# Not all is lost for anonymity – but quite a lot is.

## Coordination among users can help anonymity.

Debajyoti Das
Purdue University
das48@purdue.edu

Sebastian Meiser
Visa Research
smeiser@visa.com

Esfandiar Mohammadi
ETH Zurich
mohammadi@inf.ethz.ch

Aniket Kate
Purdue University
aniket@purdue.edu

## 1. INTRODUCTION

Anonymous communication (AC) is a fundamental building block in numerous privacy enhancing technologies and applications. While there is a successful line of research and development on anonymous communication, it is an open question whether current approaches for anonymous communication networks are optimal. As a key step towards identifying how much anonymity an optimal AC network can provide, we take a different approach: we search for inherent limitations that apply to AC networks and outline the landscape of achievable anonymity.

In a recent work [3], we have shown the first such upper bounds for anonymity, i.e., that certain combinations of bandwidth overhead, latency overhead and strong anonymity are impossible to achieve when faced with a global and passive network-level eavesdropper and node-level eavesdropper (i.e., a passive attacker). In this work, we show that the combination of secret-sharing and onion routing are able to escape our prior impossibility bounds, but we also prove novel impossibility bounds for such, more powerful protocols. For *hybrid protocols* that combine secret sharing and mix-nets techniques, the upper bounds on anonymity are significantly lower than for pure mix-nets, for the same latency and bandwidth overhead. In particular, while such hybrid protocols exhibit more resilience against compromisation than mix-net-like protocols, strong anonymity, low latency overhead and low bandwidth overhead still cannot be simultaneously achieved. Our work leaves as an open problem whether this combination of secret-sharing and onion routing is a theoretical effect or whether it can be realized in practice.

## 2. MAIN RESULTS

### 2.0.1 User Distributions, Communication Rounds, Bandwidth Overhead, and Latency

We consider two types of *user distributions*, i.e., distributions describing when users want to send anonymous messages. In the first user distribution (*synchronized*) $\mathsf{N}$ users send their messages in exactly $\mathsf{N}$ rounds. Per round, exactly one user sends a message. The protocol can additionally choose $B$ users to send noise messages in each round.

In the second user distribution (*unsynchronized*) each user independently decides whether to send a message in a round using a coin flip, with a success probability $p$. If a user does not have a real message to send, the user sends a dummy message. The number of dummy messages is bounded by $B$ per real message.

We consider synchronous communication *rounds* as in [4, 6, 7]. We model latency overhead $\ell$ as the number of rounds a message can be delayed by the protocol before being delivered. We formalize bandwidth overhead B as the number of noise messages the protocol can add per real message.

**Adversaries.** We consider global passive adversaries, that can observe all communication between protocol parties. Additionally, our adversaries can *passively* compromise some protocol parties to learn the mapping between input and output packets of those parties.

To keep the presentation concise, we focus on the bounds for sender anonymity [2, 5].

## 2.1 Preliminary Bounds

Let $\Pi$ be a protocol with $\mathsf{N}$ users, restricted by bandwidth overhead $B \geq 0$ and latency overhead $\ell \geq 0$, and the adversary can compromise $\mathsf{c}$ out of $\mathsf{K}$ protocol parties. We derive the following lower bounds for $\delta$-sender anonymity [2, 5] in the respective scenarios.

**Synchronized Users:**

$$\delta \geq \left(1 - \tfrac{B}{\mathsf{N}-1}\right) \times \left[1 - \tfrac{(\tau+1)\mathsf{N} - B\hat{\ell} - \hat{\ell}}{\mathsf{N}} \times g(\tau) - \tfrac{B\hat{\ell} + \hat{\ell} - \tau\mathsf{N}}{\mathsf{N}} \times g(\tau+1)\right]$$

where $\tau = \lfloor \frac{B\hat{\ell}+\hat{\ell}}{\mathsf{N}} \rfloor$, $\hat{\ell} = \ell + 1$, and

$$g(x) = \begin{cases} 1 & \mathsf{c} < x\hat{\ell} \leq \mathsf{K} \\ 1 & \mathsf{c} \leq \mathsf{K} \leq x\hat{\ell} \\ 1 - \binom{\mathsf{c}}{x\hat{\ell}} \Big/ \binom{\mathsf{K}}{x\hat{\ell}} & x\hat{\ell} \leq \mathsf{c} \leq K. \end{cases}$$

**Unsynchronized Users:**

$$\delta \geq \begin{cases} \left(1 - \frac{B_{\text{eff}}}{\mathsf{N}-1}\right) \left[1 - g(Z)\left(1 - (1-p)^{\hat{\ell}}\right)\right], & \mathsf{c} \geq \hat{\ell} \\ \left(1 - \frac{B_{\text{eff}}}{\mathsf{N}-1}\right) (1-p)^{\hat{\ell}-\mathsf{c}} \Bigg[1 - \Big(1 - (1-p)^{\mathsf{c}}\Big) & \\ \times \Big(\Pr[W \geq 1] + \Pr[W = 0]\left[1 - 1/\tbinom{\mathsf{K}}{\mathsf{c}}\right]\Big)\Bigg] & \mathsf{c} < \hat{\ell} \end{cases}$$

where $B_{\text{eff}} = min(B, \hat{\ell}p - 1)$, $Z = min(\hat{\ell}, 2B_{\text{eff}} + 1)$, $W$ is a random variable denoting the number of additional shares for the challenge message. $\Pr[W \geq 1]$ (or $\Pr[W = 0]$) is calculated based on the actual value of $p$.

