Mitigating the ICA Attack against Rotation-Based Transformation for Privacy Preserving Clustering

Abedelaziz Mohaisen and Dowon Hong

ABSTRACT—The rotation-based transformation (RBT) for privacy preserving data mining is vulnerable to the independent component analysis (ICA) attack. This paper introduces a modified multiple-rotation-based transformation technique for special mining applications, mitigating the ICA attack while maintaining the advantages of the RBT.

Keywords—RBT, ICA, Multiple rotations, data clustering.

I. Introduction

While it is important for data owners to publish their data to a third party to provide data mining services, the privacy of the data itself needs to be maintained. Therefore, several perturbation methods have been introduced considering potential applications. One of these methods is rotation-based transformation (RBT) in which the data is transformed geometrically while the distance between the data points is preserved. Such distance preservation is vital for providing high accuracy which incurs minimal data loss when data clustering is performed [1].

Perturbation schemes, and the RBT scheme in particular, have been studied in relation to different attack methods, including the naïve estimation-based attack, the reconstruction-based attack, and the distance inference-based attack [2]. While the first attack method is infeasible for most perturbation methods, and the second is not applicable to RBT, the distance inference-based attack has some impact. This impact was studied recently, and two statistical tools were used, namely, principle component analysis (PCA) [3] and independent component analysis (ICA) [4]. Both tools have shown efficiency in breaching privacy under some operating conditions.

In this letter, we revise the ICA attack on RBT and introduce a multiple RBT (MRBT) method that helps mitigate the ICA attack.

Technically, RBT is a method that was introduced for data perturbation to guarantee exact accuracy and maintain the privacy of geometrical and numerical data [2], [5]. The general transformation follows the model $\mathbf{Y}=\mathbf{R_\theta X}$, where $\mathbf{X}$ is the original data, $\mathbf{R_\theta}$ is a rotation matrix, and $\mathbf{Y}$ is the rotated (transformed) data to be released to a third party. The rotation matrix $\mathbf{R_\theta}$ needs to be orthogonal to satisfy the distance-invariant property. That is, $\mathbf{R_\theta R_\theta}^T=\mathbf{R_\theta}^T\mathbf{R_\theta}=\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The RBT preserves the vector length, Euclidean distance, and inner product between two vectors which are essential in many clustering algorithms.

On the other hand, the ICA is a method for independent component separation that aims to separate two or more signals with some specific properties following the form $\mathbf{Y}=\mathbf{AS}$, where $\mathbf{S}$ is the set of independent components (data), and $\mathbf{A}$ is a mixing matrix, corresponding to $\mathbf{X}$ and $\mathbf{R}$ in RBT, respectively. The ICA is subject to several restrictions on the data used, including that the data attributes need to be linearly independent and have a non-Gaussian distribution (except for one attribute at most).

The *a-priori knowledge ICA* (AK-ICA) is a simple modification of the ICA and has recently been shown to be effective on RBT [4]. Given a small portion of the original data, its transformed image, and the transformed image of all of the data, it is then possible to recover all of the data (population) using the ICA, given enough information about the distribution of the original private data. That is, the attacker first applies the

Manuscript received May 19, 2008; revised July 10, 2008; accepted Oct. 8, 2008.

This work was supported by the IT R&D program of MKE/IITA, Rep. of Korea [2005-Y-001-04, Development of next generation security technology].

Abedelaziz Mohaisen (phone: + 82 42 860 6658, email: a.mohaisen@etri.re.kr) and Dowon Hong (email: dwhong@etri.re.kr) are with the Software & Content Research Laboratory, ETRI, Daejeon, Rep. of Korea.

ICA on the whole transformed dataset then applies it again on the known private data portion. From the separated components, the attacker finds some matrix ***J*** that maximizes the mutual information between the separated components by computing

$$I(f_i, f_j^{'}) = \tfrac{1}{2} E[\int_{\Omega z} | f_i(z) - f_j^{'}(z) | dz]$$

at a point z, where f_i and $f_j^{'}$ are the density distributions of the i-th component of the original dataset and the j-th component of the reconstructed dataset, respectively. The smaller I is, the more similar the estimated data and the original data are.

The ICA attack is powerful because the known fraction of data has sufficient information about the distribution of the whole data population. In practice, it is difficult to mount such an attack by only observing a small portion of the data.

The proposed MRBT scheme to mitigate the impact of AK-ICA is based on dividing the data into sets and transforming them independently in order to harden the process of recovering the ICA components and applying the AK-ICA. Experimentally, our scheme shows reasonable efficiency in mitigating AK-ICA. We also show that AK-ICA attacks on geometric data (see [7]) produce less accuracy than the dataset in [4].

