Practical Semi-Open Group Messaging (a Proposal)

Fernando Virdia
University of Surrey

Joint work with Alex Davidson and Luiza Soezima

UK Crypto Day, September 11 2025

Secure messaging and collective action

- Online communication plays an important role in contemporary protest and activist movements [HZ15; URW18; VV18; Tre20; ZAACR21]
- Today, secure messaging offers powerful formal "end-to-end" guarantees

Confidentiality and authentication	Forward secrecy	Post-compromise security

- Yet, these protocols often fail to address other "on-the-ground" requirements
- Remote message deletion, scheduled messaging, and group polling can prove central to the use of messaging by activists [Alb+21]

Group messaging, scenario 1

- You are an activist group trying to increase your reach to plan a demonstration
- You want to use group chats, provided by the most common messaging platform in your area
- You are particularly worried by anonymity, as the adversary may penalise individual members taking part

"Closed" chat group

Admins manually invite users:

+ only invited people can see messages and identities
− vetting of candidates slows growth
− significant time commitment for the admins

"Open" group

Admins publicly share a link for people to join:

+ anyone with the link can join the chat
+ quick group growth possible
− the adversary can easily join too
  $\rightarrow$ and deanonymise

Group messaging, scenario 2

- You are a national-security leader
- You may be trying to avoid national record laws and would rather use private messaging apps
- You value action for action's sake, and don't think too much when adding a buddy to a chat

"Closed" chat groups only

Admins manually invite users:

+ only invited people can see messages and identities
− requires keeping track of who's in your phone's address book
→ always at risk of inviting a journalist to a chat about military strikes

Group onboarding is outside of model

- Today, secure messaging assumes you know who you'll talk to
- Messaging protocols do not capture user "reputation"
- Yet, measures of reputation [HZNR09] and privacy-preserving reputation schemes have received significant attention [GG21]

We ask: could we integrate messaging with reputation systems?

Our attempt: defining a notion of "semi-open" group messaging

- Assume a closed group G is initially formed among a few trusted contacts
- Then a link to join the group is openly shared
- Whenever an external user E opens the link, the in-group reputation of E among the users (G_i) is computed
  - if "high enough", E is added to the group automatically
  - if "too low", E is added to a waiting list to be vetted manually
- Think: holding an election every time an external asks to join (Scenario 1)
- Dual: regularly hold elections to kick out low-reputation users (Scenario 2)

Aim: useful

Practically deployable

Clear security guarantees

Meaningful

Practical requirements

- Adoptable into existing messaging protocols without changes
  - Single-server, no re-adding users from scratch, no GiB-sized key material
- User-interaction overhead should be kept to a minimum
  - À la Whatsapp "Block this unknown contact? Yes/No"
- Voting/rating an external can happen at any moment
  - You may meet E before any group was formed, and want to rate them
  - Reputation can be computed (tallied) even if most group members are offline

Security requirements

- Ideally, the system should offer some amount of:
  - vote confidentiality, unlinkability, integrity
  - tally auditability
- Any party should be considered adversarial
  - An **external user** may want to be included even with low reputation
  - A **group admin** may want to be able to link votes to voters
  - A **server** and a **voter** may collude to unfairly exclude a specific external user with a high reputation
  - ...
- The system should offer some security even if different parties collude

Meaningfulness

- Matching someone's "reputation" to a score is inevitably noisy
- In many cases, individuals in a group may not know each other enough to give a score

How does this affect the threat model? What could the use cases be?

Nation-state adversaries

+ Infiltration of open groups is extremely likely
+ Closed groups may require lengthy in-person vetting [Alb+21]

– A successful infiltration may be catastrophic

→ Reputation for automatic admission risky

→ Reputation for recovery from infiltration could be helpful (post-compromise security?)

