International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May2014)

Incentive Compatible Privacy-Preserving Data Analysis

Dr.M.V.Siva Prasad
Principal & Professor
Anurag Engg College, Kodad, Nalgonda

G.Srinivas Rao
Associate Professor & HOD of CSE
Anurag Engg College, Kodad, Nalgonda

L.Moulika
M. Tech Student, CSE
Anurag Engg College, Kodad, Nalgonda
moulikalingam@gmail.com

***Abstract*: The competing parties who have private data may collaboratively conduct privacy preserving distributed data analysis (PPDA) tasks to learn beneficial data models or analysis results. For example, different credit card companies may try to build better models for credit card fraud detection through PPDA tasks. Similarly, competing companies in the same industry may try to combine their sales data to build models that may predict the future sales. In many of these cases, the competing parties have different incentives. Although certain PPDA techniques guarantee that nothing other than the final analysis result is revealed, it is impossible to verify whether or not participating parties are truthful about their private input data. In other words, unless proper incentives are set, even current PPDA techniques cannot prevent participating parties from modifying their private inputs. This raises the question of how to design incentive compatible privacy-preserving data analysis techniques that motivate participating parties to provide truthful input data. In this paper, we first develop key theorems, then base on these theorem, we analyze what types of privacy-preserving data analysis tasks could be conducted in a way that telling the truth is the best choice for any participating party.**

***Keywords:* PPDA, SMC, Feasibility, Testing, NCC, DNCC**

1. INTRODUCTION

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

2. PROPOSED SYSTEM:

In design incentive compatible privacy-preserving data analysis techniques that motivate participating parties to provide truthful input data. In this paper, we first develop key theorems, then base on these theorem, we analyze what types of privacy-preserving data analysis tasks could be conducted in a way that telling the truth is the best choice for any participating party. Secure multi-party computation (SMC) has recently emerged as an answer to this problem. Advantages of proposed system are Users give their truth full data for security system, User Only Knows the answers for security questions, Users Knows the Fraud entry, Fraud could be detected.

3. SYSTEM STUDY

3.1. FEASIBILITY STUDY: The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

3.2. ECONOMICAL FEASIBILITY: This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

3.3. TECHNICAL FEASIBILITY: This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

3.4. SOCIAL FEASIBILITY: The aspect of study is to check the level of acceptance of the system by the user. This

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May2014)

includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

4. 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.

5. TYPES OF TESTS

5.1. Unit testing: Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

5.2. Integration testing: Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components

5.3. Functional test: Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.
Invalid Input : identified classes of invalid input must be rejected.
Functions : identified functions must be exercised.
Output : identified classes of application outputs must be exercised.
Systems/Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

5.4. System Test: System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

5.5. White Box Testing: White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

5.6. Black Box Testing: Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

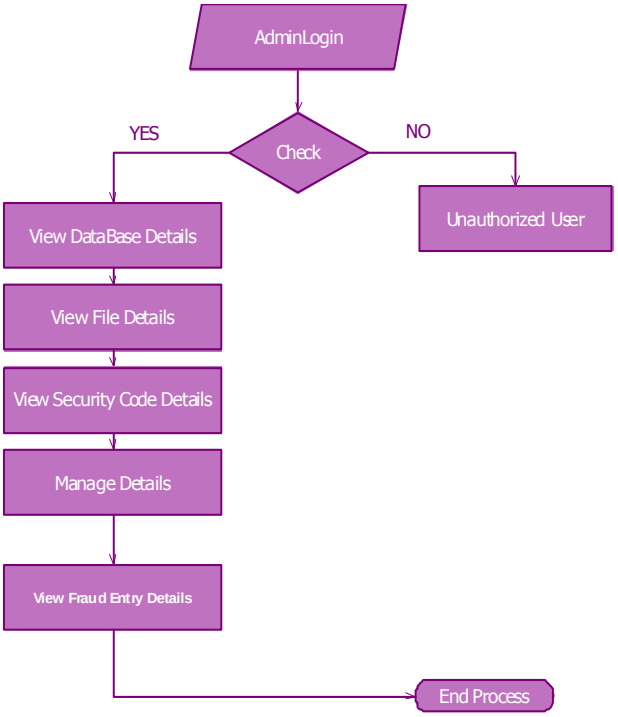
6. SYSTEM DESIGN

Data Flow Diagram: The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.

