© 2016 IJSRST | Volume 2 | Issue 4 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Preserving Privacy Using Attribute Based Access Control in Data Stream

Suyog Vilas Patil[1], Prof. K. B. Manwade[2]

[1]ME (CSE) Student,Ashokrao Mane Group of Institutons,Vathar, Kolhapur, Maharashtra, India

[2]Asst. Prof., Ashokrao Mane Group of Institutons, Vathar, Kolhapur, Maharashtra, India

ABSTRACT

Data security is a very broad area that addresses many issues, like legal and ethical issues regarding the right to access certain information. The sensitive data is accessible to authorized users only. The database security is based on the Access Control Mechanism (ACM) and the Privacy Protection Mechanism (PPM). The Access Control Mechanisms (ACM) is used to ensure that only information is available authorized to users. The authorized user can only access the authorized data. The privacy protection mechanism (PPM) is a general method used to transform the original data into some anonymous form to prevent from accessing owners sensitive information. There are numerous methods to provide the privacy for the sensitive data. In existing literature, the Access Control mechanism allows Role Based Access using Time Based Sliding Window Query. It protects the user information from the unauthorized access. PPM meets privacy requirement through k-anonymity it provides better privacy for the sensitive information which is to be shared. The privacy is achieved by the high accuracy of the user information. To protect data, the anonymization method is one of the best privacy protection mechanisms. The anonymization process will transform the sensitive information to some anonymzed form using K-anonymity, ℓ -diversity. The PPM needs to satisfy an additional constraint namely the Imprecision bound for each selection predicate. The imprecision bound reduced the delaying for publishing data stream. The challenge is to optimize the delay in publishing data stream, if the delay in publishing query is reduced then the imprecision bound is satisfied. Proposed system is an accuracy based access control using Attribute Based Access Control with Tuple Based Sliding Window Query. and PPM with the ℓ -diversity. The ℓ -diversity method is an extension of the k-anonymity method, it is more efficient than the k-anonymity method. It avoids the attacks like background knowledge attack and others in k-anonymity method.

Keywords: Access Control Mechanisms, Privacy Protection Mechanism, Information Security, Top Down Selection Mondrian, Attribute Based Access Control, Total Minimum Imprecision

I. INTRODUCTION

In the field of Information Security, the data security refers to protecting data from unauthorized users. It is based on confidentially, integrity and availability. To accessing a DataStream, it is not possible to control the order in which data arrive, nor feasible to storing data. The data owners may not be willing to exactly processing the true value of their data due to various reasons, mostly privacy consideration and accuracy of data accessing. However, the unauthorized user can also take advantages of this to accessing the data. Therefore it is need to user has a better access control mechanism to provide both security and privacy.

II. METHODS AND MATERIAL

1. Literature Review

1. To provide better Access control mechanism to ensure both security and privacy of the sensitive information, Access control mechanism for Data Stream Management System (DSMS) provide access of authorized part of the stream to each user**. Zahid Pervaiz, Arif Ghafoor, Fellow, Walid G. Aref et.al** uses a Role Based Control Policy (RBAC) [1] defines the authorized view of the data stream for each role.

2. Role based access control gives permission to the users to access data based on their role. For Relational data **Nagabhushan,Arif Ghafoor,Zahid Pervaiz et al** defines [2] selection predicates query technique is savailable to role while the privacy requirement is satisfy.

3. The stream data offers query processing over continues and sequencing data for data publishing and the windowing techniques generally emphasize on the streaming data. **T.Ghanem, A.Elmagarmid, P.Larsen and w.Aref et al**. proposed the predicate window query processing for streaming data [3].The access control uses Role based techniques to satisfy minimum Imprecision bound using the time based sliding window query [1].

4. To maintain the privacy of data it is need to minimize the imprecision of aggregate for all queries. The imprecision bound is a resulted value which determines the amount of imprecision that can be tolerated for each query. Privacy preserving mechanism needs to sum of false negative and false positive is less than imprecision bound**. Zahid Pervaiz, Arif Ghafoor, Fellow, Walid G. Aref et.al** proposed the Top Down Selection Mondrian (TDSM) used to minimize imprecision bound for rational data [2].