"Weak" adversaries ("your employer")

+ Infiltration of open groups is less likely
+ Successful infiltration potentially less catastrophic

→ Automatic admission could allow lower admin overhead

Reputation systems

- Privacy-preserving reputation systems already exist in the literature
- Many are invoked to protect online stores from spam product reviews
- A couple address online communities: AnonRep [Zha+16] and PRSONA [GG22]

An outline of AnonRep/PRSONA

- Bulletin-board systems, where time is divided into epochs
- Under a pseudonym, users can post messages and vote on other users' messages
- Periodically, a mix-net tallies votes and updates user global reputation scores

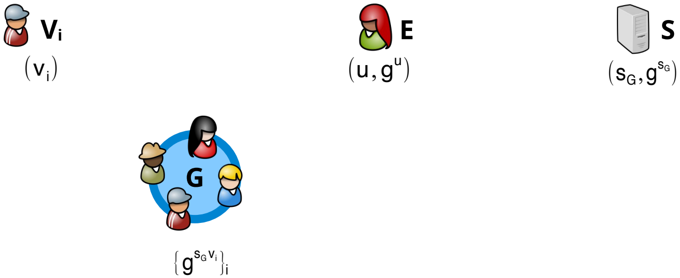
Not quite practical to "add" to (your fav protocol)

These systems require a mix-net, ring signatures, and (partially-)homomorphic encryption.

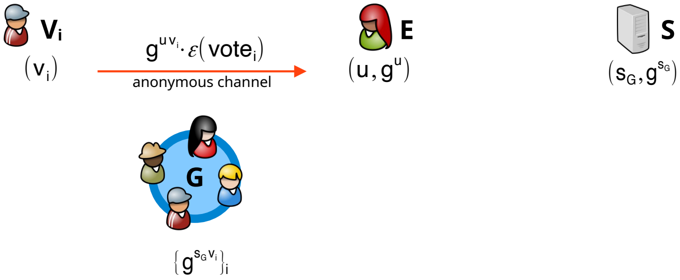
- Hard to maintain multiple secure and truly independent service providers
- Anonymous authentication is achieved via ring-signatures
  - Signers need a list of every public key in the system
  - Likely impossible with millions of users
- Partially-homomorphic encryption of feedback limits the kind of computable tally functions
- Reputation scores are global $\rightarrow$ do not capture group composition
- Provable guarantees are unclear

Our approach: let's try rolling our own crypto

Protocol overview



Protocol overview



Protocol overview

$\mathbf{V_i}$

(v_i)

E

(u, g^u)

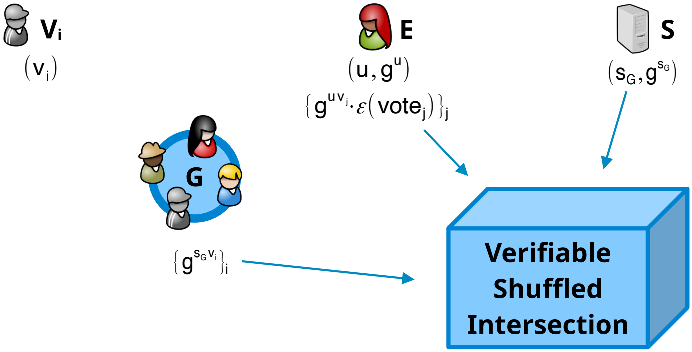
$\{g^{u v_j} \cdot \varepsilon(\text{vote}_j)\}_j$

S

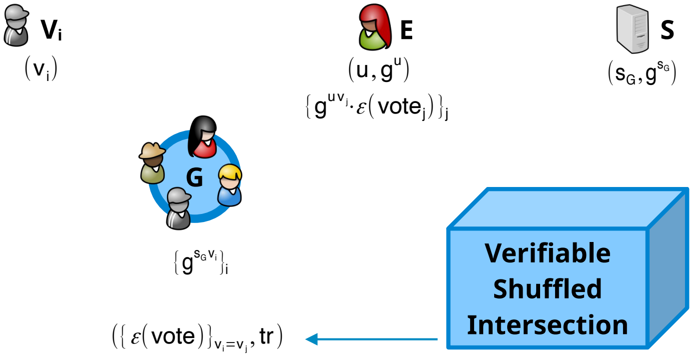
(s_G, g^{s_G})

