Secure aggregation of distributed information in multi-agent systems

or

How to share secrets securely in front of the spy

Valentin Goranko
Stockholm University
(Joint work with **David Fernández-Duque**)

Lorentz Center Workshop
"To Be Announced! Synthesis of Epistemic Protocols"
Leiden, August 18, 2015

Outline

1. Secure aggregation of distributed information (SADI) problems: basic concepts.
2. An illustrative example.
3. The general 3-agent case.
4. Reduction methods for solving multi-agent SADI problems.
5. Extensions of the method and concluding remarks.

Reference:
David Fernández Duque and Valentin Goranko,
Secure aggregation of distributed information: How a team of agents can safely share secrets in front of a spy,
Discrete Applied Mathematics, in press.
`http://www.sciencedirect.com/science/article/pii/S0166218X1500308X`
Arxiv version: `http://arxiv.org/abs/1407.7582`

Introduction:
Secure Aggregation of Distributed Information problems

- Team of agents have information distributed amongst them.
- They have to exchange and aggregate that information as common knowledge within the group.
- The exchange is over insecure communication channels and is presumed intercepted by an adversary "eavesdropper".
- The team's task: to aggregate the distributed information, following a prearranged protocol, so that during and after the exchange the adversary does not learn "important pieces" of the information.
- We are interested in *absolute* information security, based not on computationally hard to break encrypting, but on the combinatorial properties of the communication protocols.

Modelling of SADI problems with card deals

We assume that:

- the information of each agent is encoded by a set of "cards" that she holds in her hands.
- The cards are drawn from a publicly known deck.
- Every card is in the hands of exactly one agent of the team and only she can see it.

The goal of the team:

– to exchange and spread across the whole team the information about how the cards are distributed amongst the agents,

– by following a (presumably) *publicly known* protocol,

– communicating by public announcements over insecure channels,

– so that the "eavesdropper" does not learn the location of any of the cards by using the exchange protocol and analyzing the announcements exchanged in the team.

SADI problems and Generalized Russian Cards problems

Russian Cards problem: 2 agents A, B plus an eavesdropper E are dealt respectively 3,3 and 1 cards from the deck $\{1, 2, 3, 4, 5, 6, 7\}$.

A and B must exchange public announcements so as to inform each other about their hands, in such a way that E does not learn the location of any of the other cards.

Generalized Russian Cards (GRC) problems: still 2 agents plus an eavesdropper, with respectively m, n, k cards.

SADI problems: many agents.

In GRC problems the eavesdropper holds cards.
In SADI problems – not. This makes essential difference.

Distribution types, decks, deals, hands

- Distribution type: a vector $\bar{s} = (s_P)_{P \in \mathrm{Agt}}$ of natural numbers. We denote $|\bar{s}| := \sum_{P \in \mathrm{Agt}} s_P$. This is the size of $\bar{s}$.
- Deck for $\bar{s}$: a set of cards $\mathrm{Deck} = \{1, \ldots, |\bar{s}|\}$.
- A deal (distribution) of type $\bar{s}$ over Deck: partition of Deck $H = (H_P)_{P \in \mathrm{Agt}}$, such that $|H_P| = s_P$ for each agent P.

  H_P is the hand of P.

Announcements, runs and protocols

- Agents exchange info by making announcements (actions). Typically: a set of deals. Also, possibly "pass" or "end".
- Run: sequence of announcements made by the agents in turns (in a pre-defined order). Terminal and non-terminal runs.
- Protocol (for $\bar{s}$): a function π assigning to every deal H and every non-terminal run $\rho \in \mathrm{Run}$ a non-empty set of actions for the agent in turn, that only depend on her hand and ρ.
- A (terminal) execution of a protocol π is a pair (H, ρ) such that ρ is a (terminal) run consistent with π and the deal H.

Informativity and safety

The execution of a protocol is:

- informative for an agent P if at the end of the run the agent knows the precise card distribution.
- safe for a card c if at the end of the run the eavesdropper does not learn which agent holds c.

There are precise definitions of these.

A protocol π is:

I: informative if every terminating execution of π is informative for *every* agent in Agt.

S: safe if every execution of π is safe for *every* card c.

There are other informatively and safety conditions.

SADI problems formally defined

A Secure Aggregation of Distributed Information (SADI) problem is a triple $(\bar{s}, \iota, \sigma)$ consisting of:

– a distribution type $\bar{s}$,

– an informativity condition ι,

– and a safety condition σ.

