International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-4)

Research Article | **April 2015**

Impact of Known Input Output Attacks in Euclidean Distance Preserving Perturbation for Privacy Preserving Data Mining

Bhupendra Kumar Pandya, Umesh Kumar Singh, Keerti Dixit
Institute of Computer Science, Vikram University,
Ujjain, India

Abstract:

Privacy preserving Data Mining considers the problem of running data mining algorithms on confidential data that is not supposed to be revealed even to the party running the algorithm. In this technique, some statistical data that is to be released, so that it can be used for research using statistical and /or Data Mining, may contain confidential data and so is first modified so that the data does not compromise anyone's privacy and it is still possible to obtain meaningful result by running data mining algorithms on the modified data set. In this research paper we analyze an attack that allows the attacker to estimate the original data tuple associated with each perturbed tuple and calculate the probability that the estimation results in a privacy breach.

Keywords: Distance Preserving Data Perturbation, I/O Attack.

I. INTRODUCTION

Privacy preserving data mining is an important property that any mining system must satisfy. So far, if we assumed that the information in each database found in mining can be freely shared. This research paper offers an overview of distance preserving Perturbation: its definition, application scenarios, etc. Throughout this paper (unless otherwise stated), all matrices and vectors discussed are assumed to have real entries. All vectors are assumed to be column vectors and M′ denotes the transpose of any matrix M. An m × n matrix M is said
to be orthogonal if $M' M = I_n$, the n × n identity matrix. If M is square, it is orthogonal if and only if $M' = M^{-1}$ [2]. The determinant of any orthogonal matrix is either +1 or −1. Let O_n denotes the set of all n × n, orthogonal matrices.

II. DISTANCE PRESERVING PERTURBATION

2.1 Definition and Fundamental Properties

To define the distance preserving transformation, let us start with the definition of metric space. In mathematics, a metric space is a set S with a global distance function (the metric d) that, for every two points x, y in S, gives the distance between them as a nonnegative real number d(x, y). Usually, we denote a metric space by a 2-tuple (S, d). A metric space must also satisfy

1. d(x, y) = 0 iff x = y (identity),
2. d(x, y) = d(y, x) (symmetry),
3. d(x, y) + d(y, z) ≥ d(x, z) (triangle inequality).

A metric space (S_1, d_1) is isometric to a metric space (S_2, d_2) if there is a bijection T: S1 → S2 that preserves distances. That is, $d_1(x, y) = d_2(T(x), T(y))$ for all x, y ϵ S1. The metric space which most closely corresponds to our intuitive understanding of space is the Euclidean space, where the distance d between two points is the length of the straight line connecting them. In this chapter, we specifically consider the Euclidean space and define $d(x, y) = ||x - y||$, the l^2-norm of vector x − y. A function $T : R^n \rightarrow R^n$ is distance preserving in the Euclidean space if for all x, y ϵ R^n, $||x - y|| = ||T(x) - T(y)||$. Here T is also called a rigid motion. It has been shown that any distance preserving transformation is equivalent to an orthogonal transformation followed by a translation [2], . In other words, there exists M_T ϵ O_n and v_T ϵ R^n such that T equals x ϵ $R^n \rightarrow M_T x + v_T$. If T fixes the origin, T(0) = 0, then v_T = 0; hence, T is an orthogonal transformation. Henceforth we assume T is a distance preserving transformation which fixes the origin – an orthogonal transformation. Such transformations preserve the length (l^2-norm) of vectors: ||x|| = ||T(x)|| (i.e., given any M_T ϵ On, ||x|| = $||M_T x||$). Hence, they move x along the surface of the hyper-sphere centered at the origin with radius ||x||. From a geometric perspective, an orthogonal transformation is either a rigid rotation or a rotoinversion (a rotation followed by a reflection). This property was originally discovered by Schoute in 1891 [3]. Coxeter [4] summarized Schoute's work and proved that every orthogonal transformation can be expressed as a product of commutative rotations and reflections. To be more specific, let Q denote a rotation, R a reflection, 2q the number of conjugate imaginary eigenvalues of the orthogonal matrix M, and r the number of (-1)'s in the n − 2q real eigenvalues. The orthogonal transformation is expressible as $Q^q R^r (2q + r \leq n)$. Especially, in 2D space, det(M) = 1 corresponds to a rotation, while det(M) = −1 represents a reflection.

