



Detailed Investigation on Strategies Developed for Effective Discovery of Matching Dependencies

R.Santhya¹, S.Latha², Prof.S.Balamurugan³, S.Charanyaa⁴

Department of IT, Kalaignar Karunanidhi Institute of Technology, Coimbatore, TamilNadu, India^{1,2,3}

Senior Software Engineer Mainframe Technologies Former, Larsen & Tubro (L&T) Infotech, Chennai, TamilNadu,
India⁴

ABSTRACT: This paper details about various methods prevailing in literature for efficient discovery of matching dependencies. The concept of matching dependencies (MDs) has recently been proposed for specifying matching rules for object identification. Similar to the functional dependencies with conditions, MDs can also be applied to various data quality applications such as detecting the violations of integrity constraints. The problem of discovering similarity constraints for matching dependencies from a given database instance is taken into consideration. This survey would promote a lot of research in the area of information mining.

KEYWORDS: Data Anonymization, Matching Dependencies(MDs), Object, Similarity Constraints, Information Mining.

I. INTRODUCTION

Need for publishing sensitive data to public has grown extravagantly during recent years. Recent days have seen a steep rise in preserving data quality in the database community due to the huge amount of “dirty” data originated from different. These data often contain duplicates, inconsistencies and conflicts, due to various mistakes of men and machines. In addition to the cost of dealing with the huge volume of data, manually detecting and removing “dirty” data is definitely out of practice because human proposed cleaning methods may introduce inconsistencies again. Therefore, data dependencies, which have been widely used in the relational database design to set up the integrity constraints. Hence protecting privacy of individuals and ensuring utility of social network data as well becomes a challenging and interesting research topic.. In this paper we have made an investigation on the attacks by matching dependencies and possible solutions proposed in literature and efficiency of the same.

II. EFFICIENT DISCOVERY OF SIMILARITY CONSTRAINTS FOR MATCHING DEPENDENCIES

The matching dependencies concept is that to providing the certain matching rules for the object identification. The MDs works like the conditional functional dependencies and it is applied to the various application such as identifying the any violation of integrity constraints. In database community the most popular and important thing is given to data quality because of the large amount of the data “Dirty” data from the various resources .These type of the dirty data contains the duplication, conflict, inconsistencies because of the error introduced by the human and mechanism and the large amount of the data manually it will detect and remove the dirty data is only the out of the practice because the humans are well proposed the cleaning method it can also introduced the inconsistencies repeatedly. So the data dependencies which is used in relational database design to construct the integrity constraints. This will used to find the inconsistencies in the given data.

The data base instance which is discover the large volume of matching dependencies if we setting the similarity threshold on attribute. The traditional FD is the one of the special cases of the MDs.The setting of the threshold for MDs is not useful for every time.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

The MDs with confidence is used to achieve high detection accuracy and if the user wants to recall object identification. Then MDs with high supports are used. So the MDs contains both support and confidences.

Contribution: first introduced the support and confidence measures for the utility of the MDs evaluation. Then create the exact algorithms which is used to find the various similarity threshold on attribute setting. These algorithms are travelled all the data while computation. Then propose the approximate solution for that which uses some of the data.

The traditional data dependencies are used for the schema design. These are again and again revisited for the new application privacy preserving. The conditional functional dependencies is an extension of the FD used for the data cleaning. Making the FDs is the major idea of CFDs. This will valid for a certain tuples by conditions. Canonical cover of all FDs. Due to the Inherent hardness of the discovery problem. A series of strategies are used to increase the efficiency. The problem of discovering similarity threshold are not taken in the discovering FDs. Once determined the attributes X and Y in the dependencies than this will implies the equality constraints on the each attributes. So the existing techniques are not to improve the efficiency of the similarity threshold values.

There are different ways to measures the dependencies. In measures of FDs, it may hold on a relation instances. In that the minimal number of the tuples are removed from relation instances. So the most number of the measures are defined equity functions for FDs these are not applicable for MDs with similarity metrics.

Matching dependencies: The traditional FDs to identify the dependency relationship using the equity operator=, but it is not possible to identify the matching over the text values in the real world applications. So the MDs is based on the matching quality. The MDs using the matching operator declare the similarity constraints the operator denoted by till symbol such as edit distance \sim , for text values.