The above bounds show that, similar to [3], strong anonymity requires a combination of latency and bandwidth overhead even for hybrid protocols; however, hybrid protocols have to obey much more relaxed impossibility bounds than mix-nets.

**Table 1:** Impossibility Conditions for Anonymous Communication, with the number of protocol-nodes $\mathsf{K}$, number of compromised protocol parties $\mathsf{c}$, number of clients $\mathsf{N}$, latency $\ell$. In all cases we assume that $\ell < \mathsf{N}$ and $(\mathsf{N}-1) - \epsilon(\eta) \geq B \geq 1$ and $\epsilon(\eta) = 1/\eta^d$ for a positive constant $d$. We compare $\ell = x$ of mix-net type protocols with $\hat{\ell} = x$ of hybrid protocols; and we denote the case with $\ell = x$ (See Footnote 1) in the leftmost column. All other columns shows the impossibility conditions for anonymity for the combination of user distribution and protocol class. Where two rows have overlapping cases (leftmost column), if either of the conditions are true, strong anonymity is impossible.

| Cases | synchronized, mix-net | unsynchronized, mix-net | synchronized, hybrid | unsynchronized, hybrid |
|---|---|---|---|---|
| $\mathsf{c} \geq 0$ | $2\ell B < \mathsf{N} - \epsilon(\eta)$ | $2\ell p < 1 - \epsilon(\eta)$ | $2\hat{\ell} B < \mathsf{N} - \epsilon(\eta)$ | $p\hat{\ell} < 1 - \epsilon(\eta)$ |
| $B < 1$ | $2\ell B < \mathsf{N} - \epsilon(\eta)$ | $2\ell p < 1 - \epsilon(\eta)$ | $2\hat{\ell} < N - \epsilon(\eta)$ | $p\hat{\ell} < 1 - \epsilon(\eta)$ |
| $0 < \mathsf{c} \leq \ell$ | $2(\ell - \mathsf{c})B < \mathsf{N} - \epsilon(\eta)$ | $2(\ell - \mathsf{c})p < 1 - \epsilon(\eta)$ | $2(\hat{\ell} - \mathsf{c})B < \mathsf{N} - \epsilon(\eta)$ | $p(\hat{\ell} - \mathsf{c}) < 1 - \epsilon(\eta)$ |
| $\ell < \mathsf{c} \leq B\ell$ | $\ell \in O(1)$ | $\ell \in O(1)$ | $2(\hat{\ell} - \mathsf{c})B < \mathsf{N} - \epsilon(\eta)$ | $p(\hat{\ell} - \mathsf{c}) < 1 - \epsilon(\eta)$ |
| $B\ell < \mathsf{c} \leq \ell^2$ | $\ell \in O(1)$ | $\ell \in O(1)$ | $\hat{\ell} \in O(1)$ | $p(\hat{\ell} - \mathsf{c}) < 1 - \epsilon(\eta)$ |
| $\mathsf{c} > \ell^2$ | $\ell \in O(1)$ | $\ell \in O(1)$ | $\hat{\ell} \in O(1)$ | $\hat{\ell} \in O(1)$ |
| $\mathsf{K}/\mathsf{c} \in O(1)$ | $\ell \in log(\eta)$ | $\ell \in log(\eta)$ | $\hat{\ell}^2 \in log(\eta)$ | $\hat{\ell}^2 \in log(\eta)$ |

We note that the benefit from secret-sharing techniques are limited to defending against compromisation, i.e., we can show the same bounds as [3] against adversaries that do not compromise any protocol parties, with the noteworthy exception in case $\ell = 0$[1]. When $B \geq (\mathsf{N}-1)$, strong anonymity can be achieved even for $\ell = 0$ — which is not possible for mix-net protocols. Moreover, our lower bounds leave the possibility for hybrid protocols to achieve strong anonymity even if a large fraction of nodes is compromised (cf. $g(x)$ for $x\hat{\ell} \leq c \leq K$). In Table 1 we compare the impossibility conditions for anonymity for hybrid protocols with mix-net protocols. Whenever the conditions in a line in Table 1 are met, strong anonymity is impossible, e.g., for synchronized user distribution for mixnets, if $\mathsf{c} > \ell$ and either of $\ell \in O(1)$ or $2\ell B < \mathsf{N} - \epsilon(\eta)$ is true, strong anonymity is impossible.

