International Journal of Advance Research in Science and Engineering
Vol. No. 5, Issue No. 07 , July 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

# AN ACCURACY-CONSTRAINED PRIVACY-PRESERVING ACCESS CONTROL FRAMEWORK FOR RELATIONAL DATA

**Keerthana Tumula[1], Lakshmi Bala Ch[2]**

[1]*M.Tech Student, Dept of C.S.E, TPIST/JNTUK, (India)*

[2]*Asst Professor, Dept of C.S.E, TPIST/JNTUK, (India)*

## ABSTRACT

*Access control mechanisms protect sensitive information from unauthorized users. However, when sensitive information is shared and a Privacy Protection Mechanism (PPM) is not in place, an authorized user can still compromise the privacy of a person leading to identity disclosure. A PPM can use suppression and generalization of relational data to anonymize and satisfy privacy requirements, e.g., k-anonymity and l-diversity, against identity and attribute disclosure. However, privacy is achieved at the cost of precision of authorized information. In this paper, we propose an accuracy-constrained privacy-preserving access control framework. The access control policies define selection predicates available to roles while the privacy requirement is to satisfy the k-anonymity or l-diversity. An additional constraint that needs to be satisfied by the PPM is the imprecision bound for each selection predicate. The techniques for workload-aware anonymization for selection predicates have been discussed in the literature. However, to the best of our knowledge, the problem of satisfying the accuracy constraints for multiple roles has not been studied before. In this paper we proposed our formulation of the aforementioned problem, we propose heuristics for anonymization algorithms and show empirically that the proposed approach satisfies imprecision bounds for more permissions and has lower total imprecision than the current state of the art.*

***Keywords: Privacy Protection Mechanism, k-anonymity, l-Diversity, Anonymization algorithms***

## I.INTRODUCTION

Access control mechanisms protect sensitive information from unauthorized users. However, when sensitive information is shared and a Privacy Protection Mechanism (PPM) is not in place, an authorized user can still compromise the privacy of a person leading to identity disclosure. A PPM can use suppression and generalization of relational data to anonymized and satisfies privacy requirements, e.g., k-anonymity and l-diversity, against identity and attribute disclosure. However, privacy is achieved at the cost of precision of authorized information. In this paper, we propose an accuracy-constrained privacy-preserving access control framework. The access control policies define selection predicates available to roles while the privacy requirement is to satisfy the k-anonymity or l-diversity.

However, to the best of our knowledge, the problem of satisfying the accuracy constraints for multiple roles has not been studied before. In our formulation of the aforementioned problem, we propose heuristics for anonymization algorithms and show empirically that the proposed approach satisfies imprecision bounds for more permissions and has lower total imprecision than the current state of the art.The privacy protection mechanism ensures that the privacy and accuracy goals are met before the sensitive data is available to the access control mechanism. The permissions in the access control policy are based on selection predicates on the QI attributes. The policy administrator defines the permissions along with the imprecision bound for each permission/query, user-to-role assignments, and role-to-permission assignments.

### 1.1 Modules

- Access Control Enforcement
- Probabilistic Analysis
- Heuristics for partitioning

**MODULES DESCRIPTION**

**1. Access Control Enforcement**

The exact tuple values in a relation are replaced by the generalized values after the anonymization. In this case, access control enforcement over the generalized data needs to be defined. In this section, we discuss the Relaxed and Strict access control enforcement mechanisms over anonymized data. The access control enforcement by reference monitor can be of the following two types:

**A. Relaxed:** Use overlaps semantics to allow access to all partitions that are overlapping the permission.

**B. Strict:** Use enclosed semantics to allow access to only those partitions that are fully enclosed by the permission.

Both schemes have their own pros and cons. Relaxed enforcement violates the authorization predicate by giving access to extra tuples but is beneficial for applications where low cost of a false alarm is tolerable as compared to the risk associated with a missed event

**2. Probabilistic Analysis**

In this Module, the relaxed enforcement of access control is analyzed probabilistically. The access control policy administrator sets the imprecision bound BQI for each query, and requires that the imprecision bound for the least number of queries be violated by PPM. The policy administrator might revise the imprecision bounds for queries and further relax the access control policy if it is known with a high probability that a large number of queries will violate the bounds and access requests for roles will be denied.