2.2 Generation of Orthogonal Matrix

Many matrix decompositions involve orthogonal matrices, such as QR decomposition, SVD, spectral decomposition and polar decomposition. To generate a uniformly distributed random orthogonal matrix, we usually fill a matrix with independent Gaussian random entries, then use QR decomposition. Stewart [5] replaced this with a more efficient idea

that Diaconis and Shahshahani [6] later generalized as the subgroup algorithm. We refer the reader to these references for detailed treatment of this subject.

2.3 Data Perturbation Model

Orthogonal transformation-based data perturbation can be implemented as follows. Suppose the data owner has a private database $X_{n\times m}$, with each column of X being a record and each row an attribute. The data owner generates an n × n orthogonal matrix M_T , and computes

$$Y_{n\times m} = M_{Tn\times n} X_{n\times m}$$

The perturbed data $Y_{n\times m}$ is then released for future usage. Next we describe the privacy application scenarios where orthogonal transformation can be used to hide the data while allowing important patterns to be discovered without error.

Orthogonal transformation has a nice property that it preserves vector inner product and distance in Euclidean space. Therefore, any data mining algorithms that rely on inner product or Euclidean distance as a similarity criteria are invariant to orthogonal transformation. Put in other words, many data mining algorithms can be applied to the transformed data and produce exactly the same results as if applied to the original data [7-9], e.g., KNN classifier, perception learning, support vector machine, distance-based clustering and outlier detection. We refer the reader to [10] for a simple proof of rotation-invariant classifiers.

In this study we have used Students result database of Vikram University, Ujjain. I randomly selected 7 rows of the data with only 7 attributes (Marks of Foundation, Marks of Mathematics, Marks of Physics, Marks of Computer Science, Marks of Physics Practical, Marks of Computer Science Practical and Marks of Job Oriented Project).

With this data we have generated a noise matrix with the help of orthogonal transformation and this resultant noise data set is multiplied with the original data set to form the perturb data. We have evaluated Euclidean Distance of original and perturbed data with pdist() fuction of Matlab. We have plotted the graph 1 and 2 which shows the comparison between Euclidean Distances of original data and perturbed data after applying Distance Preserving Perturbation.

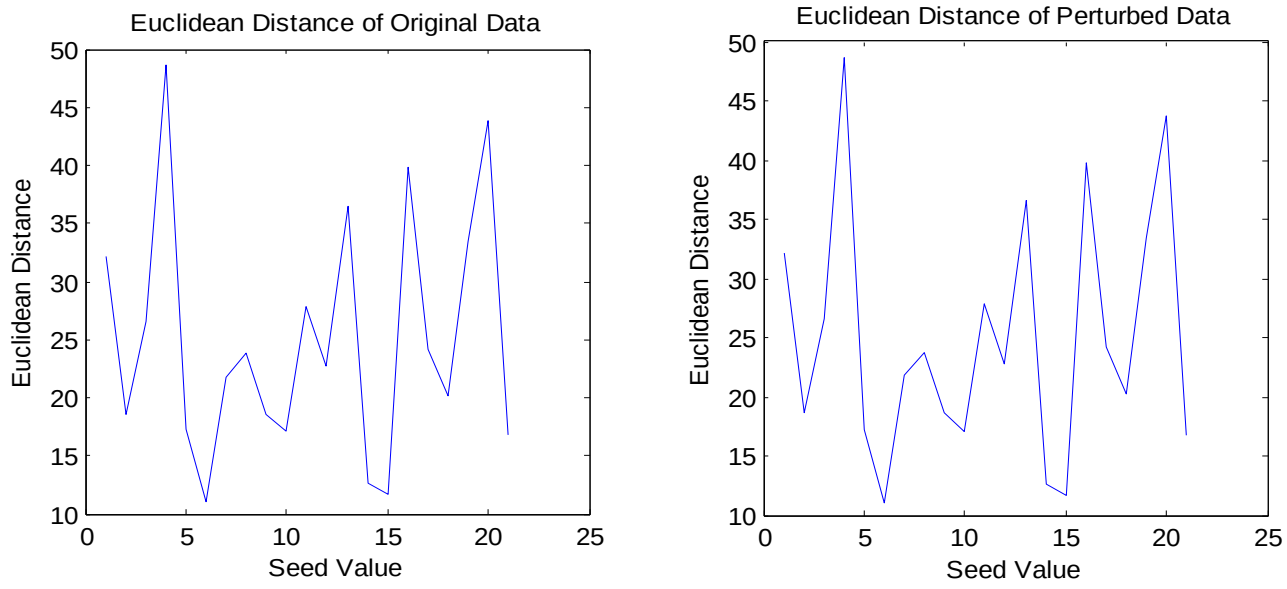

Figure 1and 2

The above graph shows that the Euclidean Distance among the data records are preserved after perturbation. Hence the data perturbed by Euclidean Distance Preserving Perturbation can be used by various data mining applications such as k-means clustering, hierarchical clustering etc. And we get the same result as obtained with the original data.