II. Multiple RBT and Applications

The main goal of our rotation scheme is to preserve the distance and the inner product between data sets *partially*, as a valid goal for different applications, while providing a high level of privacy for the transformed data. The scheme is summarized in this section, and the data is assumed to be numerical.

Step 1. The data owner normalizes the data to unity.

Step 2. According to some n chosen in advance, the data owner divides the data into n equal parts defined as

$$X^{'} = \{X_1^{'} \| X_2^{'} \| X_3^{'} \| \cdots \| X_n^{'}\} .$$

Step 3. The data owner generates n different random seeds. Using each seed i, the data owner generates an orthogonal matrix $\mathbf{R}_{\theta i}$ for rotating the corresponding part of **X**.

Step 4. The data owner transforms his data as follows:

$$Y^{'} = \{Y_1^{'} \| \cdots \| Y_n^{'}\} = \{R_{\theta 1} X_1^{'} \| \cdots \| R_{\theta n} X_n^{'}\} .$$

Step 5. The data owner releases the rotated data for public use.

The resulting rotation preserves the inner product between the corresponding tuples in the original data. Also, it preserves the inner product between two tuples falling into the same corresponding subsets. However, the inner product is not preserved for tuples other than those mentioned here. The first claim can be easily proven given that these tuples are transformed using the same matrix (follows from [1]). Similarly, we prove the second claim as follows. Consider the following rotated data sets:

$$Y^{'} = \{Y_1^{'} \| \cdots \| Y_n^{'}\} = \{R_{\theta 1} X_1^{'} \| \cdots \| R_{\theta n} X_n^{'}\} ,$$

$$Y^{''} = \{Y_1^{''} \| \cdots \| Y_n^{''}\} = \{R_{\theta 1} X_1^{''} \| \cdots \| R_{\theta n} X_n^{''}\} .$$

The inner product between these two datasets is

$$Y^{'T} Y^{''} = \begin{pmatrix} Y_1^{'T} Y_1^{''} & Y_1^{'T} Y_2^{''} & \dots & Y_1^{'T} Y_n^{''} \\ Y_2^{'T} Y_1^{''} & Y_2^{'T} Y_2^{''} & \cdots & Y_2^{'T} Y_n^{''} \\ \vdots & \vdots & \ddots & \vdots \\ Y_n^{'T} Y_1^{''} & Y_n^{'T} Y_2^{''} & \cdots & Y_n^{'T} Y_n^{''} \end{pmatrix} .$$

For the diagonal part of the above product matrix, it is easy to verify the preservation of the inner product:

$$\begin{aligned} Y_i^{'T} Y_i^{''} &= (R_{\theta i} X_i^{'})^T R_{\theta i} X_i^{''} = X_i^{'T} R_{\theta i}^T R_{\theta i} X_i^{''} \\ &= X_i^{'T} I X_i^{''} = X_i^{'T} X_i^{''}; \end{aligned}$$

therefore, the second claim is proven by this result. Also, the distance between the corresponding subsets can be easily driven by

$$d(X^{''}, X^{'}) = \sqrt{\sum\nolimits_i (x_i^{''} - x_i^{'})} = \sqrt{2 - 2X^{''T} X^{'}} .$$

The quantification of privacy in our scheme follows the same metrics as those given in [4], and we use the maximization of the differences' covariance between the original data and the rotated data [1]. The latter metric can be systemically guaranteed by setting the different rotation matrices that maximize the difference covariance per data subset.

In our MRBT scheme, the geometric shape of the data is distorted into n different parts where a simple clustering algorithm will not work correctly. However, our scheme can be applied in several promising applications based on the coordinated pair-wise distance.

Application 1: Computing the inner product of two vectors transformed using MRBT. This is a straightforward application of the proposed scheme on data subsets. After releasing the transformed data, the third party computes the inner product on the corresponding subsets.

Application 2: Computing the distance between the corresponding tuples of two private data vectors. First, the two parties perform the routine in application 1 to obtain the inner product. Then, the third party plugs the resulting inner product into the distance/inner product formula to compute the distance.

Application 3: Network diagnosis. A data owner has real data representing a site access log from which he would like to diagnose his own site by clustering different values (processing, bandwidth, and so on). In doing so, the data owner wants to

ETRI Journal, Volume 30, Number 6, December 2008 Abedelaziz Mohaisen et al. 869

measure the anomaly log that deviates from the steady state log file. The steady state log data is measured ahead of time. This application is directly transferred into a clustering problem over the resulting distance in accordance with the time as an index. To do so, the following procedure is performed.

