Coalition-Resistant Peer Rating for Long-Term Confidentiality

Giulia Traverso, Denis Butin, Johannes Buchmann
TU Darmstadt
Darmstadt, Germany
gtraverso@cdc.informatik.tu-darmstadt.de

Alex Palesandro
D2SI
Paris, France
alex.palesandro@d2-si.eu

***Abstract*—The outsourced storage of sensitive data requires long-term confidentiality guarantees. Proactive secret sharing in a distributed storage system provides such guarantees. However, some storage service providers lack in reliability or performance for proactive secret sharing to be viable, which can threaten data confidentiality. Data owners need guidance to select the best-performing storage service providers. Aggregated peer ratings with a mediator can provide such guidance. Nevertheless, providers may rate each other inaccurately to undermine competitors. This rational behaviour must be taken into account to devise performance scoring mechanisms generating accurate aggregate scores. The natural formalism to analyse the strategies of rational agents is game theory. In this paper, we introduce a game-theoretic model of the peer rating strategies of providers. Within this model, we first show that an unincentivised performance scoring mechanism results in providers reporting inaccurate ratings. We then introduce an incentivised performance scoring mechanism, modelled as an infinitely repeated game, that discourages inaccurate ratings. We prove that this mechanism leads to accurate ratings and thus to accurate performance scores for each provider, within a margin depending on coalition sizes.**

***Index Terms*—Long-term confidentiality; Reputation mechanisms; Game theory**

I. Introduction

Sensitive data often require long-term protection: their security must be guaranteed over decades. For instance, the confidentiality of an electronic health record must be protected for the entire lifetime of the patient. Currently used cryptography is unsuitable for securely storing data over lengthy periods [6]. One threat is cryptanalytic progress: currently used encryption algorithms are not immune to sudden compromise, e.g. the discovery of a significantly improved integer factorisation algorithm. Another threat is the emergence of quantum computing. Most currently used public-key cryptography (e.g., RSA) is vulnerable to Shor's algorithm [20], which can run on a sizeable quantum computer. Large quantum computers are not yet available, but significant engineering progress is ongoing [22]. Furthermore, adversaries such as intelligence agencies may store vast quantities of encrypted data now, only to decrypt them decades later once the underlying mathematical problems become tractable.

978-1-5386-7493-2/18/$31.00

Several secret sharing [5], [19], [21] schemes provide information-theoretic confidentiality, also called unconditional security: they are not vulnerable to cryptanalytic progress or quantum computing. Secret sharing is thus suitable for the long-term protection of sensitive data. In the special case of *proactive* secret sharing schemes [9], shares are renewed periodically. This category of secret sharing schemes is especially suitable for long-term confidentiality [3]. Indeed, the periodic update of shares now requires an adversary to obtain sufficiently many shares *within a given time period* in order to learn the secret.

In an outsourcing scenario, it is natural to perform proactive secret sharing on a distributed storage system consisting of several storage service providers (SSPs). If a single SSP is used instead, all shares are vulnerable to a single point of failure (the key management of the SSP), even if the shares are spread across multiple data centres. From now on, we therefore assume that a data owner wants to perform proactive secret sharing on a distributed storage system. This is a viable solution for information-theoretic long-term confidentiality in distributed storage systems from a purely cryptographic perspective. However, long-term confidentiality relies in practice on high-performing SSPs, reliably carrying out proactive secret sharing. SSP performance vary, and data owners require guidance to select the individual SSPs making up the distributed storage system, as pointed out by NIST for the special case of cloud infrastructures [10]. The performance (understood in this paper in a broad sense, notably including reliability) of the individual SSPs is the main criterion for inclusion in the distributed storage system. However, data owners do not have access to comprehensive and comparable performance figures for candidate SSPs.

Data owners can be guided in their choice if a trusted third party measures and publishes performance figures. However, for a large number of SSPs and frequent measurements, the workload becomes unwieldy for a single entity. Aggregated peer rating is an alternative approach that distributes the workload. In this approach, SSPs provide mutual ratings for each others' performances. However, SSPs benefit commercially from providing selfish ratings

to undermine competitors. Thus, the aggregate scores computed through the ratings are unreliable. For aggregate scores to yield accurate results, a performance scoring mechanism encouraging accurate ratings is needed.

In this setting, it makes more sense to view the SSPs as rational agents (trying to increase their own utility) than to classify them as "good" or "bad". This mirrors the *rational secret sharing* modelling by Halpern and Teague [8]. The natural formalism to analyse the strategies of rational agents is game theory.

a) Contributions: We address the above scenario with two main contributions:

1) We formalize, for the first time, the computation of accurate aggregate scores through mutual ratings in distributed storage systems as an economic problem and provide a game-theoretical model of the peer rating strategies of SSPs.
2) We design an incentivised performance scoring mechanism, mediated by a trusted authority (TA). This mechanism encourages accurate ratings from the SSPs. The accuracy of the computed aggregate scores is resilient to a certain amount of inaccurate ratings from *coalitions* (coordinated groups) of SSPs, assuming a honest majority of SSPs. Data owners obtain these aggregate scores from the TA, and accordingly select the best-rated SSPs for their distributed storage system. In our game-theoretic formalism, the performance scoring mechanism is modelled as an infinitely repeated game. The global accuracy guarantee is modelled by the performance scoring mechanism achieving a k-resilient equilibrium — a strong version of a Nash equilibrium.

We stress that our contribution does not aim at cryptographically improving proactive secret sharing in distributed storage systems, which is a well established solution for long-term confidentiality of outsourced data [4], [9]. Instead, our goal is decision support for the selection of high-performing SSPs storing shares, so as to enable reliable and practical long-term confidential data storage.

b) Outline: We first recall necessary notions of game theory (Sec. II), and formalize performance scoring mechanisms and rational SSP peer rating strategies in a game-theoretic model (Sec. III). We next show that aggregate performance scores are not accurate if participating (rational) SSPs are not incentivised to report performance measurements faithfully (Sec. IV). Afterwards, we present our incentivised performance scoring mechanism and prove that even in the presence of coalitions, it leads to accurate aggregate performance scores, within a margin dependent on maximal coalition size (Sec. V). A survey of related work (Sec. VI) and conclusions (Sec. VII) follow.

II. Notions of Game Theory

We now recall general notions of game theory. In Sec. III, we will use this formalism to model the peer rating strategies of SSPs, seen as *players*.

In game theory [18], a set of players $P = \{P_1, \ldots, P_n\}$, an action profile $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$, where $\mathcal{A}_i$ is the set of actions of player P_i, and a *utility function* $u_i : \mathcal{A} \to \mathbb{R}$ are the components of a *game*. A game is denoted by $\Gamma(P_i, \mathcal{A}_i, u_i)$, for $i = 1, \ldots, n$. The utility function of a player defines its preferences with respect to the actions it takes and to the actions all other players take. That is, given two distinct outcomes $\boldsymbol{a}, \boldsymbol{a}' \in \mathcal{A}$, with $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{a}' = (a'_1, \ldots, a'_n)$, if $u_i(\boldsymbol{a}) \geq u_i(\boldsymbol{a}')$, then player P_i prefers $\boldsymbol{a}$ to $\boldsymbol{a}'$.

Games consist of either one or multiple rounds and the latter are referred to as *repeated games*. Among repeated games, there are *infinitely* and *finitely* repeated games, which consist of, respectively, an infinite or a finite number of rounds. At each round, all players are asked to choose a certain action simultaneously.

Players can rely on different strategies with respect to the actions to choose. A *strategy* for player P_i is a probability distribution $\sigma_i : \mathcal{A}_i \to [0, 1]$ that determines how the actions $a_i \in \mathcal{A}_i$ are chosen, where $\sigma_i \in \mathcal{S}_i$ and $\mathcal{S}_i$ is called the *strategy profile* of player P_i. Since a strategy is a way to choose actions, it is possible to express the expected outcome of a game by using strategies. That is, a game can be denoted by $\Gamma(P_i, \sigma_i, u_i)$, for $i = 1, \ldots, n$ and $\sigma_i \in \mathcal{S}_i$. Given $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ the tuple composed of each player P_i's strategy σ_i, the utility of player P_i when strategy $\boldsymbol{\sigma}$ is played can be denoted by $u_i(\boldsymbol{\sigma})$. We refer to $\boldsymbol{\sigma}$ as the *joint strategy* of players $P_1, \ldots, P_n$.

