Brief Announcement: Space Adaptation: Privacy-preserving Multiparty Collaborative Mining with Geometric Perturbation

Keke Chen, Ling Liu
College of Computing, Georgia Institute of Technology, Atlanta, GA, USA
kekechen@cc.gatech.edu, lingliu@cc.gatech.edu

Categories and Subject Descriptors: K.4.1 [Privacy]
General Terms: Algorithms
Keywords: Privacy-Preserving Collaborative Mining

1. INTRODUCTION

The service-oriented infrastructure has become popular for collaboratively mining data distributed over organizations [3], where the participants are the data providers who submit their perturbed datasets to the designated data mining service provider (the data miner) for mining commonly interested models. Figure 1 shows the service-oriented framework for collaborative multiparty data mining. Two kinds of parties are directly involved in the computing. The mining service provider (SP) is a party independent of the data providers, which owns abundant computing power, data mining tools and talents. SP offers their data mining services to the contracted parties through certain service provision scheme. We

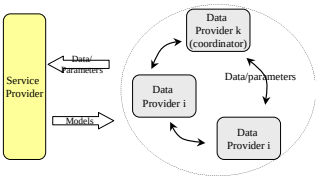

Figure 1: Service-oriented multiparty privacy preserving data mining.

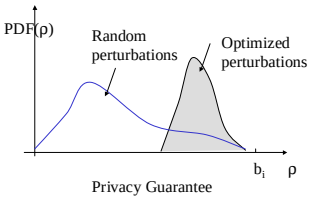

Figure 2: Optimized perturbation gives higher privacy guarantee on average.

assume a semi-honest model for all parties. Therefore, we do not consider the scenarios where either the data miner or some data providers are malicious and can collude with one another. We also assume that encryption is applied before data is transmitted on the network. In this paper, we will study the problem of privacy preserving multiparty collaborative data classification using *geometric data perturbation*.

Geometric data perturbation has unique benefits for privacy preserving data classification [1, 2]. First, many popular classifiers, such as linear classifiers and Support Vector Machine (SVM), are invariant to geometric transformation in the sense that the classifiers trained with the perturbed data through geometric rotations have almost the same accuracy as those trained with the original raw data. Second, geometric data perturbation can easily produce multiple random transformations, each of which preserves classification model accuracy for the discussed classifiers. Thus, an individual data provider needs only to select one perturbation that can provide satisfactory privacy guarantee. A randomized perturbation optimization algorithm is also developed in previous work [2] to provide high privacy guarantee with high probability (Figure 2). Comparing with other existing approaches to privacy preserving classification, geometric data perturbation significantly reduces the complexity in balancing data utility and data privacy guarantee.

The key challenge for applying geometric data perturbation to multiparty collaborative data classification is to unify the perturbations used by different data providers without sacrificing much data privacy and data utility. In this paper, We develop the Space Adaptation Protocol (SAP) for securely unifying the perturbations. SAP enables parties to anonymously submit the perturbed data and minimizes the risk of privacy breach with low cost.

2. GEOMETRIC DATA PERTURBATION

We define a geometric perturbation as a combination of random rotation perturbation, random translation perturbation, and noise addition. It can be represented as $G(X) = RX + \Psi + \Delta$. X denotes the *normalized* original dataset with N rows and d columns, R is a $d \times d$ random orthogonal matrix, and Ψ is a $d \times N$ random translation matrix [2]. $\Psi = \mathbf{t} \times \mathbf{1}'$ and $\mathbf{t}$ is randomly generated using the uniform distribution over [-1, 1]. Δ is a noise matrix with i.i.d. (independent identically distributed) elements, which is used to perturb distances.

In papers [1, 2], we defined two privacy metrics for multi-column privacy evaluation. In this paper we by default use the "Minimum Privacy Guarantee" to represent the privacy guarantee of DP_i, denoted by ρ_i. ρ_i is greater than or equal to zero and is bounded by some value, say b_i, which may be different for different datasets.

With the optimization algorithm [2], we can get a higher local privacy guarantee ρ_i compared to the randomly generated perturbation as Figure 2 illustrates. Let the mean of optimized privacy guarantee be $\bar{\rho}_i$. We use the *optimality rate* $\mathcal{O}$ to represent the efficiency of optimization for the particular dataset: $\mathcal{O}_i = \frac{\bar{\rho}_i}{b_i}$. The bound b_i is usually estimated empirically by looking at the maximum privacy guarantee of n-round optimizations, i.e., $\hat{b} = max\{\rho^{(i)}, 1 \leq i \leq n\}$.

