

Toward Transparency in Public Policy via Privacy Enhanced Social Flow Analysis with Applications to Ecological Networks and Crime

Zhenhui (Jessie) Li, College of Information Sciences and Technology

Daniel Kifer, Department of Computer Science & Engineering

Corina Graif, Department of Sociology and Criminology

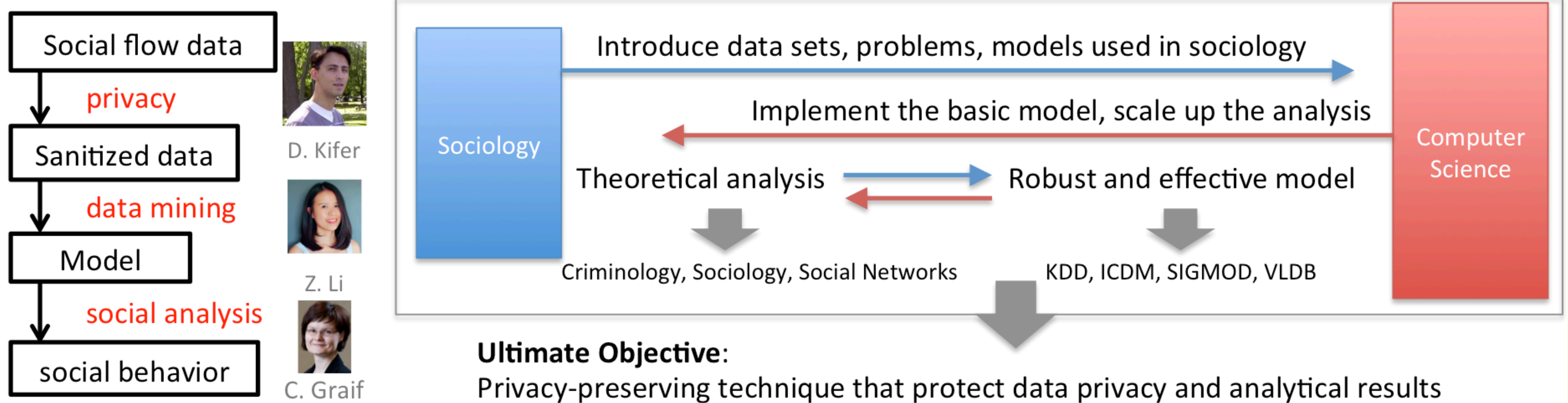
Challenge:

Techniques to "sanitize" sensitive raw data (e.g., anonymization) could have negative impact on the quality of the scientific results that use the data.

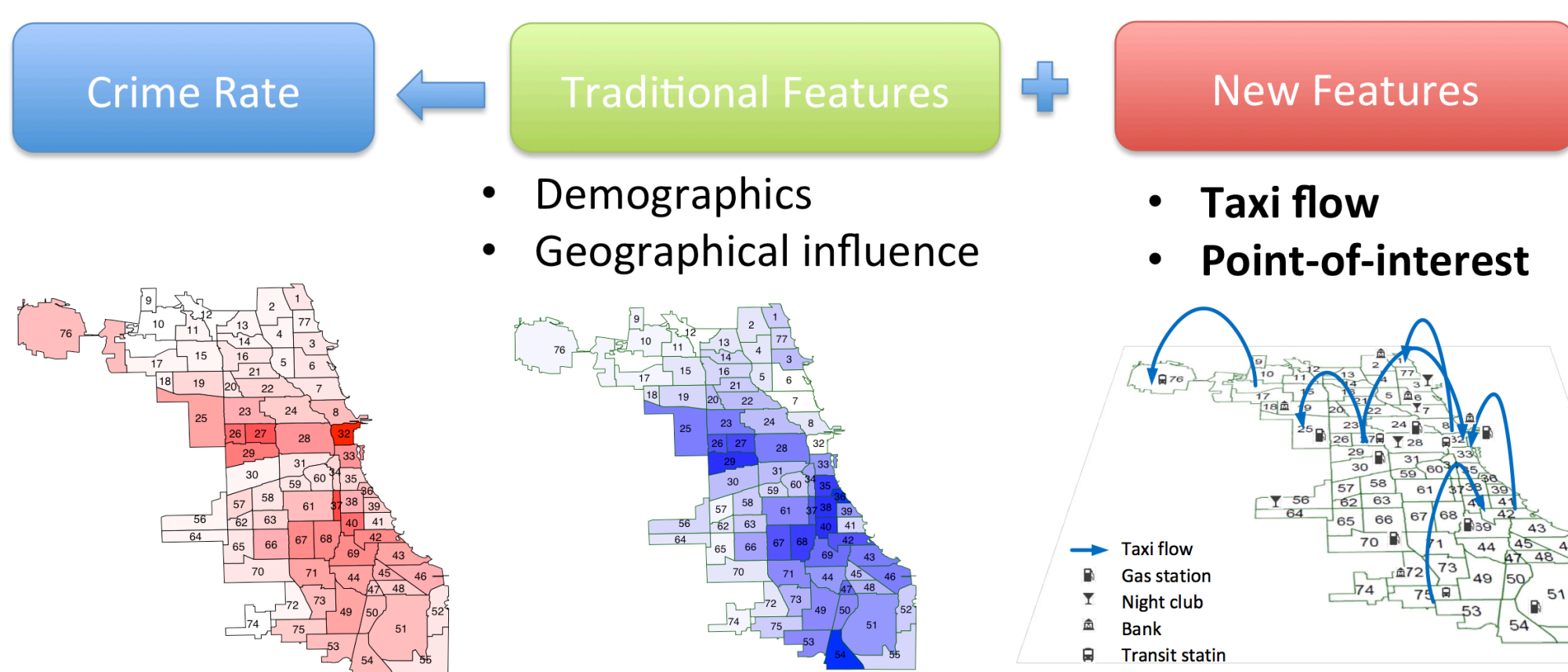
Example: Criminologists use social flow data (e.g., Census home-work commuting data) to analyze and understand crime. But such data have been distorted for privacy concern. How will the privacy preserving techniques affect analytical results?

Why social flow? Crime is not only correlated between geographically close locations, but could also be correlated between two locations with a strong social flow (e.g., commuting, taxi, migration). However, social flow data are sensitive and remain largely untapped in criminology literature.

Interdisciplinary Collaboration



Key Achievement: Improved Crime Rate Inference with Social Flow Data



Taxi flow represents a type of social flow
POI represents urban functions in a region

Demographics	Geo Influence	Taxi Flow	POI	Mean Relative Error
✓				0.335
✓	✓			0.331
✓	✓	✓		0.321
✓	✓		✓	0.303
✓	✓	✓	✓	0.269

- Negative Binomial model is used
- Results shown for year 2014 in Chicago at Community Area level

H. Wang, D. Kifer, C. Graif, Z. Li, Crime Rate Inference with Big Data, in Proc. of 2016 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'16), San Francisco, CA, Aug. 2016.

Work in progress:

- Explain effects of social flow using social theory
- Assess the impacts of privacy techniques on analytical results
- More robust crime inference model

Graduate students supported:



Alyssa Howard-Tripp
(criminology)



Hongjian Wang
(IST)

Undergrad students (REU):

B. Bowman, C. Brown, B. Eppinger, R. Schneider

Interested in meeting the PIs? Attach post-it note below!

Contact: Jessie Li, JessieLi@ist.psu.edu



National Science Foundation
WHERE DISCOVERIES BEGIN

The 3rd NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting

January 9-11, 2017

Arlington, Virginia



PennState