

Multi-Party Privacy-Preserving Record Linkage (MP-PPRL)

Thilina Ranbaduge, Dinusha Vatsalan, Peter Christen

thilina.ranbaduge@anu.edu.au, dinusha.vatsalan@anu.edu.au, peter.christen@anu.edu.au

Research School of Computer Science, College of Engineering & Computer Science, The Australian National University, Australia

What is MP-PPRL ?

- In today's Big Data era, integrating data from multiple sources is important for efficient and effective decision making.
- Due to **privacy and confidentiality concerns**, organizations are often not willing or allowed to share or reveal their sensitive data.
- Multi-party privacy-preserving record linkage (MP-PPRL)** aims to identify records in multiple databases that relate to the same entity without revealing any private information about these entities.
- Real-world applications include health, crime and fraud detection, national security, etc.

Challenges and Research Directions

Challenges :

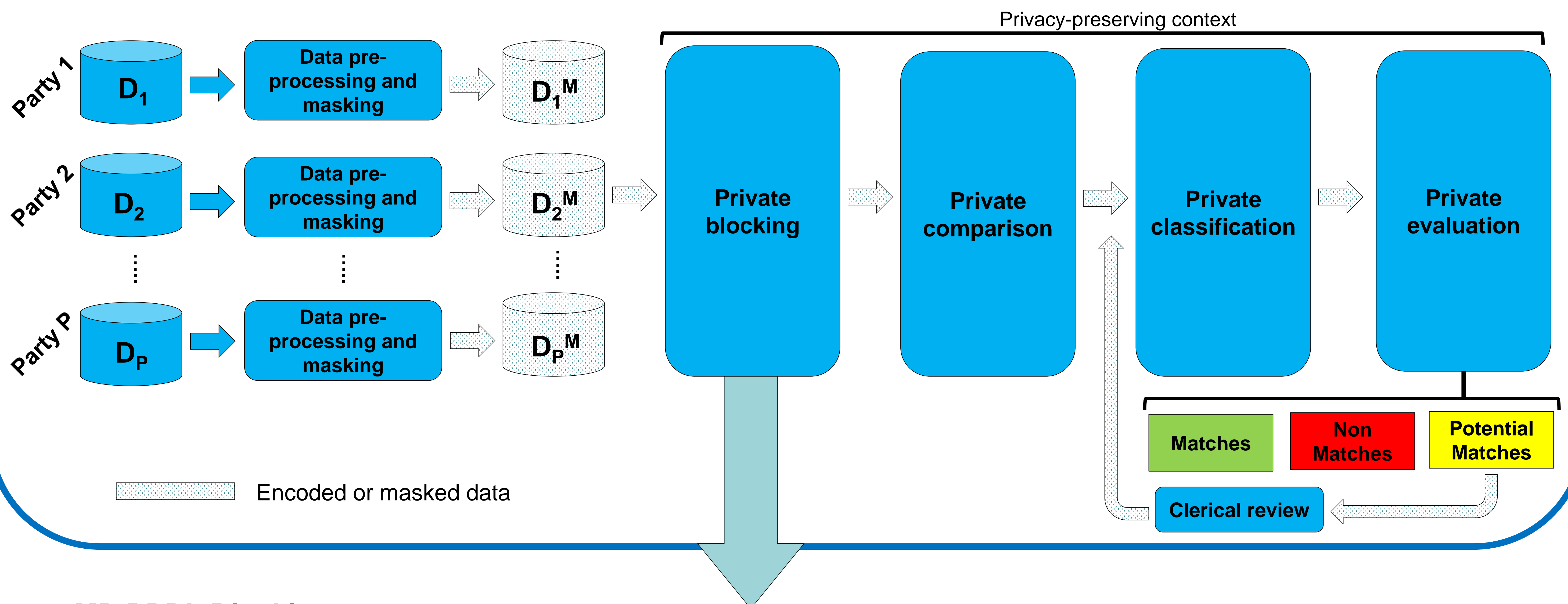
- Existing works only consider two-party linkage
- Exponential complexity with number of parties
- Collusion between participating parties

Research directions :

