Towards Trustworthy Online Voting: Distributed Aggregation of Confidential Data

PhD Thesis Defense

Robert Riemann

Inria/ENS de Lyon

18th December 2017


NATS

Complexity of Cooperation

Observations

1. size of cooperation is increasing in terms of peers & links
2. diversification and specialisation
3. overall complexity is increasing

Problems

1. How to ensure trust in cooperation?
2. How to govern large cooperations?

Good Governance Principles promoting Trust[1]

Characteristics for Trust:

- Transparency
- Participation
- Accountability

Characteristics for Scalability:

- Responsiveness
- Efficiency

[1] UNESCAP. "What is Good Governance ?". In: United Nations Economic and social Comission for Asia and the Pacific (2009).

Good Governance Principles promoting Trust[1]

Characteristics for Trust:

- Transparency
- Participation
- Accountability

Characteristics for Scalability:

- Responsiveness
- Efficiency, which includes somehow
- Convenience

[1] UNESCAP. "What is Good Governance ?". In: United Nations Economic and social Comission for Asia and the Pacific (2009).

Scalability

100k Citizens

Fig. Athens 500 BC

10 Mio Citizens

Fig. Tokyo nowadays
(© Flickr/inefekt69)

Trust in Cooperations

Personal Trust

- based on personal relationships among cooperation members

Institutional Trust

- based on organisational security
- e.g. division of power and checks and balances

Technological Trust

- based on physical security
- e.g. barriers, locks and cryptography


become a supporter
subscribe
find a job
theguardian
news sport opinion arts lifestyle
UK world business football environment tech UK politics science
iOS

Apple fixes HomeKit bug that allowed remote unlocking of users' doors

Security flaw in latest iPhone and iPad iOS 11.2 software meant hackers could potentially gain remote control of lights, cameras and locks in smart homes

Fig. Guardian Article by S. Gibbs published on 8th December 2017

Identified Problem

Common Properties of Large Cooperations:

- Large cooperations employ often authorities.
- Unverified physical security is like organisational security.

$\Rightarrow$ Cooperation is vulnerable to e.g. adversaries.

Ambition for more Trustworthiness:

- limit impact of authorities and maximise division of powers
- limit complexity of physical security for easier verification

Online Services

Online Services are among the largest cooperations. Facebook counts 2 billion monthly active users (¼ world population).

Common Properties:

- few authorities (service operators) vs many users
- assume institutional trust (in law compliance and enforcement)

Online Service emerge in all areas of life:

- Commerce (Alibaba, Amazon, eBay)
- Social Networks (Facebook, Twitter, Weibo, VK)
- Intermediary Services (AirBnB, Uber)
- eGovernment (Registries, Taxation, eParticipation)

Potential Solution

Distributed (P2P) Systems

- distributed filesharing BitTorrent[2]
- distributed currency Bitcoin[3]

[2] B. Cohen. The BitTorrent Protocol Specification. 2008.
[3] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008.

Generic Paper-based Voting

1. **Preparation Phase**
   central voter registry issues list of eligible voters,
   prints undistinguishable voting ballots
2. **Casting Phase**
   on-site, public supervision, voting station(s) run by citizens
3. **Aggregation Phase**
   tallying of casted ballots
4. **Evaluation Phase**
   computation of the voting outcome from public tally
5. **Verification Phase**
   observation during the vote (eye-sight), recounts

Challenge: Conflicting Protocol Properties

Ensure set of security properties at the same time:

- unconditional secrecy of the ballot
- universal verifiability of the tally
- eligibility of the voter

Achievable only with unrealistic assumptions[4]:

compromise required

[4] B. Chevallier-Mames et al. “On Some Incompatible Properties of Voting Schemes”. In: Towards Trustworthy Elections: New Directions in Electronic Voting. Springer, 2010.

Impact of Technology on Voting I

Fig. Digital Natives.
(Flickr/antmcneill CC by-sa)

Fig. Paper-based Voting.
(Flickr/coventrycc CC by-nc-nd)

Impact of Technology on Voting II

Impact on Expectations

- comfort on a par with other online services
- flexibility
- automation for cost efficiency

Impact on Security

- hidden body cameras
- invisible ink
- fingerprint databases
- DNA analysis

Online Voting

Online Voting

remote electronic voting

- no chain of custody verifiable per eye-sight
- electronic signals are easy to duplicate

Need for new concepts to ensure security properties.

Classical Online Voting Security Concepts

- **Trusted Authorities**
  essentially give up secrecy and correctness
- **Anonymous Voting**
  assume unlinkability of distinct communication channels
- **Random Pertubation**
  assume shuffle of encrypted votes before their decryption
- **Homomorphic Encryption**
  assume aggregation of encrypted votes before decryption

Identified Issues

- concentration of power (assumed trust)
- concentration of data

Distributed Protocols

Without consensus on trusted authorities, it is reasonable to omit authorities altogether.