Measures: To evaluate the matching dependencies we can adopt the support and confidence but it is not worth for FDs measures. However the above similarity constraints consider the matching equality of all the tuples t_1 and t_2 for R. While evaluating the measures of MDs, pre compute the pair wise tuple matching offline and store the results for reuse.

The determination of the similarity threshold for MDs with statistical distribution based problem is very new one and it is different from the FDs discovery. The X and Y in the dependencies is given for FD and it implies the similarity threshold using the MD syntax. Matching similarity threshold setting is different from FDs. Statistical distribution is processed from the various relation, similarity threshold for MDs is properly discovered and the support and confidence are satisfied.

“To discover MD, matching dependency with minimum support and confidence, what is Y? and what is matching quality requirements? these two preliminary questions are addressed by the application.”

Pruning strategies: The original algorithm must travel all the statistical distribution and candidate threshold pattern in Ct. Given support and confidence identify the relationship between the similarity threshold and avoid checking.

The result of this experiments demonstrates the pruning and approximation techniques can improve the efficiencies of MDs. The statistical distribution will increase the time cost but EPS can reduce the time costs by pruning candidates.

Two different MDs cover many different dependencies which will leads the problem of generating MDs set. The future work is to more exciting applications of MDa are expected discovery of the similarity constraints for novel dependencies like metric inclusion dependencies, conditional inclusion dependencies and multivalued dependencies.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

III. ON GENERATING NEAR-OPTIMAL TABLEAU FOR CONDITIONAL FUNCTIONAL DEPENDENCIES

In this paper the author characterizes the better pattern tableau, which is based on the support, confidence and parsimony. In this paper the problem is about generating an optimal tableau for a given FD is NP complete and it can be approximated in polynomial time using the greedy algorithm.

In tableau generation problem, the relation instances is the input problem and FDs should not hold exactly on the given data and the FDs knows to hold the data over some patterns. These patterns are given to the input, in that we will neglect the such patterns and assume to make the tableau should be make empty first. Then we will find the parsimonious set of patterns. This will gives the meaningful tableau.

Greedy algorithm states that, the set of candidate patterns with the 'P' elements, local confidence is taken into the account for eliminating the 'P' elements and find the patterns which covers all the possible sets. That includes atleast 'N' tuples from 'R'. This will yield tableau to satisfy the global support threshold. These minimum size tableau will meet the global support and local confidence. This greedy algorithm computes the support and confidence for all the candidate pattern.

Data quality and undocumented semantics are popular issues in the real world. There are many tools introduced to solve this problem such as CFDs. The CFD to capture the semantics of the data and identify the problems. In this paper CFDs is fully realized by defining the tableau pattern, based on the some properties like support, confidence and parsimony and also studying the complexity of the automatic generation of optimal tableau and to providing the approximate algorithm

IV. TEXT JOIN IN AN RDBMS FOR WEB DATA INTEGRATION

In this paper the author describes the problem of web services, data integration challenges. There are lots of global identifiers are presents in the database. So the same entity might represents different textual formats. These same entity refers that matching of strings during data integration

In this paper cosine similarity metric are used for the string matches across web sources. The data integration from heterogeneous web sources is the main important thing for application. Such data is textual string obtained from the web sources. These are all gives the semantics and performance re related challenges. To identify these integration problem, one need to match the multiple textual description.

- Erroneous information
- Abbreviated, Incomplete or missing information
- Differences in information "formatting".

V. COMPRESSION - BASED EVALUATION OF PARTIAL DETERMINATIONS

In this paper, the author determining the problem of partial determination and the compression based method are used to evaluate the above problem. This is viewed as generalizations of both FD and association rules. It will extending the measures used for evaluating the support and confidence.

Partial determinations are generalizations of functional dependencies. It can be expressed as $X \rightarrow dY$. Where d is the number. The set of x will be referred as LHS, and Y will be referred as RHS. The partial determination is used for both the $X \rightarrow dY$ and $pdx \rightarrow dY$.

