A Study of Correlation Impact on Privacy Preserving Data Mining

Foundation of Computer Science (FCS), NY, USA

Volume 129
Number 15

Year of Publication: 2015

Authors:
J. Hyma, P.V.G.D. Prasad Reddy, A. Damodaram

10.5120/ijca2015907152

Abstract

Data sharing is obvious in present day scenario of digital world, and when data is being shared among various application areas the sensitive data of the individuals is disclosed to the public. An evident awareness about this privacy violation has been created among the people now when compared to the earlier days and they are also showing a real concern towards their privacy in the technology enabled digital world. At one end several studies have been proved that privacy is a primary concern and also suggesting not to disclose too much of individual information, but at the other end people are disclosing their personal information knowingly or unknowingly through online surveys, social networks, online shopping sites, e-commerce, government agencies etc. This information sharing is obvious and it can’t be unavoidable. Consequently several techniques have been proposed to protect privacy of the individual disclosed information, but still there is an immense need of new privacy preserving techniques that can equally accommodate with the proportionate expansion of the digital data. Existing privacy techniques applied on the data set assuming all the records are independently sampled, where as in the real world data set the correlations among the records is obvious and needs to
be studied to achieve accurate privacy protection. This paper provides an overview of the development of privacy preserving models and the further enhancements to be carried out to accommodate with the diverse privacy requirements and data utilization along with the correlation study.

References

2. LjiljanaBrankovic and Vladimir Estivill-Castro, "Privacy Issues In Knowledge Discovery and Data Mining", In Proceedings of Australian Institute of Computer Ethics Conference, Melbourne, Victoria, Australia, July 1999

Index Terms

Computer Science Information Sciences
Keywords

Data Mining, Privacy Preserving Data Mining (PPDM), Correlation, Correlation constraints.