5. The Data anonymization is one of the important privacy protection techniques. It transforms the sensitive information to some anonymzed form**. J. Cao, B. Carminath, E. Ferrai and K. Tan et al.** presents the continues anonymzing data stream [4]. To anonymize data use generalization will replace the sensitive information with border range. **C. Clifton and T. Tesa et al** proposed the differential privacy model using generalization [5**]. L. Sweency** represents an anonymization method to preserve a privacy of data. i.e. k-Anonymity [6]. To prevent uncontrollable information loss and affects the accuracy of crowd sourcing database **Saiwo,Xi aloi, Sheng Wang et al** proposed K-Anonymity for crowd sourcing database [7].

2. Proposed Work

In proposed system, the access control mechanism allows only authorized queries predicates on sensitive data using Attribute Based Access Control (ABAC) policy. It defines an access control paradigm whereby access rights are granted to users through the use of policies which combine attributes together. The policies can use any type of attributes (user attributes, resource attributes, environment attribute etc.).The imprecision bound is a threshold value which determines the amount of imprecision that can be tolerated for each query. The access control policy administrator sets the imprecision bound using overlap semantics and enclosed semantics for each predicate Tuple Based Sliding Window Query. This query meets the minimum imprecision bound, means it prevents and solves the query violation problem. The proposed system formulate query generation module of the system using tuple based sliding window query that offers overlap and enclosed semantics. The proposed system refers data anonymization using the ℓ -diversity. ℓ -diversity method reduces the granularity of representation of the data, ℓ -diversity can still defend against background knowledge that is unknown to the data publisher. The ℓ -diversity method is an extension of the k-anonymity method.

The following techniques is used are as follows,

1. To use Attribute Based Access Control Policy for precession bounded access over DataStream
2. Implementation of heuristics algorithm for satisfying the imprecision bounds.
3. To use Tuple Based Sliding Window Query for accessing data to maintain accuracy.
4. To use ℓ -diversity method for anonymization of DataStream instead of k-anonymity

3. Proposed Architecture

The proposed system provides following dynamic policy management modules:

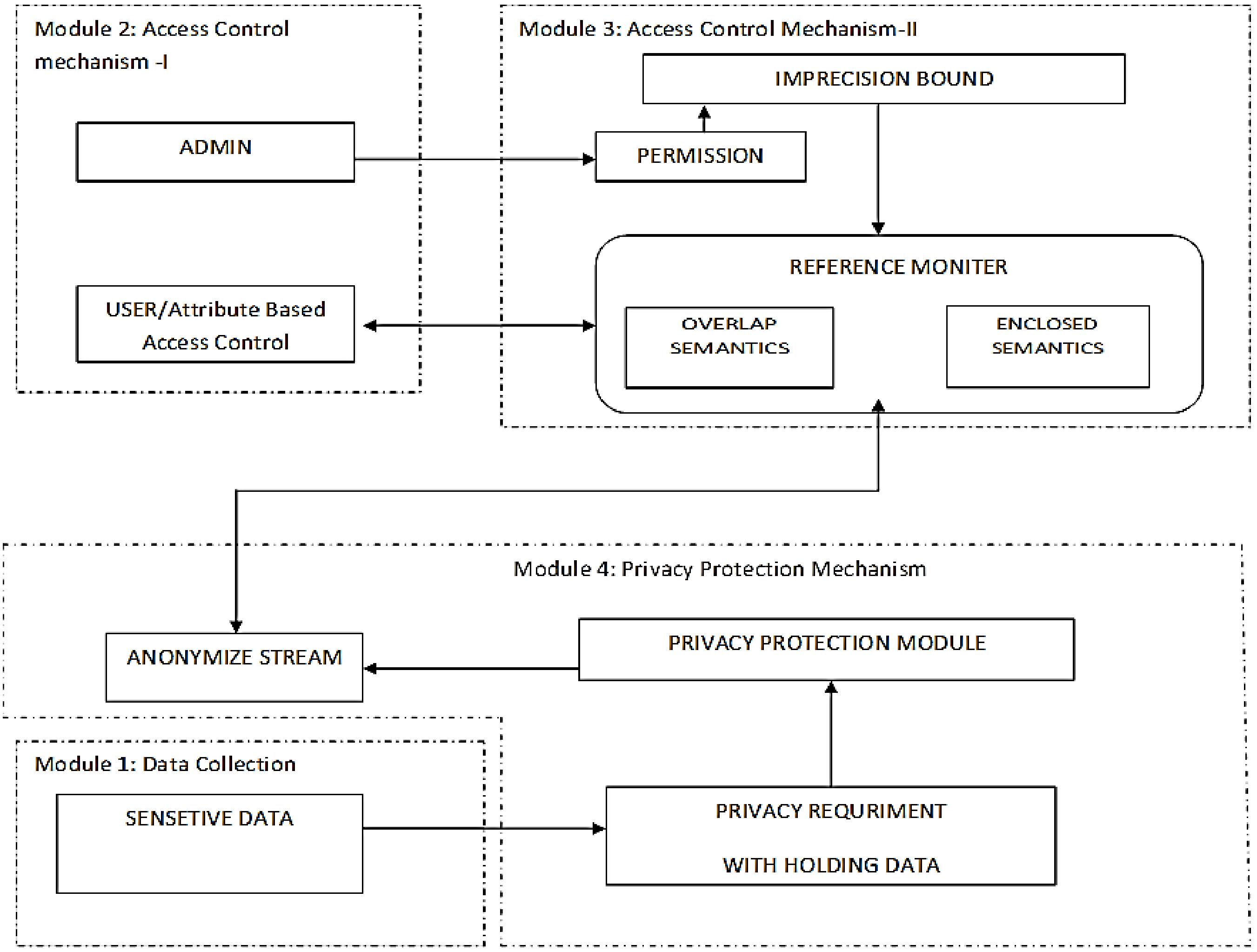

Module 2: Access Control mechanism -I
ADMIN
USER/Attribute Based Access Control
Module 3: Access Control Mechanism-II
IMPRECISION BOUND
PERMISSION
REFERENCE MONITER
OVERLAP SEMANTICS
ENCLOSED SEMANTICS
Module 4: Privacy Protection Mechanism
ANONYMIZE STREAM
PRIVACY PROTECTION MODULE
Module 1: Data Collection
SENSETIVE DATA
PRIVACY REQURIMENT
WITH HOLDING DATA

Figure 1: Architecture of Attribute based Access Control System

- ✓ **Data Collection**

In this module, the proposed system refers sensitive information in the form of streaming data. DataStream offers a continues data arrive in a system for storing and processing.

Access Control Mechanism I

The Access control mechanism-I is divide into two parts are as follows:

1**) User/attribute**: In this framework user accessing the DataStream using attribute Based Access Control (ABAC) technique. It provides a high level of flexibility that promotes security and information sharing. ABAC is composed of a set of Users, a set of attributes, and a set of Permissions.

2) **Admin Permissions**: System administrator set the permission to imprecision bound for each query, user-to-attribute assignments, and attribute-to permission assignments.

Access Control Mechanism II

The Access Control Mechanism –II is divided into two parts are as follows:

1) **Imprecision Bound**: It ensures that the authorized data has the desired level of accuracy. The imprecision bound can be used to meet the privacy requirement. The privacy protection mechanism is required to meet the privacy requirement according to the imprecision bound for each permission.

2) **Reference monitor**: A predicate Tuple Based Sliding Window Query is evaluated for a DataStream by including all the stream tuples that satisfy the query predicate. For predicate evaluation over an anonymzed DataStream will be used reference monitor, including all the tuples in equivalence classes that overlap the query predicate range.

Privacy Protection Mechanism:

The privacy protection module anonymzed the data using ℓ-diversity to meet privacy requirements and imprecision constraints on predicates set by the access control mechanism.

