Extracting Trip Destinations from GPS Trajectories with Irregular Sampling Frequency

Ying Song
Geography, Environment and Society
University of Minnesota – Twin Cities

Tianci Song
Rui Kuang
Computer Science and Engineering
University of Minnesota – Twin Cities

UNIVERSITY OF MINNESOTA
Driven to Discover

Introduction

- **Location-aware Technologies – Movement Data**

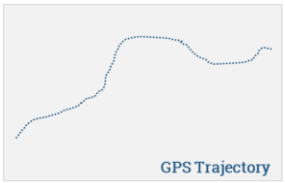
Mobile

TomTom

Garmin

Commute

- **Semantic Trajectory Data Analysis** (Yan and Spaccapietra 2009)



Introduction

❖ **Path Segmentation & Stop Detection**

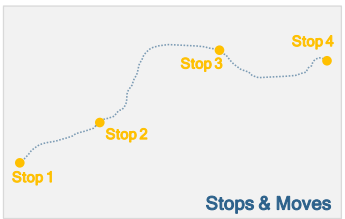
- **Stops & Moves**
  1) **Stop -** not move for a minimal time duration, activities
  2) **Move -** routes & modes

- **Stop Detection**
  1) **Duration** $t_{i+1} - t_i > \tau_t$
  2) **Speed** $d_{i,i+1}/(t_{i+1} - t_i) < \varepsilon_v$
  3) **Density** $t_{i+1} - t_i \leq \tau_t \ \ AND \ \ d_{i,i+1} \leq \sigma_d$
  4) **Direction change**

Hybrid, Domain Knowledge, Data-driven

Introduction

❖ **Path Segmentation & Stop Detection**



- Stops & Moves
- Stop Detection
  1) Duration $t_{i+1} - t_i > \tau_t$
  2) Speed $d_{i,i+1}/(t_{i+1} - t_i) < \varepsilon_v$
  3) Density $t_{i+1} - t_i \leq \tau_t \ \ AND \ \ d_{i,i+1} \leq \sigma_d$
  4) Direction change

➡ **Irregular sampling frequency $(t_{i+1} - t_i)$???**

quality of the device
atmosphere interference
signal loss/reflection/blocking
embedded signal processing methods

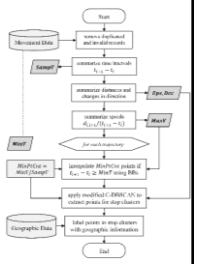
Objective

Develop analytical methods and computational procedures
to handle the irregular sampling frequency of movement trajectories
while detecting and describing stops along trajectories

Method

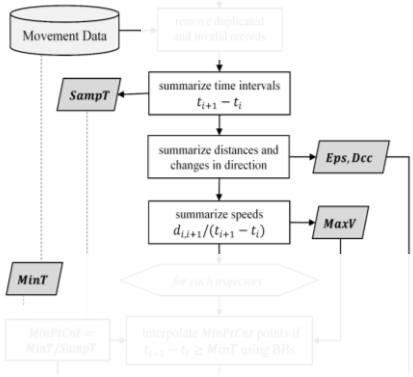
Path segmentation with space-time interpolation and density-based spatial clustering

1.	get to know your data	determine model parameters
2.	large time interval	space-time interpolation (truncated BBs)
3.	extract stops	density-based spatial clustering (DBSCAN)
4.	describe stops	label points in each stop episode



1. Parameter Setting

- **Stop duration** - ***MinT***
  - domain knowledge; application oriented
- **Sample frequency** - ***SampT***
  - most frequent time interval; data provider
- **Distances and directions** - ***Eps, Dcc***
  - neighbors for density-based clustering
- **Moving speed** - ***MaxV***
  - travel modes, type of moving object, …

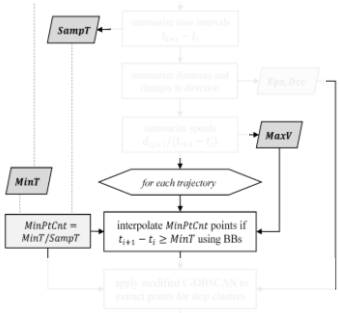


2. Space-time Interpolation

- **When?**
  - large time interval

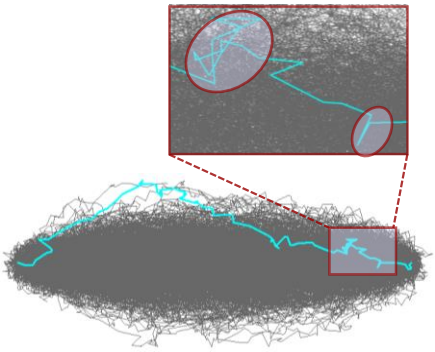
    $t_{i+1} - t_i > MinT$
- **Why?**
  - interpolate movement points between gap
  - use interpolated points in spatial clustering

    $MinPt = MinT/SampT$
- **How?**
  - **truncated Brownian bridge** (Song and Miller 2014)