We denote by $(\sigma'_i, \boldsymbol{\sigma}_{-i})$ the vector of strategies where all players maintained the same strategies of the joint strategy $\boldsymbol{\sigma}$, except for player P_i, which replaced strategy σ_i by another strategy σ'_i. That is, $(\sigma'_i, \boldsymbol{\sigma}_{-i}) = (\sigma_1, \ldots, \sigma_{i-1}, \sigma'_i, \sigma_{i+1}, \ldots, \sigma_n)$. Given a subset $C \subset P$ of players of cardinality $n_C \leq n$, we denote by $(\boldsymbol{\sigma}'_C, \boldsymbol{\sigma}_{-C})$ the vector of strategies where all players maintained the same strategies of the joint strategy $\boldsymbol{\sigma}$, except for player $P_j \in C$, which replaced strategy σ_j by another strategy σ'_j, for $j = 1, \ldots, n_C$. That is, $(\boldsymbol{\sigma}'_C, \boldsymbol{\sigma}_{-C}) = (\sigma'_1, \ldots, \sigma'_{n_C-1}, \sigma'_{n_C}, \sigma_{n_C+1}, \ldots, \sigma_n)$, where, without loss of generality, players in C are the first n_C players.

Players act as rational deciders, meaning that they always play the strategy that maximizes their utilities, which depends on the other players' actions. No player knows what strategies other players adopt at a given round, and can only form beliefs in this respect. The pay-off players get is not only subjected to the way they choose strategies, but also to the type of game that is played.

There are two types of game. A *non-cooperative* (or strategic) game deals with actions chosen by players individually and the pay-off of a player is given solely by its utility function. Instead, a *cooperative* (or coalitional) game deals with the actions a subset of players agree to take collectively and their pay-off is given by splitting the overall utility among themselves. In non-cooperative games, the concept of a *Nash equilibrium* conveys the idea that players choose a strategy by both looking at the best available

actions and by taking into account the belief of how the other players might behave. A joint strategy $\boldsymbol{\sigma}$ such that no player P_i alone has an incentive in choosing a strategy σ_i' other than σ_i, while all other players P_j stick to strategy $\sigma_{j\neq i}$ is called a Nash equilibrium.

Definition 1 (Osborne [18]). *A joint strategy $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is called a* Nash equilibrium *if $u_i(\sigma_i', \boldsymbol{\sigma}_{-i}) \leq u_i(\sigma_i, \boldsymbol{\sigma}_{-i})$, for each player P_i, where $\sigma_i' \neq \sigma_i$. It is a* strict Nash equilibrium *if $u_i(\sigma_i', \boldsymbol{\sigma}_{-i}) < u_i(\sigma_i, \boldsymbol{\sigma}_{-i})$.*

The counterpart of a Nash equilibrium for cooperative games is introduced in [1] and called a *k-resilient equilibrium*.

Definition 2 (Abraham [1]). *A joint strategy $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is a* k-resilient equilibrium *if $u_i(\boldsymbol{\sigma}'_C, \boldsymbol{\sigma}_{-C}) \leq u_i(\boldsymbol{\sigma}_C, \boldsymbol{\sigma}_{-C})$, for each subset $C \subset P$ of cardinality $n_C \leq k$, where $\sigma_i' \neq \sigma_i$, for $P_i \in C$. It is a* strongly resilient equilibrium *if it is a k-resilient equilibrium for $k \leq n-1$.*

Def. 1 and Def. 2 can also be expressed by using a game outcome $\boldsymbol{a}$ sampled from the action profile instead of the joint strategy $\boldsymbol{\sigma}$. From now on, we refer to the subset C of players deviating from the joint strategy in Def. 2 as a *coalition*. Another important concept in game theory is the one of dominated strategies (or actions). A strategy is dominated by another strategy if it always provides the player with a lower pay-off. This concept is formalized in the definition below, denoting by $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$ the strategy profile of players $P_1, \ldots, P_n$ and by $\mathcal{S}_{-i} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_{i-1} \times \mathcal{S}_{i+1} \times \cdots \times \mathcal{S}_n$ the strategy profile obtained by removing the set $\mathcal{S}_i$ of possible strategies for player P_i from $\mathcal{S}$.

Definition 3 (Osborne [18]). *Given two strategies $\sigma_i, \sigma_i' \in \mathcal{S}_i$ for player P_i, strategy σ_i is* weakly dominated *by strategy σ_i' if $u_i(\sigma_i, \boldsymbol{\sigma}_{-i}) \leq u_i(\sigma_i', \boldsymbol{\sigma}_{-i})$ for each $\boldsymbol{\sigma}_{-i} \in \mathcal{S}_{-i}$ and there exist $\boldsymbol{\sigma}'_{-i} \in \mathcal{S}_{-i}$ such that $u_i(\sigma_i, \boldsymbol{\sigma}'_{-i}) < u_i(\sigma_i', \boldsymbol{\sigma}'_{-i})$. Strategy σ_i is* strictly dominated *if $u_i(\sigma_i, \boldsymbol{\sigma}_{-i}) < u_i(\sigma_i', \boldsymbol{\sigma}_{-i})$ for each strategy $\sigma_i' \in \mathcal{S}_i$.*

III. A Game-Theoretic Model of Rating Strategies

We recall the context of our scenario. As described in Sec. I, the long-term protection of outsourced data entails a distributed storage system where multiple SSPs perform proactive secret sharing. In this scenario, SSPs are rational agents aiming to maximize their utility, which is the economic return from their business: data (here, shares) storage. SSPs thus aim to maximize the number of shares they get.

A way to obtain more shares is to obtain a comparatively high aggregate score. This can be achieved by behaving rationally (i.e., selfishly) during the performance scoring mechanism. SSPs send potentially selfish mutual ratings to the TA. The TA collects these ratings and computes the aggregate scores. Based on them, the data owner distributes the shares to high-performing SSPs. As mentioned earlier (Sec. I), we use *performance* in a broad sense, including qualities such as reliability. The performance metric used in practice is independent of our model and results. We formalize this distributed storage system scenario through a game-theoretic perspective modelling the rating strategies of the SSPs.

In game theory, the agents involved are referred to as players. In our scenario, the players are the n SSPs involved in the distributed storage system. From now on, we refer to those SSPs as players $P_1, \ldots, P_n$. Players $P_1, \ldots, P_n$ are interested in maximizing their economic return when they offer long-term storage to data owners via a distributed storage system. This entails preferences with respect to the shares that data owners may distribute or withdraw from them, depending on the aggregate scores of players $P_1, \ldots, P_n$. In other words, the aim of players $P_1, \ldots, P_n$ is to increase over time the number of shares they store because this leads to greater income.

In the following, we formalize the preferences of players $P_1, \ldots, P_n$ with respect to gaining or losing shares by defining relevant utility functions. Afterwards, we formalize the preferences of these players with respect to their aggregate scores, because gaining or losing a share depends on the aggregate score assigned to a player. In particular, it depends on the comparison between the aggregate score of a player and the aggregate scores of all other players. Further utility functions related to aggregate scores are defined for players $P_1, \ldots, P_n$.

We first formalize this scenario for non-cooperatives situations, where players act individually. Then, we formalize this scenario for cooperative situations, where players form coalitions among each other. Players in a coalition act in coordination for mutual benefit.

In a non-cooperative setting, we denote by $U_i(\boldsymbol{\sigma})$ the utility function of a player P_i with respect to gaining or losing a share when the joint strategy is $\boldsymbol{\sigma}$. The utility function $U_i(\boldsymbol{\sigma})$ is defined as follows.