The future work of this plan is to extending with other strategies like genetic algorithms and combinations of search algorithm. The new compression - based measures are used to evaluate the partial determination and this is used for the search. This partial determination is a useful form of KDD since it is more expressive. The other measures of the partial determination is MDC based functions. This will avoiding the over fitting the data.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

VI. CONCLUSION AND FUTURE WORK

This paper detailed about various methods prevailing in literature for efficient discovery of matching dependencies. The concept of matching dependencies (MDs) has recently been proposed for specifying matching rules for object identification. Similar to the functional dependencies (with conditions), MDs can also be applied to various data quality applications such as detecting the violations of integrity constraints. The problem of discovering similarity constraints for matching dependencies from a given database instance is taken into consideration. This survey would promote a lot of research in the area of information mining.

REFERENCES

1. Shaouxu Song, Lei Chen, "Efficient discovery of similarity constraints for matching dependencies", Data & Knowledge Engineering, Elsevier, 2013.
2. S. Abiteboul, R. Hull, V. Vianu, Foundations of Databases, Addison-Wesley, 1995.
3. R. Agrawal, T. Imielinski, A.N. Swami, Mining association rules between sets of items in large databases, SIGMOD Conference, 1993, pp. 207–216.
4. R. Bassée, J. Wijsen, Neighborhood dependencies for prediction, PAKDD, 2001, pp. 562–567.
5. C. Batini, M. Scannapieco, Data quality: concepts, methodologies and techniques, Data-Centric Systems and Applications, Springer, 2006.
6. L.E. Bertossi, S. Kolahi, L.V.S. Lakshmanan, Data cleaning and query answering with matching dependencies and matching functions, ICDT, 2011, pp. 268–279.
7. M. Bilenko, R.J. Mooney, W.W. Cohen, P. Ravikumar, S.E. Fienberg, Adaptive name matching in information integration, IEEE Intelligent Systems 18 (5) (2003) 16–23.
8. D. Bitton, J. Millman, S. Torgersen, A feasibility and performance study of dependency inference, ICDE, 1989, pp. 635–641.
9. P. Bohannon, W. Fan, F. Geerts, X. Jia, A. Kementsietsidis, Conditional functional dependencies for data cleaning, ICDE, 2007, pp. 746–755.
10. L. Bravo, W. Fan, F. Geerts, S. Ma, Increasing the expressivity of conditional functional dependencies without extra complexity, ICDE, 2008, pp. 516–525.
11. L. Bravo, W. Fan, S. Ma, Extending dependencies with conditions, VLDB, 2007, pp. 243–254.
12. T. Calders, R.T. Ng, J. Wijsen, Searching for dependencies at multiple abstraction levels, ACM Transactions on Database Systems 27 (3) (2002) 229–260.
13. F. Chiang, R.J. Miller, Discovering data quality rules, PVLDB 1 (1) (2008) 1166–1177.
14. W.W. Cohen, Integration of heterogeneous databases without common domains using queries based on textual similarity, SIGMOD Conference, 1998, pp. 201–212.
15. G. Cong, W. Fan, F. Geerts, X. Jia, S. Ma, Improving data quality: consistency and accuracy, VLDB, 2007, pp. 315–326.
16. A.K. Elmagarmid, P.G. Ipeirotis, V.S. Verykios, Duplicate record detection: a survey, IEEE Transactions on Knowledge and Data Engineering 19 (1) (2007) 1–16.
17. W. Fan, Dependencies revisited for improving data quality, PODS, 2008, pp. 159–170.
18. W. Fan, H. Gao, X. Jia, J. Li, S. Ma, Dynamic constraints for record matching, The VLDB Journal (2010) 1–26.
19. W. Fan, F. Geerts, L.V.S. Lakshmanan, M. Xiong, Discovering conditional functional dependencies, ICDE, 2009, pp. 1231–1234.
20. W. Fan, J. Li, X. Jia, S. Ma, Reasoning about record matching rules, PVLDB, 2009.
21. W. Fan, J. Li, S. Ma, N. Tang, W. Yu, Interaction between record matching and data repairing, SIGMOD Conference, 2011, pp. 469–480.
22. W. Fan, S. Ma, Y. Hu, J. Liu, Y. Wu, Propagating functional dependencies with conditions, PVLDB 1 (1) (2008) 391–407.
23. P.A. Flach, I. Sarnik, Database dependency discovery: a machine learning approach, AI Communications 12 (3) (1999) 139–160.
24. J. Gardezi, L.E. Bertossi, I. Kiringa, Matching dependencies with arbitrary attribute values: semantics, query answering and integrity constraints, LID, 2011, pp. 23–30.
25. C. Giannella, E.L. Robertson, On approximation measures for functional dependencies, Information Systems 29 (6) (2004) 483–507.
26. L. Golab, H.J. Karloff, F. Korn, A. Saha, D. Srivastava, Sequential dependencies, PVLDB 2 (1) (2009) 574–585.
27. L. Golab, H.J. Karloff, F. Korn, D. Srivastava, B. Yu, On generating near-optimal tableaux for conditional functional dependencies, PVLDB 1 (1) (2008) 376–390.
28. L. Gravano, P.G. Ipeirotis, N. Koudas, D. Srivastava, Text joins in an rdbms for web data integration, WWW, 2003, pp. 90–101.
29. Y. Huhtala, J. Kärkkäinen, P. Porkka, H. Toivonen, Efficient discovery of functional and approximate dependencies using partitions, ICDE, 1998, pp. 392–401.
30. Y. Huhtala, J. Kärkkäinen, P. Porkka, H. Toivonen, Tane: an efficient algorithm for discovering functional and approximate dependencies, The Computer
31. I.F. Ilyas, V. Markl, P.J. Haas, P. Brown, A. Aboulnaga, Cords: automatic discovery of correlations and soft functional dependencies, SIGMOD Conference, 2004, pp. 647–658.
32. R.S. King, J.J. Legendre, Discovery of functional and approximate functional dependencies in relational databases, JAMDS 7 (1) (2003) 49–59.
33. J. Kivinen, H. Mannila, Approximate inference of functional dependencies from relations, Theoretical Computer Science 149 (1) (1995) 129–149.
34. N. Koudas, A. Saha, D. Srivastava, S. Venkatasubramanian, Metric functional dependencies, ICDE, 2009, pp. 1275–1278.
35. S. Kramer, B. Pfahringer, Efficient search for strong partial determinations, KDD, 1996, pp. 371–374.
36. S.E. Madnick, H. Zhu, Improving data quality through effective use of data semantics, Data & Knowledge Engineering 59 (2) (2006) 460–475.
37. H. Mannila, K.-J. Räihä, Design of Relational Databases, Addison-Wesley, 1992.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015