Compare development to:

- **Bitcoin**
  gold, fiat money, online banks, Bitcoin
- **BitTorrent**
  circulating disks, FTP (web server), Bittorrent

Empowerment of Voters

Assumption of a Distributed Online Voting Protocol

- no authority
- equally privileged, equipotent voters

Promises

- reflects democratic principle of equally powerful voters
- all voters are potential voting officers
- all voters responsible to enfore policy of protocol
- with no weakest link, promise of improved resiliance against DDoS attacks
- balance of knowledge among voters

Notions of Distribution in Online Voting

1. **Degree of Specialisation**
   from equipotent voters to specialised authorities
2. **Topology** of communication/responsabilities
   from centralised over decentralised to distributed
3. **Phase**
   consider phases that are actually distributed

Notions of Distribution in Online Voting

1. **Degree of Specialisation**
   from equipotent voters to specialised authorities
2. **Topology** of communication/responsabilities
   from centralised over decentralised to distributed
3. **Phase**
   consider phases that are actually distributed

Fully distributed Protocol

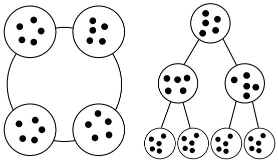
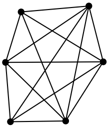
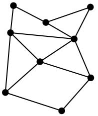
- equipotent voters, no authorities,
- distributed topology
- in all phases (but the registration)

From Centralised to Distributed Online Voting

What if all voters become authorities?

- reuse existing protocols with:
  distributed key generation and threshold decryption
- fits the purpose of small board room votings
- does not scale

Review of Distributed Online Voting




(a) DPol (b) SPP (c) SMC (d) Blockchain

- **Secure Multi-party Computation (SMC)**
  communication in $\mathcal{O}\left(n^2\right)$, for board room votings
- **Distributed Polling (DPol)**
  secret sharing scheme applied to groups aligned in a circle
- **Secure and Private Polling (SPP)**
  SMC and threshold decryption applied to groups in a tree
- **Blockchain-based Voting**
  Bitcoin to aggregate votes (coloured coins)

Taxonomy of Distributed Online Voting

Protocol	Degree of Special.	Topology	Distrib. Phases
Paper-based	none (flexible)	distributed	all
Helios,[5]	selected authorities	centralised	verification
SPP,[6]	random authorities	structured, tree	aggregation
DPol,[7]	none	structured, ring	all
Blockchain-based	none (flexible)	distributed	all

[5] B. Adida. “Helios: Web-based Open-Audit Voting.”. In: USENIX Security Symposium 17 (2008), pp. 335–348.

[6] S. Gambs et al. “Scalable and Secure Aggregation in Distributed Networks”. In: (2011). DOI: 10.1109/SRDS.2012.63.

[7] R. Guerraoui et al. “Decentralized polling with respectable participants”. In: Journal of Parallel and Distributed Computing 72.1 (Jan. 2012), pp. 13–26. DOI: 10.1016/j.jpdc.2011.09.003.

Taxonomy of Distributed Online Voting[5]

Protocol	Degree of Special.	Topology	Distrib. Phases
Paper-based	none (flexible)	distributed	all
Helios	selected authorities	centralised	verification
SPP	random authorities	structured, tree	aggregation
DPol	none	structured, ring	all
Blockchain-based	none (flexible)	distributed	all

Remarks:

- Blockchain-based protocols are most promising for their similarity with paper-based voting
- To our knowledge: no publication yet on Blockchain-based protocols

[5] R. Riemann and S. Grumbach. "Distributed Protocols at the Rescue for Trustworthy Online Voting". In: Proc. of the 3rd Intern. Conf. on Information Systems Security and Privacy (ICISSP). Porto, Feb. 2017.

Novel Protocol

Novel fully distributed Online Voting Protocol:

ADVOKAT[6]

- different compromise between secrecy and verifiability
- probabilistic definitions: confidentiality and individual verifiability
- probabilistic results: almost correct with high probability
- assume that voters are always connected (cf. IoT)
- assume trust in technology (instead of in authorities)

