September 2025: TOP 10 Read Articles in Database Management Systems Research Articles

International Journal of Database Management Systems (IJDMS)

*** WJCI Indexed***

Citations, h-index, i10-index

Citations 3689 h-index 29 i10-index 81

MAPPING COMMON ERRORS IN ENTITY RELATIONSHIP DIAGRAM DESIGN OF NOVICE DESIGNERS

Rami Rashkovits[1] and Ilana Lavy[2]

[1]Department of Management Information Systems, Peres Academic Center, Israel

[2] Department of Information Systems, Yezreel Valley College, Israel

ABSTRACT

Data modeling in the context of database design is a challenging task for any database designer, even more so for novice designers. A proper database schema is a key factor for the success of anyinformation systems, hence conceptual data modeling that yields the database schema is an essential process of the system development. However, novice designers encounter difficulties in understanding and implementing such models. This study aims to identify the difficulties in understanding and implementing data models and explore the origins of these difficulties. This research examines the data model produced by students and maps the errors done by the students. The errors were classified using the SOLO taxonomy. The study also sheds light on the underlying reasons for the errors done during the design of the data model based on interviews conducted with arepresentative group of the study participants. We also suggest ways to improve novice designer'sperformances more effectively, so they can draw more accurate models and make use of advanceddesign constituents such as entity hierarchies, ternary relationships, aggregated entities, and alike. Theresearch findings might enrich the data body research on data model design from
the students' perspectives.

KEYWORDS

Database, Conceptual Data Modelling, Novice Designers

REFERENCES

[1] Moody, D. L. & Shanks, G. G. (1998). "What Makes a Good Data Model? A Framework for Evaluating and Improving the Quality of Entity Relationship Models," Australian Computer Journal, vol. 30, pp. 97-110

[2] Codd, E. F. (1970). A relational model of data for large shared data banks. Communications of the ACM, 13(6), 377-387.

[3] Codd, E. F. (1979). Extending the database relational model to capture more meaning. ACM Transactions on Database Systems (TODS), 4(4), 397-434.

[4] Frederiks, P. J., & Van der Weide, T. P. (2006). Information modeling: The process and the required competencies of its participants. Data & Knowledge Engineering, 58(1), 4-20.

[5] Chen, P. P. S. (1976). The entity-relationship model—toward a unified view of data. ACM Transactions on Database Systems (TODS), 1(1), 9-36.

[6] Teorey, T.J., Yang, D., and Fry, J.F. (1986). A logical design methodology for relational databases using the extended entity-relationship model. Computing Surveys, Vol. 18, No. 2, pp. 197-222.

[7] Ram, S. (1995). "Deriving Functional Dependencies from the Entity Relationship Model," Communications of the ACM. Vol. 38, No. 9, pp. 95-107.

[8] Batra, D. and Antony, S (1994). Novice errors in database design. European Journal of Information Systems, Vol. 3, No. 1, pp. 57-69.

[9] Antony, S. R., & Batra, D. (2002). CODASYS: a consulting tool for novice database designers. ACM Sigmis Database, 33(3), 54-68.

[10] Batra, D. (2007). Cognitive complexity in data modeling: Causes and recommendations. Requirements Engineering 12(4), 231–244.

[11] Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Allyn & Bacon.

[12] Biggs, J. B., & Collis, K. F. (1982). Evaluation the quality of learning: the SOLO taxonomy (structure of the observed learning outcome). Academic Press.

[13] Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. New York: McKay, 20, 24.

[14] Rashkovits, R. & Lavy, I. (2020). Students difficulties in identifying the use of ternary relationships in data modeling. The International Journal of Information and Communication Technology Education (IJICTE), Vol. 16, Issue 2, 47-58.

[15] Balaban, M., & Shoval, P. (1999, November). Resolving the ―weak status‖ of weak entity types in entity-relationship schemas. In International Conference on Conceptual Modeling (pp. 369-383). Springer Berlin Heidelberg.

[16] Or-Bach, R., & Lavy, I. (2004). Cognitive activities of abstraction in object-orientation: An empirical study. The SIGCSE Bulletin, 36(2), 82-85.

[17] Liberman, N., Beeri, C., Ben-David Kolikant, Y., 2011). Difficulties in Learning Inheritance and Polymorphism. ACM Transactions on Computing Education, 11, (1), Article 4, 1-23.

[18] Chick, H. (1998). Cognition in the formal modes: Research mathematics and the SOLO taxonomy. Mathematics Education Research Journal, 10(2), 4-26.

[19] Huang, I. L. (2012). An empirical analysis of students' difficulties on learning conceptual data modeling. Journal of Management Information and Decision Sciences, 15(2), 73.

[20] Leung, F., & Bolloju, N. (2005). Analyzing the quality of domain models developed by novice systems analysts. In System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on (pp. 188b-188b). IEEE.

[21] Kesh, S. (1995). Evaluating the quality of entity relationship models. Information and Software Technology, 37(12), 681-689.

[22] Topi, H., Valacich, J. S., Wright, R. T., Kaiser, K., Nunamaker Jr, J. F., Sipior, J. C., & de Vreede, G. J. (2010). IS (2010): Curriculum guidelines for undergraduate degree programs in information systems. Communications of the Association for Information Systems, 26(1), 18.

[23] Lindland, O. I., Sindre G., and Solvberg A., (1994). "Understanding quality in conceptual modeling," IEEE Software, vol. 11, pp. 42-49.

[24] Taylor, S.J. & Bogdan, R. (1998). Introduction to Qualitative Research Methods. New York: John Wiley & Sons.

[25] Neuendorf, K. A.(2002). The Content Analysis Guidebook. Thousand Oaks, CA: Sage Publications.

[26] Batra, D., & Davis, J. G. (1992). Conceptual data modelling in database design: similarities and differences between expert and novice designers. International journal of man-machine studies, 37(1), 83-101

AUTHORS

Dr Rami Rashkovits is a senior lecturer at Peres Academic Center, head of the Department of Management Information Systems. His PhD dissertation (in the Technion – Israel Institute of Technology) focused on content management in wide-area networks using profiles concerning users' expectations. His research interests are in the fields of distributed content management as well as computer sciences education. He has published over thirty papers and research reports.

Professor Ilana Lavy is an associate professor with tenure at the Academic College of Yezreel Valley. Her PhD dissertation (in the Technion – Israel Institute of Technology) focused on the understanding of basic concepts in elementary number theory. After finishing a doctorate, she was a post-doctoral research fellow at the Education faculty of Haifa University. Her research interests are in the field of pre-service and mathematics teachers' professional development as well as the acquisition and understanding of mathematical and computer science concepts. She has published over a hundred papers and research reports (part of them is in Hebrew).

ALGORITHM FOR RELATIONAL DATABASE NORMALIZATION UP TO 3NF

Moussa Demba

Department of Computer Science & Information, Aljouf University Sakaka, Kingdom of Saudi Arabia

ABSTRACT

When an attempt is made to modify tables that have not been sufficiently normalized undesirable side effects may follow. This can be further specified as an update, insertion or deletion anomaly depending on whether the action that causes the error is a row update, insertion or deletion respectively. If a relation R has more than one key, each key is referred to as a candidate key of R. Most of the practical recent works on database normalization use a restricted definition of normal forms where only the primary key (an arbitrary chosen key) is taken into account and ignoring the rest of candidate keys.

In this paper, we propose an algorithmic approach for database normalization up to third normal form by taking into account all candidate keys, including the primary key. The effectiveness of the proposed approach is evaluated on many real world examples.

KEYWORDS

Relational database, Normalization, Normal forms, functional dependency, redundancy

REFERENCES

[1] Thomas, C., Carolyn, B. (2005) Database Systems, A Practical Approach to Design, Implementation, and Management, Pearson Fourth edition .

[2] Bahmani A., Naghibzadeh, M. and Bahmani, B. (2008) "Automatic database normalization and primary key generation", Niagara Falls Canada IEEE.

[3] Beynon-Davies, P. (2004) Database systems, Palgrave Macmillan, Third edition, ISBN 1–4039—1601–2.