- Efficient filtering and communication patterns for MP-PPRL
- Advanced classification techniques for MP-PPRL
- MP-PPRL approaches for other adversary models
- Sub-group blocking and matching

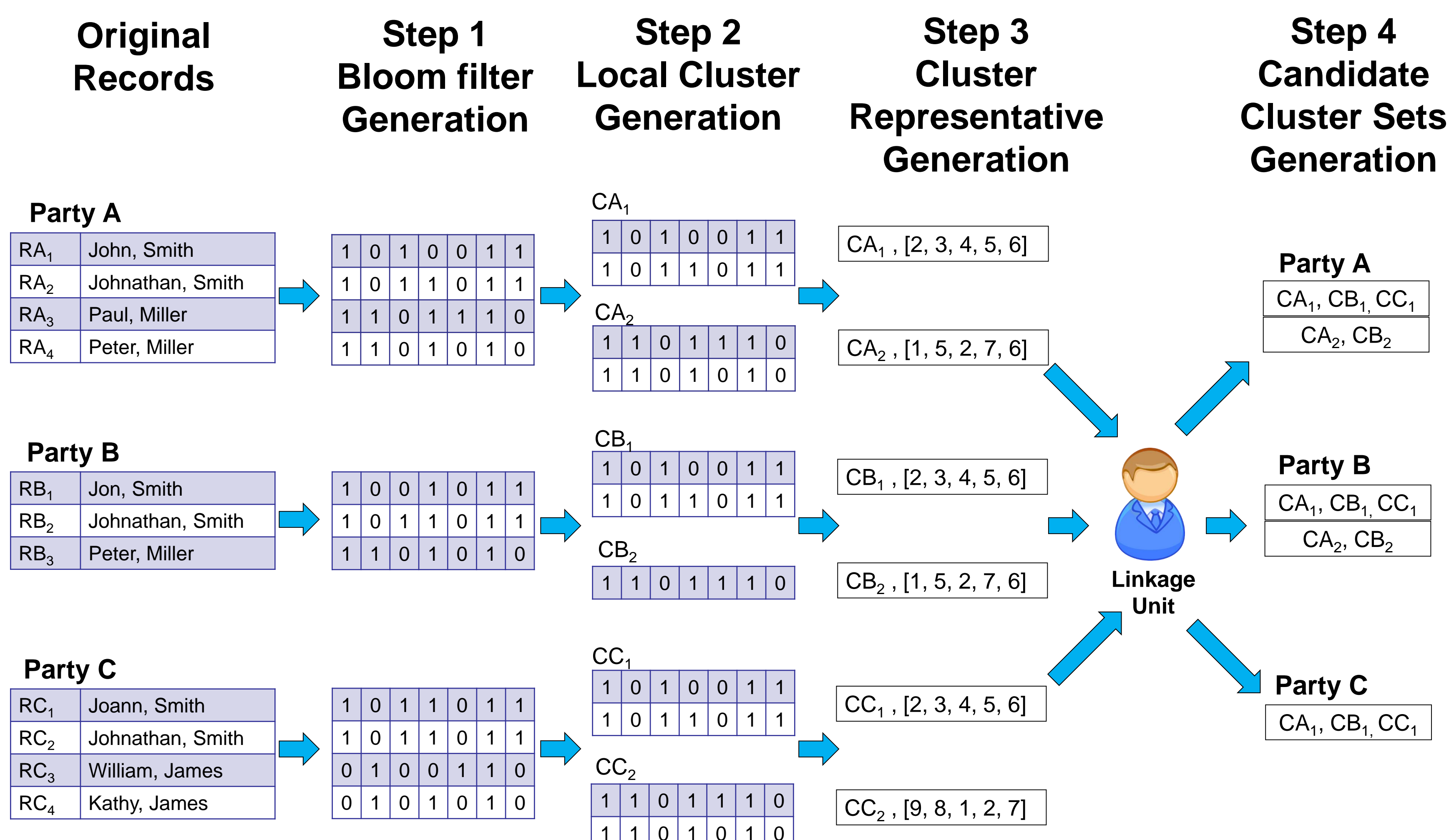
How it works ?

