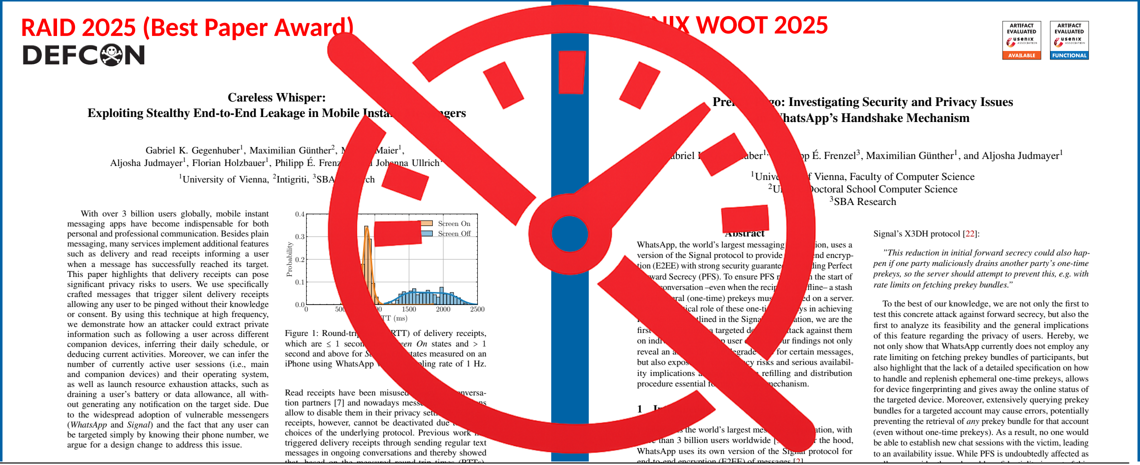
WhatsApp is the world's largest messaging application, with more than 3 billion users worldwide [31]. Under the hood, WhatsApp uses its own version of the Signal protocol for end-to-end encryption (E2EE) of messages [2].

Signal's X3DH protocol [22]:

> *"This reduction in initial forward secrecy could also happen if one party maliciously drains another party's one-time prekeys, so the server should attempt to prevent this, e.g. with rate limits on fetching prekey bundles."*

To the best of our knowledge, we are not only the first to test this concrete attack against forward secrecy, but also the first to analyze its feasibility and the general implications of this feature regarding the privacy of users. Hereby, we not only show that WhatsApp currently does not employ any rate limiting on fetching prekey bundles of participants, but also highlight that the lack of a detailed specification on how to handle and replenish ephemeral one-time prekeys, allows for device fingerprinting and gives away the online status of the targeted device. Moreover, extensively querying prekey bundles for a targeted account may cause errors, potentially preventing the retrieval of *any* prekey bundle for that account (even without one-time prekeys). As a result, no one would be able to establish new chat sessions with the victim, leading to an availability issue. While PFS is undoubtedly affected as

(Our) Previous Work and Responsible Disclosures



RAID 2025 (Best Paper Award)

DEFCON

Careless Whisper:
Exploiting Stealthy End-to-End Leakage in Mobile Insta… …gers

Gabriel K. Gegenhuber[1], Maximilian Günther[2], M… …Maier[1],
Aljosha Judmayer[1], Florian Holzbauer[1], Philipp É. Frenz… …l Johanna Ullrich…
[1]University of Vienna, [2]Intigriti, [3]SBA… …ch

With over 3 billion users globally, mobile instant messaging apps have become indispensable for both personal and professional communication. Besides plain messaging, many services implement additional features such as delivery and read receipts informing a user when a message has successfully reached its target. This paper highlights that delivery receipts can pose significant privacy risks to users. We use specifically crafted messages that trigger silent delivery receipts allowing any user to be pinged without their knowledge or consent. By using this technique at high frequency, we demonstrate how an attacker could extract private information such as following a user across different companion devices, inferring their daily schedule, or deducing current activities. Moreover, we can infer the number of currently active user sessions (i.e., main and companion devices) and their operating system, as well as launch resource exhaustion attacks, such as draining a user's battery or data allowance, all without generating any notification on the target side. Due to the widespread adoption of vulnerable messengers (*WhatsApp* and *Signal*) and the fact that any user can be targeted simply by knowing their phone number, we argue for a design change to address this issue.

Figure 1: Round-tri… …RTT) of delivery receipts, which are ≤ 1 secon… …een On* states and > 1 second and above for *S*… …tates measured on an iPhone using WhatsApp … …pling rate of 1 Hz.

