

Securing semi-open group messaging

Fernando Virdia

NOVA LINCS & Universidade NOVA de Lisboa

Joint work with Alex Davidson and Luiza Soezima

May 26 2024



NOVALINCS

LABORATORY FOR COMPUTER
SCIENCE AND INFORMATICS

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 cryptographic powerful formal “end-to-end” guarantees

Confidentiality and authentication

Forward secrecy

Post-compromise security

- Yet, these protocols often miss to address “on-the-ground” requirements
- Remote message deletion, scheduled messaging, group polling can prove central to the use of messaging by activists [[Alb+21](#)]

Group messaging, a scenario

- 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
- slow and difficult vetting of candidates
- 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
- the adversary can easily join too
⇒ and deanonymise

Group onboarding is outside of model

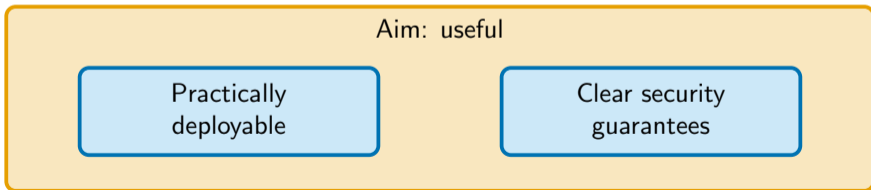
- Today, secure messaging “kind-of” assumes you know who you’ll talk to
- Messaging protocols do not capture user “reputation”
- Yet, measures of reputation [HZNR09] and privacy-preserving reputation schemes have received significant attention [GG21]

Idea: could we integrate messaging with reputation systems?

- In most settings, infiltration of open groups is extremely likely
- Against nation-state adversaries, closed groups and lengthy in-person vetting may be necessary [Alb+21]
- Against weaker adversaries, a relaxation of vetting requirements in exchange for a lower admin overhead may be of use

Our attempt: to define a notion of “semi-open” group

- Assume a group is initially formed among a few trusted contacts
- A link to join the group is openly shared
- Whenever an external user E opens the link, 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
- Can think of this as holding an election every time an external asks to join



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”
 - ▶ Only optionally more
- 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 and votees
 - ▶ A **server** and a **voter** may collude to unfairly exclude a specific external user with high reputation
 - ▶ ...
- Different parties should be allowed to collude
- Everyone contributes inputs, semi-honest security is not enough

Definitely an ambitious project, too good to be true? Where to start?

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 in 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, (partially-)homomorphic encryption.

- Multiple independent servers \sim federation
- Authentication with anonymity is obtained by ring-signatures
 - ▶ Signers need a list of every public key in the system
 - ▶ Likely impossible with millions of users
- Partial-homomorphic encryption of feedback limits the kind of tally functions
- Reputation scores are global, do not capture group composition
- Provable guarantees are unclear

Our approach: let's try rolling our own crypto






$g^{u v_i} \cdot \mathcal{E}(\text{vote}_i)$
anonymous channel



$\{g^v\}_i$

 V_i
 (v_i)

 E
 (u, g^u)
 $\{g^{uv_i} \cdot \mathcal{E}(\text{vote}_j)\}_j$

 S

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

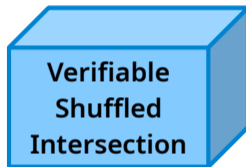
V_i
 (v_i)

E
 (u, g^u)
 $\{g^{uv_i} \cdot \mathcal{E}(\text{vote}_j)\}_j$


S



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



 V_i
 (v_i)

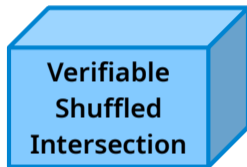
 E
 (u, g^u)
 $\{g^{uv_i} \cdot \mathcal{E}(\text{vote}_j)\}_j$

 S




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

$(\{\mathcal{E}(\text{vote})\}_{v_i=v_j}, \text{tr})$



 V_i
 (v_i)

 E
 (u, g^u)
 $\{g^{uv_i} \cdot \mathcal{E}(\text{vote}_j)\}_j$

 S



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

$(\{\mathcal{E}(\text{vote})\}_{v_i=v_j}, \text{tr})$



In / waiting list

- To argue security: ideal functionality + simulation-based proofs.
- So far: sketch-proofs under various combinations of **two** colluding parties.

Three main obstacles

1. Collusion between parties
2. Achieving malicious security
 - ▶ Requires group members to be online to check protocol execution
 - ▶ We assume they should be able to be offline (contradiction?)
3. Role-fluidity of the adversary

Let's look at some examples of what can go wrong.

“Group member + external” collusion

- Chat content can be trivially leaked
- May still want to protect anonymity of vote-on-external

Role fluidity

- If “E + X” collude and E gains access as G_i , they become “E + X + G_i ”
- Could be used to deanonymise votes
- Likely requires “ratcheting” to address
- Does this affect other protocols? Meshes?

Malicious security

- Most group members assumed to be offline
 - Hence unable to check correct protocol execution, e.g. their vote could be ignored
- ⇒ Make transcript checkable when back online: “*checkable semi-honest*”
- If server or admin misbehave, blow the whistle and hope for external incentives
 - Online parties can instead abort rather than callout ⇒ malicious security if all online

Current limitations

- Currently relies on the “WIP conjecture” (ie, no mistakes)
- During intersection, anonymous vote plaintexts are recovered
 - + Compatible with any tally function
 - No vote confidentiality, at most anonymity
- Only tolerates collusion of up to two parties
- “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](#)]

Conclusion

- Reputation in messaging systems presents interesting challenges
- We see this as an example “fine-grained cryptography” [Ros20],
 - ▶ Somewhere between semi-honest and malicious
 - ▶ Somewhere between no security and resistance to NSA-level adversary
- This functionality could be of use in some “weak-adversary” activist settings
- We attempt to give a solution with provable guarantees, eprint soon

Thank you

`fernando@fundamental.domains`

Resources I

- [HZ15] Gulizar Hacıyakupoglu 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.
- [GG21] Stan Gurtler and Ian Goldberg. “SoK: Privacy-preserving reputation systems”. In: *Proceedings on Privacy Enhancing Technologies* (2021).

- [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](https://doi.org/10.1145/3559613.3563197). URL: <https://doi.org/10.1145/3559613.3563197>.

Resources IV

- [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](https://doi.org/10.1145/1592451.1592452). URL: <https://doi.org/10.1145/1592451.1592452>.
- [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>.