[4] Dongare,Y. V., Dhabe,P. S. and Deshmukh, S. V. (2011) RDBNorma: "A semi-automated tool for relational database schema normalization up to third normal form", International Journal of Database Management Systems, Vol.3, No.1.

[5] Vangipuram, R., Velputa, R., Sravya, V. (2011) "A Web Based Relational database design Tool to Perform Normalization", International Journal of Wisdom Based Computing, Vol.1(3).

[6] Codd, E.F. (1972) "Further normalization of the data base relational model", In Database Systems, Courant Inst. Comptr. Sci. Symp. 6, R. Rustin, Ed., Prentice-Hall, Englewood Cliffs, pp. 33—64.

[7] Elmasri, R., Navathe, S.B. (2003) Fundamentals of Database Systems, Addison Wesley, fourth Edition.

[8] Date, C.J. (2004) Introduction to Database Systems (8th ed.). Boston: Addison-Wesley. ISBN 978-0- 321-19784-9.

[9] Ullman, J.D. (1982) Principe of Database Systems. Computer Science Press, Rockville, Md.

[10] Maier, D. (1983) The Theory of relational Databases. Computer Science Press, Rockville, Md.

[11] Bernstein, P.A. (1976) "Synthesizing third normal form relations from functional dependencies", ACM Transactions on database Systems, vol.1, No.4, pp.277—298.

[12] Diederich, J., Milton, J. (1988) "New Methods and Fast Algorithms for Database Normalization", ACM Transactions on database Systems, Vol.13, No.3, pp. 339—365

INTEGRATION OF A COURSE ENROLMENT AND CLASS TIMETABLE SCHEDULING IN A STUDENT INFORMATION SYSTEM

Vangel V. Ajanovski

Faculty of Computer Science, Saints Cyril and Methodius University, Skopje, Macedonia

ABSTRACT

This paper is dealing with the most important issues of two of the most significant university administrative processes – the process of term enrolment with course enrolment and the process of organization of classes and scheduling a conflict-free class timetable. In the first part of the paper an introduction to the problem is set and the impact and interrelation of the two processes is explained in detail. In the second part of the paper, the focus is on the process of class timetable scheduling, with an analysis of how severe is the risk of having conflicts in the class schedule. Conflicts in the class schedule are rarely introduced due to the lack of resources, in the usual understanding of the term, but are mostly related to the availability human resources. The analysis gives more details on the probability that the students will have a conflict in the timetables among different courses they attend. The third part of the paper discusses some of the formal complexities of the course enrolment process and gives short overview of an earlier effort to overcome such complexities with an introduction of a virtual academic adviser. The virtual adviser components, are then integrated with a recommender system that gives each student personalized course recommendations based on historical data of courses and success. Finally, a solution is given in the form of a model that is currently under development in which a single unifying process is envisioned that covers both earlier processes. The integrated process defines a continuous dialogue between the student groups, enrolment administration and scheduling administration. It is expected from the solution to streamline both processes into a fully integrated system that will improve the student satisfaction from the organization and transparency of the administration.

KEYWORDS

Administrative processes, Course Enrolments, Recommender Systems, Scheduling.

REFERENCES

[1] Ajanovski, Vangel V., (2010) "IS for Students by the Students", Proceedings of the 7th International Conference for Informatics and Information Technology (CIIT 2010), Institute of Informatics, Skopje, Macedonia, pp 129, ISBN 978-9989-668-88-3.

[2] Ajanovski, Vangel V., (2011) "Towards a Virtual Academic Adviser", Proceedings of the 8th International Conference for Informatics and Information Technology (CIIT 2011), Institute of Informatics, Skopje, Macedonia, pp. 146-149. ISBN 978-9989-668-90-6.

[3] O'Mahoney Michael and Smyth Barry (2007), A Recommender System for On-line Course Enrolment:An Initial Study, RecSys'07, Minneapolis, Minnesota, USA, ACM 978-1-59593-730-8/07/0010, pp. 136-137.

[4] Romero Cristobal and Ventura Sebastian, (2010), Educational Data Mining: A Review of the State of the Art, IEEE transactions on systems, man and cybernetics. Part C, Applications and reviews, Vol. 40, No. 6, pp 601-618.

[5] Aher Sunita and Lobo Louis, (2012), Applicability of Data Mining Algorithms for Recommendation System in E-Learning, ICACCI '12, CHENNAI, India, ACM 978-1-4503-1196-0/12/08, pp 1034- 1040.

[6] Easyrec, http://www.easyrec.org/

[7] ISISng, http://develop.finki.ukim.mk/projects/isis

AUTHORS

Vangel V. Ajanovski is a teaching assistant at the Faculty of Computer Science and Engineering. In the period 2005-2009 he was the head of the computing centre and led the development of the information system. He has been working on information system process improvements in student information systems for 20 years.

TOP NEWSQL DATABASES AND FEATURES CLASSIFICATION

Ahmed Almassabi[1], Omar Bawazeer and Salahadin Adam[2]

[1]Department of Computer Science, Najran University, Najran, Saudi Arabia
[2]Department of Information and Computer Science, King Fahad University of Petroleum and Mineral, Dhahran, Saudi Arabia

ABSTRACT

Versatility of NewSQL databases is to achieve low latency constrains as well as to reduce cost commodity nodes. Out work emphasize on how big data is addressed through top NewSQL databases considering their features. This NewSQL databases paper conveys some of the top NewSQL databases [54] features collection considering high demand and usage. First part, around 11 NewSQL databases have been investigated for eliciting, comparing and examining their features so that they might assist to observe high hierarchy of NewSQL databases and to reveal their similarities and their differences. Our taxonomy involves four types categories in terms of how NewSQL databases handle, and process big data considering technologies are offered or supported. Advantages and disadvantages are conveyed in this survey for each of NewSQL databases. At second part, we register our findings based on several categories and aspects: first, by our first taxonomy which sees features characteristics are either functional or non-functional. A second taxonomy moved into another aspect regarding data integrity and data manipulation; we found data features classified based on supervised, semi-supervised, or unsupervised. Third taxonomy was about how diverse each single NewSQL database can deal with different types of databases. Surprisingly, Not only do NewSQL databases process regular (raw) data, but also they are stringent enough to afford diverse type of data such as historical and vertical distributed system, real-time, streaming, and timestamp databases. Thereby we release NewSQL databases are significant enough to survive and associate with other technologies to support other database types such as NoSQL, traditional, distributed system, and semi- relationship to be as our fourth taxonomy-based. We strive to visualize our results for the former categories and the latter using chart graph. Eventually, NewSQL databases motivate us to analyze its big data throughput and we could classify them into good data or bad data. We conclude this paper with couple suggestions in how to manage big data using Predictable Analytics and other techniques.

KEYWORDS

NewSQL, NoSQL, RDBMs. FF, Non-FF, and Big data.

REFERENCES

[1] Ismail, M., Gebremeskel, E., Kakantousis, T., Berthou, G., Dowling, J. (2017, June). Hopsworks: Improving User Experience and Development on Hadoop with Scalable, Strongly Consistent Metadata. In Distributed Comput-ing Systems (ICDCS), 2017 IEEE 37th International Conference on(pp. 2525-2528). IEEE.

[2] Sangtani, M., D'souza, S. M., Harsh, M., Chander, S., Vijaya, P. IN-TERNATIONAL JOURNAL OF ENGINEERING SCIENCES RESEARCH TECHNOLOGY IMPLEMENTATION CHALLENGES INVOLVED IN BIG DATA ANALYTICS.

[3] Kobielus, J. (2012). Hadoop: Nucleus of the next-generation big data ware-house. IBM Data Management Magazine.

[4] Lightstone, S., Ohanian, R., Haide, M., Cho, J., Springgay, M., Steinbach, T. (2017, April). Making Big Data Simple with dashDB Local. In Data Engi-neering (ICDE), 2017 IEEE 33rd International Conference on (pp. 1195-1205). IEEE.

[5] Santos, M. Y., Costa, C., Galv~ao, J., Andrade, C., Martinho, B. A., Lima, F. V., Costa, E. (2017, July). Evaluating SQL-on-hadoop for big data warehous-ing on not-so-good hardware. In Proceedings of the 21st International Database Engineering Applications Symposium (pp. 242-252). ACM.