Read receipts have been misused… …onversa-tion partners [7] and nowadays messe… …ns allow to disable them in their privacy setti… receipts, however, cannot be deactivated due t… choices of the underlying protocol. Previous work h… triggered delivery receipts through sending regular text messages in ongoing conversations and thereby showed that, based on the measured round-trip times (RTTs)

…NIX WOOT 2025

ARTIFACT EVALUATED usenix ASSOCIATION AVAILABLE

ARTIFACT EVALUATED usenix ASSOCIATION FUNCTIONAL

Pre… …go: Investigating Security and Privacy Issues
…hatsApp's Handshake Mechanism

…abriel I… …uber[1], … …p É. Frenzel[3], Maximilian Günther[1], and Aljosha Judmayer[1]
[1]Unive… …f Vienna, Faculty of Computer Science
[2]U… …octoral School Computer Science
[3]SBA Research

Abstract

WhatsApp, the world's largest messaging… …ion, uses a version of the Signal protocol to provide… …end encryp-tion (E2EE) with strong security guarante… …ding Perfect …ward Secrecy (PFS). To ensure PFS r… …n the start of …conversation –even when the recip… …fline– a stash …eral (one-time) prekeys mus… …ed on a server. …tical role of these one-ti… …ys in achieving … …tlined in the Signa… …ation, we are the first… …a targeted de… …ttack against them on indi… …p user … …ur findings not only reveal an a… …legrade … for certain messages, but also expos… …cy risks and serious availabil-ity implications a… …e refilling and distribution procedure essential fo… …mechanism.

1 In…

…s the world's largest mes… …ation, with …e than 3 billion users worldwide [… …r the hood, WhatsApp uses its own version of the Signal protocol for end-to-end encryption (E2EE) of messages [2].

Signal's X3DH protocol [22]:

> *"This reduction in initial forward secrecy could also happen if one party maliciously drains another party's one-time prekeys, so the server should attempt to prevent this, e.g. with rate limits on fetching prekey bundles."*

To the best of our knowledge, we are not only the first to test this concrete attack against forward secrecy, but also the first to analyze its feasibility and the general implications of this feature regarding the privacy of users. Hereby, we not only show that WhatsApp currently does not employ any rate limiting on fetching prekey bundles of participants, but also highlight that the lack of a detailed specification on how to handle and replenish ephemeral one-time prekeys, allows for device fingerprinting and gives away the online status of the targeted device. Moreover, extensively querying prekey bundles for a targeted account may cause errors, potentially preventing the retrieval of *any* prekey bundle for that account (even without one-time prekeys). As a result, no one would be able to establish new chat sessions with the victim, leading to an availability issue. While PFS is undoubtedly affected as

Finding a Phone Number Input Dataset

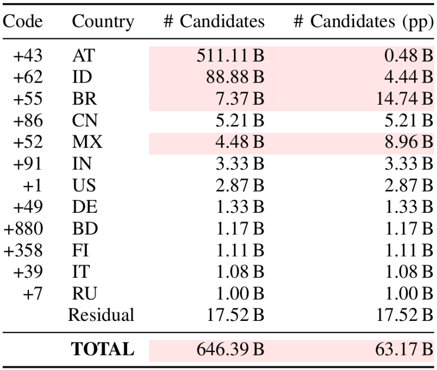
- E.g., a public phone book, previous data leak
- In 2021: dataset of **500 M Facebook users was publicly leaked** (data was scraped in 2019)
  - Containing phone number, name, email, birth date, employer, etc.
  - Once posted on the Internet, **leak cannot be undone**
- Two initial research questions
  - *RQ1: Is WhatsApp still/again **vulnerable to large-scale phone number enumeration** in 2025?*
  - *RQ2: **How persistent** are **potential harms** caused by public data leaks?*

Enumerating the 2021 Facebook Leak Dataset

- Facebook 2021 Leak (500 M) was **fully crawled in 5 hours**
- We wrote a generator for **all existing mobile numbers in the US numbering plan**
  - 2.87 B possible mobile phone numbers
  - Enumeration finished after 1 day
- Surprisingly, no accounts banned after one week
  - *RQ3: How quickly does the **service operator detect large-scale (global?) data exfiltration** incidents?*

Finding a Phone Number Input Dataset (Continued)

