SAEED SAMET, University of Ottawa, SITE, 800 King Edward Avenue Ottawa, Ontario, K1N 6N5, Canada

Privacy-Preserving Data Mining

Although, data is very valuable in every organization, it must be processed in order to be useful. Data mining is a collection of techniques which find patterns and associations in raw data, classify or cluster the items according to their attributes. Nowadays, related data is normally distributed among two or more parties in different configurations, and mining can be done in an accurate and useful way for all parties involved, on all data collections. However, privacy is often crucial in different scenarios, and organizations involved do not want to disclose their own private information to each other. Therefore, standard algorithms of data mining must be modified, or new protocols have to be designed to preserve the privacy of the parties. In the last decade, Privacy-Preserving Data Mining has received increase attention from researchers in computer science, and many protocols and techniques have been presented for different data mining methods, each of which has a level of security, efficiency and accuracy. However, there are still many open problems in this field of study in terms of security and efficiency such as developing privacy-preserving protocols for public channels and incremental algorithms, preventing sensitivity and collusion attack, reducing intermediate outputs, and balancing the distribution of the final results. In this talk we shortly review the background, present some solutions, and discuss possible future directions in this area.