A SADI problem $(\bar{s}, \iota, \sigma)$ is solvable if there exists a terminating protocol π for $\bar{s}$ that satisfies the safety condition ι and the informativity condition σ.
Every such protocol is called a solution of the SADI problem.

Hereafter we focus on the case of SADI problems for *safe and informative protocols*, i.e., of the type $(\bar{s}, \mathrm{I}, \mathrm{S})$.

Illustrative example

Consider a team of 3 agents, Alice ($\mathcal{A}$), Bob ($\mathcal{B}$) and Cath ($\mathcal{C}$).

They hold respectively $2, 3$ and 4 cards from the deck $\{1, \ldots, 9\}$.

$H_{\mathcal{A}}|H_{\mathcal{B}}|H_{\mathcal{C}}$ denotes the deal.

W.l.o.g., $\mathcal{A}$ gets $\{1, 2\}$, $\mathcal{B}$ gets $\{3, 4, 5\}$, and $\mathcal{C}$ gets $\{6, 7, 8, 9\}$.

Thus, the deal is: $1, 2 \mid 3, 4, 5 \mid 6, 7, 8, 9$.

Illustrating example: designing a protocol

Step 1. Alice chooses at random a card not in her hands, say 9.

Then she makes an announcement, saying (essentially):

My cards are among $\{1, 2, 9\}$.

After this announcement, the agent who holds the extra card (9) – in this case Cath – knows the card distribution.

Step 2. Cath makes the next announcement. (This is safe.)

The set of cards $\Gamma = \{1, 2, 9\}$ may be distributed among Alice and Cath in three possible ways: $1, 2 \mid 9$, $2, 9 \mid 1$ or $1, 9 \mid 2$.

The idea: Cath must relate each of these with a respective hand for $\mathcal{B}$ in a safe and informative way.

That is, Cath is to make an announcement of the type:

"If I hold 1 then the hand of $\mathcal{B}$ is . . . ,
and if I hold 2 then the hand of $\mathcal{B}$ is . . . ,
and if I hold 9 then the hand of $\mathcal{B}$ is . . . ."

Illustrative example: the protocol continued

Let Δ be the set of possible hands of $\mathcal{B}$. For safe announcement of the distribution, Cath must choose a mapping $f : \Gamma \to \Delta$ such that:

1. All cards that $\mathcal{B}$ could have are mentioned in the values of f.
2. No card belongs to all values of the mapping.
3. The mapping is injective.
4. Cath's actual card is mapped to Bob's actual hand.
5. All other values of the mapping are chosen at random.

We call such mappings spreads.

One such spread is:

$$f(\mathcal{C} : 9) = \mathcal{B} : \{3, 4, 5\},$$
$$f(\mathcal{C} : 2) = \mathcal{B} : \{5, 6, 7\},$$
$$f(\mathcal{C} : 1) = \mathcal{B} : \{6, 7, 8\}.$$

Illustrative example: the protocol completed

Using the spread

$$f(\mathcal{C}:9) = \mathcal{B}:\{3,4,5\},$$
$$f(\mathcal{C}:2) = \mathcal{B}:\{5,6,7\},$$
$$f(\mathcal{C}:1) = \mathcal{B}:\{6,7,8\},$$

Cath now announces that:

The actual deal belongs to the set

$$\{1,2 \mid \mathbf{3},\mathbf{4},\mathbf{5} \mid 6,7,8,\mathbf{9};\ \ 1,9 \mid \mathbf{5},\mathbf{6},\mathbf{7} \mid \mathbf{2},3,4,8;\ \ 2,9 \mid \mathbf{6},\mathbf{7},\mathbf{8} \mid \mathbf{1},3,4,5\}$$

This announcement completes the protocol.

Claim: This protocol is informative and safe.

The general 3-agent case

When does the protocol from the example work for the SADI problem with distribution type $\bar{s}_3 = (a, b, c)$?

W.l.o.g. we assume that (a, b, c) are arranged so that the agent with a cards makes the 1st announcement and that $b \leq c$. Then:

Theorem

The following conditions are necessary and sufficient for the protocol from the example to solve the SADI problem $(\bar{s}_3, \mathrm{I}, \mathrm{S})$*:*

1. $\binom{b+c-1}{c} \geq a + 1$.
2. $a(b-1) \geq c$.