- MP-PPRL follows a **stepwise process** where some steps need to be performed with **privacy-preserved**.

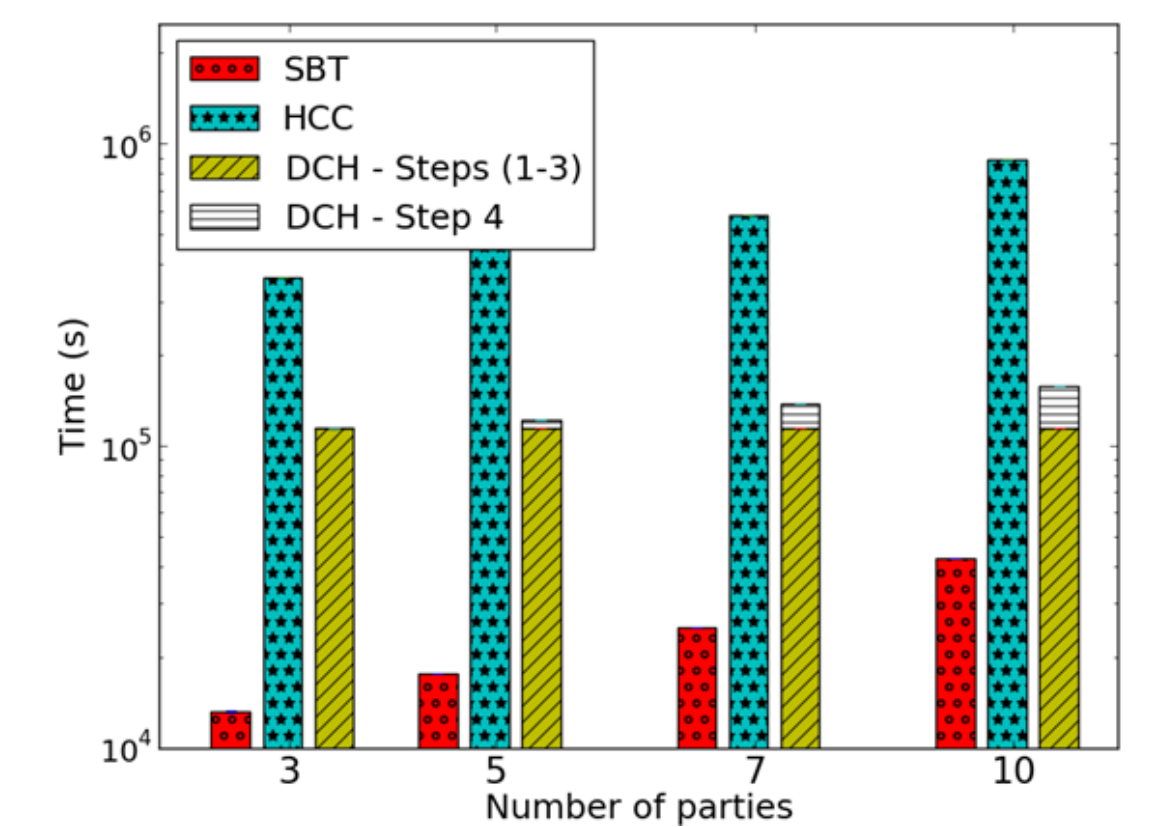
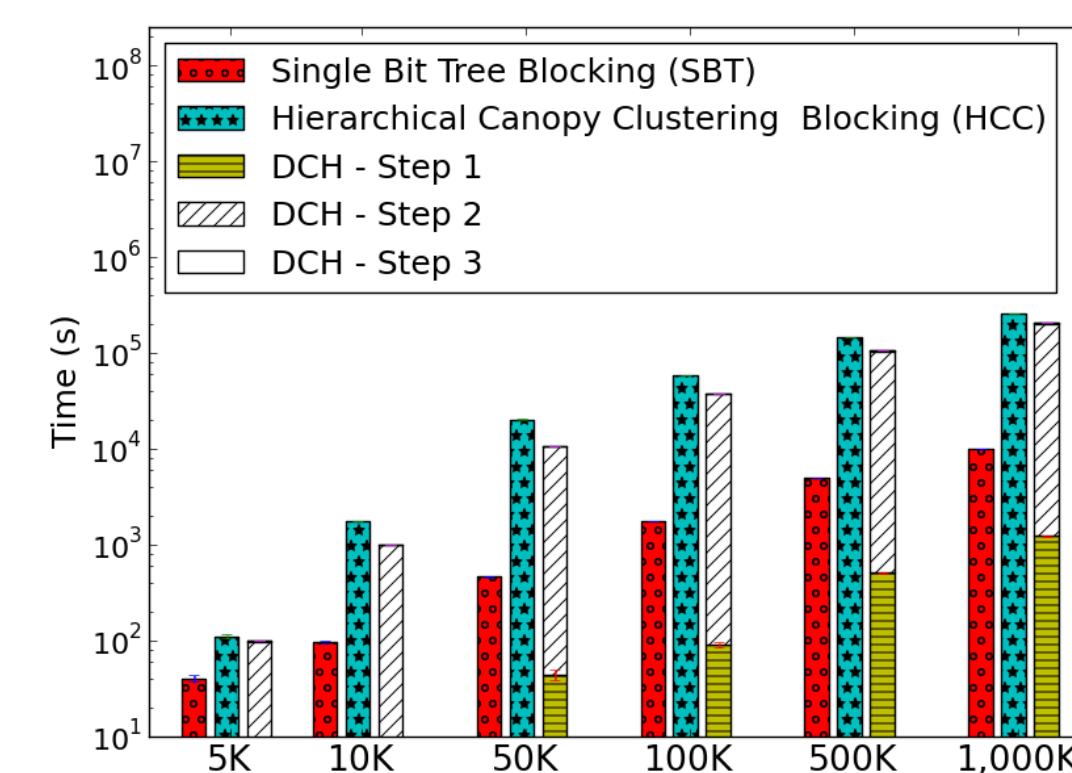


