Efficient and Secure Mining for Vertically Distributed Database

By

Krishna J. Ray
Enrollment No.:140370702526

Guided by:
Sheetal Thakar
Asst. Prof, CSE Dept.
Parul Institute of Engineering & Technology,
P.O: Limda, Ta.: Waghodia, Dist.: Vadodara

A **Thesis Report**
Submitted to
Gujarat Technological University
In Partial Fulfillment of the Requirements for
The Degree of Master of Engineering
In **Computer Engineering**

May – 2016

Computer Science and Engineering Department,
Parul Institute of Engineering & Technology
P.O: Limda, Ta.: Waghodia, Dist.: Vadodara

Efficient and Secure Mining for Vertically Distributed Database

Submitted by

Krishna Ray

Supervised by

Ms. Sheetal Thakar

M.Tech (ICT)

Assistance Professor of CSE Dept.

Parul Institute of Engineering & Technology, Limda

ABSTRACT

Privacy preserving is most popular study for the research field. Privacy means gives the protection to the private information at preserving time. In Market place, discovery of frequent item sets using association rules mining is one of most important tasks in mining. Association rules is very helpful for finding frequent item sets to predicate about which item sets purchased together in a market and generate qualitative information that is useful for decision making. For distributed environment, database may be distributed as horizontally, vertically or mixed in computer network. The main problem in secure mining with the help of association rules is that transactions are distributed as vertically and the various sites want to find frequent item sets by participating themselves without discovering their individual data. The proposed method will find frequent item sets for vertical distributed database with the help of data miner using encryption based technique. Each sites prepare matrix with the local frequent item sets as per minimum support and encrypted it than send to other sites. The Scalar product with Boolean matrix is used for finding frequent item sets with secure computation between multiple sites without disclosing private input which improved efficiency and privacy of system.