The evaluation of multiparty privacy guarantee is thus defined by two aspects: the risk of data source being identified (source identifiability) and the reduction of local privacy guarantee by using unified perturbation. We define the source identifiability π_i as the probability that the received data is indeed from the data provider DP_i, denoted by $\pi_i = Pr(DP_i|X_i)$. Next, let the locally optimized perturbation give the privacy guarantee of ρ_i for DP_i, and the global perturbation G gives the privacy guarantee ρ_i^G. We define the *satisfaction level* for the unified perturbation G by the data provider DP_i as $s_i = \frac{\rho_i^G}{\rho_i}$. Let b_i be the upper bound of the minimum privacy guarantee of data provider DP_i. We define the *Risk*

PODC'07, August 12–15, 2007, Portland, Oregon, USA.
ACM 978-1-59593-616-5/07/0008.

of Privacy Breach for DP_i, denoted by $\mathcal{R}_i^G$, as follows:

$$\mathcal{R}_i^G = \pi_i \cdot \frac{(b_i - s_i\rho_i)}{b_i} = \pi_i(1 - s_i \cdot \frac{\rho_i}{b_i}) \quad (1)$$

3. SPACE ADAPTATION PROTOCOL

Space adaptation protocol utilizes the concept of space adaptation to support de-identification of data sources. The basic idea of reducing the identifiability of data sources is to make use of secure random exchange of perturbed datasets between data providers with the help of *space adaptation*.

Let the perturbation parameters for the data provider DP_i are $G_i : (R_i, \mathbf{t}_i)$, with a common noise component Δ used by all parties. Let the original sub-dataset be X_i and Y_i be the perturbed data. Now, suppose that we want to transform Y_i to $Y_{i\to t}$ in the target space $G_t : (R_t, \mathbf{t}_t)$ that has *no noise component*. The following procedure is applied. Since $Y_i = G_i(X_i) = R_iX_i + \Psi_i + \Delta_i$, and thus $X_i = R_i^{-1}(Y_i - \Psi_i - \Delta_i)$, we can easily prove the following equation hold:

$$Y_{i\to t} = R_tR_i^{-1}Y_i + (\Psi_t - R_tR_i^{-1}\Psi_i) - R_tR_i^{-1}\Delta_i$$

This equation consists of three components. We define the first component $R_tR_i^{-1}$ as *the rotation adaptor* R_{it}. $R_tR_i^{-1}\Psi_i$ is still a translation matrix, and thus we name the second part $\Psi_t - R_tR_i^{-1}\Psi_i$ as the *the translation adaptor* Ψ_{it}. The third part involves the original noise component and we name $\Delta_{it} = R_tR_i^{-1}\Delta_i$ as *the complementary noise*. It is easy to prove that removing the complementary noise component in the target space G_t is equivalent to inheriting the noise component Δ_i from the original space G_i. Therefore, we use $< R_{it}, \Psi_{it} >$ as the space adaptor.

The protocol starts with the random selection of target perturbation, say, $G_t : (R_t, \mathbf{t}_t)$. Let each data provider, DP_i, also have a locally optimized perturbation $G_i : (R_i, \mathbf{t}_i)$. DP_i provides only the locally optimally perturbed dataset $G_i(X_i)$ to other parties.

Next, the coordinator, without loss of generality, DP_k, generates a sequence, which is a random permutation of the k data providers: $(1, \ldots, k) \leftarrow (\tau(1), \ldots, \tau(k))$. Let DP_i receive data from $DP_{\tau(i)}$. However, we do not allow the coordinator to receive any dataset, since the coordinator will also receive parameters later, which will help to recover any received perturbation. Therefore, we randomly redirect $\tau(k)$ to any j, $j \in 1 \ldots k-1$, instead. Finally, the mapping becomes $(1, \ldots, k-1, j) \leftarrow (\tau(1), \ldots, \tau(k))$ $j \neq k$. Now, each dataset $G_i(X_i)$ has a probability of $\frac{1}{k-1}$ going to any of the $k-1$ data providers. After random exchange, each data provider sends the received dataset to the data miner. By doing this, the identifiability of data source in the service provider's view is reduced to $\pi_i = \frac{1}{k-1}$.

