Privacy Preserving Record Linkage
Mihir Kumar
Mentor: Vincent Bindschaedler

Motivation and Objective
Identifying records that refer to the same individual from multiple databases without revealing the identity of the person nor the contents of the database

Overview
PPRL
Matches
Non Matches

Blocking
Field Comparison
Record Pair Comparison
Record Pair Classification

No blocking
With blocking

Binary Field Comparison
Approximate Field Comparison
Fellegi-Sunter (FS)
Winkler’s Modification to FS

Field Comparison Vector
Field Comparison Vector

Dice coefficient
= (2 * 5)/13 = 0.77

Experimental results: runtime

Experimental results: accuracy