

---

# SENSING SEMANTIC INFORMATION FROM MOBILE SOCIAL NETWORKS

---

Ting Zhang and Wen Yi  
zhan1013@purdue.edu, yi35@purdue.edu

---



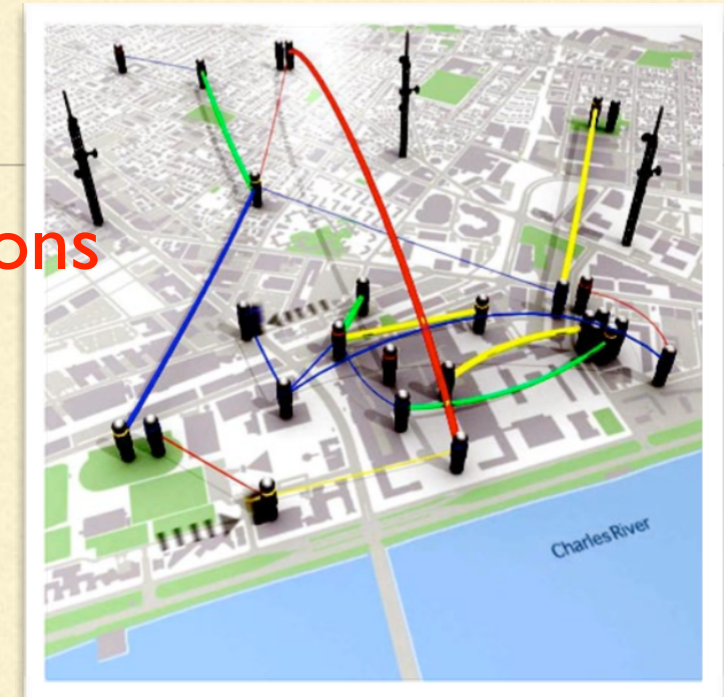
# INTRODUCTION

- Human society consists of extensive **communications and interactions** between individuals
  - mobile sensors, mobile phones, tablets and GPS

- Understand individual relations

**Promote** ↔ **Facilitate**

Interactions between individuals



- list phone contacts in semantic orders according to the **time** and **location**.
- friend **grouping suggestion** function inside the cellphone contact managing software

- Friendship Inference

- Community Detection

- Reality Mining: Smart Cellphone Based Network



---

# RELATED WORK

---

- Friendship Inference:

- **Similarity between nodes:** Liben-Nowell, D. and Kleinberg, J. (2007), The link-prediction problem for social networks.
  - – – Topology structure
- **Feature:** location, phone call records, message records

- Community Detection:

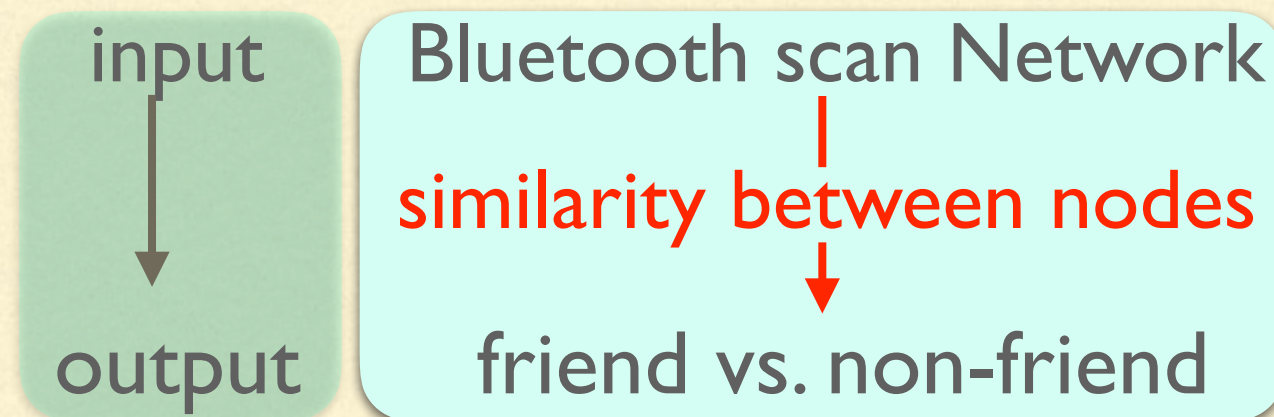
- **Top down, component breaking:** M. Girvan and M. E. J. Newman. Community structure in social and biological networks.
  - **Bottom up, Cluster merging:** M. E. J. Newman. Fast algorithm for detecting community structure in networks.
-



# MACHINE LEARNING SET UP & ALGORITHM

- Friendship Inference:

- **Topology structure:**



- **SVM with feature:** # of phone calls, duration, night call ratio

- **Hybrid:** Topology + SVM with feature

- Community Detection:

- Start from **single large** component, breaks the component by **removing edges** potentially across the community

- Evaluate the edge by the “**Betweenness**”. Weight the edge by computing the **shortest path** between each vertices pair, and adding weight to the edges in path



---

# EXPERIMENTS

---

- Friendship Inference:
    - Feature selection: **Good** night call ratio + duration of phone calls  
**Poor** # of phone calls
    - **Topology** vs. **SVM with feature** vs. **Hybrid**
    - **Non-friend accuracy**: 99.0%
      - Unbalanced data: In **1980** entries of phone record, **31** entries are between friends calls.
  - Change Bluetooth record into binary undirected relation graph
  - Building clustering tree for community detection, and evaluating after each new component breaking: all **new** components with **100%** purity; separate **all** the member of **2** school when breaking into **5** components (**100%** purity for **all** components).
-



---

# SUMMARY

---

- Feature selection appears to be an important phase in machine learning, especially for semantic inference.
    - To infer friendship relation, call in nights can indicate higher probabilities of friend than larger number of phones.
    - To detect communities between business school and media lab students, bluetooth network performs better than friendship network.
  - The face-to-face meeting behavior (tracked by Bluetooth scanning record) are similar inside the same school, although across different sub-groups; the behavior are significantly different across members from different school. Community detection algorithms would fit the general large group detection better.
-