ISSN:0975-9646

Ultimate Way to Preserve Privacy in Data Mining

Neha Agrawal, Nidhi Chaturvedi

Lecturers, Department of Computer Science and Engineering, HITS&M

Abstract:Popular matter from the listed matters of data minning researchers is Privacy preserving data mining (PPDM). To strike a balance between privacy protection and knowledge discovery in the sharing process is an important issue. Main focus point of this work is privacy preserving utility mining (PPUM) and presents t an algorithms that preserve privacy through new approach of encryption to preserve privacy in dataming i.e. this approach is to hide the sensitive itemsets by applying algorithm to perform encryption to save sensitive itemsets. That whole idea is conclude from the current approach of combinational effort that performed from the HHUIF and MSICF, to achieve the goal of hiding sensitive itemsets so that the adversaries cannot mine them from the modified database. Which also work to minimizes the impact on the sanitized database of hiding sensitive itemsets.

Keywords: UWPP, PPDM, MDTNA (Mine Data Through Neha Algo.).

INTRODUCTION:

The privacy-preserving data mining (PPDM) has become an important issue in recent years. Tzung-Pei Hong *et al*. [8] proposed a paper, a greedy-based approach for hiding sensitive item sets by inserting dummy transactions. That computes the maximal number of transactions to be inserted into the original database for totally hiding sensitive item sets. Experimental results were also performed to evaluate the performance of that proposed approach. In recent years, the wide availability of personal data has made the problem of Privacy Preserving Data Mining an important one. The increasing ability to track and collect large amounts of data with the use of current hardware technology has lead to an interest in the development of data mining algorithms which preserve user privacy. A number of methods have recently been proposed for privacy preserving data mining of multidimensional data records. The proposed algorithm presents a new and ultimate way to secure and safe approach to preserve privacy in data mining.

PROPOSED UNIQUE WAY OF ENCRYPTION TO PRESERVE PRIVACY IN DATA MINING:

Homomorphic Encryption through the integers, which is denoted as HE in this paper, consists of four algorithms *KeyGen from the data*, *Encrypt*, *Decrypt* and *Evaluate*. The size (bit length) of various integers used in the scheme is denoted by the parameters *e*, *t*, *r*, *g*, *d*, which represent the size of the secret key, number of elements in the public key, size of the noise in the public key integers, the size of each integer in the public key, size of the noise used for encryption, respectively and are polynomial in the security parameter *n*.

The parameter setting suggested in view of the homomorphism and security is, Proposed
$e = \tilde{O}(n3)$, $r = n$, $d = 2n$, $g = \tilde{O}(n5)$, and $t = g + n$. This makes the public key size as $\tilde{O}(n15)$, because, the public key consists of $t = \tilde{O}(n5)$ integers each of size $g = \tilde{O}(n5)$.
KeyGen(*n*): Choose a uniform pattern integer from the right open interval [i, 2^(i+1)] as the secret key P. For i = 0,1,…..,*text size* , Choose a random integer from initial charector Qi from index 0, another integer Ri from the open interval random interwal (*r*, 2*r*), and compute chepher text = ch_int + 2 pow (nway+1)upto the size of text until the conditions Xi > Xi+1,….., X*t* , X0 mod 2 = 1, and (X0 mod P) mod 2 = 0 are satisfied. Output the public key PK = (X0, X1,….., X*t*) and the secret key SK = P.
Encrypt(PK, M ∈{0, 1}): Choose an integer B from (*r*, 2*r*) as noise for encryption. Output the ciphertext as C = ch_int + 2 pow (nway+1)+ascii ch_int, nway = 0,1,2,3.....n give the n different way to encrypt data.
Decrypt (SK, C): Output the plane text as P = ch_int - 2 pow (nway+1)-ascii ch_int
Evaluate
Multiplication or addition of such near multiples results in another near multiple. Therefore in decryption, C mod P results in N, and N mod 2 gives the plaintext bit M. For every multiplication during the *Evaluate*, the size of the resulting noise equals the sum of the sizes of the multiplicand noises, which crosses the size of P after certain number of multiplications, resulting in incorrect decryption. This makes the scheme an SHE and to make it an FHE, the transformation based on Gentry's blueprint [2][3] is used by [5].
Proposed approach targets only the underlying EPKG. The method being proposed is denoted as PKG, in which the public key consists of only one big integers 'n'. n is an exact multiple of the integer P and n is an approximate multiple, i.e., multiple of P containing some additive error R. To encrypt a plaintext bit n, the result is added to the plaintext bit and the final sum is reduced modulo the error-free integer X0 in the public key. For homomorphic evaluation of a function, the addition and multiplication operations in the corresponding arithmetic circuit are performed over ciphertexts, modulo the error free integer X0 in the public key.