38. H. Mannila, K.-J. Räihä, Algorithms for inferring functional dependencies from relations, *Data & Knowledge Engineering* 12 (1) (1994) 83–99.
39. A. McCallum, K. Nigam, L.H. Ungar, Efficient clustering of high-dimensional data sets with application to reference matching, *KDD*, 2000, pp. 169–178.
40. G. Navarro, A guided tour to approximate string matching, *ACM Computing Surveys* 33 (1) (2001) 31–88.
41. B. Pfahringer, S. Kramer, Compression-based evaluation of partial determinations, *KDD*, 1995, pp. 234–239.
42. T. Scheffer, Finding association rules that trade support optimally against confidence, *Intelligent Data Analysis* 9 (4) (2005) 381–395.
43. J.C. Schlimmer, Efficiently inducing determinations: a complete and systematic search algorithm that uses optimal pruning, *ICML*, 1993, pp. 284–290.
44. S. Song, L. Chen, Discovering matching dependencies, *CIKM*, 2009, pp. 1421–1424.
45. S. Song, L. Chen, Differential dependencies: reasoning and discovery, *ACM Transactions on Database Systems* 36 (4) (2011).
46. S. Song, L. Chen, P.S. Yu, On data dependencies in dataspaces, *ICDE*, 2011, pp. 470–481.
47. U. ul Hassan, S. O'Riain, E. Curry, Leveraging matching dependencies for guided user feedback in linked data applications, *Proceedings of the Ninth International Workshop on Information Integration on the Web, IIWeb '12*, ACM, New York, NY, USA, 2012, pp. 5:1–5:6.
48. H. Wang, R. Liu, Privacy-preserving publishing microdata with full functional dependencies, *Data & Knowledge Engineering* 70 (3) (2011) 249–268.
49. C.M. Wyss, C. Giannella, E.L. Robertson, Fastfids: a heuristic-driven, depth-first algorithm for mining functional dependencies from relation instances—extended abstract, *DaWaK*, 2001, pp. 101–110.
50. B.Powmeya , Nikita Mary Ablett ,V.Mohanapriya,S.Balamurugan,,"An Object Oriented approach to Model the secure Health care Database systems,"In proceedings of International conference on computer , communication & signal processing(IC³SP)in association with IETE students forum and the society of digital information and wireless communication,SDIWC,2011,pp.2-3
51. Balamurugan Shanmugam, Visalakshi Palaniswami, "Modified Partitioning Algorithm for Privacy Preservation in Microdata Publishing with Full Functional Dependencies", *Australian Journal of Basic and Applied Sciences*, 7(8): pp.316-323, July 2013
52. Balamurugan Shanmugam, Visalakshi Palaniswami, R.Santhya, R.S.Venkatesh "Strategies for Privacy Preserving Publishing of Functionally Dependent Sensitive Data: A State-of-the-Art-Survey", *Australian Journal of Basic and Applied Sciences*, 8(15) September 2014.
53. S.Balamurugan, P.Visalakshi, V.M.Prabhakaran, S.Chranya, S.Sankaranarayanan, "Strategies for Solving the NP-Hard Workflow Scheduling Problems in Cloud Computing Environments", *Australian Journal of Basic and Applied Sciences*, 8(15) October 2014.
54. Charanyaa, S., et. al., , A Survey on Attack Prevention and Handling Strategies in Graph Based Data Anonymization. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(10): 5722-5728, 2013.
55. Charanyaa, S., et. al., Certain Investigations on Approaches for Protecting Graph Privacy in Data Anonymization. *International Journal of Advanced Research in Computer and Communication Engineering*, 1(8): 5722-5728, 2013.