MP-PPRL Blocking

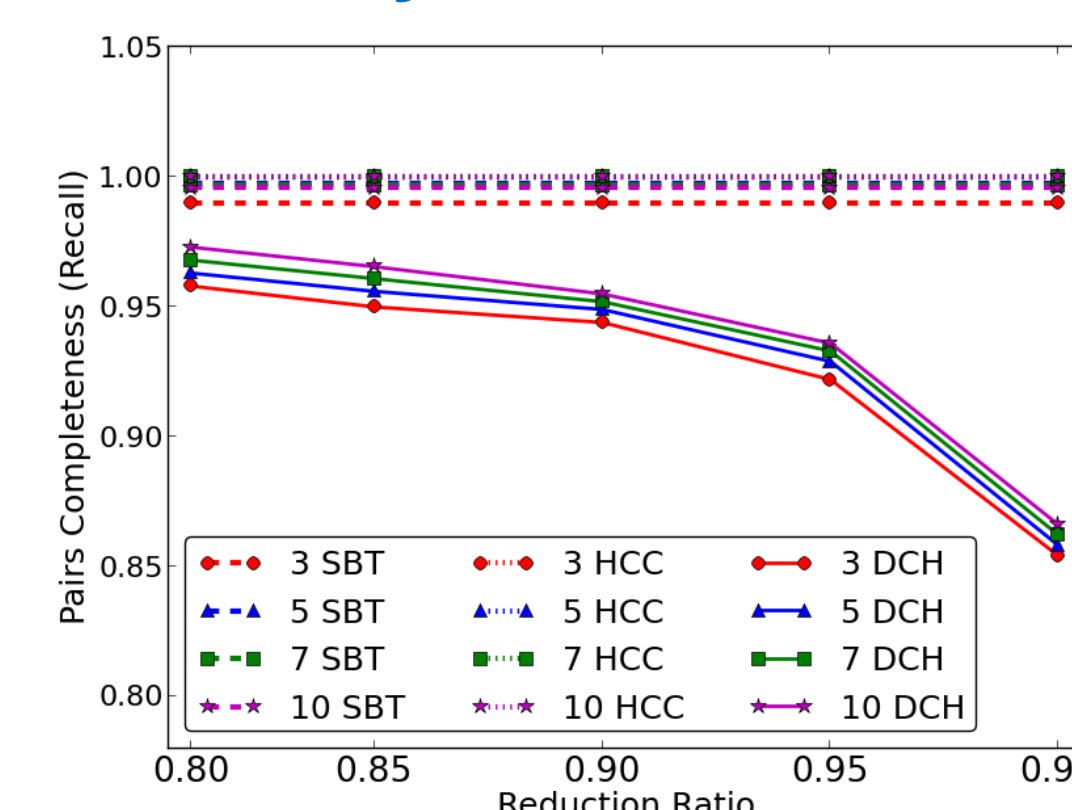
Distributed Clustering and Hashing (DCH) based blocking