III. PRIVACY BREACH

Orthogonal transformation-based data perturbation has the nice property that many data mining algorithms can be applied to the perturbed data and produce exactly the same results as if applied to the original data. We are assuming the role of an attacker with prior information regarding the original data. We examine how well the attacker can recover the original data from the perturbed data and prior information.

Liu has given the definition of *privacy breach*. [11] An attacker will have X and Y and that Y was produced from X by an orthogonal transformation. The attacker will also have prior knowledge. The attacker will produce $\hat{x} \in R^n$ and $1 \le \hat{\imath} \le m$, where $\hat{x}$ is the attacker's estimate of $x_{\hat{\imath}}$, the $\hat{\imath}^{th}$ data tuple (column) in X.

Definition 3.1 (∈-Privacy Breach) For any ∈ > 0, we say that an ∈-privacy breach occurs if $\| \hat{x} - x_{\hat{\imath}} \| \le \| x_{\hat{\imath}} \| \in$.

Informally stated, an ∈-privacy breach occurs if the attacker's estimate is wrong with relative error no more than ∈. We further define the probability of privacy breach as follows:

Definition 3.2 (Probability of ∈-Privacy Breach) The probability $\rho(x_{\hat{\imath}}, \in)$ that an ∈-privacy breach occurs given that the attacker chose $\hat{\imath}$, i.e., $\rho(x_{\hat{\imath}}, \in) = \text{Prob}\{\| \hat{x} - x_{\hat{\imath}} \| \le \| x_{\hat{\imath}} \| \in\}$.

IV. PRIOR KNOWLEDGE

Let the n×m matrix X denote a private dataset, with each column of X being a record and each row an attribute. We assume that the attacker knows that transformation function T is an orthogonal transformation and knows the perturbed data $Y = M_T X$. In most realistic scenarios, the attacker has some additional *prior knowledge* which can potentially be used effectively for breaching privacy. We consider three types of prior knowledge.

Known input-output
The attacker knows some collection of linearly independent private data records. In other words, the attacker has a set of linearly independent input-output pairs. In this scenario, we can use an attack algorithm based on linear algebra and statistics theory.
Known sample The attacker knows that the original dataset arose as independent samples of some n-dimensional random vector V with unknown p.d.f. Also the attacker has another collection of independent samples from V. For technical reasons, we make a mild additional assumption: the covariance matrix of V has distinct eigenvalues. In this scenario, we can use a principal component analysis (PCA)-based attack algorithm.
Independent signals Each data attribute can be thought of as a time-varying signal. All the signals, at any given time, are statistically independent and all the signals are non-Gaussian with the exception of one. In this scenario, we can use an independent component analysis (ICA)-based attack algorithm.

V. KNOWN INPUT-OUTPUT ATTACK

Consider the perturbation model
$Y = M_T X \Leftrightarrow$
$(Y_k\ Y_{m-k}) = M_T\ (X_k\ X_{m-k})$.
Let X_k denote the first k columns of X and X_{m-k} the remainder (likewise for Y). We assume that columns of X_k are all linearly independent and X_k is known to the attacker (Y is, of course, also known). The attacker will produce $\hat{x}$ and $1 \leq \hat{\imath} \leq m - k$ such that $\hat{x}$ is a good estimate of $x_{\hat{\imath}}$, the $\hat{\imath}^{th}$ column in X_{m-k} (the $(k + \hat{\imath})^{th}$ column in X). If k = n, then the attacker can recover any column in X_{m-k} perfectly as $X_{m-k} = (Y_k X_k^{-1})' Y_{m-k}$. Thus, we assume k < n. Based on known information, the attacker can narrow down the space of possibilities for M_T to $M(X_k, Y_k) = \{M \in On : MX_k = Y_k\}$.
Because the attacker has no additional information, any of these matrices is equally likely to have been M_T. The attacker chooses $\hat{M}$ uniformly from $M(X_k, Y_k)$ and chooses index $1 \leq \hat{\imath} \leq m-k$ based on $\rho(x_{\hat{\imath}}, \epsilon)$ (the probability that an ϵ-*privacy breach* occurs given that $\hat{\imath}$ was chosen), then produces $\hat{x} = \hat{M}'\ y_{\hat{\imath}} = \hat{M}' M_T\ x_{\hat{\imath}}$. Later we will show how the attacker can compute $\rho(x_{\hat{\imath}}, \epsilon)$ for all $1 \leq \hat{\imath} \leq m - k$ from ϵ and Y (known information). Note that $M(X_k, Y_k)$, in most cases, is uncountable. As such, more precise definitions are needed for "choosing $\hat{M}$ uniformly from $M(X_k, Y_k)$" and "the probability that $\| \hat{M}' M_T x - x \| \leq \|x\| \epsilon$ ".
The goal of the attacker is to use the perturbed data tuples and known original data tuples to produce good estimates of unknown original data tuples along with links to their perturbed counterparts. To achieve this, we can use an attack technique called the known input attack which proceeds in three steps.
1. The attacker links as many of the known original data tuples (columns in X) to their corresponding perturbed counterparts (columns in Y).
2. For each unlinked perturbed data tuple, the attacker computes the breach probability of the associated unknown original data tuple. This is the probability that the following stochastic procedure will result in an accurate enough estimate of the associated unknown original data tuple to be considered a privacy breach (the probability calculation is done by applying a closed-form expression we derive later).
(a) A Euclidean distance-preserving transformation is uniformly chosen from the space of such transformations that satisfy the original-perturbed (input-output) constraints from step 1.
(b) The inverse of the chosen transformation is used to estimate original data tuples from their perturbed counterparts.
3. The attacker chooses the perturbed data tuples which are most vulnerable to breach based their probabilities from step 2, e.g. chooses the one with the maximum probability or chooses all whose probability exceeds a threshold, and generates estimates of their associated known original data tuples.