**3. Heuristics for partitioning**

In this we proposed three algorithms based on greedy heuristics are proposed. All three algorithms are based on kd-tree construction. Starting with the whole tuple space the nodes in the kd-tree are recursively divided till the partition size is between kand2k. The leaf nodes of the kd-tree are the output partitions that are mapped to equivalence classes in the given table. Heuristic 1 and 2 have time complexity. Heuristic 3 is a modification over Heuristic 2 to have complexity, which is same as that of TDSM. The proposed query cut can also be used to split partitions using bottom-up ($R^+$-tree) techniques.

## II PROPOSED MODEL

### 2.1 Existing System

The concept of privacy-preservation for sensitive data can require the enforcement of privacy policies or the protection against identity disclosure by satisfying some privacy requirements. Investigate privacy-preservation from the anonymity aspect. The sensitive information, even after the removal of identifying attributes, is still susceptible to linking attacks by the authorized users.

**Disadvantages of Existing System**

- Minimize the imprecision aggregate for all queries.
- The imprecision added to each permission/query in the anonymized micro data is not known.
- Not satisfying accuracy constraints for individual permissions in a policy/workload.

### 2.2 Proposed System

- The heuristics proposed in this paper for accuracy constrained privacy-preserving access control are also relevant in the context of workload-aware anonymization.
- The framework is a combination of access control and privacy protection mechanisms.
- The access control mechanism allows only authorized query predicates on sensitive data.
- The privacy preserving module anonymizes the data to meet privacy requirements and imprecision constraints on predicates set by the access control mechanism.

**Advantages of Proposed System**

- Formulate the accuracy and privacy constraints.
- Concept of accuracy-constrained privacy-preserving access control for relational data.
- Approximate the solution of the k-PIB(k-anonymous Partitioning with Imprecision Bounds) problem and conduct empirical evaluation.

### 2.3 Design Analysis

Software design sits at the technical kernel of the software engineering process and is applied regardless of the development paradigm and area of application. Design is the first step in the development phase for any engineered product or system. The designer's goal is to produce a model or representation of an entity that will later be built. Beginning, once system requirement have been specified and analyzed, system design is the first of the three technical activities -design, code and test that is required to build and verify software. The importance can be stated with a single word "Quality". Design is the place where quality is fostered in software development. Design provides us with representations of software that can assess for quality. Design is the only way that we can accurately translate a customer's view into a finished software product or system. Software design serves as a foundation for all the software engineering steps that follow. Without a strong design we risk building an unstable system – one that will be difficult to test, one whose quality cannot be assessed until the last stage. During design, progressive refinement of data structure, program structure, and procedural details are developed reviewed and documented. System design can be viewed from either technical or project management perspective. From the technical point of view, design is comprised of four activities – architectural design, data structure design, interface design and procedural design.

### 2.4 UML Diagrams

The Unified Modeling Language allows the software engineer to express an analysis model using the modeling notation that is governed by a set of syntactic semantic and pragmatic rules.

A UML system is represented using five different views that describe the system from distinctly different perspective. Each view is defined by a set of diagram, which is as follows.

- **User Model View**
  i. This view represents the system from the users perspective.
  ii. The analysis representation describes a usage scenario from the end-users perspective.
- **Structural model view**
  i. In this model the data and functionality are arrived from inside the system.
  ii. This model view models the static structures.
- **Behavioral Model View**
  It represents the dynamic of behavioral as parts of the system, depicting the interactions of collection between various structural elements described in the user model and structural model view.
- **Implementation Model View**
  In this the structural and behavioral as parts of the system are represented as they are to be built.
- **Environmental Model View**

  In this the structural and behavioral aspects of the environment in which the system is to be implemented are represented.

International Journal of Advance Research in Science and Engineering
Vol. No. 5, Issue No. 07 , July 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

UML is specifically constructed through two different domains they are:

- ✓ UML Analysis modeling, this focuses on the user model and structural model views of the system.
- ✓ UML design modeling, which focuses on the behavioral modeling, implementation modeling and environmental model views.