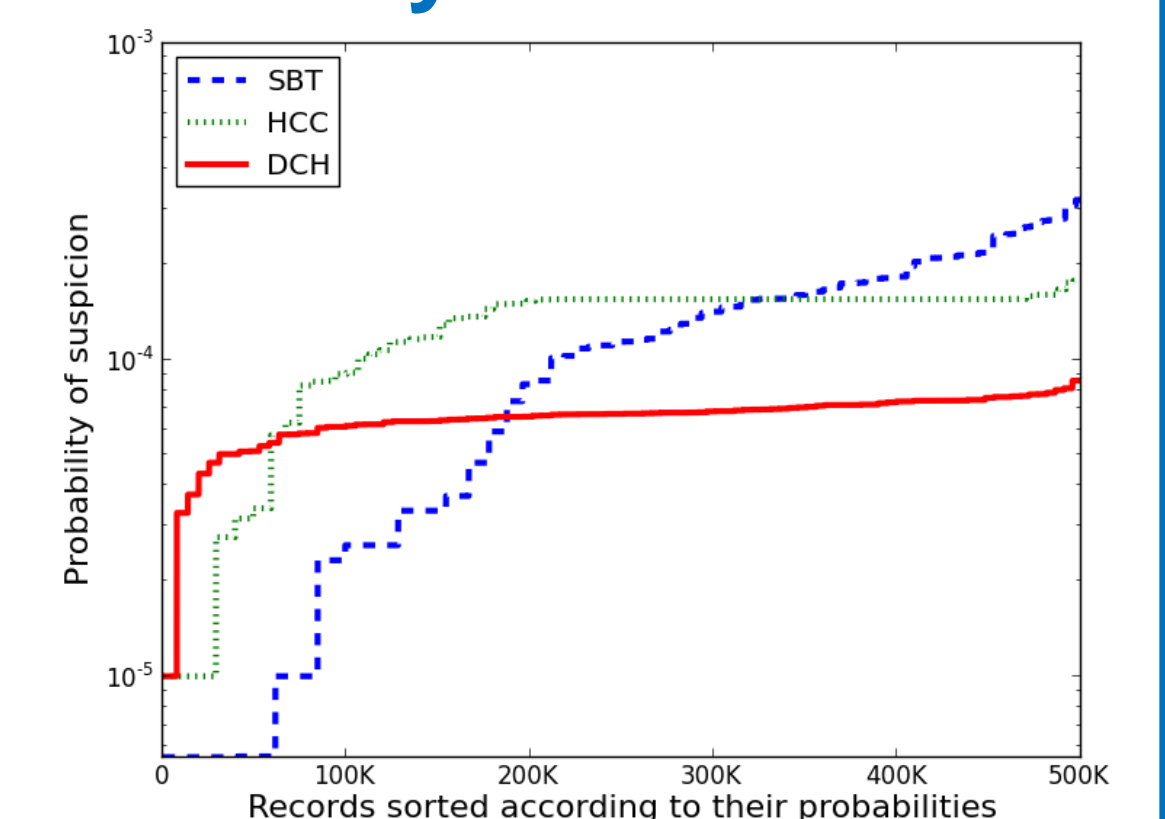
Scalability



Quality



Privacy



T. Ranbaduge, D. Vatsalan, P. Christen, V. Verykios : Hashing-based distributed multi-party blocking for privacy-preserving record linkage. In: PAKDD (2016)

This research is funded by the Australian Research Council under Discovery Project DP130101801.