DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF SOUTHERN DENMARK, ODENSE

COMPUTER SCIENCE COLLOQUIUM

Speeding Up Privacy Preserving Record Linkage

Alexandros Karakasidis
University of Cyprus, Cyprus

Thursday, 17 March, 2016 at 10:15
U56

Abstract:

With the data explosion we have been experiencing recently, the problem of linking records that represent the same real world entity across various and often heterogeneous sources has become increasingly important for a number of application areas such as data integration, business intelligence, web mining and recommendation systems. When dealing with sensitive and personal user data, the process of record linkage raises privacy issues. Thus, privacy preserving record linkage has emerged with the goal of identifying matching records across multiple data sources while preserving the privacy of the individuals these records describe. This task is very resource demanding, considering the abundance of available data, which, in addition, are often dirty. Blocking techniques are deployed prior to matching to prune out unlikely to match candidate records so as to reduce processing time. However, when scaling to large datasets, such methods often result in quality loss. In this talk, a privacy preserving blocking technique will be presented which provides high fault-tolerance, maintains result quality and scales linearly with respect to the dataset size.

Host: Yongluan Zhou