



Hiding Sensitive XML Association Rules With Supervised Learning Technique

PDF (Size: 705KB) PP. 219-229 DOI: 10.4236/iim.2011.36027

Author(s)

Khalid Iqbal, Dr. Sohail Asghar, Dr. Abdulrehman Mirza

ABSTRACT

In the privacy preservation of association rules, sensitivity analysis should be reported after the quantification of items in terms of their occurrence. The traditional methodologies, used for preserving confidentiality of association rules, are based on the assumptions while safeguarding susceptible information rather than recognition of insightful items. Therefore, it is time to go one step ahead in order to remove such assumptions in the protection of responsive information especially in XML association rule mining. Thus, we focus on this central and highly researched area in terms of generating XML association rule mining without arguing on the disclosure risks involvement in such mining process. Hence, we described the identification of susceptible items in order to hide the confidential information through a supervised learning technique. These susceptible items show the high dependency on other items that are measured in terms of statistical significance with Bayesian Network. Thus, we proposed two methodologies based on items probabilistic occurrence and mode of items. Additionally, all this information is modeled and named PPDM (Privacy Preservation in Data Mining) model for XARs. Furthermore, the PPDM model is helpful for sharing markets information among competitors with a lower chance of generating monopoly. Finally, PPDM model introduces great accuracy in computing sensitivity of items and opens new dimensions to the academia for the standardization of such NP-hard problems.

KEYWORDS

XML Document, Association Rules, Bayesian Network, PPDM Model, NP-Hard, K2 Algorithm

Cite this paper

K. Iqbal, D. Asghar and D. Mirza, "Hiding Sensitive XML Association Rules With Supervised Learning Technique," *Intelligent Information Management*, Vol. 3 No. 6, 2011, pp. 219-229. doi: 10.4236/iim.2011.36027.

References

- [1] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, " From Data Mining to Knowledge Discovery in Databases," *AI Magazine*, Vol. 17, No. 3, 1996, pp. 37-54.
- [2] T. Porter, B. Green, " Identifying Diabetic Patients: A Data Mining Approach, Americas Conference on Information Systems," *Proceedings of the Fifteenth Americas Conference on Information Systems*, San Francisco, 2009.
- [3] C. Apte, B. Liu, E. P. D. Pednault and P. Smyth, " Business Applications of Data Mining, Communications of the ACM," Vol. 45, No. 8, 2002, pp. 49-53. doi:10.1145/545151.545178
- [4] L. Chen, T. Sakaguchi and M. N. Frolick, " Data Mining Methods," *Applications and Tools, Information Systems Management*, Vol. 17, No. 1, 2002, pp. 65-70.
- [5] A. Berson, S. Smith and K. Thearling, " Building Data Mining Applications for CRM," *McGraw-Hill Companies*, p. 510.
- [6] T. H. Davenport, J. G. Harris and A. K. Kohli, " How Do They Know Their Customers So Well?" *MIT Sloan Management Review*, Vol. 42, No. 2, pp. 63-73.

- [Open Special Issues](#)
- [Published Special Issues](#)
- [Special Issues Guideline](#)

[IIM Subscription](#)
[Most popular papers in IIM](#)
[About IIM News](#)
[Frequently Asked Questions](#)
[Recommend to Peers](#)
[Recommend to Library](#)
[Contact Us](#)

Downloads:	144,654
------------	---------

Visits:	362,657
---------	---------

[Sponsors >>](#)