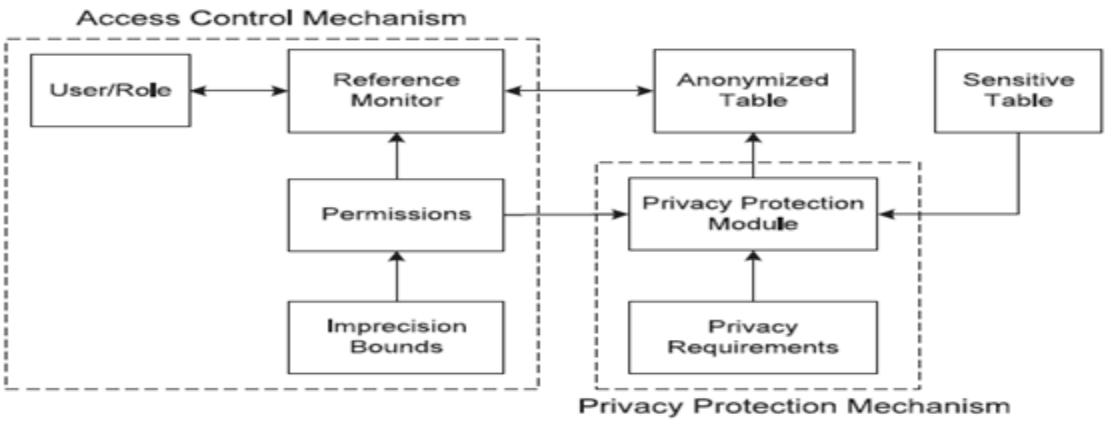


**Figure 2.4.1**

### 3. System Testing

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement

## IV RESULTS


**Fig: 4.1 Home Page**

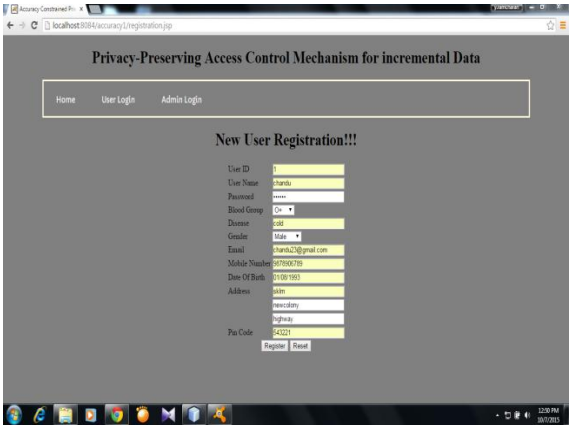


**Fig: 4.2 User Registration**

International Journal of Advance Research in Science and Engineering
Vol. No. 5, Issue No. 07 , July 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

