

Privacy-Preserving Data Analytics

School of Information Systems

Associate Professor Yingjiu Li



CENTRAL CONCERNS & QUESTIONS

The major concern in analyzing a large collection of user data is to preserve each user's privacy in individual data items while enabling data analytics to be performed at aggregation level.

RESEARCH HIGHLIGHTS

- A set of formulae is provided for deriving the bounds of protected values in data tables from available marginal ranges.
- The concept of interval based inference is proposed, together with the algorithms to control such inference in query process.
- New architecture and protocol are designed to resolve the bottleneck of traditional inference control by distributing inference control from server side to user side.

EMERGING IDEAS & INITIATIVES

The basic idea to address the privacy concern in data analytics is to block direct access to privacy-disclosing data items and control indirect access to such items (i.e., inference channels) while minimizing the impact on the scope and accuracy of data analytics.

SELECTED PUBLICATIONS

- Yingjiu Li, Haibing Lu: <u>Privacy Risk Assessment</u> with Bounds Deduced from Bounds. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (IJUFKS), 19(4), 685-715, World Scientific, 2011.
- Yingjiu Li, Haibing Lu: Disclosure analysis and control in statistical databases. <u>13th European</u> <u>Symposium on Research in Computer Security</u> (<u>ESORICS</u>), pages 146-160, LNCS 5283, Malaga, Spain, October 6-8, 2008.
- 3. Lingyu Wang, Yingjiu Li, Sushil Jajodia, Duminda

FUNDING AND PEOPLE

- FUNDING: Privacy Preserving Techniques for Accessing Network Services. Submitted to SMU Office of Research, Mar2012 - Jan 2013.
- FACULTY / RESEARCH STUDENTS / RESEARCH STAFF: Yingjiu Li, Robert Deng, Kevin Chiew, Haibing Lu, Bing Liang
- EXTERNAL COLLABORATORS: Sushil Jajodia, Sean X. Wang, Xintao Wu, Lingyu Wang

Wijesekera: Parity-Based Inference Control for Multi-Dimensional Range Sum Queries. Journal of <u>Computer Security</u>, 15(5):417-445, IOS Press, 2007
4. Yingjiu Li, Haibing Lu, Robert Deng: Practical Inference Control for Data Cubes (Extended Abstract). <u>IEEE Symposium on Security and Privacy</u> (S&P'06), pages 115-120, Oakland, USA, May 21-24, 2006.

 Xintao Wu, Ying Wu, Yongge Wang, Yingjiu Li: Privacy Aware Market Basket Data Set Generation: A Feasible Approach for Inverse Frequent Set Mining. <u>2005 SIAM International Conference on Data Mining</u> (SDM'05), pages 103-114, Newport Beach, CA, USA, April 21-23, 2005.