- U1) If player P_i gains a share, then $U_i(\boldsymbol{\sigma}) = 1$.
- U2) If player P_i neither gains nor loses a share, then $U_i(\boldsymbol{\sigma}) = 0$.
- U3) If player P_i loses a share, then $U_i(\boldsymbol{\sigma}) = -1$.

Utility $U_i(\boldsymbol{\sigma})$ is directly related to the economic pay-off of player P_i, because the amount of shares players store and manage is proportional to their economic return. However, the amount of shares gained or lost by players ultimately depends on the given aggregate score. The data owner periodically checks the aggregate scores of $P_1, \ldots, P_n$ and accordingly arranges how the shares are distributed among them.

Periodically, the aggregate scores are updated through the aggregation of the players' mutual ratings. Let us denote by r the last round where the aggregate scores are updated before the data owner checks them. We denote the aggregate scores of players $P_1, \ldots, P_n$ at round r by $\tau_1^r, \ldots, \tau_n^r$, where $0 \leq \tau_i^r \leq 1$, for $i = 1, \ldots, n$. Player P_i

eventually gains a share if its aggregate score τ_i^r is high enough to convince the data owner to do so. We formalize this idea of "high enough" as having an aggregate score that is on average higher than all other aggregate scores. Vice versa, having an aggregate score "low enough" to convince the data owner to withdraw a share means that a player P_i is assigned an aggregate score τ_i^r that is on average lower than the aggregate scores of all other players. We formalize this by defining the utility function $u_i(r)$ with respect to the aggregate scores (where the joint strategy $\boldsymbol{\sigma}$ is omitted for simplicity) of a player P_i as follows, for $i = 1, \ldots, n$.

u1) If $\tau_i^r > \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \tau_j^r$, then $u_i(r) = 1$.
u2) If $\tau_i^r = \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \tau_j^r$, then $u_i(r) = 0$.
u3) If $\tau_i^r < \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \tau_j^r$, then $u_i(r) = -1$.

Utility $u_i(r)$ is not directly related to the economic pay-off of player P_i, because having $u_i(r) = 1$ at round r does not necessarily imply that one share is distributed to player P_i. Also, even if its aggregate score is higher than all other aggregate scores at a given round, it does not mean that all other aggregate scores are so low as to convince the data owner to rearrange the shares' distribution. However, consistently getting high aggregate scores is the only way to obtain additional shares. The output of aggregate scores $\tau_1^r, \ldots, \tau_n^r$ at round r is the result of a repeated non-cooperative game among players $P_1, \ldots, P_n$ of r rounds.

Instead, if players $P_1, \ldots, P_n$ form coalitions, then the computation of the aggregate scores $\tau_1^r, \ldots, \tau_n^r$ at round r is the result of a repeated cooperative game of r rounds. In this case, the pay-off from gaining shares is split among the members of a coalition. Thus, the goal for player P_i is that at least one of his coalition partners gains a share and none of them loses any share. We assume that the coalitions are at most as numerous as the players and that each player belongs at most to one coalition. For a coalition C_k, we denote $J = \{j \mid P_j \in C_k \ \wedge \ U_j(\boldsymbol{\sigma}) = 1\}$, and $J' = \{j' \mid P_{j'} \in C_k \ \wedge \ U_{j'}(\boldsymbol{\sigma}) = -1\}$. That is, we see the indexes of players in the coalition C_k as the union of two subsets, J and J', where J are the indexes of the subsets of players in C_k that gained a share and J' are the indexes of the subset of players in C_k that lost a share. The utility function $U_i'(\boldsymbol{\sigma})$ of player $P_i \in C_k$ with respect to gaining or losing shares within the coalition is defined as follows.

U1') If $\sum_{j\in J} U_j(\boldsymbol{\sigma}) > \sum_{j'\in J'} U_{j'}(\boldsymbol{\sigma})$, then $U_i'(\boldsymbol{\sigma}) = 1$.
U2') If $\sum_{j\in J} U_j(\boldsymbol{\sigma}) = \sum_{j'\in J'} U_{j'}(\boldsymbol{\sigma})$, then $U_i'(\boldsymbol{\sigma}) = 0$.
U3') If $\sum_{j\in J} U_j(\boldsymbol{\sigma}) < \sum_{j'\in J'} U_{j'}(\boldsymbol{\sigma})$, then $U_i'(\boldsymbol{\sigma}) = -1$.

In other words, utility $U_i'(\sigma)$ for a player $P_i \in C_k$ is positive if the amount of players within coalition C_k that gained a share is greater than the amount of players within C_k that lost a share. Vice versa, utility $U_i'(\sigma)$ is negative if the amount of players within coalition C_k that gained a share is smaller than the amount of players within C_k that lost a share. The definition of utility $U_i'(\boldsymbol{\sigma})$ with respect to shares shapes the definition of utility $u_i'(r)$ with respect to the aggregate scores, mirroring what holds for non-cooperative games. The goal of player $P_i \in C_k$ is to maximize the average of the aggregate scores assigned to the players within the same coalition. That is, having a "high enough"aggregate score for a player $P_i \in C_k$ means that the average of the aggregate scores of the players within coalition C_k is higher than the average of the aggregate scores of players outside C_k. Vice versa, having a "low enough"aggregate score means that the average of the aggregate scores of the players within coalition C_k is lower than the average of the aggregate scores of players outside C_k. We denote $M = \{m \mid P_m \in C_k\}$, and $L = \{l \mid P_l \notin C_k\}$. The utility function $u_i(r)$ for a player $P_i \in C_k$ with respect to aggregate scores, where coalition C_k has cardinality n_k, is defined as follows, for $i = 1, \ldots, n$.

u1') If $\frac{1}{n_k} \sum_{m\in M} \tau_m^r > \frac{1}{n-n_k} \sum_{l\in L} \tau_l^r$, then $u_i'(r) = 1$.
u2') If $\frac{1}{n_k} \sum_{m\in M} \tau_m^r = \frac{1}{n-n_k} \sum_{l\in L} \tau_l^r$, then $u_i'(r) = 0$.
u3') If $\frac{1}{n_k} \sum_{m\in M} \tau_m^r < \frac{1}{n-n_k} \sum_{l\in L} \tau_l^r$, then $u_i'(r) = -1$.

We defined the utility of gaining and losing shares as, respectively, 1 and -1 for the sake of simplicity. Instead, defining the utility of gaining or losing shares as the number of, respectively, gained or lost shares is also possible but entails additional formalism.

In the remainder of this paper, for simplicity, we use shorthand notation such as $i \in C_k$ to denote $i \in \{j \mid P_j \in C_k\}$ (and similarly for $i \notin C_k$).

IV. Score Inaccuracy for Unincentivised Ratings

Recall that our goal is to guide data owners in selecting SSPs with the highest possible performance to establish a distributed storage system. To this end, SSPs must be assigned aggregate scores, meant to reflect performance, and this can be achieved through a performance scoring mechanism based on peer rating. We assume that SSPs act rationally: they are prepared to give ratings that do not reflect reality if this can help them increase their utility, i.e. the number of shares they are assigned.

We now show that an *unincentivised* performance scoring mechanism (which does not reward accuracy) does not yield accurate aggregate scores. In an unincentivised performance scoring mechanism, the TA computes the aggregate score of player P_i by simply taking into account the rating that each player $P_{j\neq i}$ gives for player P_i, for $i = 1, \ldots, n$. The TA computes the aggregate scores through these ratings (usually by averaging them) and outputs them to the data owner upon request. On the basis of those aggregate scores, the data owner decides whether to rearrange the distribution of the shares and whether to exclude players from the distributed storage system. However, the unincentivised aggregation and average of these ratings do not lead to accurate aggregate scores. In the following, we show this both when single players issue selfish ratings (Sec. IV-A) and when they form coalitions (Sec. IV-B).