VI. KNOWN INPUT-OUTPUT ATTACK ALGORITHM

As stated earlier, the adversary chooses $\hat{M}$ uniformly from $M(X_k, Y_k)$ and $1 \leq \hat{\imath} \leq m - k$ to maximize $\rho(x_{\hat{\imath}}, \epsilon)$.
Algorithm Known Input-Output Attack Technique
Inputs: X_k, an set of linearly independent columns from X known to the attacker and$Y = M_T X$, known to the attacker, where $M_T \in On$ is an unknown, and $\epsilon \geq 0$, known to the attacker.
Outputs $1 \leq \hat{\imath} \leq m-k$ which maximizes $\rho(x_{\hat{\imath}}, \epsilon)$ and $\hat{x} \in R^n$ the corresponding estimate of $x_{\hat{\imath}}$.
1: Compute V_k an n × k, orthogonal matrix where $Col(V_k) = Col(Yk)$ from Y_k using the Gram-Schmidt process.
2: For each $1 \leq j \leq m - k$ do
3: Compute $d(y_j, Y_k) = \|V_k V'_k y_j - y_j\|$ and $\|y_j\| \epsilon$.
4: Compute $\rho(x_j, \epsilon)$ using Equation 4.3.
5: End For.
6: Set $\hat{\imath} \leftarrow \max_{1 \leq j \leq m-k} \{\rho(x_j, \epsilon)\}$.
7: Choose $\hat{M}$ uniformly from $M(X_k, Y_k)$.
8: Set $\hat{x} \leftarrow \hat{M}' y_{\hat{\imath}}$.

Experimental Result: We have already taken the student record of Vikram University. We have applied the Input/output attack on the Perturbed data. With this data we have generated an orthogonal matrix with help of Gram- Schmidt process. After this we have calculated the inverse of orthogonal matrix and applied on the perturbed data. We have plotted the graph with original data, perturbed data and recovered data. The graph 3 shows original data and perturbed data and the graph 4 shows the original data and recovered data.