✓ **Recursive ℓ diversity:**

Recursive (c, ℓ)-Diversity. In a given q⋆-block, let ri denotethe number of times the ith most frequent sensitive value appears in that q⋆-block. Given a constant c, the q⋆-block satisfies recursive (c, ℓ)-diversity if $r_1 < c(r_\ell + r_{\ell+1} + \cdot\ \cdot\ \cdot + r_m)$. A table T ⋆ satisfies recursive (c, ℓ)-diversity if every q⋆-blocksatisfies recursive ℓ-diversity. We say that 1-diversity is always satisfied.Now, both entropy and recursive ℓ-diversity may be too restrictive. To see whylet us first look at entropy ℓ-diversity. Since −x log(x) is a concave function, it canbe shown that if we split a q⋆-block into two sub-blocks qa and q⋆b then entropy(q⋆) ≥ min(entropy(q⋆a), entropy(q⋆b)). This implies that in order for entropy ℓ-diversity tobe possible, the entropy of the entire table must be at least log(ℓ). This might notbe the case, especially if one value of the sensitive attribute is very common – for example, if 90% of the patients have “heart problems” as the value for the “Medical Condition” attribute.This is also a problem with recursive ℓ-diversity. It is easy to see that if 90% of the patients have “heart problems” as the value for the “Medical Condition” attribute then there will be at least one q∗-block where “heart problems” will have frequency of at least 90%. Therefore if we choose c < 9 in Definition 4.2, no generalization of the base table will satisfy recursive (c, ℓ)-diversity

4. Implementation Steps

The implementation steps are as follows:

Step 1. The proposed system uses a continues and sequencing data . In step I the system uses DataStream.

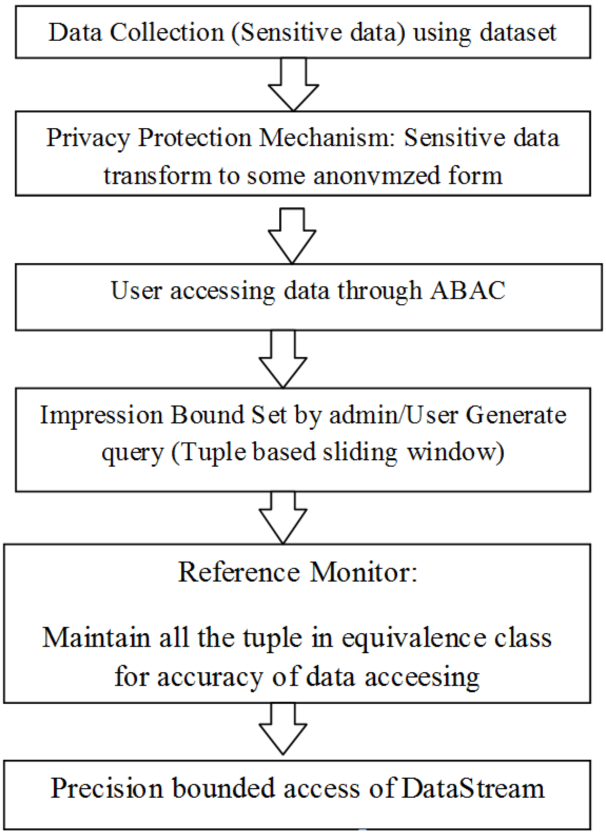
Step 2. In Step II the privacy protection mechanism deals with sensitive data.it transfer the sensitive data into some anonymize form using ℓ -diversity.

Step 3. In access control mechanism -I user access the accurate data using attribute based access control.

Step 4. In access control mechanism –II generate the tuple based sliding window query for precision access of data and admin set the permission to maintain imprecision bound.

Step 5. The reference Monitor maintains all the tuples in equivalence class for accuracy of data accessing.

Step 6. In this step ,attributed based access control using tuple based sliding window query maintain precision bound access of the DataStream.


Data Collection (Sensitive data) using dataset
Privacy Protection Mechanism: Sensitive data transform to some anonvmzed form
User accessing data through ABAC
Impression Bound Set by admin/User Generate query (Tuple based sliding window)
Reference Monitor:
Maintain all the tuple in equivalence class for accuracy of data acceesing
Precision bounded access of DataStream

5. Scope

Role Based Access Control policy cannot be ensures permission on sequence of operation need to be controlled. In proposed system, Attribute-based access control (ABAC), provides a high level of flexibility that promotes security and information sharing.

The Total Minimum Imprecision (TIM) algorithm use to meet the desired level accuracy for relational data accessing. In proposed system, the heuristics algorithm will be used for streaming data to satisfy precision bounded access.

The purpose of access control is to ensure that each user access only the authorized information. This semantic suggested in reference monitor, the Role Base Access Control (RBAC) uses the time based sliding window query. To prevent an overlapping data and maximum response time of query evolution, the proposed system refers Tuple Based Sliding Window Query with use of overlap and enclosed semantics.