**Improved bounds for mix-net protocols.** As a byproduct of our new bounds for hybrid protocols, we also derive improved bounds for mix-net protocols in case of unsynchronized user distribution $U_P$. For $\mathsf{c} < \ell$ we get:
$\delta \geq 1 \times (1-p)^{\hat{\ell}-\mathsf{c}} \left[1 - (1-(1-p)^{\mathsf{c}}) \left[1 - 1/\binom{\mathsf{K}}{\mathsf{c}}\right]\right]$.
Similarly for $\mathsf{c} \geq \ell$: $\delta \geq 1 - \left(1 - (1-p)^{\ell}\right) \left[1 - \binom{\mathsf{c}}{\ell} / \binom{\mathsf{K}}{\ell}\right]$.

## 2.2 Insights from new results

**Relaxed impossibility bounds.** We can show that for the same values of $B$, $\ell$, $\mathsf{K}$ and $\mathsf{c}$ we get a lower (lower-)bound on the adversary's advantage $\delta$ for hybrid protocols compared to mix-net protocols. For instance, for unsynchronized user distribution consider $\mathsf{N} = \eta^2, \mathsf{K} = \eta, \ell = \mathsf{K}/4, B = \eta, p = 0.5, \mathsf{c} = \mathsf{K}-1$. For $\eta = 128$, $\delta$ is lower bounded by $2.32 * 10^{-10}$ for hybrid protocol. On the contrary, for mix-net protocols $\delta$ is lower bounded by 0.75. This huge difference in the lower bound of $\delta$ for $\mathsf{c} = \mathsf{K} - 1$ shows that hybrid protocols can give much more hope to protocol designers trying to design AC protocols with any-trust assumption.

**Purely bandwidth overhead suffices for anonymity.** For mix-net-like protocols, – no matter the bandwidth overhead – strong anonymity is categorically impossible unless $\ell$ compensates for the number of compromised parties. If, as an example, $\mathsf{c} > \ell$ and $\ell \in O(1)$, strong anonymity is impossible even for $B = \mathsf{N}$. Even when $\mathsf{c} = 0$ and $B \geq \mathsf{N}$, we need at least $\ell = 1$ for mix-net-like protocols. However, for protocols that use secret sharing techniques, there is always the possibility of strong anonymity if $B \geq \mathsf{N}$, as done in DC-nets, even with $\ell = 0$ and $\mathsf{c} = \mathsf{K}$.

**Impossibility of strong anonymity only for very high compromisation.** The adversary needs to compromise a lot more protocol parties to break anonymity against hybrid protocols than our mix-nets scenario. Let $\mathsf{N} = \eta^3, \mathsf{K} = \eta^2, \ell = \eta, p = 0.1, B = \eta$, and consider the unsynchronized user distribution. With hybrid protocols strong anonymity is possible here for $\mathsf{c} < 0.1\eta^2$, while it is impossible for mix-net protocols for any $\mathsf{c} \geq \eta$. This mismatch indicates that an adversary needs to compromise more protocol parties to break strong anonymity against hybrid protocols.

For interested readers, we refer to [1] for our tech-report and more details about the work.

# 3. REFERENCES

[1] Anonymity Trilemma Project Webpage. https://freedom.cs.purdue.edu/anonymity/trilemma/.

[2] M. Backes, A. Kate, P. Manoharan, S. Meiser, and E. Mohammadi. AnoA: A Framework For Analyzing Anonymous Communication Protocols. In *Proc. 26th IEEE Computer Security Foundations Symposium (CSF 2013)*, pages 163–178, 2013.

[3] D. Das, S. Meiser, E. Mohammadi, and A. Kate. Anonymity trilemma: Strong anonymity, low bandwidth, low latency—choose two. Cryptology ePrint Archive, Report 2017/954, 2017.

[4] R. Gennaro, M. O. Rabin, and T. Rabin. Simplified VSS and fact-track multiparty computations with applications to threshold cryptography. In *Proc. ACM PODC*, pages 101–111, 1998.

[5] M. Backes, A. Kate, P. Manoharan, S. Meiser, and E. Mohammadi. AnoA: A Framework For Analyzing Anonymous Communication Protocols. *Journal of Privacy and Confidentiality (JPC)*, 7(2)(5), 2016.

[6] T. Ruffing, P. Moreno-Sanchez, and A. Kate. P2P Mixing and Unlinkable Bitcoin Transactions. In *NDSS*, 2017.

[7] T. K. Srikanth and S. Toueg. Simulating authenticated broadcasts to derive simple fault-tolerant algorithms. *Distributed Computing*, 2(2):80–94, 1987.

[1] We approximate noise generated by internal nodes of latency $\ell$ with user noise of latency $\hat{\ell} = \ell + 1$. That also allows protocols with only user noise to have latency $\hat{\ell}$. It is unfair to compare them with mix-net protocols with latency $\ell$. Moreover, when $\ell = 0$, there is no intermediate party, so there is no internal noise.