A. Score Inaccuracy in Non-Cooperative Games

If the aggregate scores are computed through repeated non-cooperative games, the players $P_1, \ldots, P_n$ do not form coalitions. Thus, here, *selfish* ratings means that players are giving low ratings even to high-performing players. The data owner checks the aggregate scores $\tau_1^r, \ldots, \tau_n^r$, at round, say, r. It is *known* to all players that r is the last round the performance scoring mechanism is run before the data owner eventually reallocates the shares. We denote by $\tau_1^{r'}, \ldots, \tau_n^{r'}$ the aggregate scores of players $P_1, \ldots, P_n$ at round r', which occurs strictly earlier than round r, and assume without loss of generality that $\tau_1^{r'} \geq \ldots \geq \tau_n^{r'}$. We assume that the aggregate score $\tau_i^{r'}$ of player P_i is the t^{th} highest score. Player P_i can choose one among the following (mixed) strategies:

- $\sigma 1$) At all rounds, give low ratings to all other players $P_{j \neq i}$.
- $\sigma 2$) At all rounds, give low ratings to the players $P_1, \ldots, P_{t-1}$ with the 1^{st}, 2^{nd}, $\ldots$, $(t-1)^{\text{th}}$ highest aggregate scores $\tau_1^{r'} \geq \ldots \geq \tau_{t-1}^{r'}$.
- $\sigma 3$) From round r' on, give low ratings to all other players $P_{j \neq i}$.
- $\sigma 4$) From round r' on, give low ratings to players $P_1, \ldots, P_{t-1}$ with the 1^{st}, 2^{nd}, $\ldots$, $(t-1)^{\text{th}}$ highest aggregate scores $\tau_1^{r'} \geq \ldots \geq \tau_{t-1}^{r'}$.
- $\sigma 5$) Always give ratings reflecting the actual measured performance of all other players $P_{j \neq i}$.

Theorem 1 shows that strategy $\sigma 1$ weakly dominates all other strategies for each player P_i when the goal is to maximize its aggregate score τ_i^r at round r.

Theorem 1. *Let $u_1, \ldots, u_n$ be the utilities of, respectively, players $P_1, \ldots, P_n$ satisfying u1)–u3) when the aggregate scores $\tau_1^r, \ldots, \tau_n^r$ at last round r are computed. Then, $\sigma 1$ weakly dominates all other available strategies, i.e. $u_i(\sigma_i', \boldsymbol{\sigma}_{-i}) \leq u_i(\sigma 1, \boldsymbol{\sigma}_{-i})$, for $\sigma_i' \neq \sigma 1$ and $i = 1, \ldots, n$.*

We first outline the structure of the proof. In short, we first show that strategy $\sigma 1$ is weakly dominant for player P_i with the t-th highest aggregate score $\tau_i^{r'}$. This is due to the fact that each other player $P_{j \neq i}$ is assumed to be rational as well and strategy $\sigma 1$ has to be played by player P_i to contrast the undermining effect of all other players. Second, it is proven that this holds even for the case where player P_i has the lowest aggregate score or the highest aggregate score.

We now formally prove Theorem 1 by showing that strategy $\sigma 1$ always dominates a less selfish strategy in the following three steps.

If P_i is given the t^{th} highest aggregate score $\tau_i^{r'}$ by the TA at round $r' < r$, then we have $\tau_i^{r'} \leq \tau_j^{r'}$, for $j = 1, \ldots, t-1$. Depending on the aggregate scores $\tau_1^{r'}, \ldots, \tau_{t-1}^{r'}$ and on t, $u1), u2), u3)$ can all occur. In particular, if either $u2)$ or $u3)$ occurs, then utility $u_i(r') \neq 1$. A way for player P_i to prevent this is to attempt to lower the aggregate scores of players $P_1, \ldots, P_{t-1}$ from round $r'+1$ on to increase the chances that $u1)$ occurs at the final round r. That is, strategy $\sigma 4$ is for player P_i weakly dominant with respect to $\sigma 5$, assuming player $P_{j \neq i}$ plays strategy $\sigma 5$. Furthermore, the more rounds player P_i gives low ratings for players $P_1, \ldots, P_{t-1}$, the more their aggregate scores decrease over time and, thus, the higher the chance that $u1)$ occurs at round r. A weakly dominant strategy for player P_i is to start giving low ratings as early as possible and thus $\sigma 2$ weakly dominates $\sigma 4$, assuming player $P_{j \neq i}$ plays $\sigma 5$. Getting the t^{th} highest aggregate score means that player P_i is the t^{th} best-performing player and that player $P_{j \neq i}$ played $\sigma 5$. In case player $P_{j \neq i}$ plays a strategy other than $\sigma 5$, then player P_i might be given by the TA an aggregate score lower than the t^{th} highest aggregate score $\tau_i^{r'}$, which increases the chances that $u3)$ occurs. Strategy $\sigma 5$ is weakly dominated by $\sigma 2$ and $\sigma 4$ for players $P_{t+1}, \ldots, P_n$, which are, respectively, given the lowest aggregate scores $\tau_{t+1}^{r'}, \ldots, \tau_n^{r'}$. Hence players $P_{t+1}, \ldots, P_n$ give low ratings for players $P_1, \ldots, P_t$, including player P_i itself. In order to compensate low ratings from worse-performing players, player P_i has to give low ratings for players $P_{t+1}, \ldots, P_n$ as well. However, because player P_i already plays $\sigma 2$, it eventually assigns low ratings to each player $P_{j \neq i}$. In case players $P_{t+1}, \ldots, P_n$ play $\sigma 4$, then player P_i can respond by playing $\sigma 3$. However, to maximize the desired effect, it is more effective for player P_i to give low ratings as soon as possible. Thus, it plays $\sigma 1$. In conclusion, player P_i plays $\sigma 1$ because this increases the chances that, at round r, $u1)$ occurs and $u_i(r) = 1$. Therefore, $\sigma 1$ weakly dominates $\sigma 2 - \sigma 5$.

If P_i is given the lowest aggregate score $\tau_i^{r'}$ by the TA at round $r' < r$, then $\tau_i^{r'} < \tau_j^{r'}$, for each player $P_{j \neq i}$. Thus, $u3)$ occurs for player P_i. In order to avoid this, a possible option for player P_i is to try to lower the aggregate scores of all the players that are given higher aggregate scores, which means to give low ratings for player $P_{j \neq i}$ from round $r'+1$ on. Thus, $\sigma 3$, which in this case is identical to $\sigma 4$, weakly dominates $\sigma 5$. However, in order to decrease the probability of $u3)$ to occur, it is better for player P_i to give low ratings to all other players as early as possible. Thus $\sigma 1$ weakly dominates $\sigma 3$. Since $\sigma 1$ is here identical to $\sigma 2$, $\sigma 1$ weakly dominates $\sigma 2$ (by Def. 3). Due to the transitivity of weak strategy dominance, $\sigma 1$ also weakly dominates $\sigma 5$. In conclusion, player P_i plays $\sigma 1$ regardless of what each other player $P_{j \neq i}$ plays because $\sigma 1$ weakly dominates $\sigma 2 - \sigma 5$. This case where player P_i is the worst-performing player can be obtained by induction from *1)* where P_i is given the t^{th} highest aggregate score, with t increasing towards n.

If P_i is given the highest aggregate score $\tau_i^{r'}$ by the TA at round $r' < r$, then $\tau_i^{r'} > \tau_j^{r'}$, where player $P_{j \neq i}$ played $\sigma 5$. So for player P_i, sticking to $\sigma 5$ is sufficient for $u1)$ to occur at round r. However, for the same reasons discussed in *1)*, player $P_{j \neq i}$ is rational and gives low ratings to the players given higher aggregate scores, which always includes player P_i. Thus, $u1)$ is unlikely to occur. To

balance out this effect, player P_i's can give low ratings to the lower performing players, which in this case means to give low ratings to all other players. Thus, player P_i plays $\sigma 3$. However, for reasons discussed in *1)*, $\sigma 1$ weakly dominates $\sigma 3$ in case each other player $P_{j\neq i}$ plays $\sigma 2$ rather than $\sigma 4$. In conclusion, if player P_i is given the highest aggregate score, then $\sigma 1$ weakly dominates $\sigma 2 - \sigma 5$.
Because the above analysis can be applied to each player P_i, for $i = 1, \ldots, n$, we conclude that strategy $\sigma 1$ weakly dominates strategy $\sigma 2 - \sigma 5$ for all players $P_1, \ldots, P_n$. Thus, $u_i(\sigma_i', \boldsymbol{\sigma}_{-i}) \leq u_i(\sigma 1, \boldsymbol{\sigma}_{-i})$, for $\sigma_i' \neq \sigma 1$ and $i = 1, \ldots, n$. □

Since each player P_i plays strategy $\sigma 1$, and thus gives low ratings to each other player $P_{j\neq i}$ at all rounds, the following corollary ensues.