[6] Ismail, M., Gebremeskel, E., Kakantousis, T., Berthou, G., Dowling, J. (2017, June). Hopsworks: Improving User Experience and Development on Hadoop with Scalable, Strongly Consistent Metadata. In Distributed Com-puting Systems (ICDCS), 2017 IEEE 37th International Conference on (pp. 2525-2528). IEEE.

[7] Thusoo, A., Shao, Z., Anthony, S., Borthakur, D., Jain, N., Sen Sarma, J., Liu, H. (2010, June). Data warehousing and analytics infrastructure at face-book. In Proceedings of the 2010 ACM SIGMOD International Conference on Management of data (pp. 1013-1020). ACM.

[8] Barkhordari, M., Niamanesh, M. (2017). Atrak: a MapReduce-based data warehouse for big data. The Journal of Supercomputing, 1-15.

[9] Tankard, C. (2012). Big data security. Network security, 2012(7), 5-8.

[10] Corbett, J. C., Dean, J., Epstein, M., Fikes, A., Frost, C., Furman, J. J., Hsieh, W. (2013). Spanner: Google's globally distributed database. ACM Transactions on Computer Systems (TOCS), 31(3), 8.

[11] Song, L., Smola, A., Gretton, A., Borgwardt, K. M., Bedo, J. (2007, June). Supervised feature selection via dependence estimation. In Proceedings of the 24th international conference on Machine learning (pp. 823-830). ACM.

[12] Huang, S. H. (2015). Supervised feature selection: A tutorial. Arti cial Intelligence Research, 4(2), 22.

[13] Erturk, E., Jyoti, K. (2015). Perspectives on a Big Data Application: What Database Engineers and IT Students Need to Know. Engineering, Technology Applied Science Research, 5(5), pp-850.

[14] Davenport, T. H., Barth, P., Bean, R. (2012). How big data is di erent. MIT Sloan Management Review, 54(1), 43.

[15] [Book] Iafrate, F. (2015). From big data to smart data (Vol. 1). John Wiley Sons.

[16] Russom, P. (2013). Managing big data. TDWI Best Practices Report, TDWI Research, 1-40.

[17] Cox, M., Ellsworth, D. (1997, August). Managing big data for scienti c visualization. In ACM Siggraph (Vol. 97, pp. 21-38).

[18] https://www.predictiveanalyticstoday.com/voltdb/

[19] https://www.predictiveanalyticstoday.com/tibco-activespaces/

[20] https://www.predictiveanalyticstoday.com/percona-tokudb/

[21] https://www.predictiveanalyticstoday.com/pivotal-gem re-xd/

[22] http://www.actordb.com

[23] http://trafodion.apache.org/

[24] https://www.predictiveanalyticstoday.com/nuodb/

[25] https://www.nuodb.com/product/durable-distributed-cache

[26] https://www.whoishostingthis.com/resources/ansi-sql-standards/

[27] https://www.predictiveanalyticstoday.com/memsql/

[28] https://www.memsql.com/product/

[29] https://www.predictiveanalyticstoday.com/cockroachdb/

[30] https://www.cockroachlabs.com/

[31] https://www.predictiveanalyticstoday.com/clustrixdb/
[32] https://www.clustrix.com
[33] https://www.predictiveanalyticstoday.com/c-treeace/
[34] https://www.faircom.com/products/c-treeace
[35] https://www.predictiveanalyticstoday.com/altibase/
[36] http://altibase.com
[37] http://epubs.siam.org/doi/abs/10.1137/1.9781611972771.75
[38] https://www.tibco.com/products/tibco-activespaces
[39] https://www.predictiveanalyticstoday.com/what-is-predictive-analytics/
[40] http://databasemanagement.wikia.com/wiki/Real-timeDatabase
[41] https://en.wikipedia.org/wiki/Real-time_database
[42] https://aws.amazon.com/streaming-data/
[43] https://businessintelligence.com/dictionary/historical-data/
[44] Zahid, A., Masood, R., & Shibli, M. A. (2014, June). Security of sharded NoSQL databases: A comparative analysis. In Information Assurance and Cyber Security (CIACS), 2014 Conference on (pp. 1-8). IEEE.
[45] P. Kadebu, I. Mapanga, A security requirements perspective towards a secured nosql database environment.
[46] databases2, in: 2011 IEEE Workshop on Electronics, Computer and Applications, Changsha, 2011.
[47] Pavlo, M. Aslett, What's really new with newsql?, ACM Sigmod Record 30 45 (2) (2016) 45–55.
[48] Moniruzzaman, A. B. M. (2014). Newsql: towards next-generation scalable rdbms for online transaction processing (oltp) for big data management. arXiv preprint arXiv:1411.7343.
[49] Aslett, M. (2011). How will the database incumbents respond to NoSQL and NewSQL.San Francisco, The, 451, 1-5.
[50] Stonebraker, M. (2012). Newsql: An alternative to nosql and old sql for new oltp apps. Communications of the ACM. Retrieved, 07-06.
[51] Hoff, T. (2012). Google Spanner's Most Surprising Revelation: NoSQL is Out and NewSQL is In. Retrieved 2012-10-07.
[52] Aslett, M. (2010). What we talk about when we talk about NewSQL.
[53] Jain, H. T. S. C. V. Rise of NewSQL.
[54] https://www.predictiveanalyticstoday.com/newsql-database

EXPERIMENTAL EVALUATION OF NOSQL DATABASES

Veronika Abramova[1], Jorge Bernardino[1,2] and Pedro Furtado[2]

[1] Polytechnic Institute of Coimbra - ISEC / CISUC, Coimbra, Portugal
[2]University of Coimbra – DEI / CISUC, Coimbra, Portugal

ABSTRACT

Relational databases are a technology used universally that enables storage, management and retrieval of varied data schemas. However, execution of requests can become a lengthy and inefficient process for some large databases. Moreover, storing large amounts of data requires servers with larger capacities and scalability capabilities. Relational databases have limitations to deal with scalability for large volumes of data. On the other hand, non-relational database technologies, also known as NoSQL, were developed to better meet the needs of key-value storage of large amounts of records. But there is a large amount of NoSQL candidates, and most have not been compared thoroughly yet. The purpose of this paper is to compare different NoSQL databases, to evaluate their performance according to the typical use for storing and retrieving data. We tested 10 NoSQL databases with Yahoo! Cloud Serving Benchmark using a mix of operations to better understand the capability of non-relational databases for handling different requests, and to understand how performance is affected by each database type and their internal mechanisms

KEYWORDS

NoSQL databases, SQL databases, performance evaluation, database models, YCSB

For More Details : **https://airccse.org/journal/ijdms/papers/6314ijdms01.pdf**

Volume Link : **https://airccse.org/journal/ijdms/current2014.html**

REFERENCES

[1] Fayech, I. and Ounalli, H.: Towards a Flexible Database Interrogation. International Journal of Database Management Systems (IJDMS) Vol.4, No.3, June 2012 .

[2] http://nosql-database.org/. International Journal of Database Management Systems (IJDMS) Vol.6, No.3, June 2014 15

[3] Vimala, S., Khanna Nehemiah, H., Bhuvaneswaran, R. S., and Saranya, G.: Design Methodology for Relational Databases: Issues Related to Ternary Relationships in Entity- relationship Model and Higher Normal Forms.

[4] Stonebraker, M.: SQL databases vs. NoSQL databases. Communications of the ACM, Vol. 53 No. 4, Pages 10-11.

[5] Gajendran, S.: A Survey on NoSQL Databases, 2012, http://ping.sg/story/A-Survey-on-NoSQLDatabases---Department-of-Computer-Science.

[6] Cooper, B., Silberstein, A., Tam, E., Ramakrishnan, R., and Sears, R.: Benchmarking cloud serving systems with YCSB. In Proceedings of the 1st ACM Symposium on Cloud Computing (SoCC '10). ACM, New York, NY, USA, 143-154.

