Abstract

Over the years, t-closeness has been dealt with in great detail in Privacy Preserving Data Publishing and Mining. Other methods like k-anonymity fail in terms of attribute disclosure and background knowledge attack as demonstrated by many papers in this field. l-diversity also fails in case of skewness attack. t-closeness takes care of all these shortcomings and is the most robust privacy model known till date. However, till now t-closeness was only applied upon a single sensitive attribute. Here, a novel way in determining t and applying t-closeness for multiple sensitive attributes is presented. The only information required beforehand is the partitioning classes of Sensitive Attribute(s). Since, t-closeness is generally applied on anonymized datasets, it is imperative to know the t values beforehand so as to unnecessarily anonymize data beyond requirement. The rationale of using the measure of determining t is discussed with conclusive proof and speedup achieved is also shown.

References

Determining $t$ in t-closeness using Multiple Sensitive Attributes


- ARX – Powerful Data Anonymization http://arx.deidentifier.org/


Keywords
Privacy Preserving Data Mining  Privacy Preserving Data Publishing  t-closeness  Multiple Sensitive Attributes