Corollary 1. *When computed through unincentivised ratings, the aggregate scores $\tau_1^r, \ldots, \tau_n^r$ the data owner checks at last round r are inaccurate, i.e. they do not describe the actual performance of, respectively, players $P_1, \ldots, P_n$.*

In particular, in this case where no coalitions are formed, all aggregate scores $\tau_1^r, \ldots, \tau_n^r$ are low. In terms of SSPs and shares, the data owner might believe that all SSPs are underperforming and eventually withdraw all shares from all SSPs. This implies that $U_i(r) = -1$, for $i = 1, \ldots, n$, which is the worst possible situation for an SSP.

B. Score Inaccuracy in Cooperative Games

In a performance scoring mechanism modelled as a cooperative game, players can form coalitions to cooperatively obtain high aggregate scores and coordinate on players to undermine. Here, *selfish* rating means that players give high ratings to coalition partners and low ratings to all other players, regardless of the actual witnessed performance. The strategies a player P_i can adopt combine the strategies listed in Sec. IV-A with the possibility of increasing the aggregate scores of fellow coalition members by giving them high ratings. As for Theorem 1, one can show that the weakly dominant strategy (adapted from $\sigma 1$) is the one where each player simultaneously gives high ratings to fellow coalition members and low ratings to all other players at all rounds. The proof, omitted here due to space constraints, is similar to the one of Theorem 1. Thus, Corollary 1 holds also for the case where coalitions among players are formed.

Theorem 1 and Corollary 1 state the impossibility of getting accurate aggregate scores in unincentivised and finitely repeated games. They mirror the impossibility result for rational secret sharing presented in [8]. The performance scoring mechanism described in this section is a finitely repeated game because all players know that round r is the last one. However, the same impossibility result holds for infinitely repeated games, where it is unknown to the players that round r is the last one. The proof of Theorem 1 is not tied to round r because it is weakly dominant for the players to play strategy $\sigma 1$ as soon as possible. This is true in particular when players do not know which round is the final one, because they possibly have even less chances to cope with undermining ratings of the other players. This problem is solved in the next section, where it is shown that in order to obtain accurate aggregate scores, the performance scoring mechanism has to be modelled *both* as an *incentivised* and *infinitely repeated* game.

V. Score Accuracy for Incentivised Ratings: A Performance Scoring Mechanism Resilient to Coalitions

As just shown, accurate aggregate scores cannot be obtained without incentives for the participating SSPs. We now introduce a novel performance scoring mechanism for distributed storage systems (Sec. V-A) and prove that it leads to accurate aggregate scores (within a margin depending on coalition sizes) and is resilient to coalitions of bounded size, assuming an honest majority of SSPs (Sec. V-B).

A. A New Performance Scoring Mechanism

We design a performance scoring mechanism modelled as an *incentivised*, *infinitely repeated*, and *cooperative*[1] game with a mediator (the TA). Each round of the repeated game consists of running a performance scoring mechanism distributedly. The TA runs the performance scoring mechanism at random intervals, *independently* of the intervals in which share renewal for proactive secret sharing is performed. We emphasise that the performance scoring mechanism is entirely separate from the proactive secret sharing operations (such as share renewal), which are performed as usual. The players do not know how many rounds the performance scoring mechanism is run before the data owner checks the aggregate scores. Thus, even if the performance scoring mechanism is run a finite amount of times, it can be modelled as an infinitely repeated game. This enables the TA to not only collect and processes ratings, but also encourages the submission of accurate ratings and penalizes the players submitting selfish ratings. This allows the definition of a strategy leading to a k-Nash equilibrium and pushes the players to follow it, because incentives and penalties, respectively, positively and negative influence the final aggregate score of the players.

We now provide a high-level description of our performance scoring mechanism. Each aggregate score τ_i^r at round r is computed by the performance scoring mechanism as a convex combination of a first component τ_i', a second component $\tau_i^{''}$, and a third component τ_i^{r-1}, for $i = 1, \ldots, n$. The first component τ_i' is computed through steps 1)–3) of the mechanism and is the aggregate score of all ratings submitted by all players for the player being rated for the current round r. The second component $\tau_i^{''}$

[1]It is sufficient to model the mechanism as a cooperative game, because non-cooperative games are a particular case where coalitions have only one player each.

is computed through steps 4)–6) of the mechanism and is the aggregate score of all incentives and penalties given to player P_i by the TA for, respectively, the accurate and selfish ratings submitted for the current round r. For the first and the second component, the round r is omitted to simplify notation. The third component τ_i^{r-1} is the aggregate score of player P_i at the previous round $r-1$. Thus, it was previously computed and is an input of the performance scoring mechanism run at the current round r. The third component keeps track of the evolution of the performance of a player over time and increasingly rewards good performance and penalizes poor performance. The final aggregate score τ_i^r at round r is computed during step 7) of the mechanism.

We now show how to compute the aggregate scores $\tau_1^r, \ldots, \tau_n^r$ at round r for players $P_1, \ldots, P_n$, where round r is not necessarily the round where the data owner checks the aggregate scores. We defined each aggregate score τ_i^r as a real number between 0 and 1, i.e. $0 \leq \tau_i^r \leq 1$ for $i = 1, \ldots, n$. The same holds for the first and the second component denoted by, respectively, τ_i' and $\tau_i^{''}$, and for the aggregate score τ_i^{r-1} of player P_i at round $r-1$. The rating submitted by player P_i to evaluate player P_j for the computation of its aggregate score τ_j^r is denoted by $\rho_{i,j}^r$, where $0 \leq \rho_{i,j}^r \leq 1$. We denote by $\overline{\tau}_1', \ldots, \overline{\tau}_n'$ the targeted first components, i.e. they are the first components $\tau_1', \ldots, \tau_n'$ when computed through ratings that perfectly match the actual performance of players $P_1, \ldots, P_n$, respectively. We now provide a definition of accurate ratings.

Definition 4. *Let $\overline{\tau}_1', \ldots, \overline{\tau}_n'$ be the targeted first components of the aggregate scores of players $P_1, \ldots, P_n$ and let $t_\varepsilon > 0$. A rating $\rho_{i,j}^r$ submitted by player P_i to evaluate player P_j is called* t_ε-accurate *if $\overline{\tau}_i' - t_\varepsilon \leq \rho_{i,j}^r \leq \overline{\tau}_i' + t_\varepsilon$. We refer to t_ε as the* accuracy threshold. *In cases where t_ε is clear from the context, we simply write* accurate.

We now present the performance scoring mechanism with the TA as a mediator for the computation of the aggregate scores $\tau_1^r, \ldots, \tau_n^r$ of players $P_1, \ldots, P_n$ at round r.