[7] Strozzi, C.: NoSQL – A relational database management system, 2013, http://www.strozzi.it.

[8] Chang, F., Jeffrey, D., Ghemawat, S., Hsieh, W., Wallach, D., Burrows, M., Chandra, T., Fikes, A. and Gruber, R.: Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems, 26(2), Article 4.

[9] Decandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Vosshall, P., and Vogels, W.: Dynamo: amazon's highly available key-value store. In Proceedings of twenty-first ACM SIGOPS Symposium on Operating Systems principles (SOSP '07). ACM, New York, NY, USA, 205-220.

[10] Notes from the NoSQL Meetup, 2013, http://developer.yahoo.com/blogs/ydn/posts/2009/06/nosql_meetup.

[11] Hecht, R. and JABLINSKI, S.: NoSQL Evaluation A Use Case Oriented Survey. Proceedings International Conference on Cloud and Service Computing, pp. 12-14.

[12] Han, J.: Survey on NOSQL Databases. Proceedings 6th International Conference on Pervasive Computing and Applications, pp. 363-366.

[13] Leavitt, N.: Will NoSQL Databases Live up to Their Promise?. Computer Magazine, Vol. 43 No. 2, pp. 12-14.

[14] Floratou, A., Teletia, N., Dewitt, D., Patel, J. and Zhang, D.: Can the elephants handle the NoSQL onslaught?. Proc. VLDB Endow. 5,1712-1723.

[15] Tudorica, B.G. and Bucur, C.: A comparison between several NoSQL databases with comments and notes. Roedunet International Conference (RoEduNet), pp.1-5.

[16] Pritchett, D.: BASE: An Acid Alternative. ACM Queue 6(3), 48-55.

[17] Cook, J. D.: ACID versus BASE for database transactions, 2009, http://www.johndcook.com/blog/2009/07/06/brewer-cap-theorem-base.

[18] Browne, J.: Brewer's CAP Theorem, 2009, http://www.julianbrowne.com/article/viewer/brewers-captheorem.

[19] Indrawan-Santiago, M.: Database Research: Are We at a Crossroad? Reflection on NoSQL. NetworkBased Information Systems (NBiS), 15th International Conference on Network-Based Information Systems, pp.45-51.

[20] Zhang, H. and Tompa, F.W.: Querying XML documents by dynamic shredding. InProceedings of the 2004 ACM symposium on Document engineering (DocEng '04). ACM, New York, NY, USA, 21-30.

[21] Crockford, D.: JavaScript: The Good Parts. Sebastopol, CA: O'Reilly Media.

[22] Lamb, C.: Oracle NoSQL Database in 5 minutes, 2013, https://blogs.oracle.com/charlesLamb/entry/oracle_nosql_database_in_5.

[23] Armstrong, T., Ponnekanti, V., Dhruba, B., and Callaghan, M.: LinkBench: a database benchmark based on the Facebook social graph. In Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data (SIGMOD '13). ACM, New York, NY, USA, 1185-1196.

[24] Dayarathna, M. and Suzumura, T.: XGDBench: A benchmarking platform for graph stores in exascale clouds. Cloud Computing Technology and Science (CloudCom), IEEE 4th International Conference on, Taipei, 363 – 370.

AUTHORS

Veronika Abramova is currently a researcher at CISUC – Centre for Informatics and Systems of the University of Coimbra. Previously she has received a bachelor degree at Instituto Superior Engenharia de Coimbra (ISEC). She is currently working in an industrial project, with an electricity sector company, focused on transferring part of the company's data to the non-relational storage. She has evaluated and studied different NoSQL databases, focusing on their performance comparison and characteristics. Her main research fields are business intelligence, big data, database knowledge management, NoSQL and SQL databases performance evaluation.

Jorge Bernardino received the PhD degree in computer science from the University of Coimbra in 2002. He is a Coordinator Professor at ISEC (Instituto Superior de Engenharia de Coimbra) of the Polytechnic of Coimbra, Portugal. His main research fields are big data, data warehousing, business intelligence, open source tools, and software engineering, subjects in which he has authored or co-authored dozens of papers in refereed conferences and journals. Jorge Bernardino has served on program committees of many conferences and acted as referee for many international conferences and journals. He was President of ISEC from 2005– 2010. Actually, he is serving as General Chair of IDEAS conference and visiting professor at Carnegie Mellon University (CMU).

Pedro Furtado is Professor at University of Coimbra, Portugal, where he teaches courses in both Computer and Biomedical Engineering. His main research interests are data scalability and bigdata, data mining, service-oriented systems and real time systems. Lately, his research has focused both on scalable and real time warehousing, and also on middleware for wireless sensors in industrial and healthcare applications. He has more than 100 papers published in international conferences and journals, books published and several research collaboratio ns with both industry and academia. Besides a PhD in Computer Engineering from University Coimbra (UC) in 2000, Pedro Furtado also holds an MBA from Universidade Catolica Portuguesa (UCP)

HIGH CAPACITY DATA HIDING USING LSB STEGANOGRAPHY AND ENCRYPTION

Shamim Ahmed Laskar[1] and Kattamanchi Hemachandran[2]

Department of Computer ScienceAssam University, Silchar, Assam, India

ABSTRACT

The network provides a method of communication to distribute information to the masses. With the growth of data communication over computer network, the security of information has become a major issue. Steganography and cryptography are two different data hiding techniques. Steganography hides messages inside some other digital media. Cryptography, on the other hand obscures the content of the message. We propose a high capacity data embedding approach by the combination of Steganography and cryptography. In the process a message is first encrypted using transposition cipher method and then the encrypted message is embedded inside an image using LSB insertion method. The combination of these two methods will enhance the security of the data embedded. This combinational methodology will satisfy the requirements such as capacity, security and robustness for secure data transmission over an open channel. A comparative analysis is made to demonstrate the effectiveness of the proposed method by computing Mean square error (MSE) and Peak Signal to Noise Ratio (PSNR). We analyzed the data hiding technique using the image performance parameters like Entropy, Mean and Standard Deviation. The stego images are tested by transmitting them and the embedded data are successfully extracted by the receiver. The main objective in this paper is to provide resistance against visual and statistical attacks as well as high capacity.

KEYWORDS

Steganography, Cryptography, plain text, encryption, decryption, transposition cipher, Least Significant Bit, Human Visual System, Mean square error and Peak Signal to Noise Ratio.

For More Details : https://airccse.org/journal/ijdms/papers/4612ijdms05.pdf

Volume Link : https://airccse.org/journal/ijdms/current2012.html

REFERENCES

[1]   Anderson , R. J. and Petitcolas, F. A.P. (1998) "On The Limits of Steganography", IEEE Journal of Selected Areas in Communications, Vol.16 No.4, pp.474-481, ISSN 0733-8716.

[2]   Petitcolas, F.A.P., Anderson, R. J. and Kuhn, M.G. (1999) "Information Hiding -A Survey", Proceedings of the IEEE, Special issue on Protection of Multimedia Content, vol. 87, no. 7, pp.1062- 1078. International Journal of Database Management Systems (IJDMS) Vol.4, No.6, December 2012 67

[3]   Johnson, N.F. and Jajodia, S. (1998) "Exploring Steganography: Seeing the Unseen", IEEE, Computer, vol. 31, no. 2, pp. 26-34.

[4]   Raphael, A. J. and Sundaram, V. "Cryptography and Steganography – A Survey", Int. J. Comp. Tech. Appl., Vol 2 (3), pp. 626-630 , ISSN:2229-6093.

[5]   Gutte, R. S. and Chincholkar, Y. D. (2012) "Comparison of Steganography at One LSB and Two LSB Positions", International Journal of Computer Applications, Vol.49,no.11, pp.1-7.

[6]   Laskar, S.A. and Hemachandran, K. (2012), "An Analysis of Steganography and Steganalysis Techniques", Assam University Journal of Sscience and Technology, Vol.9, No.II, pp.83-103, ISSN: 0975-2773.

[7]   Younes, M.A.B. and Jantan, A. (2008), "Image Encryption Using Block-Based Transformation Algorithm," International Journal of Computer Science, Vol. 35, Issue.1, pp.15-23.