IMPROVEMENT IN BIT COMPLEXITY

As discussed earlier, the public key of the HE contains $\tilde{O}(n)$ elements each of which is $\tilde{O}(n)$ bits long. This will take $\tilde{O}(n)$ computations for complete key generation. Also, in that scheme the bit length of a fresh cipher text that encrypts a single bit is $\tilde{O}(n)$, leading to an expansion ratio of n . The public key in the scheme EPKG consists of only

two elements of Õ(*n*) bits long. This makes the complexity of key generation as Õ(*n*). This is a considerable improvement over the somewhat homomorphic schemes of [5] and [8]. Also, the encryption of an Õ(*n*) bit plaintext, which involves a multiplication of Õ(*n*5). Õ (*n*) and a modular reduction of this with Õ(*n*) bit X0 takes Õ(*n*) steps. Similarly, the bit complexity of decryption is roughly Õ(*n*). Therefore, the overall complexity of the proposed variant EPKG is Õ(*n*). Similarly, a single plaintext bit is embedded in a cipher text of Õ(n) bits making the expansion ratio also comparatively less which is *n*. With these drastic improvements in bit complexity and ciphertext expansion, this conceptually simple somewhat homomorphic scheme will be suitable for many practical applications that involve simple functions for homomorphic evaluation.

RESULT SNAPSHOT & GRAPH

Result:

This approach is designed intentionally to preserve more privacy in data mining, This approach is faster, more secure & more reliable than combinational effort putted by HHUIF and MSICF For Plane Text or data matter or sensitive data which is to be safely min in data mining, sensitive data could be save after applying the our proposed algorithm.

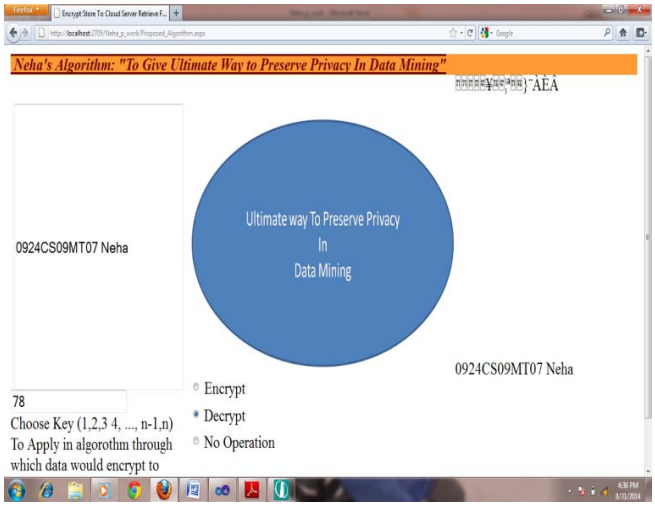
Let Encrypt Sensitive Data Like

Roll No. 0924CS09MT07

Name Neha

TA1:

Randomly choose Numeric Value To Encrypt Sensitive Data	Corresponding Cipher Text Which Could Safely Mine through less effort.
10	=C=DTa?MbfCO9d| ~
25	LRLScpN\quR^Hs
45	`f`gw bp fr\ §¡
56	kqkr m{ q}g ª²¬
78	¥ ¹ª ¡ }¨ÀÈÂ

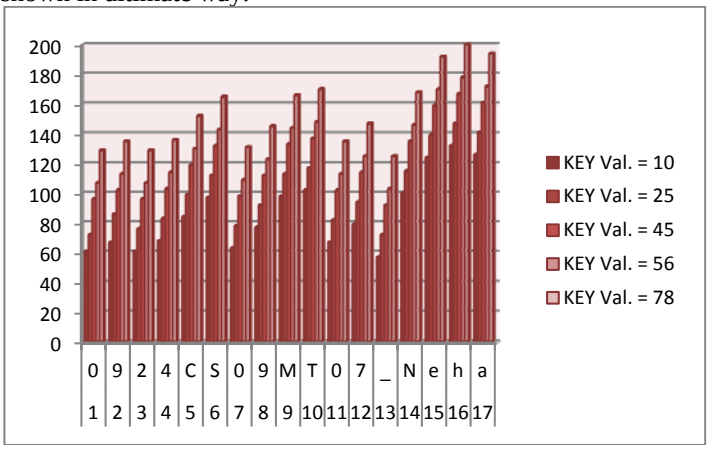


Snapshot SS1

In above snapshot SS1 it is clear that we may encrypt from written number (That would be any one from 1,2,3, …..n) Below mentioned text and the important fact is that only that number can decrypt the text through which it is encrypted.