- According to international phone number format, **up to 15 digits allowed**
  - One quadrillion combinations, **exhaustive enumeration not feasible**
- Google's *libphonenumber* to write a global phone number generator
  - Some countries required post-processing, leading to 63 B candiate numbers



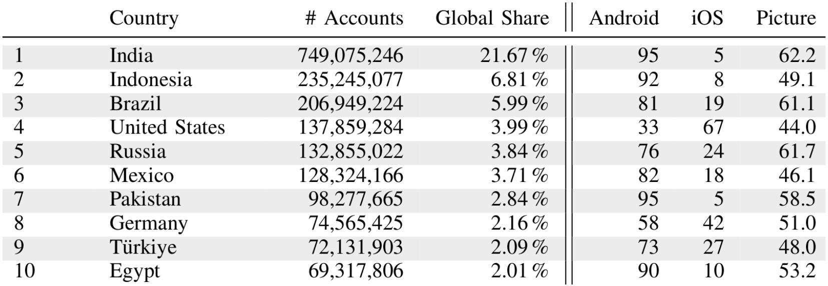
Code	Country	# Candidates	# Candidates (pp)
+43	AT	511.11 B	0.48 B
+62	ID	88.88 B	4.44 B
+55	BR	7.37 B	14.74 B
+86	CN	5.21 B	5.21 B
+52	MX	4.48 B	8.96 B
+91	IN	3.33 B	3.33 B
+1	US	2.87 B	2.87 B
+49	DE	1.33 B	1.33 B
+880	BD	1.17 B	1.17 B
+358	FI	1.11 B	1.11 B
+39	IT	1.08 B	1.08 B
+7	RU	1.00 B	1.00 B
	Residual	17.52 B	17.52 B
	TOTAL	646.39 B	63.17 B

TABLE II

libphonegen YIELDS 63.17 B PHONE NUMBERS FOR ENUMERATION. HIGHLIGHTED COUNTRIES REQUIRED POST-PROCESSING (PP).

Global User Enumeration: 3.5 B Active Accounts

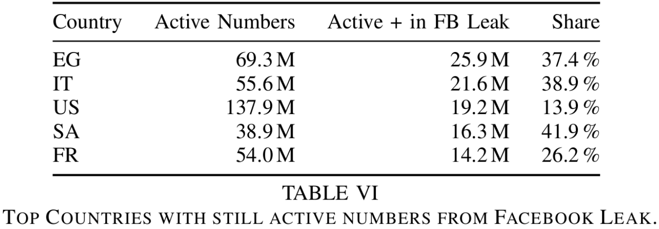
- In total, **3.5 B active WhatsApp users**
  - 63 B numbers could be queried in less than a month
- Substantial share of world population; WhatsApp especially popular in South America and Europe
- Individual + macroscopic insights: Android: 81% vs. iOS: 19%; Public profile picture: 57%



	Country	# Accounts	Global Share	Android	iOS	Picture
1	India	749,075,246	21.67 %	95	5	62.2
2	Indonesia	235,245,077	6.81 %	92	8	49.1
3	Brazil	206,949,224	5.99 %	81	19	61.1
4	United States	137,859,284	3.99 %	33	67	44.0
5	Russia	132,855,022	3.84 %	76	24	61.7
6	Mexico	128,324,166	3.71 %	82	18	46.1
7	Pakistan	98,277,665	2.84 %	95	5	58.5
8	Germany	74,565,425	2.16 %	58	42	51.0
9	Türkiye	72,131,903	2.09 %	73	27	48.0
10	Egypt	69,317,806	2.01 %	90	10	53.2

Active Numbers from Facebook 2021 Leak

- The majority (58 %) of the phone numbers were still (or again) active (> 6 years after being scraped)
  - Not possible to *undo* a leak
  - Leak contains additional detail (e.g., employer information)
    - Could be abused for targeted scams

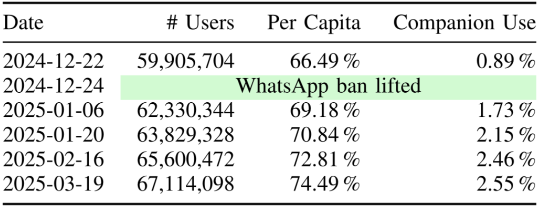


Country	Active Numbers	Active + in FB Leak	Share
EG	69.3 M	25.9 M	37.4 %
IT	55.6 M	21.6 M	38.9 %
US	137.9 M	19.2 M	13.9 %
SA	38.9 M	16.3 M	41.9 %
FR	54.0 M	14.2 M	26.2 %

TABLE VI
TOP COUNTRIES WITH STILL ACTIVE NUMBERS FROM FACEBOOK LEAK.

