ABSTRACT
With the upsurge in the amount of data collected today: collaborative data mining becomes increasingly useful. However this does require the preservation of individual privacy of the participants. Secret Sharing has been proposed for secure multiparty computation (SMC) for different scenarios for Privacy Preserving Distributed Association Rule Mining (PPDARM) and Distributed Clustering for semi-honest parties. The issue is that in real life parties tend to be selfish and rational. Hence we propose a new realistic approach with novel punishment policies for the secret sharing scheme among rational parties for SMC in distributed data mining. This approach is based on the game theoretic framework. We propose two novel punishment policies for this approach and also analyse and evaluate our scheme.

Categories and Subject Descriptors

General Terms
Security

Keywords
Rational Secret Sharing: Privacy: Distributed Data Mining: Game Theory

1. MOTIVATION
The rate at which data is collected today has increased to a large extent which in turn increases the rate to decipher useful information from this data. This abundance of knowledge extracted through various data mining methods carries an inherent risk of loss of privacy. Catering to these vulnerabilities; a new genus of privacy preserving approaches have been proposed in Distributed Data Mining assuming the parties to be semi-honest. Secret Sharing for the PPDARM semi honest model has been proposed recently by [1][2]. It has also been proposed for clustering. However in a realistic formulation of Privacy preserving distributed data mining (PPDDM) these parties are rational in nature and may prefer to just attain their own gain and if possible not contribute to the system.

Hence as an extension to the Secret Sharing scheme applied to PPDDM; we have proposed a novel secret sharing approach in PPDDM to find the sum privately among agents that are neither good nor entirely malicious but rational and work in their own interest. Hence according to [3][4] these rational agents will have a tendency to not broadcast their shares as each of them would first prefer getting the secret and secondly prefer that the fewer of the other agents that get it, the better.

There have been a few approaches proposed for rational secret sharing using the game theory, that explain rational secret sharing with mediators [3][4] without cheap talk. Game theory to achieve Nash equilibrium in a PPDDM model was introduced by [5] for the Secure Sum protocol in a ring topology. However a game theoretic approach for the Secret Sharing scheme that uses a mesh topology in PPDDM has not yet been proposed.

[m,m] secret sharing has been used in PPDMM by [1][2] to decipher the sum privately using a mesh topology for a restricted semi honest model. None of them consider the realistic scenario where the participants involved are neither entirely good nor entirely bad (which means they don’t want to sabotage the protocol by sending wrong values) but are rational. Hence our proposed algorithm is a novel algorithm that provides an alternative perspective as it models Secret Sharing in PPDDM as a repeated game. It uses cheap talk without using mediators modeled for rational parties. We are currently proposing this model along with two effective punishment policies for Distributed Association Rule Mining as a case study.

The primary advantage of this model is that it avoids the use of expensive techniques like homomorphic encryption and zero knowledge proofs. In the secret sharing scheme for PPDDM applied in [1][2]; the Nash equilibrium would be where the party refrains from sending its sum of shares. We model Secret Sharing in PPDMM as a game where the Nash equilibrium is shifted to when all parties send their shares and attain a non-collusive behavior. We propose novel punishment strategies which aim at giving all the parties maximum utilities in a state where they cooperate and penalize them when they do not. Our punishment policies aim at getting the maximum possible participants involved in the game so that they can attain maximum utilities.

2. PROBLEM FORMULATION, METHODOLOGY AND ANALYSIS
The problem that we resolve in this paper considers a co-operative scenario of vertically or horizontally partitioned databases where PPDARM is required where there are ‘p’ rational parties collaborating to find the secure global sum in their data privately. This problem extends the [m,m] threshold scheme proposed for rational agents. We model secret sharing as a repeated game among rational agents denoted by Γ(m,m) where the same set of players come to play the same game repeatedly. without mediators. The players are connected together in a mesh topology and follow a simultaneous broadcast of messages. We assume that the players are generally concerned about their future utility.