[8]   Walia, E., Jain, P. and Navdeep. (2010), " An Analysis of LSB & DCT based Steganography", Global Journal of Computer Science and Technology, Vol. 10 Issue 1 , pp 4-8.

[9]   Khare, P., Singh, J. and Tiwari, M. (2011), "Digital Image Steganography", Journal of Engineering Research and Studies, Vol. II, Issue III, pp. 101-104, ISSN:0976-7916.

[10]   Sokouti, M., Sokouti, B. and Pashazadeh, S. (2009), "An approach in improving transposition cipher system", Indian Journal of Science and Technology, Vol.2 No. 8, pp. 9-15, ISSN: 0974-6846.

[11]   Kharrazi, M., Sencar, H. T. and Memon, N. (2006), "Performance study of common image steganography and steganalysis techniques", Journal of Electronic Imaging, SPIE Proceedings Vol. 5681.15(4), 041104 pp.1-16.

[12]   R., Chandramouli, and Nasir Memon.(2001), "Analysis of LSB based image steganography techniques." In Image Processing, 2001. Proceedings. 2001 International Conference on, IEEE, vol. 3, pp. 1019-1022.

[13]   Giddy, J.P. and Safavi- Naini, R. (1994), " Automated Cryptanalysis of Transposition Ciphers", The Computer Journal, Vol.37, No.5, pp. 429-436.

[14]   Johnson, N. F. and Katzenbeisser, S. (2000), "A survey of steganographic techniques", In Information Hiding, Artech House, Norwood, MA, pp. 43-78.

[15]   Chandramouli, R. and Menon, N. (2001), "Analysis of LSB based image steganography techniques", IEEE Proceedings on Image Processing, Vol.3, pp.1019-1022.

[16]   Carvajal-Gamez , B.E., Gallegos-Funes, F. J. and Lopez-Bonilla, J. L. (2009), " Scaling Factor for RGB Images to Steganography Applications", Journal of Vectorial Relativity, Vol. 4, no. 3, pp.55-65.

[17]   Ulutas, G., Ulutas, M. and Nabiyev,V. (2011), "Distortion free geometry based secret image sharing", Elsevier Inc, Procedia Computer Science 3, pp.721–726.

[18]   Tiwari, N. and Shandilya, M. (2010), "Evaluation of Various LSB based Methods of Image Steganography on GIF File Format", International Journal of Computer Applications (0975 – 8887) Vol. 6, no.2, pp.1-4.

[19]   Rabah, K. (2004), "Steganography – The Art of Hiding Data", Information Technology Journal, Vol.3, no.3, pp. 245-269.

[20]   Deshpande, N., Kamalapur, S. and Daisy, J. (2006), "Implementation of LSB steganography and Its Evaluation for Various Bits", 1st International Conference on Digital Information Management, pp.173-178.

[21]   Karen, Bailey, and Kevin Curran.(2006) "An evaluation of image based steganography methods" Multimedia Tools and Applications, Springer Vol.30, no. 1, pp. 55-88.

[22]   Celik, M. U., Sharma, G., Tekalp, A.M. and Saber, E. (2005), "Lossless Generalized-LSB Data Embedding", IEEE Transaction on Image Processing, Vol. 14, No. 2, pp. 253-266.

[23] Huang, Y. S., Huang, Y. P., Huang, K.N. and Young, M. S. (2005), "The Assessment System of Human Visual Spectral Sensitivity Curve by Frequency Modulated Light", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp. 263-265. International Journal of Database Management Systems (IJDMS) Vol.4, No.6, December 2012 68

[24] Chan, Chi-Kwong, and L. M. Cheng. (2004), "Hiding data in images by simple LSB substitution." Pattern Recognition Vol. 37, no. 3, pp. 469-474.

[25] Brisbane, G., Safavi-Naini, R. and Ogunbona, P. 2005. "High-capacity steganography using a shared colour palette", IEEE Proceedings on Vision, Image and Signal Processing, Vol.152, No.6, pp.787- 792.

[26] Curran, K. and Bailey, K. (2003), "An Evaluation of Image Based Steganography Methods", International Journal of Digital Evidence Fall 2003, Volume 2, Issue 2, www.ijde.org.

[27] Dickman, S.D. (2007), "An Overview of Steganography", JMU-INFOSEC-TR-2007-002, http:// citeseerx.ist.psu.edu/viewdoc/summary? doi=10.1.1.137.5129. [28]

[28] Dunbar, B. (2002). "A detailed look at Steganographic Techniques and their use in an Open-Systems Environment", SANS Institute 2002, pp.1-9, http://www.sans.org.

[29] Lee, Y-K. ; Bell, G., Huang, S-Y., Wang, R-Z. and Shyu, S-J. (2009), "An Advanced LeastSignificant-Bit Embedding Scheme for Steganographic Encoding",PSIVT 2009, LNCS 5414, Springer, pp. 349–360.

[30] Smith, C. (2001), "Basic Cryptanalysis Techniques", SANS Institute 2001, GSEC Version 1.2f, http://www.sans.org.

[31] Kaur, R., Singh, B. and Singh, I. (2012), "A Comparative Study of Combination of Different Bit Positions In Image Steganography", International Journal of Modern Engineering Research, Vol.2, Issue.5, pp-3835-3840.

[32] Kruus, P., Caroline, S., Michael, H. and Mathew, M. (2002), "A Survey of Steganographic Techniques for Image Files", Advanced Security Research Journal, Network Associates Laboratories, pp.41-51.

[33] Kharrazi, M., Sencar, H. T. and Memon, N. (2004), "Image Steganography: Concepts and Practice", WSPC/Lecture Notes Series: 9in x 6in, pp.1-31.

[34] B, Li., J, He. and J, Huang. (2011), "A Survey on Image Steganography and Steganalysis", Journal of Information Hiding and Multimedia Signal Processing, Vol. 2, No. 2, pp. 142-172.

[35] Friedman, W.F. (1967), "Cryptology", Encyclopedia Britannica, Vol. 6, pp. 844-851, 1967.

[36] Kahate, A. (2008), "Cryptography and Network Security", 2nd Edition, Tata McGraw-Hill.

[37] Gonzalez, R. C. and Woods, R. E. (2002), "Digital Image Processing", 2nd edition, Prentice Hall, Inc

AUTHORS

Shamim Ahmed Laskar received his B.Sc. and M.Sc. degrees in Computer Science in 2006 and 2008 respectively from Assam University, Silchar, where he is currently doing his Ph.D. His research interest includes Image Processing, Steganography, Information Retrieval and Data Security

Prof. Kattamanchi Hemachandran obtained his M.Sc. Degree from Sri Venkateswara University, Tirupati and M.Tech and Ph.D Degrees from Indian School of Mines, Dhanbad. Presently, he is serving as Head, Department of Computer Science, Assam University, Silchar. He is associated with this department since 1998. He is supervising many research scholars. His areas of research interest are Image Processing, Software Engineering and Dist ributed Computing.

DIAGRAMMATIC APPROACH FOR COMPLETE AUTOMATION OF RELATIONAL DATABASE NORMALIZATION AT CONCEPTUAL LEVEL

Dr. M. S. Patwardhan, P. S. Dhabe, Asavari A. Deshpande, Sandhya G. Londhe, M. L. Dhore, and H. K. Abhyankar

Department of Computer Engineering, Vishwakarma Institute of Technology, Pune-411037, Maharashtra State, India.