Countries Banning WhatsApp

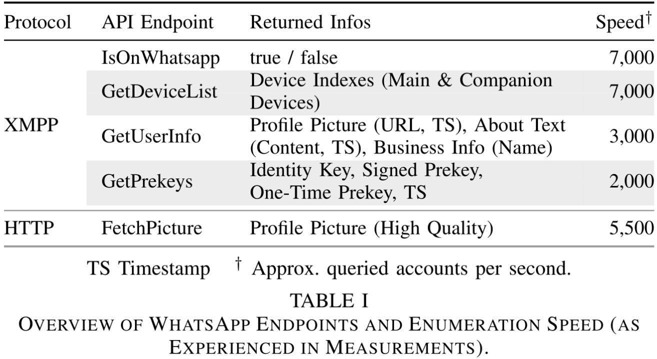
- WhatsApp bans in China, Iran, Myanmar, and North Korea
- Active accounts for all mentioned countries (ranging from 2.3 M in CN to 5 accounts in KP)
  - Considerable personal risks for users
- Iran lifted the ban during our measurements



Date	# Users	Per Capita	Companion Use
2024-12-22	59,905,704	66.49 %	0.89 %
2024-12-24	WhatsApp ban lifted		
2025-01-06	62,330,344	69.18 %	1.73 %
2025-01-20	63,829,328	70.84 %	2.15 %
2025-02-16	65,600,472	72.81 %	2.46 %
2025-03-19	67,114,098	74.49 %	2.55 %

TABLE IV
USERS IN IRAN (POPULATION 90.1 M) DESPITE WHATSAPP BAN.

Testing Different API Endpoints



Protocol	API Endpoint	Returned Infos	Speed†
XMPP	IsOnWhatsapp	true / false	7,000
	GetDeviceList	Device Indexes (Main & Companion Devices)	7,000
	GetUserInfo	Profile Picture (URL, TS), About Text (Content, TS), Business Info (Name)	3,000
	GetPrekeys	Identity Key, Signed Prekey, One-Time Prekey, TS	2,000
HTTP	FetchPicture	Profile Picture (High Quality)	5,500

TS Timestamp † Approx. queried accounts per second.

TABLE I

OVERVIEW OF WHATSAPP ENDPOINTS AND ENUMERATION SPEED (AS EXPERIENCED IN MEASUREMENTS).

Guessing Account Age/Activity via Timestamps

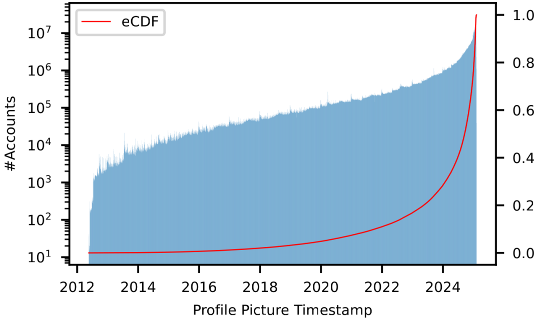


Fig. 7. Distribution of profile picture timestamps. Most users have recently updated their picture.

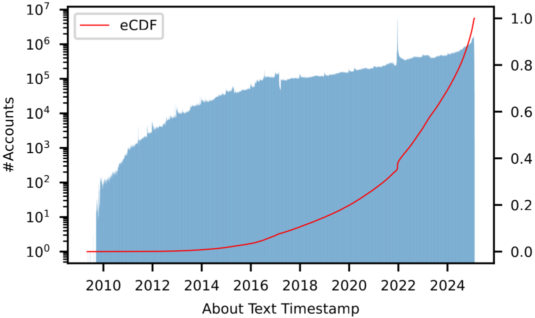


Fig. 8. Distribution of about status update timestamps. About messages are updated less frequently on average compared to profile pictures.