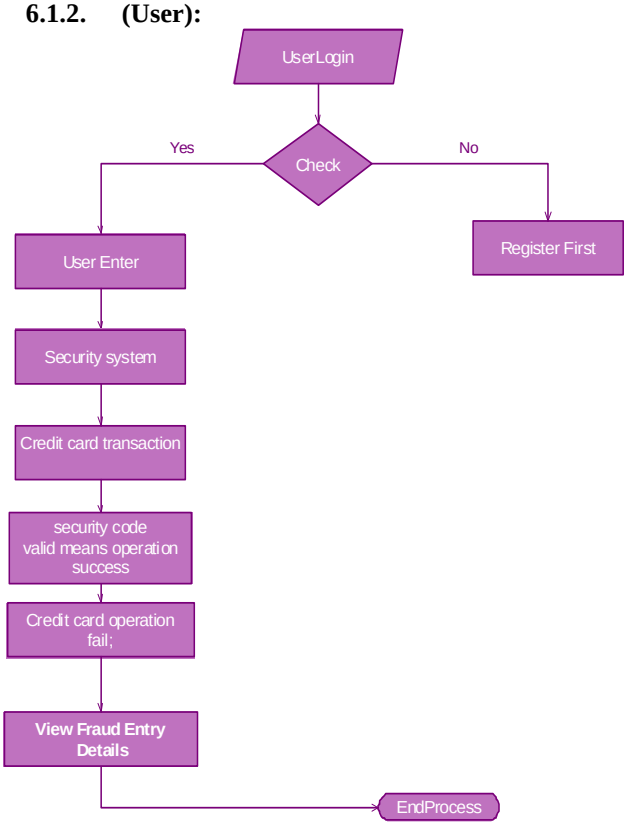
***6.1.* Data Flow Diagram:**

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May2014)

6.1.1. (Admin)

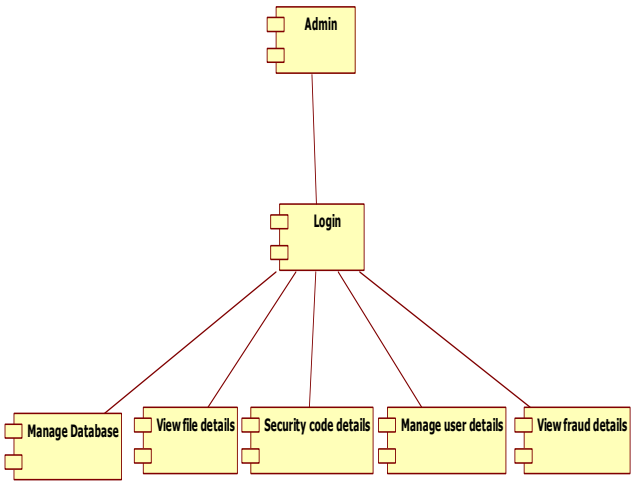


6.1.2. (User):

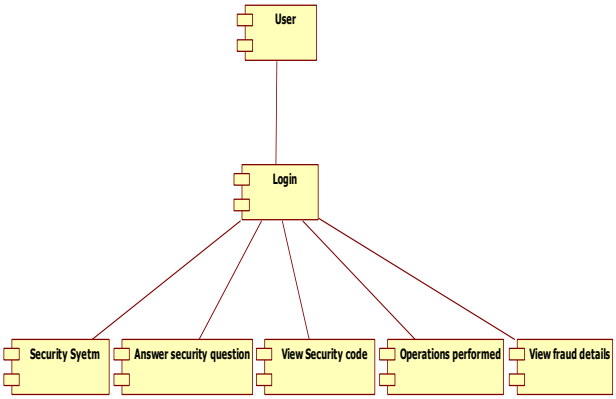


6.2. Component Diagram:

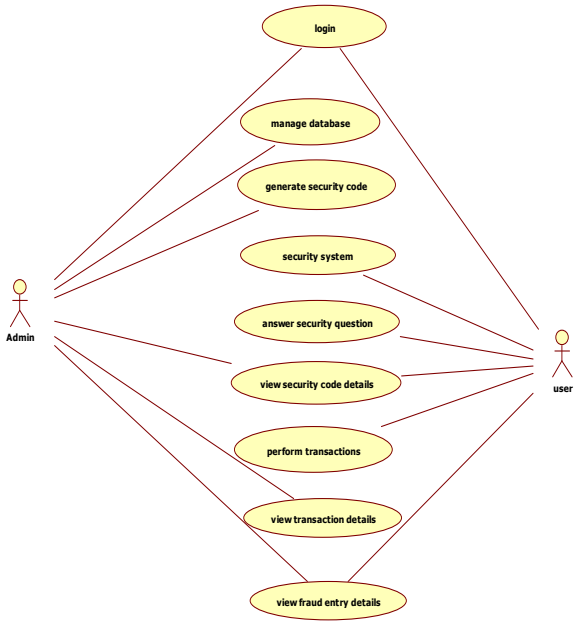
6.2.1. Admin:



6.2.2. User:



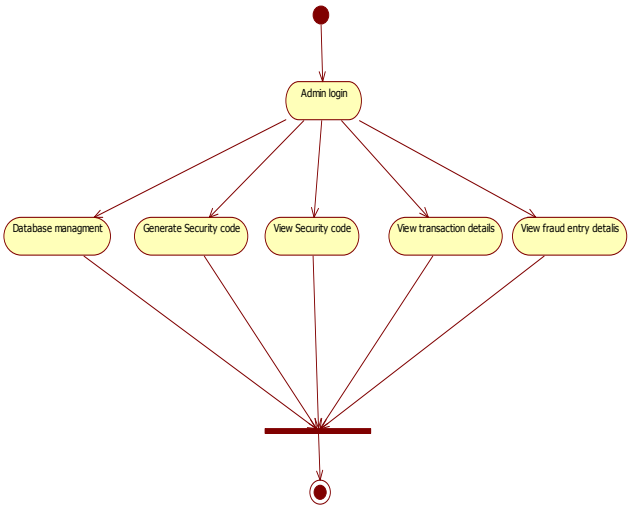
6.3. Use case Diagram:



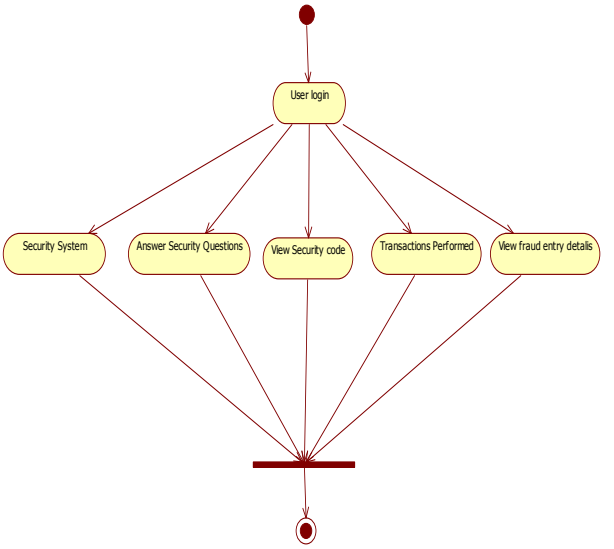
International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May2014)

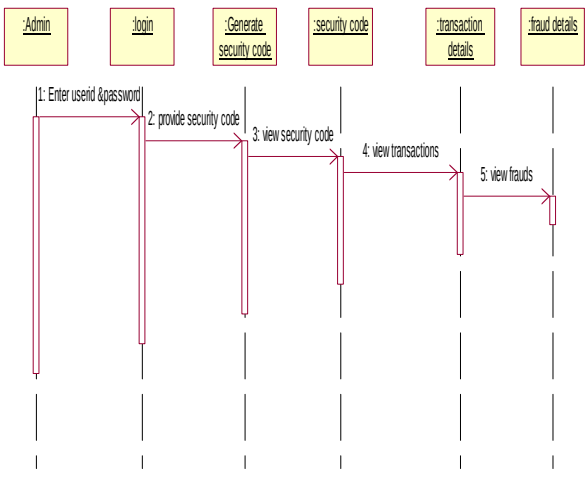
6.4. ACTIVITY DIAGRAM:

6.4.1. (Admin)