[6] Aggregation for distributed voting online using the Kademlia DHT

ADVOKAT Tree

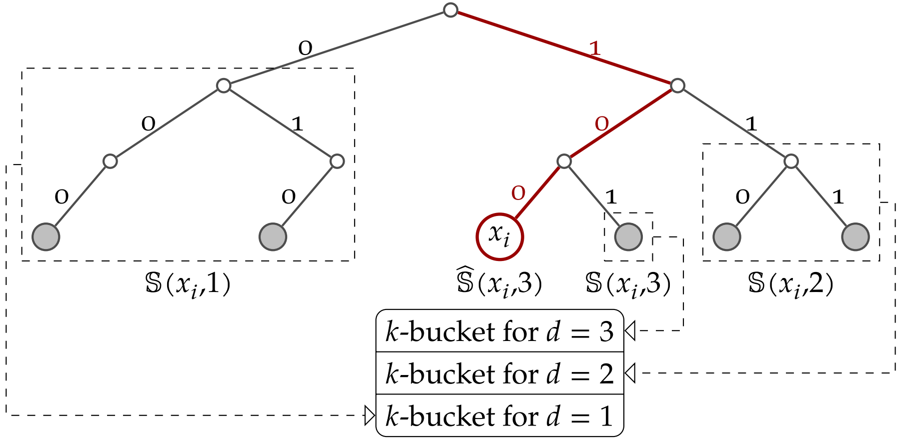


Fig. Kademlia Tree

Signatures

A	Authority
P_i	Voter, i-th out of n
a_i	Vote of P_i
$\sigma_i(m)$	P_i's signature scheme using its key pair (pk_i, sk_i)
$\sigma_A(m)$	Authority's signature scheme
$\chi(m, r)$	Blinding technique with random number r
$\delta(s, r)$	Retrieving technique of blind signature

- P_i provides $b_i = \chi(pk_i, r_i)$ to A
- A provides once for P_i the blinded signature $s_i = \sigma_A(b_i)$
- P_i retrieves authorisation token $t_i = \delta(s_i, r_i)$

Eligibility

Proof of Eligiblity

pk_i and its signature t_i from A

Proving Aggregate Authorship of a:

- generate signature for a_i and its properties $p(a)$:
  $s_a = \sigma_i(\eta(a), p(a))$ with hashing function $\eta(\cdot)$

Proof of Auhorship

a, $p(a)$, s_a, and proof of eligibility pk_i, t_i

Dealing with Dishonest Peers

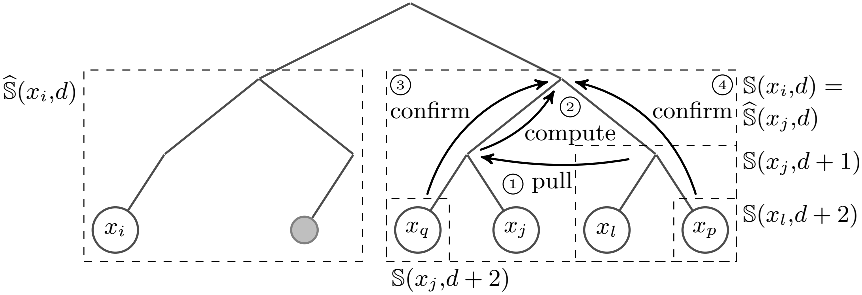
What if peers provide manipulated aggregates?

Assumption

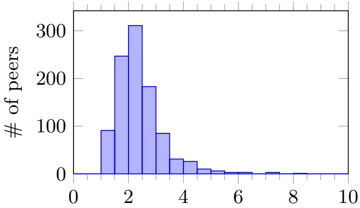
The majority of peers is honest.

- conflicting signatures of P_i constitute proof of deviation
- proofs lead to ban of peers and are stored in the DHT
- signature conflicts:
  - signatures of two distinict initial aggregates from same peer
  - signatures on parent aggregates not based on signed child aggregates
- in case of diverging aggregates:
  take aggregate with most signatures after sampling

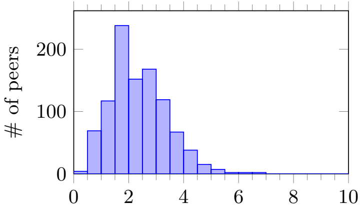
Confirmation Requests



knowledge Distribution



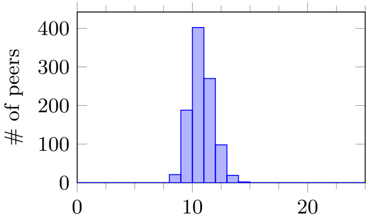
(a) Histogram of leaked information L_i.



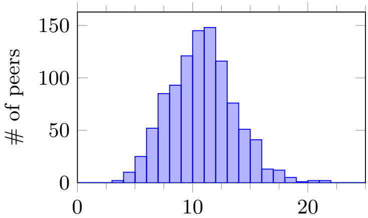
(b) Histogram of received information R_i.

In a simulation with $n = 1000$, peers leak (a), respectively receive (b), information on initial aggregates depending on the global distribution of peers on the binary Kademlia tree. L_i peaks close to the theoretical value 2 of an optimally balanced tree. Only few peers leak significantly more. While the mean for R_i is the same, the distribution is slightly different.

Load Distribution



(a) Histogram of # of received responses.



(b) Histogram of # of given responses.

In a simulation with n $= 1000$, the number of given (b) and received (a) responses has been recorded for every peer. While the distribution of received responses is very sharp, the distribution for given responses is twice as broad. In the Kademlia routing tables, some peers are more often represented than others.

Read more about ADVOKAT

Grumbach, S., & Riemann, R. (2017). Secure and trustable distributed aggregation based on Kademlia. In F. Martinelli & S. De Capitani di Vimercati (Eds.), IFIP Advances in Information and Communication Technology (Vol. 502, pp. 171–185). Rome: Springer. doi:10.1007/978-3-319-58469-0_12

Open Access: `https://hal.inria.fr/hal-01529326`