Downloading Profile Pictures

- Large-scale image download (via HTTP) also possible
  - Requirement: public visibility of profile picture (57% of all accounts)
  - Again, no considerable rate-limiting
  - Image download was faster than our disk writing speed (we had to rate-limit our queries)
- Adversary could abuse data to build a **reverse-phone book (using facial recognition)**

Proof for Active Accounts

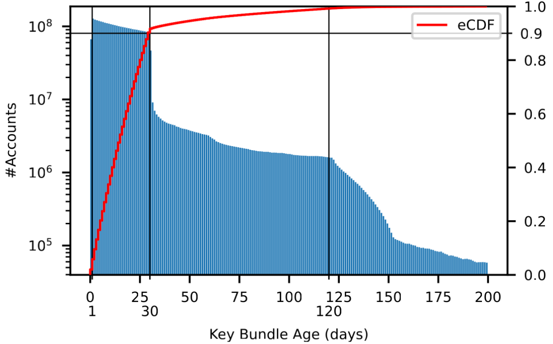
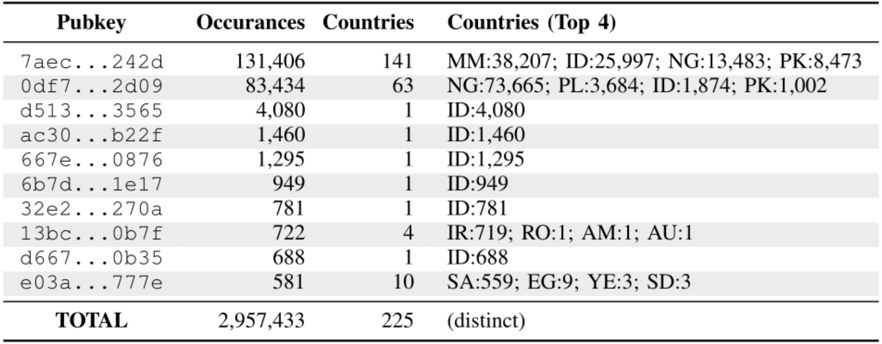


Fig. 5. Distribution of prekey bundle age across our retrieved data. Over 90 % of users have updated their keys within the past month. Consistent with WhatsApp's account deletion policy, most accounts with keys older than 120 days have been automatically removed.

Key Collisions for Potential Scam Accounts



Pubkey	Occurances	Countries	Countries (Top 4)
`7aec...242d`	131,406	141	MM:38,207; ID:25,997; NG:13,483; PK:8,473
`0df7...2d09`	83,434	63	NG:73,665; PL:3,684; ID:1,874; PK:1,002
`d513...3565`	4,080	1	ID:4,080
`ac30...b22f`	1,460	1	ID:1,460
`667e...0876`	1,295	1	ID:1,295
`6b7d...1e17`	949	1	ID:949
`32e2...270a`	781	1	ID:781
`13bc...0b7f`	722	4	IR:719; RO:1; AM:1; AU:1
`d667...0b35`	688	1	ID:688
`e03a...777e`	581	10	SA:559; EG:9; YE:3; SD:3
TOTAL	2,957,433	225	(distinct)

TABLE VII
Top 10 of reoccurring public keys.

Ethical Considerations

- We did not attempt to disguise ourselves
  - All experiments done from a **single server/IP address**
  - Reachable via abuse email, reverse DNS pointer to measurement website
- No threat to WhatsApp’s backend infrastructure
  - Scraping was **not noticed by the platform operator**
- **Empirical evaluation is the only way to estimate the success of such attacks** in the real world
  - E.g., exposing key collision is only possible with comprehensive data
- User data was deleted when our anaylsis was done
- Findings were disclosed in multiple responsible disclosure tickets (before, during, and after completing the study)

Responsible Disclosure Timeline

- Multiple Tickets (starting September 5, 2024) pointing out **missing rate-limiting and potential for enumeration attacks**
  - No actual/human response for a year
  - Specific **account enumeration ticket closed as *non applicable***
- On August 22, 2025 they finally reached out to us
  - **Reports were misstriaged** (“not used to detailed academic reports containing full research papers”)
- Assisted in developing and testing countermeasures during September, October, and November 2025
  - First fixes were rolled out in early October 2025

Discussion and Conclusion

- In 2025, large-scale enumeration was possible
  - Rate-limiting and other countermeasures discussed in the paper
- **Centralization** in messaging apps
  - Even worse when service is proprietary
- Long-lasting effects of data leaks
- Far-reaching survey possibilities
- Key collisions unnoticed by the platform operator
- Phone numbers have little entropy

Questions?

- Contact
  - Mail: gabriel.gegenhuber@univie.ac.at
  - Twitter: @GGegenhuber
  - Bluesky: @ggegenhuber.bsky.social

github.com/sbaresearch/whatsapp-census