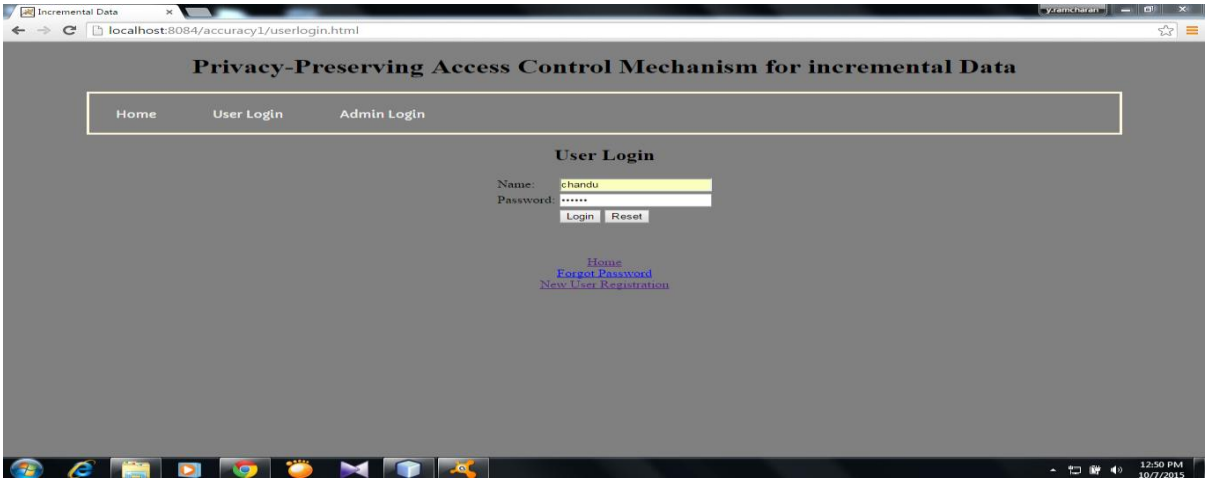


**Fig: 4.3 User Log-In**

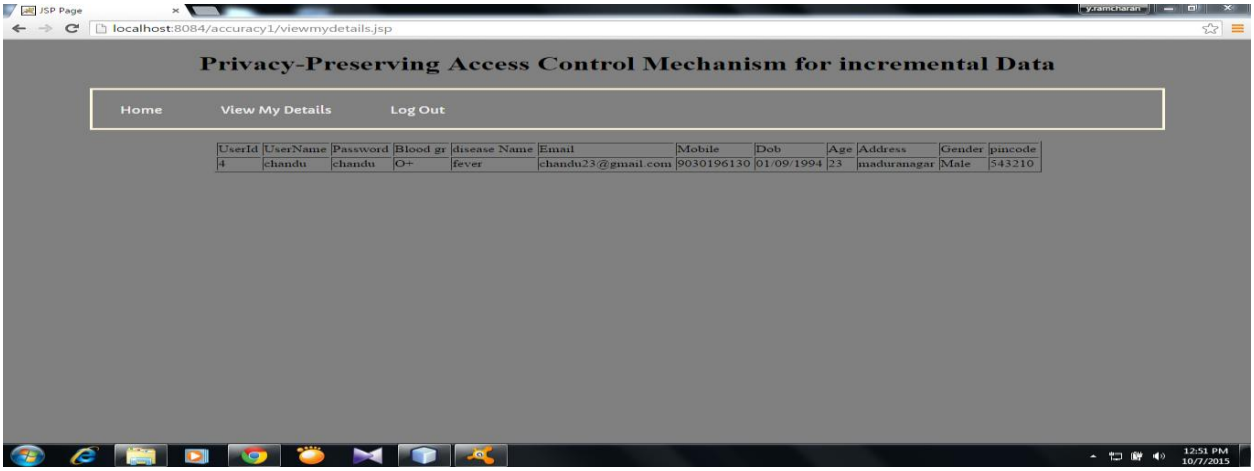


**Fig: 4.4 Checking User Details**

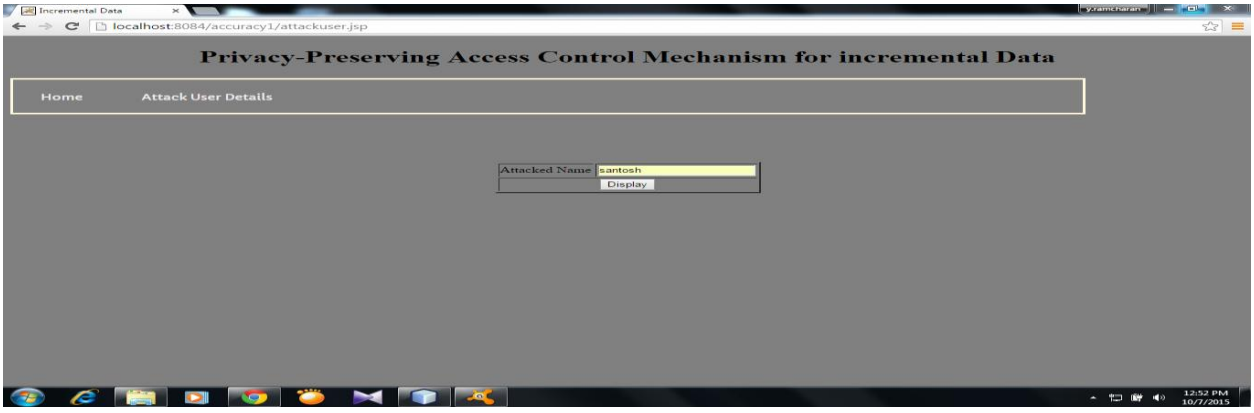


**Fig: 4.5 Attacker Log-In**

International Journal of Advance Research in Science and Engineering
Vol. No. 5, Issue No. 07 , July 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