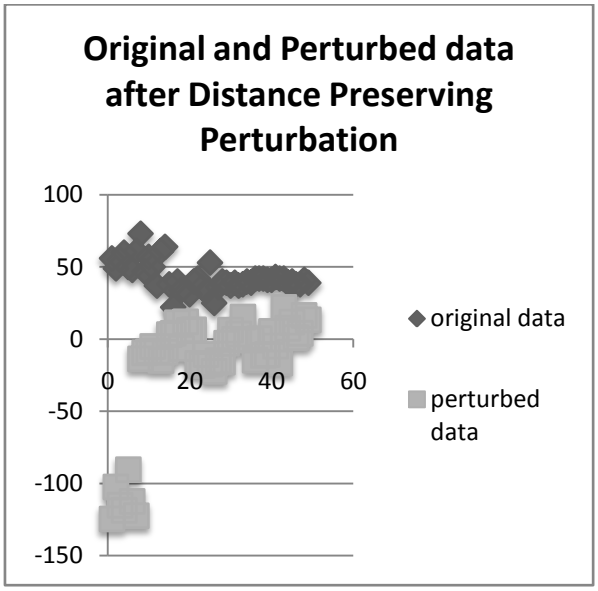

Figure 3

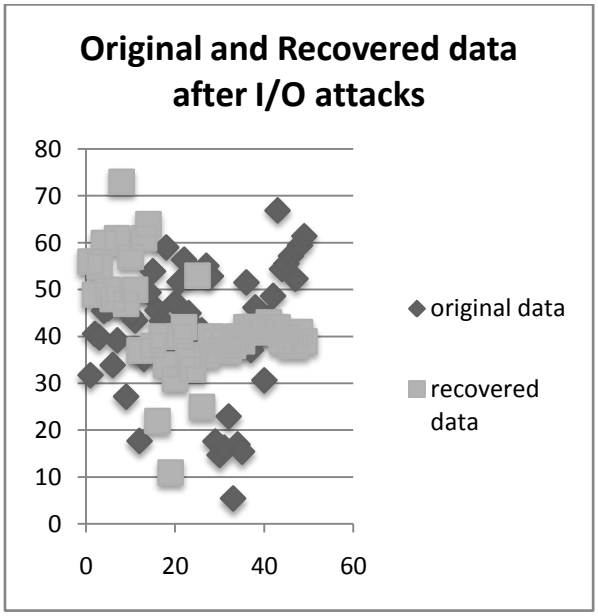

Figure 4

VII. DISCUSSION

It is proved by the above graph that in the Euclidean distance preserving data perturbation technique the euclidean distance among data records are preserved after perturbation but after applied input/output attack on the perturbed data the attacker can get the expected recovered data.

VIII. CONCLUSION

In this research paper we have analyzed the effectiveness of Euclidean distance preserving perturbation technique for privacy preserving data mining.on the one hand ,this technique is quite useful as it allows many interesting data mining algorithm to be applied directly to the perturbed data and produce an error free result, e.g., K-means clustering and K-nearest neighbour classification.

On the other hand ,by considering the prior knowledge ,an attacker may use attack technique to recover original data .Hence the privacy of original data is vulnerable.Our analysis explicitly illuminates scenario where privacy can be seriously breached.As such , valuable information is gained into the distance preserving perturbation for privacy preserving data mining.

REFERENCES

[1] M. Artin, Algebra. Prentice Hall, 1991.
[2] P. H. Schoute, "Le d´eplacement le plus g´en´eral dans l'espace `a n dimensions," Annales de l'´Ecole Polytechnique de Delft, vol. 7, pp. 139–158, 1891.
[3] H. S. M. Coxeter, Regular Polytopes, 2nd ed., 1963, ch. XII, pp. 213–217.
[4] G. W. Stewart, "The efficient generation of random orthogonal matrices with an application to condition estimation," SIAM Journal of Numerical Analysis, vol. 17, no. 3, pp. 403–409, 1980.
[5] P. Diaconis and M. Shahshahani, "The subgroup algorithm for generating uniform random variables," Probability in Engineering and Information Sciences, vol. 1, pp. 15–32, 1987.
[6] K. Chen and L. Liu, "Privacy preserving data classification with rotation perturbation," in Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05), Houston, TX, November 2005, pp. 589–592.
[7] B. Pandya,U.K.Singh and K. Dixit, "An Analysis of Euclidean Distance Preserving Perturbation for Privacy Preserving Data Mining" International Journal for Research in Applied Science and Engineering Technology, Vol. 2, Issue X, 2014.
[8] B. Pandya,U.K.Singh and K. Dixit, "Performance of Euclidean Distance Presrving Perturbation for K-Means Clustering" International Journal of Advanced Scientific and Technical Research, Vol. 5, Issue 4, pp 282-289, 2014.
[9] B. Pandya,U.K.Singh and K. Dixit, "Performance of Euclidean Distance Presrving Perturbation for K-Nearest Neighbour Classification" International Journal of Computer Application, Vol. 105, No. 2, pp 34-36, 2014.
[10] J. Han and M. Kamber. Data Mining Concepts and Techniques. Morgan Kaufmann Publishers, San Diego, CA 92101-4495,USA, 2001.[And73] Michael R. Anderberg, Cluster Analysis for Applications, Academic Press, New York and London (1973)
[11] K. Liu, C. Giannella, H. Kargupta, A survey of attack techniques on privacy-preserving data perturbation methods, in: Privacy Preserving Data Mining: Models and Algorithms, Vol. 53 of Advances in Information Security, Springer Verlag, 2008