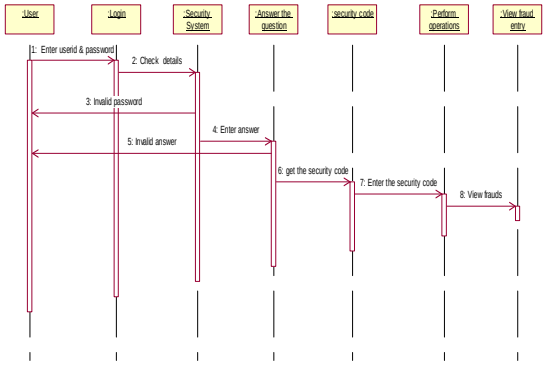


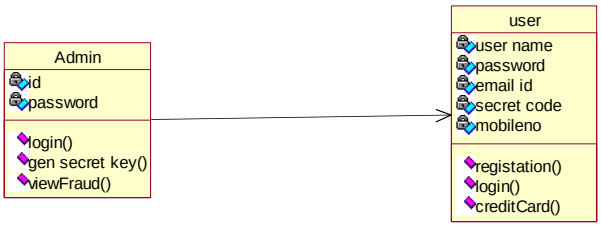
6.4.2. (User)





6.5. Sequence Diagram:





7. CONCLUSION:

Even though privacy-preserving data analysis techniques guarantee that nothing other than the final result is disclosed, whether or not participating parties provide truthful input data cannot be verified. In this paper, we have investigated what kinds of PPDA tasks is incentive compatible under the NCC model. Based on our findings, there are several important PPDA tasks that are incentive driven. As a future work, we will investigate incentive issues in other data analysis tasks, and extend the proposed theorems under the probabilistic NCC model. The PPDA tasks analyzed in the paper can be reduced to evaluation of a single function. Now, the question is how to analyze whether a PPDA task is in DNCC if it is reduced to a set of functions. In other words, is the composition of a set of DNCC functions still in DNCC? We will formally answer this question in the future. Another important direction that we would like to pursue is to create more efficient SMC techniques tailored towards implementing the data analysis tasks that are in DNCC

REFERENCES:

[1]. M. Belkin, P. Niyogi, and V. Sindhwani. Manifold regularization:A geometric framework for learning from labeled and unlabeled examples. Journal of Machine Learning Research, 7(Nov):2399–2434, 2006.

[2]. J. Blitzer, R. Mcdonald, and F. Pereira. Domain adaptation with structural correspondence learning. In EMNLP ’06: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Pro-cessing, pages 120–128. Association for Computational Linguistics,July 2006.

[3]. C. J. C. Burges, R. Ragno, and Q. V. Le. Learning to rank with nonsmooth cost functions. In NIPS '06: Advances in Neural Information Processing Systems, pages 193–200. MIT Press, Cambridge, MA, 2006.
[4]. C. J. C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender. Learning to rank using gradient descent. In ICML '05: Proceedings of the 22th International Conference on Machine Learning, 2005.
[5]. Z. Cao and T. yan Liu. Learning to rank: From pairwise approach to listwise approach. In ICML '07: Proceedings of the 24th Interna-tional Conference on Machine Learning, pages 129–136, 2007.
[6]. J. Cui, F. Wen, and X. Tang. Real time google and live image search re-ranking. In ACM Multimedia, pages 729–732, 2008.
[7]. W. Dai, Q. Yang, G.-R. Xue, and Y. Yu. Boosting for transfer learning. In ICML '07: Proceedings of the 24th international conference on Machine learning, pages 193–200, 2007.
[8]. H. Daume, III and D. Marcu. Domain adaptation for statistical classifiers. Journal of Artficial Intelligence Research, 26:101–126, 2006.
[9]. Y. Freund, R. Iyer, R. E. Schapire, Y. Singer, and G. Dietterich. An efficient boosting algorithm for combining preferences. Journal of Machine Learning Research, 4:933–969, 2003.
[10]. B. Geng, L. Yang, C. Xu, and X.-S. Hua. Ranking model adaptation for domain-specific search. In CIKM '09: Proceeding of the 18th ACM conference on Information and knowledge management, pages 197–206, 2009.

About the Authors:

Dr. M.V. Siva Prasad, Principal of Anurag Engineering College .He received B.E. [CSE] from Gulbarga University, M.Tech. [SE] from VTU, Belgaum and He was awarded Ph.D from Nagarjuna University, Guntur. He published number of papers in International & National journals. He is a Life member of ISTE M. No. : LM 53293 / 2007. His research interests are Information Security, Web Services, Mobile Computing, Data mining and Knowledge.

G.Srinivas Rao Master of Technology [CSE] from JNTU-H. His research interests are Network Security, Web Services, Cloud Computing, Data mining and Knowledge.

L. Moulika pursuing Master of Technology [Computer Science] from JNTU-H, She received B-Tech [CSE] from Swami Ramananda Thirtha Institute of Technology, Nalgonda. Her research interests are Data mining and knowledge, Information Security and Web Services.