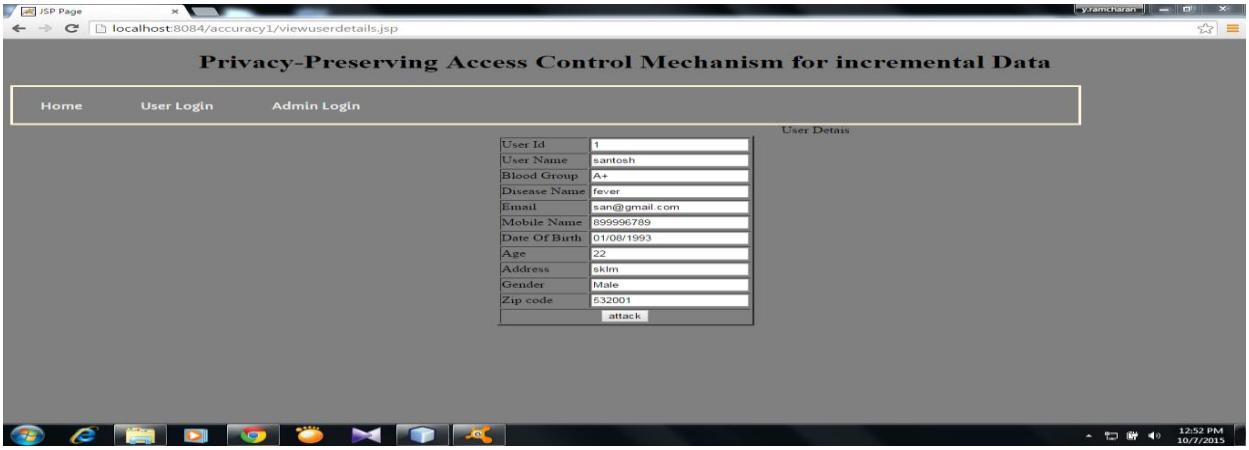


**Fig: 4.6 Attacker Modify Data**

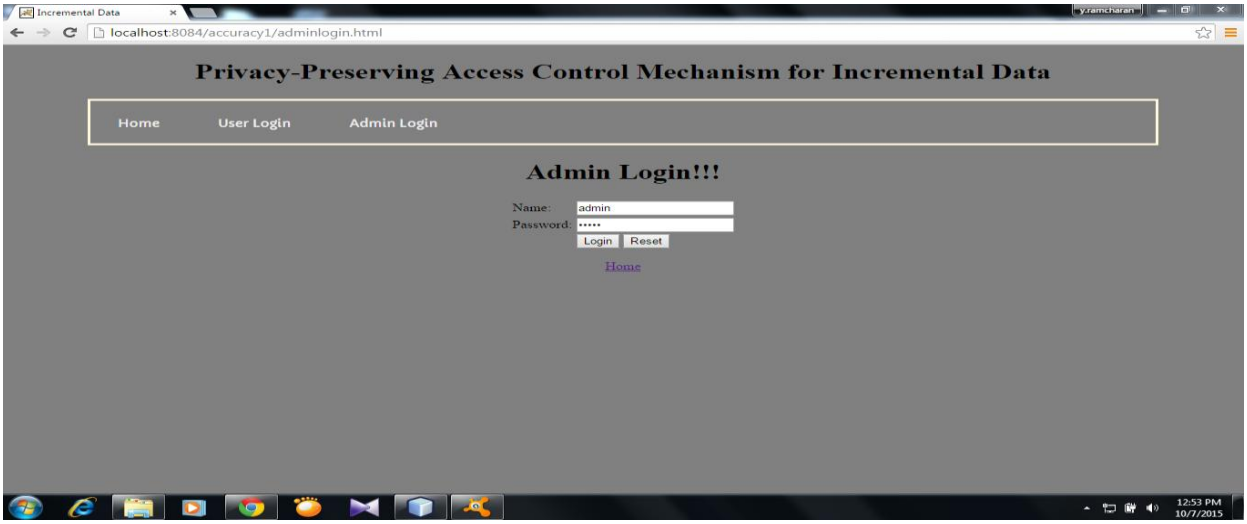


**Fig: 4.7 Admin Log-In**

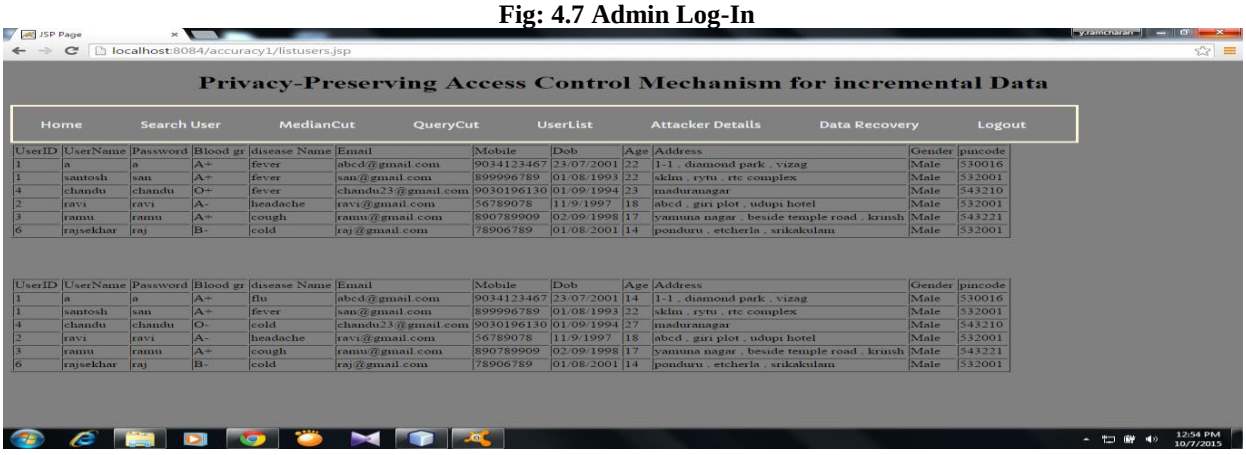


**Fig: 4.8 Updated/Change Details**

International Journal of Advance Research in Science and Engineering
Vol. No. 5, Issue No. 07 , July 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