ABSTRACT

AER (Articulated Entity Relationship) diagrams proposed by us, are the extension of Entity Relationship (ER) diagrams to accommodate functional dependencies (FDs) as their integral part. Accommodation of FDs in ER diagram helps to achieve total and unconditional automation of relational database normalization. In this paper we have implemented a software Integrated Development Environment (IDE), called AER IDE, designed by extending the Graphical Editing Framework (GEF) and Draw2d plug-in functionality of Eclipse. The AER IDE helps to draw, validate, store and normalize AER diagrams. An AER IDE has five modules. AER Editor allows creation and edition of AER diagrams using Model View controller (MVC) approach. AER diagrams are translated and stored as an xml file. AER Validator provides the facility of AER diagram validation based on a set of proposed rules. AER Normalizer allows normalization of an AER diagram, in one go, with the help of the proposed AER normalization rules and algorithms. Thus, it allows the total and unconditional automation of relational database normalization up to a given normal form; as an integral part of AER IDE. It also serves as a visual aid for the normalization process which is always easy to understand and interpret rather than theoretical approach. AER-XML Bidirectional Translator provides the import and export functionality of AER diagram, to and from an XML template making it compatible with the other toolset. AER-SQL Generator generates DDL scripts in the SQL schema format. The AER IDE is tested with ten distinct AER diagrams with all possible combinations of AER features and validation rules.

KEYWORDS

Entity Relationship Diagram, Automatic Normalization, Database Schema refinement, Relational Database Design.

For More Details : https://airccse.org/journal/ijdms/papers/1110ijdms11.pdf

Volume Link : https://airccse.org/journal/ijdms/curissue10.html

REFERENCES

[1]   Du Hongbo & Wery Laurent, (1999) "Micro: A normalization tool for relational database Designers", Journal of Network and Computer Applications 22, pp. 215-232.

[2]   Chen Peter & Pin-SHAN, (1976). "The entity-relationship model-toward a unified view of data", ACM Trans. on Database Systems, Vol. No. 1, Pages 9-36.

[3]   P.S.Dhabe,M.S.Patwardhan & Asavari Deshpande, (2010) "Articulated Entity Relationship (AER) Diagram For Complete Automation Of Relational Database Normalization", International Journal Of Database Management Systems(IJDMS),Vol.2,No.2.

[4]   Yazici Ali & Karakaya Ziya.,(2007) "JMathNorm: A database normalization tool using mathematica", Springer-Verlag Berlin Heidelberg, pp. 186193

[5]   NABIL ARMAN,(2006). "Normalizer: A case Tool to normalize relational database schemas", Information Technology Journals (2) 329-331 .ISSN 1812-5638.

[6]   Kung Hsiang-Jui.,(2006) "A web-based tool to enhance teaching/learning database normalization",.
Proceedings of the Southern Association for Information Systems Conference 251-258

[7]   Mitovic Antonija,(2002) "NORMIT: Web-Enabled Tutor for Database Normalization", Proceedings of IEEE International Conference on Computer Education(ICCE-02).

[8]   DAWSON KATHYN S.& PARKER LORRAINE M. PURGAILIS,.(1998) "From Entity Relationship Diagram to Fourth Normal Form: A Pictorial Aid to Analysis", The Computer Journal, Volume 31, Number 3

[9]   TAUQEER HUSSAIN*, SHAFAY SHAMAIL, & MIAN M. AWAIS(2003) " Eliminating process of normalization in relational database design", Proceedings IEEE INMIC CODD E. F.,(1970) "A relational model of data for large shared data banks", Communications of the ACM, vol. 13, No.6, pp. 377-387.

[10] CODD E. F.,(1970) "A relational model of data for large shared data banks", Communications of the ACM, vol. 13, No.6, pp. 377-387.

[11] Daum Berthold, (2004). "Professional Eclipse 3 For Java Developers", Wiley Publication

[12] Gamma, E., Kent & B(2004). "Contributing to Eclipse: Principles, Patterns, and Plug-Ins", Addison- Wesley.

[13] Riel Bart van.,(2008) "MVC your graphicswith Eclipse GEF", Internet Document http://www.tutorials.tfo-eservices.eu

[14] Moore Bill, Dean DAVID, GERBER ANNA, WAGENKNECHT GUNNAR, VANDERHEYDEN Philippe. IBM Redbooks _ GEF and EMF,(2004), [online] http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246302.html

[15] ConceptDrawPro 8.0.7.4, http://www.conceptdraw.com/en/products/cd5/main.php.

[16] SmartDraw2010,[online], http://www.smartdraw.com/product/upgrade/.

[17] Edraw Max 5.2, [online] http://www.edrawsoft.com/.

[18] Eclipse ERD, [online] http://eclipse-erd.sourceforge.net/

[19] DDSPro2.21.1,   http://database-design-studio-professional-dds-pro.chilli-source.qarchive.org/.

[20] SHUYUN XU, YU LI & SHIYONG LU. "ERDraw: An XML-based ER-diagram Drawing and Translation Tool", DBLP conference, 143-146.

[21] BAGUI SIKHA & ERAP RICHARD. "Database Design Using Entity-Relationship Diagrams",ISBN: 0849315484, Auerbach Publications

[22] Date C. J,(2002). "An introduction to database systems", Pearson Education, Seventh Edition

[23] FAGIN RONALD. "A normal form for relational databases that is based on domains and keys", ACM Trans. on Database Systems, Vol. 6, No. 3, Pages 387-415

[24] Hsiang-Jui Kung,, (2007) "COMPARING TWO BOTTOM-UP DATABASE DESIGN METHODS", Proceedings of the 2007 Southern Association for Information Systems Conference

AUTHORS

Dr. M. S. Patwardhan (manasi.kelkar@gmail.com) has perused a Ph.D. in Computer Science from University of Tulsa, OK, USA, in year 2007.She is presently working as associate professor in computer engineering at Vishwakarma Institute of technology, Pune from 2009.She was instrumental in playing a role of a researcher in KBSI, USA (2007-2009). Her areas of interest are database normalization, software modelling and computer security and parallel computing.
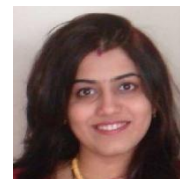
P. S. Dhabe (dhabeps@gmail.com), has completed ME in computer technology from S.G.G.S college of engineering and technology, Nanded, Maharashtra, India in year 2002. He is presently perusing his Ph.D. in systems and control engineering, IIT Mumbai. He is presently working as assistant professor in computer engineering at Vishwakarma Institute of technology, Pune from 2005. He is principal investigator of research project funded by university of pune, Maharashtra, India, titled "Database normalization tool". His areas of interest are database normalization, fuzzy neural networks and pattern recognition.

Asavari A. Deshpande (asavari.deshpande@gmail.com) has perused B.E in Computer Science and Engineering from M.G.M. college of engineering , Nanded, Maharashtra, India in year 2004.She is presently perusing Masters in Computer engineering from 2008.Her area of interest is database normalization.

Sandhya G Londhe (londhesandhya@gmail.com) has perused B.E in Information Technology from Government college of engineering, Karad, Maharashtra, India in year 2006. She is presently perusing Masters in Computer engineering from 2009. Her area of interest is database normalization.

M.L.Dhore (manikrao.dhore@vit.edu) has completed ME in computer engineering from NITR, Chandigarh, India in 1998.Currently he is working as head and associate professor in computer engineering at Vishwakarma Institute of technology, Pune. Presently he is perusing his Ph.D. from University of Solapur, Maharashtra, India, in Web localization. His areas of interest are Webpage Localization and Computer Networking.

H.K.Abhaynkar (director@vit.edu) has completed ME in control systems from Shivaji University, Kolhapur, Mahatashtra, India in 1989.He is presently working as Director and Professor at Vishwakarma Institute of technology, Pune since 1994. He was member of Board of studies at Pune University and Shivaji University, Kolhapur, Mahatashtra ,India. He has awarded by ISTE as a best Engineering college Principal at National Level in 1999.He is well recognized academic leader by various organizations in India and abroad. Recently he has been awarded by PRAJ industries as a "MAHA Intrepreneur" in 2009.

KNN CLASSIFIER AND NAÏVE BAYSE CLASSIFIER FOR CRIME PREDICTION IN SAN FRANCISCO CONTEXT

Noora Abdulrahman and Wala Abedalkhader

Department of Engineering Systems and Management, Masdar Institute of Science and Technology, Abu Dhabi, the United Arab Emirates

ABSTRACT

In this paper we propose an approach for crime prediction and classification using data mining for San Francisco. The approach is comparing two types of classifications: the K-NN classifier and the Naïve Bayes classifier. In the K-NN classifier, two different techniques were performed uniform and inverse. While in the Naïve Bayes, Gaussian, Bernoulli, and Multinomial techniques were tested. Validation and cross validation were used to test the result of each technique. The experimental results show that we can obtain a higher classification accuracy by using multinomial Naïve Bayes using cross validation.