1) The TA selects integers $a, b, c > 0$ such that $a+b+c = 1$. These are the weights associated to $\tau_i', \tau_i^{''}, \tau_i^{r-1}$, respectively, for the computation of the aggregate score τ_i at round r. These weights are chosen by the TA depending on whether it is considered more important to be high-performing at the current round or to have a consistent performance over time. They will be used in the computation of the aggregate score in step 7).
2) The TA receives ratings $\rho_{j,i}^r$ from each player $P_{j \neq i}$ to evaluate player P_i. In case the TA does not receive all the expected ratings, it broadcasts a complaint message and aborts the protocol. This is to make sure that all players participate at each round and that no player can benefit from not submitting a rating.
3) The TA computes the first component τ_i' for player P_i as follows:
   - The weight $w_{j,i}^r$ that player P_j has with respect to the evaluation of player P_i is $w_{j,i}^r = \frac{\tau_j^{r-1}}{\sum_{l \neq i} \tau_l^{r-1}}$, with $\sum_{j \neq i} w_{j,i}^r = 1$;
   - $\tau_i' = \sum_{j \neq i} w_{j,i}^r \cdot \rho_{j,i}^r$.

   It computes all first components $\tau_1', \ldots, \tau_n'$ before continuing. $w_{i,i}^r$ is implicitly undefined because player P_i cannot rate itself.
4) The TA computes a parameter ε depending on the accuracy threshold t_ε and the aggregate scores $\tau_1^{r-1}, \ldots, \tau_n^{r-1}$ of players $P_1, \ldots, P_n$ at round $r-1$ as follows. We denote by C_k the biggest *admissible coalition*, with cardinality K and members $P_1, \ldots, P_K$ (relabelling the players if necessary, without loss of generality). That is, C_k has the largest cardinality and the weights $w_{1,*}^r, \ldots, w_{K,*}^r$ of its members are not smaller than the weights of any coalition outsider, i.e. $w_{i,*}^r \geq w_{j,*}^r$, for $j \notin C_k$, $1 \leq i \leq K$. We assume an honest majority of players, i.e. $K < \frac{n}{2}$. We denote by m the index of a player $P_m \notin C_k$ being rated. Then:
$$\varepsilon = \max_{m \notin C_k} \left\{ \frac{t_\varepsilon}{\sum_{j \notin C_k, j \neq m} w_{j,m}^r} - t_\varepsilon \right\}. \tag{1}$$
   ε is referred to as the *weight threshold* because it limits the weight of a coalition against all the others. The index m that maximizes (1) corresponds to the player $P_m \notin C_k$ with the lowest aggregate score τ_m^{r-1}.
5) The TA selects the incentive α, with $\frac{1}{3} t_\varepsilon < \alpha \leq t_\varepsilon$ and the penalty β, with $-t_\varepsilon \leq \beta < -\frac{1}{3} t_\varepsilon$. Incentive α and penalty β are defined in this way to, respectively, raise and lower the aggregate score of a player as much as possible while remaining within the accuracy interval (Def. 4). This way, α and β shift the aggregate score, respectively, above and below the score corresponding to the performance while still providing a description of the performance itself.
6) The TA computes the second partial score $\tau_i^{''}$ for player P_i as $\tau_i^{''} = \frac{1}{n-1} \sum_{j=1}^{n-1} o_{i,j}$, where $o_{i,j}$ is a score defined as follows:
$$o_{i,j} = \begin{cases} \alpha, & \text{if } |\tau_j' - \rho_{i,j}^r| \leq t_\varepsilon \\ 0, & \text{if } t_\varepsilon < |\tau_j' - \rho_{i,j}^r| \leq 2t_\varepsilon \\ \beta & \text{if } |\tau_j' - \rho_{i,j}^r| > 2t_\varepsilon \end{cases} \tag{2}$$
   According to Def. 4, a t_ε-accurate rating $\rho_{i,j}^r$ varies around the targeted first partial score $\overline{\tau}_j'$, which is only a reference point. A t_ε-accurate first component τ_j' may shift the centre of the accuracy interval above or below $\overline{\tau}_j'$. A rating accurate with respect to the targeted first component $\overline{\tau}_j'$ has to be compared to the computed first component τ_j' now. Thus, rating $\rho_{i,j}^r$ might be at distance greater than t_ε from τ_j', while still being accurate. Hence we considered above three cases instead of two, where in the second case the rating is considered neutral. The performance scoring mechanism is run multiple times before the

data owner checks the aggregate scores. When this is iterated several times, as in a repeated game, the penalties and incentives based on a misleading accuracy interval are balanced out and mitigate the eventual misrepresentation of τ''.

7) The TA computes the aggregate score τ_i^r of player P_i at round r as $\tau_i^r = a \cdot \tau_i' + b \cdot \tau_i'' + c \cdot \tau_i^{r-1}$.

We present a strategy σ_i for player $P_i \in C_\ell$ to use when the performance scoring mechanism is run. This strategy provides criteria for player P_i to decide whether to submit an accurate rating $\rho_{i,j}^r$ for player $P_{j \neq i}$, and it holds for any formed coalition C_ℓ with cardinality n_ℓ.

- Player $P_i \in C_\ell$ submits to the TA an accurate rating $\rho_{i,j}^r$ for player P_j if one of the following two conditions holds:
  1) $P_j \in C_\ell$ and $\frac{\sum_{i \in C_\ell, i \neq j} w_{i,j}^r}{\sum_{l \notin C_\ell} w_{l,j}^r} \leq \varepsilon$
  2) $P_j \notin C_\ell$ and $\frac{\sum_{i \in C_\ell} w_{i,j}^r}{\sum_{l \notin C_\ell, l \neq j} w_{l,j}^r} \leq \varepsilon$
- Otherwise, it submits a selfish rating $\rho_{i,j}^r$ to the TA.

B. Resilience Against Coalitions

Let us denote by $\Gamma(P_i, \sigma_i, u_i')$, for $i = 1, \ldots, n$, the repeated game, where at every round the performance scoring mechanism described in Sec. V-A is run and where player P_i has utility $u'(r)$ (see Sec. III) and plays strategy σ_i defined in Sec. V-A, for $i = 1, \ldots, n$. We now prove that game $\Gamma(P_i, \sigma_i, u_i')$ can cope with any coalition of up to K members, where K is the cardinality of the largest admissible coalition in the sense of step 4) of the above performance scoring mechanism. That is, it computes accurate aggregate scores even if a coalition of K players deviates and submits selfish ratings. We assume that at most n coalitions are formed among players $P_1, \ldots, P_n$, and that each player belongs to at most one coalition. Let C_k be the biggest admissible coalition (see step 4 of the performance scoring mechanism in Sec. V-A). We denote by m the index of a player $P_m \notin C_k$ with the lowest aggregate score τ_m^{r-1} among the players not in C_k.

Theorem 2. *Let $\varepsilon > 0$ be a weight threshold and let C_k be the biggest admissible coalition for ε, with $|C_k| = K$. The infinitely repeated cooperative game $\Gamma(P_i, \sigma_i, u_i')$, for $i = 1, \ldots, n$, where u1')–u3') are satisfied and the mechanism in Sec. V-A is run at every round, reaches a K-resilient equilibrium for the computations of aggregate scores $\tau_1^r, \ldots, \tau_n^r$ if*

$$\frac{\sum_{i \in C_k} w_{i,m}^r}{\sum_{j \notin C_k, j \neq m} w_{j,m}^r} \leq \varepsilon.$$

Proof. The above condition on the weights can be seen as a requirement of balanced weights among coalitions. We show how this implies the accuracy of the first components (and thus aggregate scores reflecting actual performance), and the coalition-resistance of the performance scoring mechanism. The proof is composed of two steps.