Some cases where no ordering of (a, b, c) satisfies these conditions:

- $(1, b, c)$ for any b, c.
- $(2, b, c)$ for any b, c, such that $c > 2b - 2$.
  E.g., $(2, 2, 3)$, $(2, 3, 5)$, etc.

Complete analysis of the 3-agent case: still open.

Three "simple" exercises

Determine whether the SADI problem for each of the following distribution types is solvable, by either designing a provably safe and informative solution protocol or proving that one does not exist.

1. **(1,1,1)**
   (Note the correction of the claim in the original talk slides.)
2. **(1,1,2)**
3. **(1,1,3)**

An extension of the spread-based method

An extension of the method behind the presented protocol:

- One of the agents, say $\mathcal{A}$, chooses not just one extra card, but a proper (and not too large) superset S of her hand and announces that her hand is included in S.
- The protocol continues with announcements by the other agents, using generalised spreads, taking into account all possibilities for $\mathcal{A}$'s cards.
- Thus, the SADI problem is reduced to another problem with one agent less.
- The protocol evolves recursively.

General reduction method

The idea: an agent chooses a subset $S \subset \text{Deck}$ and then all agents make a round of announcements declaring how many cards from S they hold.

This splits the SADI problem to two smaller problems:
one with distribution within S and the other – within $\text{Deck} \setminus S$.

If each of the reduced problems is solvable, then the original problem is solvable, too.

For this reduction to work, the player who selects the set S must take into account some safety conditions.

Sometimes, one attempt for a reduction may not work, and then another choice of a splitting set maybe be needed.

This raises more safety concerns.

The complete analysis of the reduction method is still open.

Refinement: reduction by fusion

The general reduction method is not guaranteed to succeed.

A refinement:
Split the problem in two sub-problems, Σ_1 and Σ_2, each involving a subset of the team, such that the two sub-teams share an agent $\mathcal{A}$.

Suppose safe and informative protocols are found for each of Σ_1 and Σ_2. Once both sub-teams learns their respective deals, $\mathcal{A}$ has to communicate them safely to the whole team.

The idea: to use generalised spreads and fuse two spreads with the same number of deals into one, which $\mathcal{A}$ announces to the team.

This idea can be applied recursively to the sub-problems, until they are reduced to immediately solvable ones.

A simple example that can be solved using such reduction: the SADI problem for distribution type $\bar{s} = (k-1, k, k, \ldots, k)$.

The complete analysis of this method is still open.

Diffusions and k-solvability

A diffusion for a card distribution type $\bar{s}$ is a set of deals $\Delta \subseteq \mathrm{Deal}(\bar{s})$ such that:

1. If $H, H' \in \Delta$ are such that $H \neq H'$ and P is any agent then $H_P \neq H'_P$,
2. For every card $c \in \mathrm{Deck}$ there are $H, H' \in \Delta$ and an agent P such that $c \in H_P$ but $c \notin H'_P$.

If $\#\Delta = k$, we say that Δ is a k-diffusion or Δ has *size* k.

Let $\Sigma = (\bar{s}, \mathrm{I}, \mathrm{S})$ be a SADI problem and let $k \in \mathbb{N}$. A protocol (j, π) is a k-solution for Σ if whenever (H, ρ) is a terminal execution of π, there is a k-diffusion Δ such that $H \in \Delta$.

Σ is k-solvable if it has a k-solution.

Solving multi-agent SADI problems by iterated reduction: an example

Consider the SADI problem $\mathcal{P}$ with distribution type $(2, 3, 3, 3)$. Let Deck $= \{0, 1, \ldots, 9, 10\}$ and let the set of agents be $\{\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3\}$.

W.l.o.g. we can assume the deal is

$$H = A_0|A_1|A_2|A_3 = 0,1 \mid 2,3,4 \mid 5,6,7 \mid 8,9,10$$

Protocol for SADI(2, 3, 3, 3)

The deal again: $H = A_0|A_1|A_2|A_3 = 0,1 \mid 2,3,4 \mid 5,6,7 \mid 8,9,10$

Step 1. The agent with 2 cards (here, agent $\mathcal{A}_0$) chooses randomly an additional card x_0 and announces
"All my cards are in the set $A^0 = A_0 \cup \{x_0\}$".