KEYWORDS

Classification, K-NN Classifier, Naïve Bayes, Data Mining, Gaussian, Bernoulli, Multinomial, Uniform, Inverse & Python

REFERENCES

[1] D. K. Tayal, A. Jain, S. Arora, S. Agarwal, T. Gupta, and N. Tyagi, "Crime detection and criminal identification in India using data mining techniques," Ai & Society, vol. 30, pp. 117-127, 2014.

[2] S. V. Nath, "<Crime Pattern Detection Using Data Mining.pdf>."

[3] D. Wang, W. Ding, H. Lo, T. Stepinski, J. Salazar, and M. Morabito, "Crime hotspot mapping using the crime related factors—a spatial data mining approach," Applied Intelligence, vol. 39, pp. 772-781, 2012.

[4] G. C. Oatley and B. W. Ewart, "Crimes analysis software: 'pins in maps', clustering and Bayes net prediction," Expert Systems with Applications, vol. 25, pp. 569-588, 2003.

[5] C.-c. Sun, C.-l. Yao, X. Li, and K. Lee, "<Detecting Crime Types Using Classification Algorithms.pdf>."

[6] T. H. Grubesic, "<Detecting Hot Spots Using Cluster Analysis and GIS.pdf>."

[7] T. Almanie, R. Mirza, and E. Lor, "Crime Prediction Based on Crime Types and Using Spatial and Temporal Criminal Hotspots," International Journal of Data Mining & Knowledge Management Process, vol. 5, pp. 01-19, 2015.

[8] K. Leong and A. Sung, "<A review of spatio-temporal pattern analysis approaches on crime analysis.pdf>."

[9] W. Chang, D. Zeng, and H. Chen, "A stack-based prospective spatio-temporal data analysis approach," Decision Support Systems, vol. 45, pp. 697-713, 2008.

[10] S. Sathyadevan and S. Gangadharan., "<Crime Analysis and Prediction Using Data Mining.pdf>."

[11] M. Sharma, "Z-CRIME: A data mining tool for the detection of suspicious criminal activities based on decision tree," in Data Mining and Intelligent Computing (ICDMIC), 2014 International Conference on, 2014, pp. 1-6.

[12] Malathi and S. S. Baboo, "An enhanced algorithm to predict a future crime using data mining," 2011.

[13] S. Yamuna and N. S. Bhuvaneswari, "Datamining Techniques to Analyze and Predict Crimes," The International Journal of Engineering And Science, vol. 1, pp. 243-247, 2012.

[14] T. Wang, C. Rudin, D. Wagner, and R. Sevieri, "Learning to detect patterns of crime," in Machine Learning and Knowledge Discovery in Databases, ed: Springer, 2013, pp. 515-530.

[15] R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. S. Panahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," Indian Journal of Science and Technology, vol. 6, pp. 4219-4225, 2013.

[16] G. Oatley, J. Zeleznikow, and B. Ewart, "Matching and predicting crimes," Applications and Innovations in Intelligent Systems XII, pp. 19-32, 2005.

[17] C.-H. Yu, M. W. Ward, M. Morabito, and W. Ding, "Crime forecasting using data mining techniques," in Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on, 2011, pp. 779-786.

[18] B. Chandra and M. Gupta, "A novel approach for distance-based semi-supervised clustering using functional link neural network," Soft Computing, vol. 17, pp. 369-379, 2013.

[19] H. Chen, W. Chung, Y. Qin, M. Chau, J. J. Xu, G. Wang, et al., "Crime data mining: an overview and case studies," in Proceedings of the 2003 annual national conference on Digital government research, 2003, pp. 1-5.

[20] S. Chakravorty, "Data mining techniques for analyzing murder related structured and unstructured data," American Journal Of Advanced Computing, vol. 2, 2015.

A Study on Challenges and Opportunities in Master Data Management

Tapan kumar Das[1] and Manas Ranjan Mishra[2]

[1]SITE, VIT University, Vellore, TN, India
[2] IBM India Pvt .Ltd, Bangalore, India

ABSTRACT

This paper aims to provide a data definition of one master data for cross application consistency. The concepts related to Master data management in broader spectrum has been discussed. The current challenges companies are facing while implementing the MDM solutions are outlined. We have taken a case study to highlight why Master Data Management is imperative for the enterprises in optimizing their business Also we have identified some of the long term benefits for the enterprises on implementing MDM.

For More Details : http://airccse.org/journal/ijdms/papers/3211ijdms09.pdf

Volume Link : https://airccse.org/journal/ijdms/current2011.html

REFERENCES

[1]   Berson, A. and Dubov, L. (2007), Master Data Management and Customer Data Integration for aGlobal Enterprise, McGraw-Hill, New York, NY.

[2]   Boyd, M. (2006), "Product information management – forcing the second wave of data quality",available at: ww.thecopywritingpro.com/pages/samples_assets/2nd-wave-DQ.pdf(accessed 27 April 2010).

[3]   Breuer, T. (2009), "Data quality is everyone's business – designing quality into your datawarehouse – part 1", Journal of Direct, Data and Digital Marketing Practice, Vol. 11,pp. 20-9.

[4]   Butler,David., Stackowiak,Bob., "Master Data Management", Oracle Corporation.available at www.oracle.com.

[5]   Dayton, M. (2007), "Strategic MDM: the foundation of enterprise performance management",Cutter IT Journal, Vol. 20 No. 9, pp. 13-17.

[6]   Dreibelbis,Allen , Hechler,Eberhard, Milman ,Ivan(2009), "Enterprise Master Data Management",Pearson Education.

[7]   Dumas, M., Aalst, W. and Ter Hofstede, A. (2005), Process-aware Information Systems: BridgingPeople and Software Through Process Technology, Wiley, Hoboken, NJ.Managing onemaster data161

[8]   Gartner MDM Summit,(2011),UK

[9]   Knolmayer, G. and Ro¨thlin, M. (2006), "Quality of material master data and its effect on theusefulness of distributed ERP systems", Lecture Notes in Computer Science, Vol. 4231,pp. 362-71.

[10]   Lee, Y.W., Pipino, L.L., Funk, J.D. and Wang, R.Y. (2006)," Journey to Data Quality", MIT Press, Cambridge, MA.

[11]   Loshin, D. (2009), Master Data Management, Morgan Kaufmann, Burlington,MA.

[12]   Malcolm, Chisholm(Dec 2010),"The Governance Challenge for Master Data Management", Data Governance Conference, Orlando, Florida

[13]   "MDM Fundamentals and Best practices",www.elearningcurve.com

[14]   Moss, L. (2007), "Critical success factors for master data management", Cutter IT Journal, Vol. 20No. 9, pp. 7-12.

[15]   Rachuri, S., Subrahmanian, E., Bouras, A., Fenves, S., Foufou, S. and Sriram, R. (2008),"Information sharing and exchange in the context of product lifecycle management: role ofstandards", Computer-Aided Design, Vol. 40 No. 7, pp. 789-800.

[16]   Rittman,Mark."Introduction to Master Data Management".www.rittmanmead.com

[17]   Russon,Philip. (2006), Master Data Management – TDWI Best practice Report available at www.tdwi.org.

[18]   Schumachar, Scott (Oct 2010),"MDM: Realizing the same benefits through different implementations", www.initiatesystems.com

[19]   Toronto MDM Summit.(2008),"MDM Challenges and solutions from the real world" available at www.adastracorp.com

[20]   White, A., Newman, D., Logan, D. and Radcliffe, J. (2006), "Mastering master datamanagement", available at: http://kona.kontera.com/IMAGE_DIR/pdf/MDM_gar060125_MasteringMDMB.pdf (accessed 12 April 2010).

[21]   Wolter,Roger.,Haselden,Kirk.(2006),A White paper on MDM, Microsoft Corporation.