a) Accuracy of the first components $\tau_1', \ldots, \tau_n'$: weight $w_{i,m}^r$ of player P_i when evaluating player P_m is by definition computed from the aggregate scores $\tau_1^{r-1}, \ldots, \tau_n^{r-1}$ of players $P_1, \ldots, P_n$ at round $r-1$. Thus, the condition on the weights implies that at round $r-1$ the aggregate scores $\tau_1^{r-1}, \ldots, \tau_n^{r-1}$ have been computed so that they preserve the equilibrium among the coalitions for the next round r where the aggregate scores are updated. Moreover, we assume that they are accurate, i.e. $\overline{\tau}_i^{r-1} - t_\varepsilon \leq \tau_i^{r-1} \leq \overline{\tau}_i^{r-1} + t_\varepsilon$, for $i = 1, \ldots, n$. Since K is the cardinality of the biggest admissible coalition C_k, this means that for every other coalition C_ℓ of cardinality $\ell \leq K$, we have

$$\varepsilon \geq \frac{\sum_{i \in C_\ell} w_{i,m}^r}{\sum_{j \notin C_\ell, j \neq m} w_{j,m}^r},$$

which is the condition of strategy σ_i for player $P_i \in C_\ell$ to submit a selfish rating $\rho_{i,j}^r$ for player $P_j \notin C_\ell$. Furthermore, denoting by m' the index of a player $P_{m'} \in C_\ell$ with respect to which the weights $w_{i,m'}^r, w_{j,m'}^r$, for $i \in C_\ell, j \notin C_\ell$ are computed,

$$\frac{\sum_{i \in C_\ell} w_{i,m}^r}{\sum_{j \notin C_\ell, j \neq m} w_{j,m}^r} = \frac{\sum_{i \in C_\ell} \frac{\tau_i^{r-1}}{\sum_{l \neq m} \tau_l^{r-1}}}{\sum_{j \notin C_\ell, j \neq m} \frac{\tau_j^{r-1}}{\sum_{l \neq m} \tau_l^{r-1}}}$$

$$\geq \frac{\sum_{i \in C_\ell} \tau_i^{r-1}}{\sum_{j \notin C_\ell} \tau_j^{r-1}} \geq \frac{\sum_{i \in C_\ell, i \neq m'} \tau_i^{r-1}}{\sum_{j \notin C_\ell} \tau_j^{r-1}} = \frac{\sum_{i \in C_\ell, i \neq m'} w_{i,m'}^r}{\sum_{j \notin C_\ell} w_{j,m'}^r}.$$

Thus,

$$\varepsilon \geq \frac{\sum_{i \in C_\ell, i \neq m'} w_{i,m'}^r}{\sum_{j \notin C_\ell} w_{j,m'}^r}, \tag{3}$$

which is the condition of strategy σ_i for player $P_i \in C_\ell$ to submit an accurate rating $\rho_{i,j}^r$ for player $P_j \in C_\ell$. Thus each player $P_i \in C_\ell$ following strategy σ_i always submits an accurate rating for player P_j, regardless if player P_j belongs to the same coalition as player P_i or not. The reason is that the assumption on the weights equilibrium implies that the conditions for player P_i to give high/low ratings for player P_j in case it is, respectively, in or out of coalition C_ℓ or give low ratings for player P_j are never verified. Since $\overline{\tau}_i' - t_\varepsilon \leq \rho_{j,i}^r \leq \overline{\tau}_i' + t_\varepsilon$ and the first partial score τ_i' is defined as a convex combination of ratings $\rho_{j,i}^r$, for $j \neq i$, it follows that $\overline{\tau}_i' - t_\varepsilon \leq \tau_i' \leq \overline{\tau}_i' + t_\varepsilon$, which means accuracy of the first partial score τ_i'. Thus, also the aggregate score τ_i^r is accurate because it is defined as the convex combination of accurate terms.

b) Resilience against coalition C_k: if each player $P_i \in C_\ell$ deviates from strategy σ_i while all other players do not, then it submits inaccurate ratings $\rho_{i,j}^r$ for player P_j. However, the first component τ_j' is still accurate, i.e. $\overline{\tau}_j' - t_\varepsilon \leq \tau_j' \leq \overline{\tau}_j' + t_\varepsilon$. This holds for every coalition C_ℓ because of the assumption

$$\sum_{i \in C_k} w_{i,m}^r \leq \varepsilon \Big(\sum_{j \notin C_k, j \neq m} w_{j,m}^r \Big).$$

By definition of ε,

$$\sum_{i\in C_k} w^r_{i,m} \le \left(\frac{t_\varepsilon}{\sum_{j\notin C_k, j\neq m} w^r_{j,m}} - t_\varepsilon\right) \sum_{j\notin C_k, j\neq m} w^r_{j,m}$$
$$= t_\varepsilon - \sum_{j\notin C_k, j\neq m} w^r_{j,m}\cdot t_\varepsilon.$$

If $\rho^r_{j,m}$ is an accurate rating, then $|\rho^r_{j,m} - \overline{\tau}'_m| \le t_\varepsilon$, for $j \notin C_k$ and $j \neq m$. On the contrary, if $\rho^r_{i,m}$ is a selfish rating, then $|\rho^r_{i,m} - \overline{\tau}'_j| > t_\varepsilon$ and $|\rho^r_{i,m} - \overline{\tau}'_j| \le 1$, for $j \in C_k$ and $j \neq m$. Thus,

$$\begin{aligned} t_\varepsilon &\ge \sum_{i\in C_k} w^r_{i,m} + \sum_{j\notin C_k, j\neq m} w^r_{j,m} t_\varepsilon \ge \sum_{i\in C_k} w^r_{i,m}|\rho^r_{i,m} - \overline{\tau}'_m| \\ &\quad + \sum_{j\notin C_k, j\neq m} w^r_{j,m}|\rho^r_{j,m} - \overline{\tau}'_m| \\ &\ge \left| \sum_{i\in C_k} w^r_{i,m}(\rho^r_{i,m} - \overline{\tau}'_m) + \sum_{j\notin C_k, j\neq m} w^r_{j,m}(\rho^r_{j,m} - \overline{\tau}'_m)\right| \\ &= \left| \sum_{i\in C_k} w^r_{i,m}\rho^r_{i,m} + \sum_{j\notin C_k, j\neq m} w^r_{j,m}\rho^r_{j,m} - \sum_{j=1, j\neq m}^{n} w^r_{j,m}\overline{\tau}'_m\right| \\ &= \left| \sum_{i\in C_k} w^r_{i,m}\rho^r_{i,m} + \sum_{j\notin C_k, j\neq m} w^r_{j,m}\rho^r_{j,m} - \overline{\tau}'_m\right| = |\tau'_m - \overline{\tau}'_m|. \end{aligned} \tag{4}$$

We have $|\tau'_m - \overline{\tau}'_m| \le t_\varepsilon$, hence $\overline{\tau}'_m - t_\varepsilon \le \tau'_m \le \overline{\tau}'_m + t_\varepsilon$, which is the definition of accuracy for the first component τ'_m. Furthermore, (3) implies that K-resilience holds even if the player to be evaluated belongs to the coalition, i.e. $P_m \in C_k$. □

VI. Related Work

Our approach bridges two research fields. For one, it is a reputation mechanism where aggregate scores are computed by peers. It also formalizes the interaction of rational players, which enable data protection through secret sharing. We survey both angles in turn.

a) Peer-to-Peer Reputation Mechanisms: We review mechanisms promoting accurate ratings in peer-to-peer systems, where accurate/inaccurate ratings are rewarded/penalized. In particular, we focus on mechanisms where agents interact with one or multiple external third parties, either to learn the reputation of their peers or to aggregate their ratings. These are referred to as *third-party-aided peer reputation mechanisms* and our mechanism falls into this category. Jurca et al. [12] consider a scenario involving peers as well as broker agents where incentives are issued through payment for reputation information. Afterwards, an agent decides whether to engage with peers. This model uses a game-theoretic approach, but addresses peer interaction and not how the reputation information of peers is aggregated and computed. The same authors also investigate how the lack of incentives leads to biased recommendations and ratings on online reviews platforms such as TripAdvisor [13]. The scoring system discussed by Miller et al. [16] is referred to as a peer-prediction model and is most suitable for rating commercial items in online platforms. It involves a centre common to all peers that processes the ratings forwarded by the peers themselves and has no independent information. The model uses a peer's rating to update a probability distribution for the rating of a different peer. Based on these ratings, the centre rewards or penalizes the targeted peer. The problem of coalitions is mentioned but not solved. Our work differs from the ones listed above in the following ways. Our performance scoring mechanism is the first to be provably resilient against peer coalitions. Moreover, it provides a relation between the admissible variations of the performance scores and the size of coalitions. This way, incentives and penalties are implemented before the performance scores are computed (and not later), leading to performance scores that match the desired accuracy level. Among the (more traditional) Bayesian approaches, it is worth mentioning Subjective Logic, the prominent trust mechanism proposed by Ismail and Jøsang in [11]. This mechanism uses the beta probability density function to estimate the future behavior of the parties given both direct interactions and indirect ratings. However, unlike our proposed mechanism, this approach is unable to filter out inaccurate ratings.