Step 1. The two sites generate a rotated version of their own log data using the method in application 1 and release the rotated data to the third party.

Step 2. The third party computes the access log difference for the corresponding days, resulting in the matrix of differences.

Step 3. The third party performs the clustering algorithm (*k*-mean) on the resulting set of distances.

Note that the matrix of differences is preserved regardless of the method of rotation. Although the single day's access log data is rotated using different angles, the computed difference is maintained for the single day.

In addition to these applications, it is easy to recall and extend other applications from the literature such as those in [7] and [8] in which the privacy of clustering data or location information can be of great benefit.

III. Impact of AK-ICA on the MRBT

To study the impact of the AK-ICA on the MRBT, we performed experiments to reconstruct original data by observing the rotated data and the original data fraction known to the attacker. We use the same dataset used in [4]. In the first experiment, the used data had a normal distribution. A fraction of it as small as 10% could, using the Gaussian kernel density estimation, extrapolate the distribution of the whole dataset with accuracy as high as 94%. In the second experiment, the accuracy of the reconstructed data was estimated according to the reconstruction error measure by the Frobenius norm. Figure 1 shows the reconstruction accuracy versus the fraction of data known to the attacker for different *n*. As we assume the boundaries of the data (minimum and maximum) are known to the attacker, he or she can rescale the reconstructed data and solve the ambiguity of reconstruction amplitude noted in [2]. In the third experiment, the UJI Pen Characters dataset [6] was used. It represents more realistic geographical data. We achieved accuracy in a range between 45.7% and 89.2% in applying the AK-ICA attack when the known private data fraction was 10%.

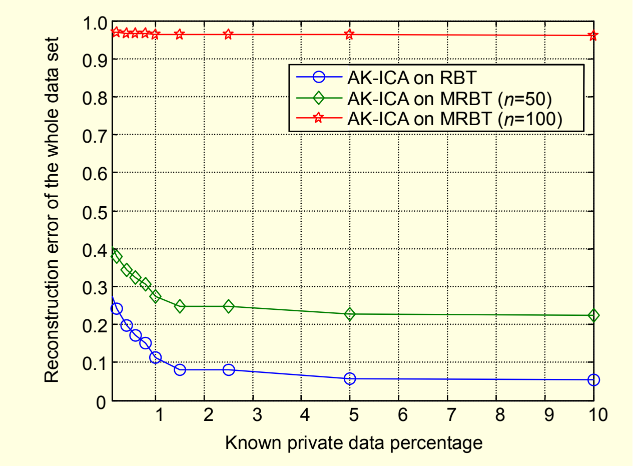

Fig. 1. Reconstruction error of MRBT with various *n* values.

IV. Conclusion

We introduced a scheme to reduce the impact of the ICA attack on RBT, which is commonly used for privacy preserving data clustering. Our results show a relative reduction of the ICA impact. Our scheme may be used in several applications, maintaining the coordinated distance and inner product between corresponding subsets of transformed data.

References

[1] K. Chen and L. Liu, "Privacy Preserving Data Classification with Rotation Perturbation," *Proc. ICDM*, 2005, pp. 589-592.

[2] K. Chen, G. Sun, and L. Liu, "Towards Attack-Resilient Geometric Data Perturbation," *Proc. SDM* , 2007, pp. 89-94.

[3] K. Liu, H. Kargupta, and J. Ryan, "Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining," *IEEE Trans. Knowl. Data Eng*., vol. 18, no. 1, Dec. 2006, pp. 92-106.

[4] S. Guo and X. Wu, "Deriving Private Information from Arbitrarily Projected Data," *Proc. PAKDD*, 2007, pp. 84-95.

[5] S.R.M. Oliveira and O.R. Zaiane, "Privacy Preservation when Sharing Data for Clustering," *Proc. SDM*, 2004, pp. 67-82.

[6] "UJI Pen Characters Data Set," UCI Machine Learning Repository. Available at http://archive.ics.uci.edu/ml/datasets/UJI+Pen+Characters.

[7] T. Nhan Vu, J. Lee, and K. Ryu, "Spatiotemporal Pattern Mining Technique for Location-Based Service System," *ETRI Journal*, vol. 30, no. 3, June 2008, pp. 421-431.

[8] S. Kang et al., "A Semantic Service Discovery Network for Large-Scale Ubiquitous Computing Environments," *ETRI Journal*, vol. 29, no. 5, Oct. 2007, pp. 545-558.