[22]   Yang, X., Moore, P.R., Wong, C.-B., Pu, J.-S. and Chong, S.K. (2007), "Product lifecycle information acquisition and management for consumer products", Industrial Management & Data Systems, Vol. 107 No. 7, pp. 936-56.

DESIGN METHODOLOGY FOR RELATIONAL DATABASES: ISSUES RELATED TO TERNARY RELATIONSHIPS IN ENTITY-RELATIONSHIP MODEL AND HIGHER NORMAL FORMS

S Vimala[1], H Khanna Nehemiah[2], R S Bhuvaneswaran[2] and G Saranya[1]

[1]Research Scholar, Ramanujan Computing Centre, Anna University, Chennai.
[2]Associate Professor, Ramanujan Computing Centre, Anna University, Chennai, India.

ABSTRACT

Entity-Relationship (ER) modeling plays a major role in relational database design. The data requirements are conceptualized using an ER model and then transformed to relations. If the requirements are well understood by the designer and then if the ER model is modeled and transformed to relations, the resultant relations will be normalized. However the choice of modeling relationships between entities with appropriate degree and cardinality ratio has a very severe impact on database design. In this paper, we focus on the issues related to modeling binary relationships, ternary relationships, decomposing ternary relationships to binary equivalents and transforming the same to relations. The impact of applying higher normal forms to relations with composite keys is analyzed. We have also proposed a methodology which database designers must follow during each phase of database design.

KEYWORDS

Entity-Relationship Model, Database Design, Binary Relationship, Ternary Relationship, Normalization, Functional Dependencies.

For More Details : http://airccse.org/journal/ijdms/papers/5313ijdms02.pdf

Volume Link : https://airccse.org/journal/ijdms/current2013.html

REFERENCES

[1] P. P. Chen, "The Entity-Relationship Model - Toward a Unified View of Data," ACM Transactions on Database Systems, vol. 1, no.1, pp. 9-36, 1976.

[2] T. J. Teorey, D. Yang and J.P. Fry, "A logical design methodology for relational databases using the extended entity-relationship model," ACM Computing Surveys, vol. 18, no. 2, pp. 197-222, 1986.

[3] T. H. Jones, and Il.-Y. Song, "Binary equivalents of ternary relationships in E-R modeling: A logical decomposition approach," Journal of Database Management, vol. 11, no. 2, pp. 12–19, 2000.

[4] S. Vimala, H. Khanna Nehemiah, G. Saranya and A. Kannan, "Analysis and Modeling of Multi valued Attributes in Entity Relationship Modeling: An Approach for Improved Database Design," (Accepted for publication), CRL Publishing -Computer Systems Science and Engineering.

[5] R. Camps, "From Ternary Relationship to Relational Tables: A Case against Common Beliefs," SIGMOD, vol. 31, no. 2, pp. 46-49, 2002.

[6] A. Badia, "Entity-Relationship Modeling Revisited," SIGMOD, vol. 33, no. 1, pp. 77-82, 2004.

[7] D. B. Bock, "Entity-Relationship Modeling and Normalization Errors," Journal of Database Management, vol. 8, no. 1, pp. 1-9, 1997.

[8] J. Dullea, Il.-Y. Song and I. Lamprou, "An analysis of structural validity in Entity-Relationship modeling," Data and Knowledge Engineering, vol. 47, no. 2, pp. 167-205, 2003.

[9] D. Dey, V. C. Storey and T. M. Barron, "Improving Database Design through the Analysis of Relationships," ACM Transactions on Database Systems, vol. 24, no. 4, pp. 453–486, 1999.

[10] S. Jajodia, A. NG. Peter and F. N. Springsteel, "The Problem of Equivalence for Entity-Relationship Diagrams," IEEE Transaction on Software Engineering, vol. 9, no. 5, pp. 617-630, 1983.

[11] S. Hartmann, "Reasoning about Participation Constraints and Chen's Constraints," in Proceedings of the 14th Australian Database Conference, Conferences in Research and Practice in Information Technology, (ADC2003), Adelaide, Australia Vol. 17(pp. 105-113).

[12] D. Cuadra, A. I. Maqueda, E. Castro and P. M. Fernandez, "Educational Experiences Detecting, Using and Representing Ternary Relationships in Database Design," IEEE Transaction on Education, vol. 53, no. 3, pp. 358–364, 2010.

[13] D. Cuadra, P. Martínez, E. Castro and H. Al-Jumaily, "Guidelines for representing complex cardinality constraints in binary and ternary relationships," Software and System Model, 1-19, 2012.

[14] M. Genero, G. Poels and M. Piattini, "Defining and Validating Metrics for Assessing the Understandability of Entity-Relationship Diagrams," Data & Knowledge Engineering, vol. 64, no. 3, pp. 534–557, 2008.

[15] J. McAllister and D. Sharpe, "An Approach for Decomposing N-ary Data Relationships," Software-Practice and Experience, vol. 28, no. 2, pp. 125–154, 1998.

[16] H. Darwen, C. J. Date and R. Fagin, "A Normal Form for Preventing Redundant Tuples in Relational Databases", ICDT 2012, March 26–30, 2012, Berlin, Germany.

[17] C. J. Date, An Introduction to Database Systems, Addison-Wesley, 8th edition, 2004.

[18] R. Normann, "Minimal Lossless Decompositions and Some Normal Forms between 4NF and PJ/NF", Information Systems, vol. 23, no. 7, pp. 509-516, 1998.

[19] M. W. Vincent, "Redundancy Elimination and A New Normal Form for Relational Database Design," InL. Libkin and B. Thalheim, editors, Semantics in Databases, volume 1358 of Lecture Notes in Computer Science,

[20] pp. 247-264. Springer, 1995.

[21] M. W. Vincent, "A Corrected 5NF Definition for Relational Database Design," Theoretical Computer Science, vol. 185, no. 2, pp. 379-391, 1997.

[22] T. H. Jones, Il-Y. Song and E.K. Park, "Ternary Relationship Decomposition and Higher Normal Form Structures Derived from Entity Relationship Conceptual Modelling," in Proceedings of the 24th ACM annual conference on Computer Science, Philadelphia, PA USA, pp. 96-104, 1996.

AUTHORS

S Vimala completed B.E in Electronics and Communication Engineering in Government College of Technology, Coimbatore, Tamil Nadu, India in year 1978. She completed M.E. in Computer Science and Engineering in P.S.G.Technology, Coimbatore, Tamil Nadu, India in year 1980. During 1980-81 she worked as senior programmer. From 1981 to 1987 she worked as Data Processing Manager in Metro Water, Chennai in the capacity of Executive Engineer. From 1987 to 1999 she worked as Deputy General Manager in Tamil Nadu Cooperative Oilseeds Federation Board. Since June 1999, she is working in Computer Science and Engineering Department, R M K Engineering College, Chennai. She is currently pursuing her Ph.D at Ramanujan Computing Centre, Anna University, Chennai, Tamil Nadu, India. Her areas of interests are Database Management Systems and Software Engineering.

H **Khanna Nehemiah** completed B.E in Computer Science and Engineering from University of Madras in 1997 and M.E in Computer Science and Engineering from University of Madras in 1998. He completed Ph.D from Anna University, Chennai, India in 2007. He is currently working as an Associate Professor at Ramanujan Computing Centre, Anna University, Chennai, Tamil Nadu, India. His research interests are software engineering, data mining, Database Management Systems, Clinical Databases, Bioinfomatics and Image Processing.

R S Bhuvaneswaran completed Master of Technology in Computer Science and Engineering from Pondicherry University, India in 1996 and Ph.D. in Computer Science and Engineering from Anna University, Chennai, India in 2003. He is a post doctoral fellow (2004–2006) of JSPS, Japan. Presently, he is with Anna University as Associate Professor. His research interests include Distributed Systems, Mobile Computing, Network Security and Fault Tolerant Systems.

G Saranya completed Bachelor of Science in Mathematics from University of Madras in 2006 and Master of Computer Application from University of Madras in 2009. She is currently pursuing Ph.D at Ramanujan Computing Centre, Anna University, Chennai, Tamil Nadu, India. Her research interests are Software Engineering and Database Management Systems.