Graphical Analysis:

Here Safe cipher value is drawn with corresponding to plane text word position as shown below GA1 and in table TA2. Where word position cipher value cipher text all shown in ultimate way.



GA1: Word Position With Safe Cipher Value

Through above analytic representation it is clear that higher numeric value reduce the possibility of attack and help to safe sensitive data to preserve privacy in data mining.

TA2

Word Position	Word in cipher text.	
	Safe cipher Value of sensitive data	Text over value
1	61	=
2	67	C
3	61	=
4	68	D
5	84	T
6	97	A
7	63	?
8	77	M
9	98	B
10	102	F
11	67	C
12	79	O
13	57	9
14	100	D
15	124	|
16	132	
17	126	~

CONCLUSION:

The entire effort of this paper is to preserve privacy in data mining by reducing the complexity of previously proposed complex structured algorithm which were putting their effort by hiding the sensitive information of data which has to be save in another where or it was necessary to remember their location id in data mining. In next effort flexibility of encryption will move at high label i.e. that will be more difficult to perform decryption by anonymous user i.e. in next effort our approach will be more easier and more effective the currently proposed algorithm to encrypt data to preserve privacy in data mining.

REFERENCE:

1. Yehuda Lindell and Benny Pinkas, "Secure Multiparty Computation for Privacy-Preserving Data Mining," The Journal of Privacy and Con¯dentiality, vol. 1, no. 1, pp. 59-98, 2009.
2. Marina Blanton, "Achieving Full Security in Privacy-Preserving Data Mining," In Proc. of the 2011 ieee third international conference on social computing (socialcom) Privacy, security, risk and trust (passat), Dame, IN, pp. 925-934, Oct 2011.
3. Bin Yang, Hiroshi Nakagawa, Issei Sato, and Jun Sakuma, "Collusion-Resistant Privacy-Preserving Data Mining," In Proc. of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 483–492, 2010.
4. Li Liu, Murat Kantarcioglu, and Bhavani Thuraisingham, "The applicability of the perturbation based privacy preserving data mining for real-world data," Journal of Data & Knowledge Engineering, vol. 65, pp. 5–21, 2008.
5. Stanley R. M. Oliveira, and Osmar R. Zaiane, "Revisiting Privacy Preserving Clustering by Data Transformation," Journal of Information and Data Management, vol. 1, no. 1, 2010.
6. Dhyanendra Jain, Amit sinhal, Neetesh Gupta, Priusha Narwariya, Deepika Saraswat, and Amit Pandey, "Hiding Sensitive Association Rules without Altering the Support of Sensitive Item(s),"International Journal of Artificial Intelligence & Applications (IJAIA), Vol. 3, No. 2, pp. 75-84, Mar 2012.
7. Elena Dasseni, Vassilios S. Verkios, Ahmed K. Elmagarmid, and Elisa Bertino, "Hiding Association Rules by Using Confidence and Support," Computer Science Technical Reports, 2000.
8. Tzung-Pei Hong, Chun-Wei Lin, Chia-Ching Chang, and Shyue-Liang Wang, "Hiding Sensitive Itemsets by Inserting Dummy Transactions," In Proc. of the IEEE International Conference on Granular Computing (GrC), Kaohsiung, Taiwan, pp. 246-249, 2011.
9. Dr. K. Duraiswamy, Dr. D. Manjula, and N. Maheswari, "Advanced Approach in Sensitive Rule Hiding," Modern Applied Science, vol. 3, no. 2, pp. 98-107, Feb 2009.
10. Jieh-Shan Yeh and Po-Chiang Hsu, "HHUIF and MSICF: Novel algorithms for privacy preserving utility mining," Jounal of Expert Systems with Applications, vol. 37, pp. 4779–4786, 2010.

AUTHORS

Neha Agrawal received her B.E. degree in Computer Science and Engineering from Gwalior Engineer College, Gwalior, in 2012. She is pursuing M.E. degree from MPCT college in Computer Science and Engineering, Gwalior. Her research interests include Ultimate way to preserve privacy in data mining. At present, she is engaged in Hindustan institute of technology and management as Lecturer in Computer Science and Engineering department.

Nidhi Chaturvedi received her B.E. degree from Gwalior Engineering College in Computer Science and Engineering in 2012. She is working as Lecturer in HITS&M. Her research interests include Ultimate way to preserve privacy in data mining.