Finally, each data provider DP_i sends the space adaptor $A_{it} =< R_{it}, \Psi_{it} >$ to the coordinator. The coordinator maps the adaptors to the right target by the permutation sequence, $(DP_1 : A_{\tau(1),t}, \ldots, DP_j : A_{\tau(j),t}, DP_j : A_{\tau(k),t}, \ldots, DP_{k-1} : A_{\tau(k-1),t})$, and sends this sequence of space adaptors to the data miner.

Let the unified perturbation G_t gives satisfaction level s_i to DP_i. Therefore, the overall risk of privacy breach for DP_i, from the view of both data providers and the data miner, is:

$$\mathcal{R}_i^{SAP} = \max\{\frac{(b_i - \rho_i)}{b_i}, \frac{(b_i - s_i\rho_i)}{b_i} \times \frac{1}{k-1}\} \quad (2)$$

4. EXPERIMENTAL RESULT

Twelve UCI machine learning datasets are used in the experiments. Each dataset is split into several randomly sized sub-datasets, simulating the distributed datasets from the data providers.

Characteristics of SAP First of all, we study the relationship between the factors: k, the number of parties, b_i, the upper bound of possible privacy guarantee for DP_i, ρ_i, the locally optimized privacy guarantee, and s_i^0: the satisfaction level that each party expects. The rate $\frac{\rho_i}{b_i}$ is approximated by the optimality rate $\frac{\bar{\rho}_i}{\bar{b}_i}$, which can be estimated on sample local optimization results. In experiments, we use the algorithm and attack models discussed in the paper [2] to study the factor $\frac{\bar{\rho}_i}{\bar{b}_i}$. Figure 3 shows the estimated rates for the three typical datasets in 100 rounds. We choose the maximum value in the series of experiments for each dataset as the upper bound. Figure 4 shows the relationship between k, s_i^0 and the optimality rate.

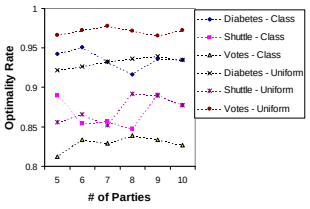

Figure 3: Some sample optimality rates $\max\{\rho_i/b_i\}$

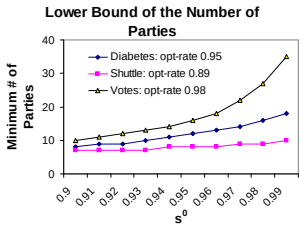

Figure 4: # of parties, optimality rate, and satisfaction level.

Effect on Model Accuracy We finalize the experiments with the study of model accuracy for two representative classifiers: KNN classifier and SVM classifier with RBF kernel. Certainly, geometric perturbation can be applied to much more classifiers as discussed in previous work [1]. We also study the effect of "partition distribution" to the model accuracy. A local dataset is treated as a sample of the pooled dataset. Uniform partition distribution means the local datasets are almost uniform sample sets of the pooled dataset, while any skewed partition distribution does not have such property. The numbers in Figure 5 and 6 show the deviation from the standard accuracy which is generated on the original unperturbed dataset. A negative number means that the actual accuracy is reduced.

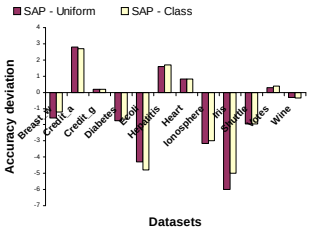

Figure 5: The average deviation of model accuracy for KNN classifier.

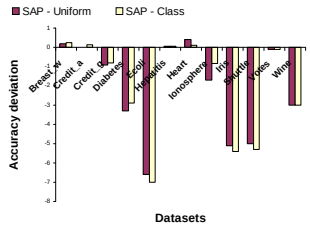

Figure 6: The average deviation of model accuracy for SVM(RBF) classifier.

5. CONCLUSION

In this paper, we proposed the space adaptation protocol to securely unify multiple geometric perturbations, analyzed the features of the protocol, and studied the relationship between the main factors and tradeoffs theoretically and empirically.

6. REFERENCES

[1] CHEN, K., AND LIU, L. A random rotation perturbation approach to privacy preserving data classification. *Proc. of Intl. Conf. on Data Mining (ICDM)* (2005).

[2] CHEN, K., AND LIU, L. Towards attack-resilient geometric data perturbation. *SIAM Data Mining Conference* (2007).

[3] ZHANG, N., WANG, S., AND ZHAO, W. A new scheme on privacy-preserving data classification. *Proc. of ACM SIGKDD Conference* (2005).