The k-anonymity is the anonymization techniques convert the sensitive information to some anonymzed form using generalization and suppression. The proposed system uses ℓ -diversity method. ℓ - diversity is a form of group based anonymization that is used to preserve privacy in data sets by reducing the granularity of a data representation.

6. Methodology

There are following techniques and algorithm in proposed methodology:

- ✓ **Data Collection**: The proposed system stores the sensitive data in the form of DataStream. Accessing a DataStream is concerned with extracting knowledge represented in non-stopping, continues and ordered sequence of information.
- ✓ **Access Control Mechanism I:** The framework of Access Control Mechanism -I uses Attribute-based access control (ABAC), provides a high level of flexibility that promotes security and information sharing. ABAC also overcomes some of the problems associated with Role Based Access Control (RBAC).RBAC cannot be ensure permission on sequence of operation need to be controlled. Using ABAC policy mining algorithm we prevent privileges unwieldy.
- ✓ **Access Control Mechanism II:** The purpose of Access Control Mechanism-II is used to Tuple Based Sliding Window Query technique. Privacy is achieved using cost of accuracy and imprecision is introduces in the authorized information under an access control policy. In proposed system a heuristic algorithm set the minimum imprecision bound. Query impression bound is the total impression acceptable for query predicate and present by the access control administrator.

4. Privacy Protection Mechanism:

Privacy protection mechanism includes data anonymization. The data anonymization is the process that transforming sensitive data to some anonymzed form. The proposed system used ℓ-diversity, it is techniques better than k-anonymity. It has strong background knowledge and maintains lack of diversity. The approach for preserving privacy is based on data anonymization.

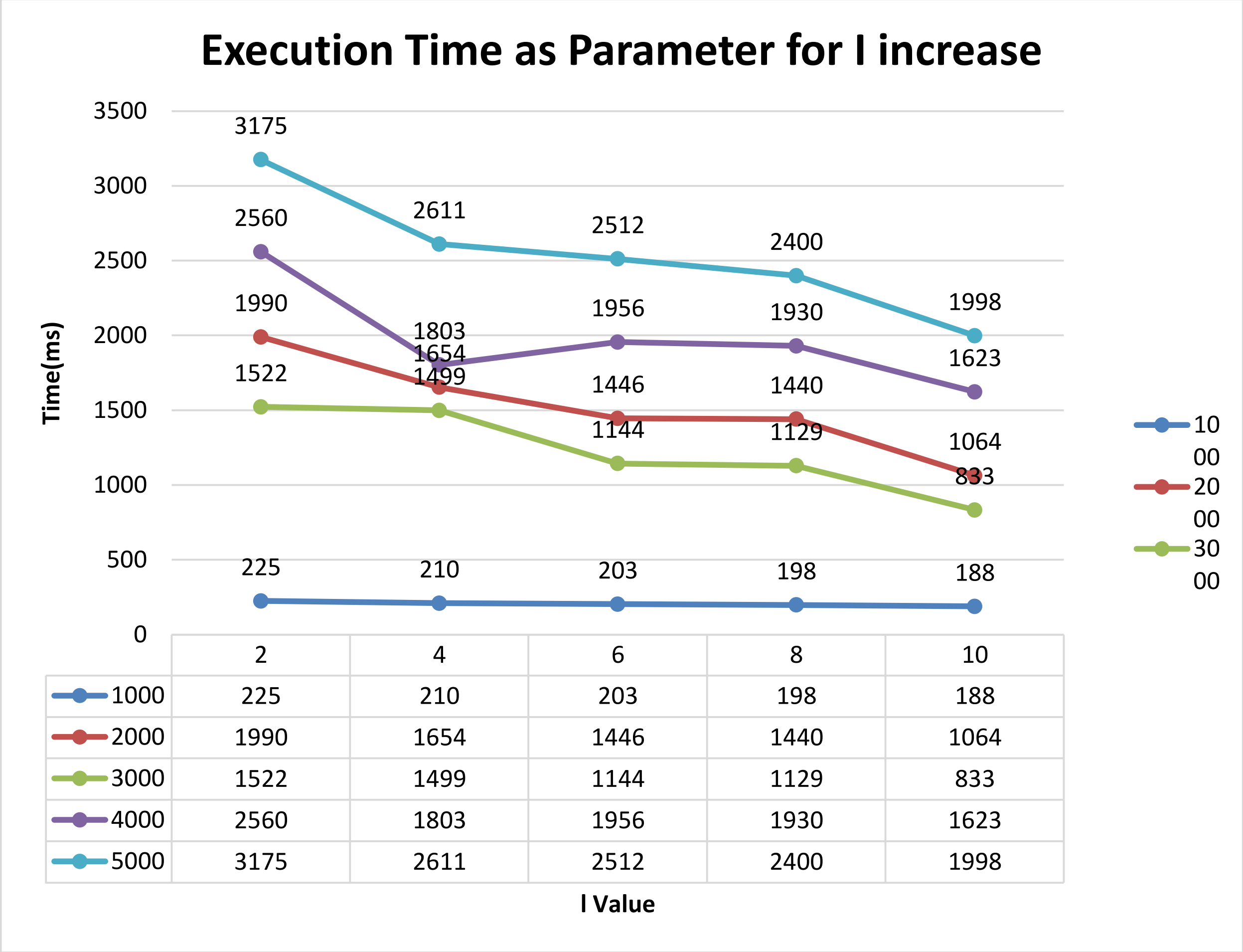
Experimental Setup & Implementation:

1. To implement above system some pre work to do.
2. To Fetch the dataset from a file and import to the database.
3. A privacy preserving module preserves the data using l- diversity.
4. The reference monitor uses the top down heuristics algorithm for accuracy of data.
5. The ABAC policy uses the data accessing through attribute.

III. RESULT AND DISCUSSION

We fetch and perform really insertion of record into database around then into frame work the normal execution time means what number of time used to insertion of record (computed in millisecond) of the both l-distinct and l – recursive technique. Framework of l-recursive in graph 1.

In a general view, the insertion of record having two different background procedure of database. In this checking methodology checker don't have a clue about the substance of client's tuple. In second stage, framework really redesigns the database relies on upon the consequence of the user. Sometimes the insertion or updating fizzled in l recursive database then it holds up until l quality gets to be sure and different tuples fail the insertion.



	2	4	6	8	10
1000	225	210	203	198	188
2000	1990	1654	1446	1440	1064
3000	1522	1499	1144	1129	833
4000	2560	1803	1956	1930	1623
5000	3175	2611	2512	2400	1998

IV. CONCLUSION

In this, we focus on a privacy preserving of l-diversity database. We have presented two secure techniques l –distinct and l recursive for database anonymization techniques for protecting individual's privacy. The Authorized user accessing the database using ABAC policy for better accuracy.

V. REFERENCES

[1] Zahid Pervaiz, Arif Ghafoor, Walid G. Aref, "Precision-Bounded Access Control Using Sliding-Window Query Views for Privacy-Preserving Data Streams", IEEE Trans. Knowl. Data Eng, July 2015.

[2] Z.Pervaiz,W.G.Aref, A.Ghafoor,andN. Prabhu, "Accuracy constrained privacy-preserving access control mechanism for relational data", IEEE Trans. Knowl. Data Eng., April 2014.

[3] T. Ghanem, A. Elmagarmid, P. Larson, and W. Aref, "Supporting views in data stream management systems," ACM Trans. Database Syst., 2010.

[4] J. Cao, B. Carminati, E. Ferrari, and K. Tan, "Castle: Continuously anonymizing data streams," IEEE Trans. Dependable Secure Comput. May/Jun. 2011.

[5] C. Clifton and T. Tassa, "On syntactic anonymity and differential privacy," in Proc. IEEE Int. Conf. Data Eng. Workshop Privacy-Preserving Data Publication Anal., 2013.

[6] B. Zhou, Y. Han, J. Pei, B. Jiang, Y. Tao, and Y. Jia, "Continuous privacy preserving publishing of data streams," in Proc. 12th Int. Conf. Extending Database Technol.: Adv. Database Technol., 2009.

[7] Sai Wu,Xiaoli Wang,Sheng Wang ,Zhenjie Zhang,"k-Anonymity for crowd sourcing Database" IEEE Trans. Knowl. Data Eng, sept 2014.