b) Protection Against Rational Parties in Secret Sharing: We review mechanisms for coping with rational players in secret sharing schemes. The notion of *rational secret sharing* was first introduced by Halpern and Teague [8]. Here, the utility of the players is tied to their goal of being the only ones to know the secret, while rationality captures their unwillingness to collaborate with other players to reconstruct the secret. The paradox of rational secret sharing is that the secret is lost because players will never actively collaborate by providing their own share during reconstruction. Halpern and Teague solve this problem by using game theory to incentivise the players to collaborate and they show how their solution reaches a Nash equilibrium if the players are not allowed to form coalitions. Later, Gordon and Katz proposed [7] a protocol also supporting two players only, which was not possible before. Abraham et al. introduced [1] the notion of a k-resilient Nash equilibrium and a secret sharing scheme reaching this equilibrium is designed. As in [1], we consider only players that act rationally according to their ultimate goal and arbitrarily behaving players with unknown utilities are not considered. Instead, this type of players are discussed by Lysyanskaya and Triandopoulos [15] and by Asharov and Lindell [2]. All so far mentioned protocols for rational secret sharing assume that communication happens simultaneously, either through a broadcast channel or through secure private channels. Kol and Naor proposed [14] a rational secret sharing scheme with a non-simultaneous broadcast channel that is also coalition-resistant. In all preceding cases, rational secret sharing protocols are treated as infinitely repeated games,

as in our approach. This allows a reward mechanism that forces the players to not behave selfishly and reach a social optimum. Nojoumian and Stinson proposed [17] a model where players are associated with a score recording the number of times they provided their shares during the reconstruction of the secret. A high score means that players might be included in future secret sharing schemes. Our model differs from the approaches discussed in this paragraph because we consider rationality in terms of economic return for SSPs, which in our scenario boils down to maximising share storage. We do not focus on the performance of the players during the reconstruction of the secret shared data. Instead, we focus on the general performance of the players, with measurements uncoupled from share renewal operations. Our performance scoring mechanism helps the data owner to select high performing storage servers in the first place. All approaches discussed above tackle the threats of private secret acquisition and uncooperative behaviour during secret reconstruction. In contrast, uniquely, we counter the threat of inaccurate peer rating to undermine storage competitors.

VII. Conclusions

The long-term confidential storage of sensitive data can be realised through proactive secret sharing, performed in a distributed storage system consisting of multiple SSPs. Data owners can be guided in the selection of high-performing SSPs through performance scoring mechanisms based on mutual peer ratings.

In this paper, we addressed the problem of modelling performance scoring mechanisms such that they are resilient against coalitions of rational SSPs. We first modelled SSPs as rational agents aiming at maximising their shares storage and formalized their rating strategies. Second, we showed that performance scoring mechanisms output aggregate scores that do not reflect actual performance if the players are not incentivised to submit accurate ratings. Third, we introduced a novel performance scoring mechanism modelled as an infinitely repeated and cooperative game with a TA as a mediator. The TA incentivises accurate ratings and detects and penalizes inaccurate ratings with respect to a margin that depends on coalition sizes. Using our game-theoretic formalism, we proved that such a performance scoring mechanism outputs accurate aggregate scores. It thus provides viable guidance for the selection of high-performing SSPs in distributed storage systems.

We plan to experimentally validate practical instantiations of our mechanism, in order to estimate the average number of rounds required for the aggregate scores to converge to accurate values.

Acknowledgments

This work was partially funded by the European Commission through grant agreement no 644962 (PRISMACLOUD). It also received funding from the DFG as part of project S6 within the CRC 1119 CROSSING.

References

[1] Ittai Abraham, Danny Dolev, Rica Gonen, and Joseph Y. Halpern. Distributed Computing Meets Game Theory: Robust Mechanisms for Rational Secret Sharing and Multiparty Computation. In *PODC*, pages 53–62. ACM, 2006.

[2] Gilad Asharov and Yehuda Lindell. Utility Dependence in Correct and Fair Rational Secret Sharing. In *CRYPTO*, volume 5677 of *LNCS*, pages 559–576. Springer, 2009.

[3] Johannes Braun, Johannes Buchmann, Ciaran Mullan, and Alex Wiesmaier. Long Term Confidentiality: a Survey. *Designs, Codes and Cryptography*, 71(3):459–478, 2014.

[4] Johannes Braun, Johannes A. Buchmann, Denise Demirel, Matthias Geihs, Mikio Fujiwara, Shiho Moriai, Masahide Sasaki, and Atsushi Waseda. LINCOS: A Storage System Providing Long-Term Integrity, Authenticity, and Confidentiality. In *AsiaCCS*, pages 461–468. ACM, 2017.

[5] Ernest F. Brickell. Some Ideal Secret Sharing Schemes. In *EUROCRYPT*, volume 434 of *LNCS*, pages 468–475. Springer, 1989.

[6] Johannes Buchmann, Alexander May, and Ulrich Vollmer. Perspectives for Cryptographic Long-Term Security. *Communications of the ACM*, 49(9):50–55, 2006.

[7] S. Dov Gordon and Jonathan Katz. Rational Secret Sharing, Revisited. In *SCN*, volume 4116 of *LNCS*, pages 229–241. Springer, 2006.

[8] Joseph Y. Halpern and Vanessa Teague. Rational Secret Sharing and Multiparty Computation: Extended Abstract. In *STOC*, pages 623–632. ACM, 2004.

[9] Amir Herzberg, Stanislaw Jarecki, Hugo Krawczyk, and Moti Yung. Proactive Secret Sharing Or: How to Cope With Perpetual Leakage. In *CRYPTO*, volume 963 of *LNCS*, pages 339–352. Springer, 1995.

[10] Michael Hogan, Fang Liu, Annie Sokol, and Jin Tong. NIST Cloud Computing Standards Roadmap (NIST-SP 500-291). *NIST Special Publication*, 35, 2011.

[11] Audun Jøsang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, 2007.

[12] Radu Jurca and Boi Faltings. An Incentive Compatible Reputation Mechanism. In *CEC*, pages 285–292. IEEE, 2003.

[13] Radu Jurca, Florent Garcin, Arjun Talwar, and Boi Faltings. Reporting Incentives and Biases in Online Review Forums. *TWEB*, 4(2):5, 2010.

[14] Gillat Kol and Moni Naor. Cryptography and Game Theory: Designing Protocols for Exchanging Information. In *TCC*, volume 4948 of *LNCS*, pages 320–339. Springer, 2008.

[15] Anna Lysyanskaya and Nikos Triandopoulos. Rationality and Adversarial Behavior in Multi-party Computation. In *CRYPTO*, volume 4117 of *LNCS*, pages 180–197. Springer, 2006.

[16] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting Informative Feedback: The Peer-Prediction Method. *Management Science*, 51(9):1359–1373, 2005.

[17] Mehrdad Nojoumian and Douglas R. Stinson. Socio-Rational Secret Sharing as a New Direction in Rational Cryptography. In *GameSec*, volume 7638 of *LNCS*, pages 18–37. Springer, 2012.

[18] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[19] Adi Shamir. How to Share a Secret. *Commun. ACM*, 22(11):612–613, 1979.

[20] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comp.*, 26(5):1484–1509, 1997.

[21] Tamir Tassa. Hierarchical Threshold Secret Sharing. *J. Cryptology*, 20(2):237–264, 2007.

[22] T. Watson et al. A programmable two-qubit quantum processor in silicon. *Nature*, 2018.