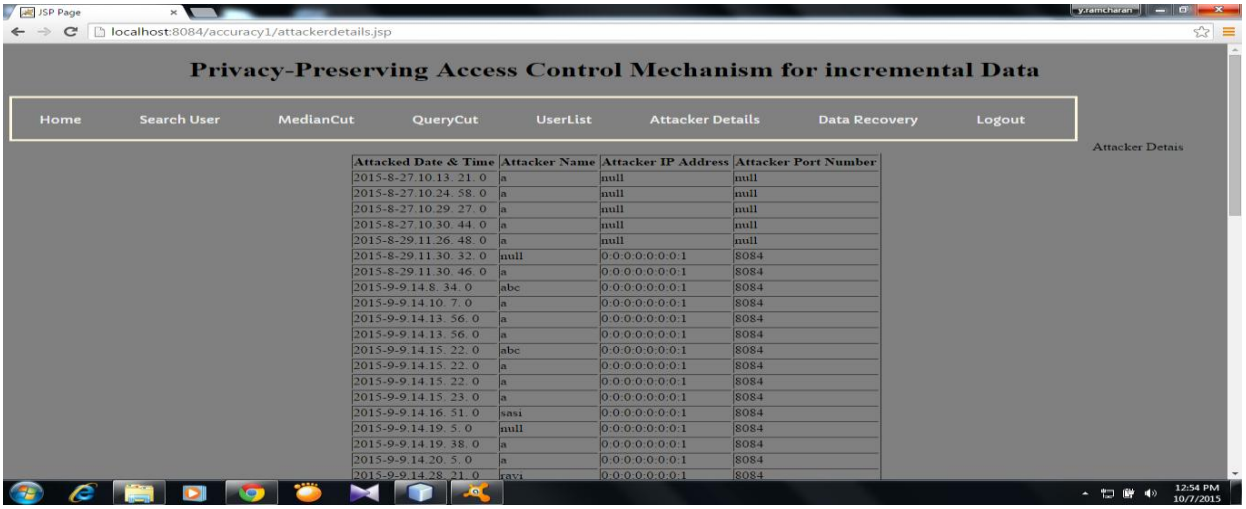


**Fig: 4.9 Attacker Details**

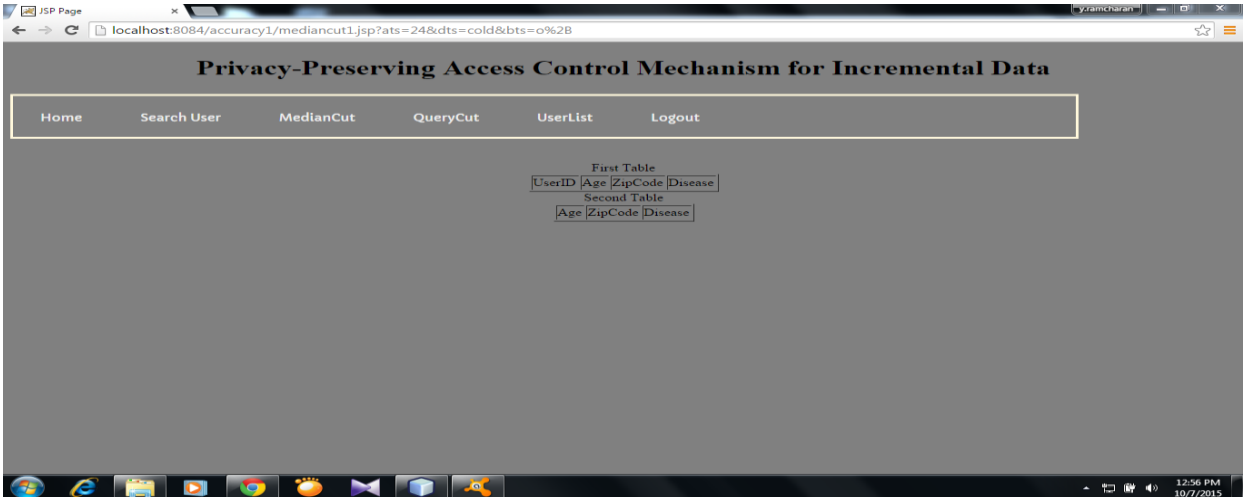


**Fig: 4.10 Data Recovery**

## V CONCLUSION

An accuracy-constrained privacy-preserving access control framework for relational data has been proposed. The framework is a combination of access control and privacy protection mechanisms. The access control mechanism allows only authorized query predicates on sensitive data. The privacy preserving module anonymizes the data to meet privacy requirements and imprecision constraints on predicates set by the access control mechanism. We formulate this interaction as the problem of k-anonymous Partitioning with Imprecision Bounds (k-PIB). We give hardness results for the k-PIB problem and present heuristics for partitioning the data to the satisfy the privacy constraints and the imprecision bounds. In the current work, static access control and relational data model has been

assumed. For future work, we plan to extend the proposed privacy-preserving access control to incremental data and cell level access control.

## REFERENCES

[1] E. Bertino and R. Sandhu, "Database Security-Concepts, Approaches, and Challenges," IEEE Trans. Dependable and Secure Computing, vol. 2, no. 1, pp. 2-19, Jan.-Mar. 2005.