Among the 3 main attacks possible in [m,m] secret sharing, which involve: sending of wrong shares by malicious agents, not sending
their shares at all and colluding with other selective parties; we resolve the last two attacks by an adversary who is not disrupting the protocol by sending wrong shares but is trying to selfishly get his own gain. We resolve these latter attacks using punishment strategies which acts as a motivation for the players to collaborate.

The previously proposed punishment strategies [5] for secure sum protocol does not consider repeated games and permanently removes the party or increases the computation and communication cost for all parties which is a gross penalty to all other good parties. Similar to this is the grim trigger strategy in [4] where the game is halted until all parties send their shares. Hence we improve on the punishment strategies considering [m, m] Secret Sharing so that all parties will have the maximum utility and attain the Nash equilibrium state.

**Policy 1:** Incremental Punishment Strategy for repeated rounds up to c times: If a party defaults; remove him from the game for 'y' rounds. He is again given a chance to enter the setup after y rounds and if he defaults again; he is removed for an incremental 'g' rounds. This defaulting behavior is allowed for 'c' attempts.

**Policy 2:** Punishment strategy allowing 'c' attempts: Give the party c maximum chances to enter the setup and perform honestly after which he is permanently ousted.

The rational parties may default in the game of secret sharing by withholding their share or sum of shares and sending it just to the coalition. Let us assume uₙ(S_i), uₚ(R_i) be the utility or payoff where the iᵣ (0 < i ≤ p) party sends and receives corresponding iᵗʰ share or sum of share. Now since multiparty PPDDM is a strategic game as mentioned in [5]; we consider the strategy as follows where the utilities of each party are measured considering the performance of the other parties:

\[ u_i(\sigma_i, \sigma_{-i}) = w_i u_i(S_i) + w_0 u_i(R_i) \]  --Equation 1

This is the formalization of the game in our scenario where we consider secret sharing in a multiparty PPDDM game. We base our game on the goal of getting maximum shares and not on the communication cost. Also without any penalty; the gain would be maximum when no sum of shares is sent for a party. Hence we enforce the punishment strategies mentioned before so that the Nash equilibrium is attained when the parties behave honestly. Hence cooperating becomes the optimal strategy for everybody.

Now consider there are a set of N items which denote the frequent itemsets or association rules at each of the rational partitions (where 0 < i ≤ p). Now we consider D as the no. of defaulters and C as the no. of colluding nodes who send the shares or sum of shares only in the colluding group. As far as the collusion analysis is concerned; if C < D; then the defaulting parties will not get the result back and so that scenario has no meaning. However if D = 1; then the party can get the sum of values of the other parties. Finally the general case when C≥D; then the colluding parties can definitely get the value of the sum polynomial Sum(xᵣ) and hence can predict the value of the non-collusive party and have a high utility as they are not sending any shares. However this can be avoided if we follow our punishment strategy so that fewer parties will collude so as to attain the maximum gain.

3. RESULTS AND IMPLICATIONS

We show a comparison between the no. of rounds of secret sharing and the percentage of bad nodes for our Policy 2. We have evaluated our results with a total of 10 nodes. We consider that 50% of the bad nodes after being penalized prefer to join the game in the coming round of which 50% would behave honestly. Also we consider the uₙ(S_i) and uₚ(R_i) as the number of shares sent and received where the uₙ(S_i) is negative as none of the rational parties want to send their shares willingly and the w₀ is twice that of sending wᵢₙ, according to the preference of the rational agents.

Our experimental results show that irrespective of the initial percentage of bad nodes; with the help of the punishment policy all the bad nodes converge and we get a Nash equilibrium at a state where all the nodes are honest. If the bad nodes do not change their behavior; they are ousted from the game permanently and still we get all honest nodes eventually. This graph would show changes if the nodes behaved in a different manner than the assumptions but it would still converge to zero bad nodes.

Hence we conclude that our protocol works in the interest of all the rational parties to attain a state where all parties are honest eventually as that is the state where each party has the maximum utility considering that Secret Sharing is a repeated game and that they have to take their future utilities in account.

4. REFERENCES