$\{g^{s_G v_i}\}_i$

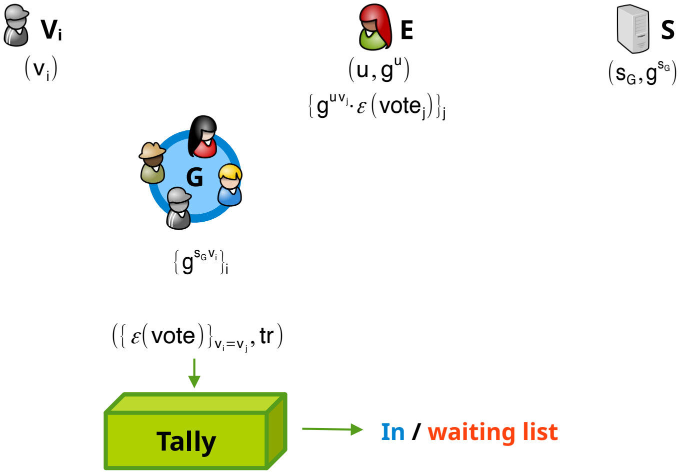
Protocol overview



Protocol overview



Protocol overview



Simulation-based security definition via an ideal functionality

- Group member V_i inputs a score $x_i \in D$ on E
- An overall admission decision $b \in \{0, 1\}$ is computed as a function of $\{x_i\}_i$
- Server and external user only learn b
- Group members learn b and the set $\{x_i\}_i$, but not what vote comes from whom
- Our definition covers a single "join-session", but our design targets multiple sessions

Intuitive guarantees

- Vote confidentiality: from E's point of view, encrypted votes are $\approx$ random, except for leakage from b
- Ballot unlinkability: assuming ballots are delivered via an *anonymous channel*, ballots are unlinkable to voters, except for leakage from the date/time of casting
- Tally integrity: by keeping a transcript of the protocol run and of user inputs and zk proofs of correct computation, a group member can recompute the tally independently

Assumptions and security model

- We work in the ROM, assuming DDH is hard
- We assume a robust internal group transcript, to be provided by the messaging protocol
- We assume the existence of one or more group *admins*
- Honest parties check the transcript, and abort the protocol if malicious behaviour is detected
  - Offline parties can only check retrospectively!
  - Reasoning: server and group admin want to protect reputation; external user can be kicked out.
- We prove results against different combinations of actively malicious colluding parties

Protocols and results

- We define two protocols P_1 and P_+, based on the number of group admins
- We prove security of:
  - P_1 against any set of malicious colluding parties excluding the server
  - P_1 against a malicious server alone
  - P_+ against a malicious server colluding with one of $\{$group members, group admins, external user$\}$ assuming *at least one honest group admin*

Proof-of-concept implementation

- We implemented a local version of the protocol in C++ / libsodium
- We use SHA2 and SHAKE as random oracles, and Ristretto255 as prime-order group
- We instantiated the required proof systems with soundness error 2^{-128}
- We run single-core simulations of the protocol on a MacBook Air M3 CPU, given:
  - A vote domain of size $|D| = 10$
  - A total number of $n + t/2 + 1$ users and 1 server
  - A group G of n users (voters)
  - One external user E (votee)
  - t users having voted on E, of which $t/2$ belonging to G
- Shuffle computation takes $O(n)$ and ballot intersection $O(n \cdot t \cdot |D|)$, both trivially parallelizable

Benchmarks 1/2

Parameters	Phase	Runtime (s) mean	st. dev.	Bandwidth (KiB)
	total	3.3	0.2	1312.2
$n = 50$	VE.Eval & check	¡ 0.1	¡ 0.1	2.6
$t = 40$	VEP.Eval & check (U)	1.2	0.1	653.2
$\|D\| = 10$	VEP.Eval & check (S)	1.2	0.1	653.2
	ballot intersection	0.9	0.1	1.2
	total	6.4	0.4	2620.1
$n = 100$	VE.Eval & check	¡ 0.1	¡ 0.1	2.6
$t = 40$	VEP.Eval & check (U)	2.2	¡ 0.1	1306.3
$\|D\| = 10$	VEP.Eval & check (S)	2.3	0.4	1306.3
	ballot intersection	1.9	0.0	1.2

Benchmarks 2/2

Parameters	Phase	Runtime (s) mean	st. dev.	Bandwidth (KiB)
	total	12.7	0.2	5235.7
$n = 200$	VE.Eval & check	¡ 0.1	¡ 0.1	2.6
$t = 40$	VEP.Eval & check (U)	4.5	0.1	2612.5
$\|D\| = 10$	VEP.Eval & check (S)	4.5	0.2	2612.5
	ballot intersection	3.7	0.0	1.2
	total	16.3	0.2	5239.4
$n = 200$	VE.Eval & check	¡ 0.1	¡ 0.1	5.1
$t = 80$	VEP.Eval & check (U)	4.5	0.1	2612.5
$\|D\| = 10$	VEP.Eval & check (S)	4.4	0.1	2612.5
	ballot intersection	7.4	0.1	2.5

Open questions

Utility / Usability

- Is this a useful primitive?
  - For what group sizes?
  - For what group formation dynamic (Scenario 1 or 2 or . . .)?

