1 Introduction

In the urban space, explain the data variatey, volume, and velocity. They pose challenges and opportunities in addressing real urban computing problem.

Challenges:

- Data sparse and nosise
- Heterogeneous data type
- Data is huge

Problems with existing methods:

- model of interaction is ad-hoc
- the spatial non-stationarity

My reserach focus is to model complicated interactions in the urban space. It could be interactions of regions, people, and features. And I want to propose a unified graphical model to do this.

1.1 Research questions include:

- How does the crime (or any other feature) in one region correlate and influence that variable of other regions.
- What influence crime in focal region, the crime in its neighbor, or the demographics in its neighbor.
- Measure the similarity of two person, given their observations.
- How would the traffic change, if build a new transportation center at location X.

Understanding the interactions can help us solve three fundamental problems in urban space.

- Inference (single source from others) problem. [Prediction]
- Structural learning. Does two variable correlated? How?
- Partition regions (address misalignment issue). community detection based on high-dimension similarity.

1.2 Why use grapihcal model?

A graphical model support three kinds of learning problem:

1. Infer unobserved values.
2. Learn parameters. Given graph, learn dependency.
3. Structural learning. Given data, the graph is too complicated to build. Learn the graph.

Therefore, a graphical model is a unified approach.