- [7] A. Scime, " Web Mining: Applications and Techniques," Idea Group Publishing, 2004, pp. 1-442. doi: 10.4018/978-1-59140-414-9
- [8] W.-M. Ouyang and Q.-H. Huang, " Privacy Preserving Association Rules Mining Based on Secure Two-Party Com-putation," D.-S. Huang, K. Li and G. W. Irwin, Eds., Springer-Verlag Berlin Heidelberg, 2006, pp. 969-975,.
- [9] C. Clifton and D. Marks, " Security and Privacy Implications of Data Mining," Proceedings of ACM Workshop Research Issues in Data Mining and Knowledge Discovery, Montreal, 1996.
- [10] S. R. M. Oliveira and O. Za?ane, " Toward Stan-dardization in Privacy-Preserving Data Mining," Proceedings of the 3rd Workshop on Data Mining Standards, 2004.
- [11] M. Atallah, E. Bertino, A. Elmagarmid, M. Ibrahim and V. Verykios, " Disclosure Limitation of Sensitive Rules," Proceedings of the 1999 Workshop on Knowledge and Data Engineering Exchange, IEEE Computer Society Washington, DC.
- [12] S.-L. Wang, B. Parikh and A. Jafari, " Hiding Informative Association Rule Sets," Expert Systems with Applications: An International Journal, Pergamon Press, Inc., Tarrytown, Vol. 33, No. 2, 2007, pp. 316-323.
- [13] M. Gupta and R. C. Joshi, " Privacy Preserving Fuzzy Association Rules Hiding in Quan-titative Data," International Journal of Computer Theory and Engineering, Vol. 1, No. 4, 2009, pp. 1793-1820.
- [14] Y. Sayg?n, V. S. Verykios and A. K. Elmagarmid, " Privacy Preserving Association Rule Mining," IEEE Computer Society, Washington, DC, 2002, p. 151.
- [15] E. Dasseni, V. S. Verykios, A. K. Elmagarmid and E. Bertino, " Hiding Associa-tion Rules By Using Confidence and Support," IBM, Almaden Research Center, San Jose, 2000.
- [16] Dr. K. Duraiswamy, Dr. D. Manjula and N. Maheswari, " A New Approach to Sensitive Rule Hiding," Computer and Information Journal, Vol. 1, No. 3, 2008, pp. 107-111.
- [17] R. Agrawal, T. Imielinski and A. Swami, " Mining Association Rules between Sets of Items in Large Databases," Proceedings of ACM SIGMOD International Conference on Management of Data, Washington DC, 1993.
- [18] A. F. A. Dafa-Alla, G. Sohn, K. H. Ryu, " Employ-ing PRBAC for Privacy Preserving Data Publishing," Seoul, 2009.
- [19] C.-C. Weng, S.-T. Chen and H.-C. Lo, " A Novel Algorithm for Completely Hiding Sensitive Association Rules," Eighth International Conference on Intelligent Systems Design and Applications, 2008, Vol. 3, pp. 202-208. doi:10.1109/ISDA
- [20] Y.-H. Guo, " Reconstruction-Based Association Rule Hiding," Proceedings of SIGMOD2007 Ph.D. Workshop on Innovative Database Research, Beijing, 2007.
- [21] Y. H. Guo, Y. H. Tong, S. W. Tang and D. Q. Yang, " A FPtree-Based Method for Inverse Frequent Set Mining," Proceedings of the 23Prd P British National Conference on Databases, Springer-Verlag 2006, pp. 152-163.
- [22] J. W. Han, J. Pei and Y. W. Yin, " Mining Frequent Patterns without Candidate Generation," C. Weidong and F. Jeffrey, Eds., Proceed-ings of the ACM SIGMOD Conference on Management of Data, Dallas, ACM Press, 2000, pp. 1-12. doi: 10.1145/342009.335372
- [23] R. R. Rajalaxmi and A. M. Natarajan, " Hybrid Conflict Ratio for Hiding Sensitive Patterns with Minimum Information Loss," International Journal of Computer Theory and Engineering, Vol. 1, No. 4, 2009, pp. 1793-8201.
- [24] G. V. Krishna and P. R. Krishna, " A Novel Approach for Statistical and Fuzzy Association Rule Mining on Quantitative Data," Journal of Scientific and Industrial Research, Vol. 67, 2008, pp. 512-517.
- [25] R. Srikant and R. Agrawal, " Mining Quantitative Association Rules in Large Relational Tables," Proceedings of ACM SIGMOD, New York, 1996, pp. 1-12.
- [26] G.-X. Wu, " A Study on the Mining Algo-rithm of Fast Association Rules for the XML Data," Interna-tional Conference on Computer Science and Information Technology, 2008. doi:10.1109/ICCSIT.2008.89
- [27] A. Abazeed, A. Mamat and M. Nasir, " Hamidah Ibrahim, " Mining Association Rules from Structured

XML Data," 2009 International Conference on Electrical Engineering and Informatics, Selangor, 2009. doi:10.1109/ICEEI.2009.5254708

- [28] C. Combi, B. Olboni and R. Rossato, " Querying XML Documents by Using Association Rules," Proceedings of the 16th International Workshop on Database and Expert Systems Applications, 2005.
- [29] Y.-J. Bei, G. Chen, L.-H. Yu, F. Shao and J.-X. Dong, " XML Query Recommendation Based On Association Rules," Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007. doi:10.1109/SNPD.2007.378
- [30] X.-W. Wang and C.-J. Cao, " Mining Association Rules from Complex and Irregular XML Documents Using XSLT and Xquery," International Conference on Advanced Language Processing and Web Information Technology, 2008. doi:10.1109/ALPIT.2008.48
- [31] X.-Y. LI, J.-S. YUAN, Y.-H. KONG, " Mining Association Rules from XML Data with Index Table," Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 2007.
- [32] J. Shin, J. Paik and U. Kim, " Mining Association Rules from a Collection of XML Documents Using