Technical

- During intersection, anonymous vote plaintexts are recovered
  - \+ Compatible with any tally function
  - − No vote confidentiality *from other group members*, at most anonymity
- "Reputation hacking" likely inevitable
  - Similarly to MPC, the protocol is cryptographic, the Tally function being evaluated isn't
  - What is the most "resilient" Tally function is unclear [HZNR09]
- Supporting multiple identities and vote updates is somewhat cumbersome

Conclusion

- We consider the use of reputation systems within group messaging
- We propose a family of practical, provably secure, single-server, collusion-resistant, reputation protocols
- We see them as an example “fine-grained cryptography” [Ros20],
  - Somewhere between semi-honest and malicious
  - Somewhere between no security and resistance to an NSA-level adversary

Thank you

Resources I

[HZ15] Gulizar Haciyakupoglu and Weiyu Zhang. "Social media and trust during the Gezi protests in Turkey". In: *Journal of computer-mediated communication* 20.4 (2015), pp. 450–466.

[URW18] Temple Uwalaka, Scott Rickard, and Jerry Watkins. "Mobile social networking applications and the 2012 Occupy Nigeria protest". In: *Journal of African Media Studies* 10.1 (2018), pp. 3–19.

[VV18] Augusto Valeriani and Cristian Vaccari. "Political talk on mobile instant messaging services: A comparative analysis of Germany, Italy, and the UK". In: *Information, Communication & Society* 21.11 (2018), pp. 1715–1731.

[Tre20] Emiliano Treré. "The banality of WhatsApp: On the everyday politics of backstage activism in Mexico and Spain". In: *First Monday* 25 (2020).

Resources II

[ZAACR21] Homero Gil de Zúñiga, Alberto Ardèvol-Abreu, and Andreu Casero-Ripollés. "WhatsApp political discussion, conventional participation and activism: exploring direct, indirect and generational effects". In: *Information, communication & society* 24.2 (2021), pp. 201–218.

[Alb+21] Martin R Albrecht et al. "Collective Information Security in {Large-Scale} Urban Protests: the Case of Hong Kong". In: *30th USENIX security symposium (USENIX Security 21)*. 2021, pp. 3363–3380.

[HZNR09] Kevin Hoffman, David Zage, and Cristina Nita-Rotaru. "A survey of attack and defense techniques for reputation systems". In: *ACM Comput. Surv.* 42.1 (2009). ISSN: 0360-0300. DOI: 10.1145/1592451.1592452. URL: https://doi.org/10.1145/1592451.1592452.

[GG21] Stan Gurtler and Ian Goldberg. "SoK: Privacy-preserving reputation systems". In: *Proceedings on Privacy Enhancing Technologies* (2021).

Resources III

[Zha+16] Ennan Zhai et al. "AnonRep: Towards Tracking-Resistant Anonymous Reputation". In: *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*. Santa Clara, CA: USENIX Association, Mar. 2016, pp. 583–596. ISBN: 978-1-931971-29-4. URL: https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/zhai.

[GG22] Stan Gurtler and Ian Goldberg. "PRSONA: Private Reputation Supporting Ongoing Network Avatars". In: WPES'22. Los Angeles, CA, USA: Association for Computing Machinery, 2022, 55–68. ISBN: 9781450398732. DOI: 10.1145/3559613.3563197. URL: https://doi.org/10.1145/3559613.3563197.

[Ros20] Alon Rosen. *Fine-Grained Cryptography: A New Frontier?* Cryptology ePrint Archive, Paper 2020/442. https://eprint.iacr.org/2020/442. 2020. URL: https://eprint.iacr.org/2020/442.