Step 2. W.l.o.g. let $x_0 = 2$, so the agent who has x_0 is $\mathcal{A}_1$. Now, $\mathcal{A}_1$ knows the hand of $\mathcal{A}_0$ and solving SADI(2, 3, 3, 3) is reduced to solving the following two simpler SADI problems:

1. $\Sigma_1 = \mathrm{SADI}(2, 1, 0, 0)$, for the deal

$$A_0|x_0| \cdot |\cdot$$

   Essentially, this is a SADI problem of type $(2, 1)$ involving only $\mathcal{A}_0$ and $\mathcal{A}_1$. It is immediately 3-solvable, using the 3-diffusion

$$\Delta_1 = \{0, 1 \mid 2 \mid \cdot \; ; \; 1, 2 \mid 0 \mid \cdot \; ; \; 2, 0 \mid 1 \mid \cdot\}$$

2. $\Sigma_2 = \mathrm{SADI}(0, 2, 3, 3)$, for the deal

$$\cdot|A_1 \setminus \{x_0\}|A_2|A_3.$$

Protocol for SADI(2, 3, 3, 3) continued

Now, the protocol essentially calls itself recursively for the problem $\Sigma_2 = \text{SADI}(2, 3, 3)$ on the deal $H_1 = A_1' | A_2 | A_3$ where $A_1' = A_1 \setminus \{2\}$.

Step 2.1 Agent $\mathcal{A}_1$ chooses randomly an additional card x_1 from the current deal H_1 and announces: "All my cards, excluding the card mentioned in A^0, are in the set $A^1 = A_1' \cup \{x_1\}$".

Step 2.2 Suppose again w.l.o.g., that $x_1 = 5$ and hence the agent who has the card x_1 is $\mathcal{A}_2$. Now agent $\mathcal{A}_2$ knows the hand A_1' of agent $\mathcal{A}_1$ in the deal H_1 (and therefore the entire deal H_1).

The problem Σ_2 is now reduced to solving the following two simpler SADI problems:

1. $\Sigma_{21} = \text{SADI}(2, 1, 0)$, for the deal $(A_1' \mid \{x_1\} \mid \emptyset)$.
2. $\Sigma_{22} = \text{SADI}(0, 2, 3)$, for the deal $H_2 = (\emptyset \mid A_2 \setminus \{x_1\} \mid A_3)$.

This is a base case, as both problems are immediately 3-solvable.

Protocol for SADI(2, 3, 3, 3) continued

The only 3-diffusion for Σ_{21} is

$$\Delta_{21} = 3,4 \mid 5 \mid \cdot;\ 4,5 \mid 3 \mid \cdot;\ 5,3 \mid 4 \mid \cdot$$

A random 3-diffusion for Σ_{22} involving the actual deal H_2 is e.g.:

$$\Delta_{22} = \cdot \mid 6,7 \mid 8,9,10;\ \cdot \mid 8,9 \mid 10,6,7;\ \cdot \mid 8,10 \mid 6,7,9.$$

Now, in order for the only agent involved in both problems, $\mathcal{A}_2$, to communicate the deal H_1 to $\mathcal{A}_1$ and $\mathcal{A}_3$, she "fuses" Δ_{21} and Δ_{22} by choosing a bijection $f : \Delta_{21} \to \Delta_{22}$, e.g.:

$$f(3,4 \mid 5 \mid \cdot) = \cdot \mid 6,7 \mid 8,9,10$$
$$f(4,5 \mid 3 \mid \cdot) = \cdot \mid 8,9 \mid 10,6,7$$
$$f(5,3 \mid 4 \mid \cdot) = \cdot \mid 8,10 \mid 6,7,9.$$

The fusion $\Delta_{21} \oplus_f \Delta_{22}$ of Δ_{21} and Δ_{22} through f is the 3-diffusion

$$\Delta_2 = \{3,4 \mid 5,6,7 \mid 8,9,10;\ 4,5 \mid 3,8,9 \mid 6,7,10;\ 3,5 \mid 4,8,10 \mid 6,7,9\}.$$

Protocol for SADI(2, 3, 3, 3) continued

Now, agent $\mathcal{A}_2$ announces: "The deal H_1 belongs to the set Δ_2".

This announcement completes the exchange for Σ_2.

