Abstract

Today’s world of universal data exchange, there is a need to manage the risk of unintended information disclosure. Publishing the data about the individuals, without revealing sensitive information about them is an important problem. K-anonymization is the popular approach used for data publishing. The limitations of K-anonymity were overcome by methods like L-diversity, T-closeness, (alpha, K) anonymity; but all of these methods focus on universal approach that exerts the same amount of privacy preservation for all persons against linking attack, which result in high loss of information. Privacy was also not guaranteed 100% because of proximity and divergence attack. Our approach is to design micro data sanitization technique to preserve privacy against proximity and divergence attack and also to preserve the utility of the data for any type of mining task. The proposed approach, apply a graded grouping transformation on numerical sensitive attribute and a mapping table based transformation on categorical sensitive attribute. We conduct experiments on adult data set and compare the results of original and transformed table to show that the proposed task independent technique preserves privacy, information and utility.

Reference

Privacy and Utility Preserving Task Independent Data Mining

Mining Algorithms”, ACM PODS Conference, 2002
[26] R.C. W. Wong, J. Li, A. W. C. Fu, K. Wang “(alpha,k)-anonymity: an enhanced k-anonymity model for privacy preserving data publishing”, ACM SIGKDD, pp.139-150,2006
[27] Xiaokui xiao, Yufei Tao, “Personalized Privacy Preservation”, SIGMOD, 2006

Index Terms

Computer Science

Data Mining

Key words

Anonymization

Data Publishing

Data utility

Privacy management

micro data sanitization