[2] P. Samarati, "Protecting Respondents' Identities in Microdata Release," IEEE Trans. Knowledge and Data Eng., vol. 13, no. 6, pp. 1010-1027, Nov. 2001.

[3] B. Fung, K. Wang, R. Chen, and P. Yu, "Privacy-Preserving Data Publishing: A Survey of Recent Developments," ACM Computing Surveys, vol. 42, no. 4, article 14, 2010.

[4] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-Diversity: Privacy Beyond k-anonymity," ACM Trans. Knowledge Discovery from Data, vol. 1, no. 1, article 3, 2007.

[5] K. LeFevre, D. DeWitt, and R. Ramakrishnan, "Workload-Aware Anonymization Techniques for Large-Scale Datasets," ACM Trans. Database Systems, vol. 33, no. 3, pp. 1-47, 2008.

[6] T. Iwuchukwu and J. Naughton, "K-Anonymization as Spatial Indexing: Toward Scalable and Incremental Anonymization," Proc. 33rd Int'l Conf. Very Large Data Bases, pp. 746-757, 2007.

[7] J. Buehler, A. Sonricker, M. Paladini, P. Soper, and F. Mostashari, "Syndromic Surveillance Practice in the United States: Findings from a Survey of State, Territorial, and Selected Local Health Departments," Advances in Disease Surveillance, vol. 6, no. 3, pp. 1- 20, 2008.

[8] K. Browder and M. Davidson, "The Virtual Private Database in oracle9ir2," Oracle TechnicalWhite Paper, vol. 500, 2002.

[9] A. Rask, D. Rubin, and B. Neumann, "Implementing Row-and Cell-Level Security in Classified Databases Using SQL Server 2005," MS SQL Server Technical Center, 2005.