56. Charanyaa, S., et. al., Proposing a Novel Synergized K-Degree L-Diversity T-Closeness Model for Graph Based Data Anonymization. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(3): 3554-3561, 2014.
57. Charanyaa, S., et. al., , Strategies for Knowledge Based Attack Detection in Graphical Data Anonymization. *International Journal of Advanced Research in Computer and Communication Engineering*, 3(2): 5722-5728, 2014.
58. Charanyaa, S., et. al., Term Frequency Based Sequence Generation Algorithm for Graph Based Data Anonymization *International Journal of Innovative Research in Computer and Communication Engineering*, 2(2): 3033-3040, 2014.
59. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Certain Investigations on Strategies for Protecting Medical Data in Cloud", *International Journal of Innovative Research in Computer and Communication Engineering* Vol 2, Issue 10, October 2014
60. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Investigations on Remote Virtual Machine to Secure Lifetime PHR in Cloud ", *International Journal of Innovative Research in Computer and Communication Engineering* Vol 2, Issue 10, October 2014
61. V.M.Prabhakaran, Prof.S.Balamurugan, S.Charanyaa," Privacy Preserving Personal Health Care Data in Cloud" , *International Advanced Research Journal in Science, Engineering and Technology* Vol 1, Issue 2, October 2014
62. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, "Investigations on Evolution of Strategies to Preserve Privacy of Moving Data Objects" *International Journal of Innovative Research in Computer and Communication Engineering*, 2(2): 3033-3040, 2014.
63. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Certain Investigations on Securing Moving Data Objects" *International Journal of Innovative Research in Computer and Communication Engineering*, 2(2): 3033-3040, 2014.
64. P.Andrew, J.Anish Kumar, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Survey on Approaches Developed for Preserving Privacy of Data Objects" *International Advanced Research Journal in Science, Engineering and Technology* Vol 1, Issue 2, October 2014
65. S.Jeevitha, R.Santhya, Prof.S.Balamurugan, S.Charanyaa, " Privacy Preserving Personal Health Care Data in Cloud" *International Advanced Research Journal in Science, Engineering and Technology* Vol 1, Issue 2, October 2014.
66. K.Deepika, P.Andrew, R.Santhya, S.Balamurugan, S.Charanyaa, "Investigations on Methods Evolved for Protecting Sensitive Data", *International Advanced Research Journal in Science, Engineering and Technology* Vol 1, Issue 4, December 2014.
67. K.Deepika, P.Andrew, R.Santhya, S.Balamurugan, S.Charanyaa, "A Survey on Approaches Developed for Data Anonymization", *International Advanced Research Journal in Science, Engineering and Technology* Vol 1, Issue 4, December 2014.
68. S.Balamurugan, S.Charanyaa, "Principles of Social Network Data Security" LAP Verlag, Germany, ISBN: 978-3-659-61207-7, 2014
69. S.Balamurugan, S.Charanyaa, "Principles of Scheduling in Cloud Computing" Scholars' Press, Germany., ISBN: 978-3-639-66950-3, 2014
70. S.Balamurugan, S.Charanyaa, "Principles of Database Security" Scholars' Press, Germany, ISBN: 978-3-639-76030-9, 2014