It is clearly informative for all agents involved in it, i.e., $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, because the first deal in Δ_2 is the only one consistent with their hands. It is safe, too, because of the properties of diffusions. Indeed, every execution of the protocol for Σ_2 is card-safe because:

- after the announcement of $\mathcal{A}_0$ the eavesdropper $\mathcal{E}$ does not learn the ownership of any card amongst $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$;
- $\mathcal{A}_1$'s announcement leaves each deal in Δ_2 possible for $\mathcal{E}$;
- for every card c in H_1 there are two deals in the diffusion Δ_2 announced by $\mathcal{A}_2$ which send that card in different hands.

Thus, $\mathcal{E}$ does not learn the distribution of any card in H_1.

Protocol for SADI(2, 3, 3, 3) completed

Step 3. Likewise, $\mathcal{A}_1$, as the only agent involved in the problems Σ_1 and Σ_2, knows the entire deal H. In order to communicate it to the others, she constructs the fusion of the 3-diffusions Δ_1 and Δ_2 randomly ordered so as to keep the actual deals aligned, to obtain a 3-diffusion for the original problem Σ:

$$\Delta = \Delta_1 \oplus \Delta_2 = \Big\{ 0,1 \mid 2,3,4 \mid 5,6,7 \mid 8,9,10;$$
$$1,2 \mid 0,4,5 \mid 3,8,9 \mid 6,7,10;$$
$$2,0 \mid 1,3,5 \mid 4,8,10 \mid 6,7,9 \Big\}$$

Finally, agent $\mathcal{A}_1$ announces: *"The deal H belongs to the set Δ".*

This completes the execution of the protocol for Σ_2.

It is clearly informative for all agents $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, because the first deal in Δ is the only one consistent with their hands, and it is safe, because of the properties of diffusions and the construction.

Solvability for 'normal' distrubutions

Definition

A distribution type is k-normal if there are at least two agents, and there is an agent $\mathcal{A}$ such that

1. $s_{\mathcal{A}} \equiv -1 \pmod{k}$
2. if $P \neq \mathcal{A}$, $s_P \equiv 0 \pmod{k}$
3. if P is any agent, $s_P \leq (k-1)^2$.

Theorem

Given $k > 2$ and any k-normal distribution $\bar{s}$, the SADI *problem* $(\bar{s}, \mathrm{I}, \mathrm{S})$ *is k-solvable.*

Two general solvability results

Theorem (Restricted solvability)

Given $m > 2$ and $k > 2m$ there exists N such that whenever $\bar{s}$ is a distribution for m players such that $|\bar{s}| > N$ and for each P, $k^2 \leq ks_P \leq (k-1)|\bar{s}|$, then $(\bar{s}, \mathrm{S}, \mathrm{I})$ is k-solvable.

Theorem (Unrestricted solvability)

Given m there is N such that whenever $|\bar{s}| > N$ is a distribution over at most m players and each player holds at least $\frac{1}{2}\sqrt{|\bar{s}|/m}$ cards then $(\bar{s}, \mathrm{S}, \mathrm{I})$ is solvable.

Further extension: pseudo-randomisation

Sometimes purely deterministic solutions are impossible.

Then, some pseudo-randomisation can be applied on the order of the agents to act and on the choices of their actions, so as not to reveal critical info.

This is still an early development.

Practical considerations and scaling up the method

While formally safe, the solutions of SADI problems may not be practically very secure as the eavesdropper may be able to guess the distribution from few possibilities left after the exchange.

There are several ways to make the method practically secure, with low probability of guessing, set in advance. For instance:

- Choosing a large enough k for k-solvability, so as to increase the number of possibilities for the eavesdropper to consider.
- Adding extra 'sugar' / irrelevant information in the exchange.
- Splitting the information into many bits and running exchange protocols for each of them independently.

Summary and concluding remarks

We have introduced and studied the SADI problems modelling the secure exchange and aggregation of distributed information in multi-agent systems by public announcements.

SADI problems are about synthesising epistemic protocols, satisfying both positive (informatively) and negative (safety) objectives.

We are after absolute info security, based not on hard to break encrypting but on the combinatorial properties of the protocols.

We have developed methods for solving such problems directly, as well as by iterated reduction to simpler ones.

Our methods solve a large class of problems, but there is no complete analysis and description of all solvable SADI problems yet.

Potential applications for distribution of large amounts of distributed information.

Follow-up work: secure aggregation amongst competing teams.

The End