BIOGRAPHY

R.Santhya and S.Latha are currently pursuing their B.Tech. degree in Information Technology at KalaignarKarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India. Their areas of research interests include Network Security, Cloud Computing and Database Security.



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 1, January 2015



Prof.S.Balamurugan obtained his B.Tech degree in Information Technology from P.S.G. College of Technology, Coimbatore, Tamil Nadu, India and M.Tech degree in Information Technology from Anna University, Tamil Nadu, India respectively. He is currently working towards his PhD degree in Information Technology at P.S.G. College of Technology, Tamil Nadu, India. At present he holds to his credit **65 papers International Journals and IEEE/ Elsevier International Conferences**. He is currently working as Assistant Professor in the Department of Information Technology, Kalaignar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India affiliated to Anna University TamilNadu, India. He is **State Rank holder** in schooling. He was **University First Rank holder** M.Tech. Semester Examinations at Anna University, Tamilnadu, India. He served as a Joint Secretary of IT Association, Department of Information Technology,

PSG College of Technology, Coimbatore, Tamilnadu, India. He is the **recipient of gold medal and certificate of merit** for best journal publication by his host institution **consecutively for 3 years**. Some of his professional activities include invited Session Chair Person for two Conferences. He has guided 16 B.Tech projects and 2 M.Tech. projects. He has won a best paper award in International Conference. His areas of research interest accumulate in the areas of Data Privacy, Database Security, Object Modeling Techniques, and Cloud Computing. He is a life member of ISTE,CSI. **He has authored a chapter in an International Book "Information Processing" published by I.K. International Publishing House Pvt. Ltd, New Delhi, India, 978-81-906942-4-7. He is the author of 3 books titled "Principles of Social Network Data Security", ISBN: 978-3-659-61207-7, "Principles of Scheduling in Cloud Computing" ISBN: 978-3-639-66950-3, and "Principles of Database Security", ISBN: 978-3-639-76030-9.**



S.Charanyaa obtained her **B.Tech** degree in Information Technology and her **M.Tech** degree in Information Technology from Anna University Chennai, Tamil Nadu, India. She was **gold medalist** in her B.Tech. degree program. She has to her credit **27 publications in various International Journals and Conferences**. Some of her outstanding achievements at school level include **School First Rank holder in 10th and 12th grade**. She was working as Software Engineer at Larsen & Turbo Infotech, Chennai for 3 years where she got promoted as Senior Software Engineer and worked for another 2 years. She worked at different verticals and worked at many places including Denmark, Amsderdam handling versatile clients. She is also the recipient of **best team player award for the year 2012 by L&T**. Her areas of research interest accumulate in the areas of

Database Security, Privacy Preserving Database, Object Modeling Techniques, and Cloud Computing. **She is the author of 3 books titled "Principles of Social Network Data Security", ISBN: 978-3-659-61207-7, "Principles of Scheduling in Cloud Computing" ISBN: 978-3-639-66950-3, and "Principles of Database Security", ISBN: 978-3-639-76030-9.**