2. Space-time Interpolation

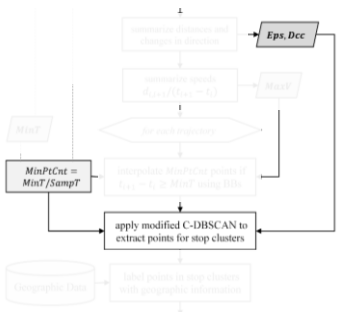
- When?
  - large time interval
    $t_{i+1} - t_i > MinT$
- Why?
  - interpolate movement points between gap
  - use interpolated points in spatial clustering
    $MinPt = MinT/SampT$
- How?
  - truncated Brownian bridge (Song and Miller 2014)
    - ✓ continuous-time stochastic process anchored with start & end values
    - ✓ temporal constraints $SampT$
    - ✓ speed constraints $MaxV$



simulated 50,000 BB trajectories

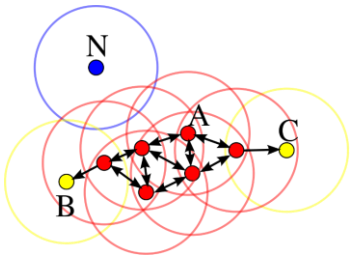
3. Density-based Spatial Clustering

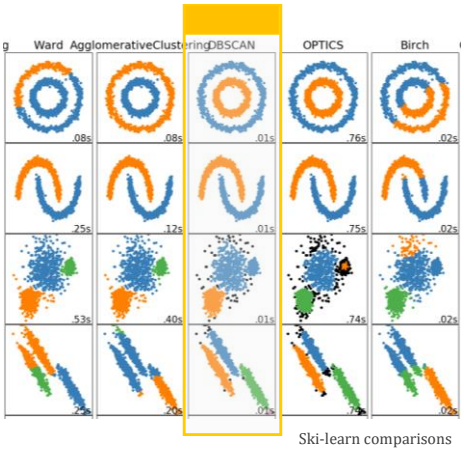
- DBSCAN
  - Density-Based Spatial Clustering of Applications with Noise
  - Cluster of different shapes
- Constraints-DBSCAN (Gong et al. 2015)



3. Density-based Spatial Clustering

❖ **DBSCAN**





Ski-learn comparisons

1) Define neighbors – circle ***Eps***
2) Get core points – red
3) Link neighboring points – red & yellow
4) Check density – ***MinPtCnt***

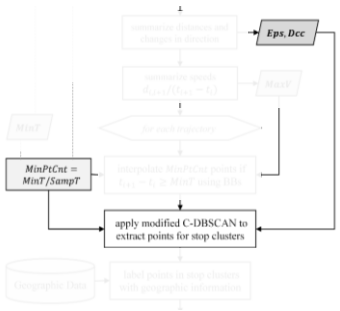
3. Density-based Spatial Clustering

❖ **DBSCAN**

- Density-Based Spatial Clustering of Applications with Noise
- Cluster of different shapes

❖ **Constraints-DBSCAN** (Gong et al. 2015)

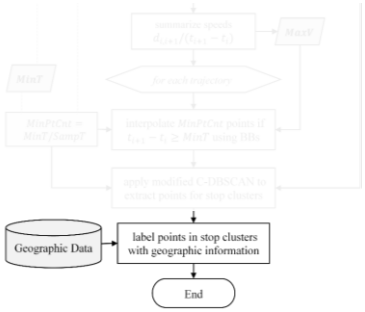
- **Sequential order** of points along a trajectory (break into sub-clusters and check density)
- **Significant direction changes** not too frequent (deal with ramp, signal reflection etc.)



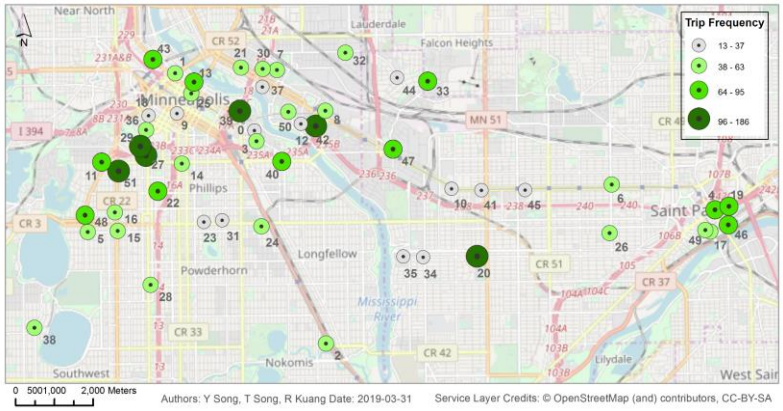
4. Contextualize Stops

- **Underlying Geographic Information**
  - **General**
    e.g. land-use parcels, neighborhoods, …
  - **Specific**
    points of interests (POIs)

- **Describe and Analyze Stop Episodes**
  - **Each stop episode**
    e.g. inferred activities
  - **Each trajectory**
    e.g. stop chaining, trip scheduling
  - **All trajectories**
    e.g. frequent types/locations of stops



HOURCAR Study Case



HOURCAR

- Twin Cities, MN, US
- 58 vehicles
- 55 service stations

Vehicle Trajectories

- Aug. to Oct. 2017
- 2,546 reservations
- 159,562 records
- Avg. 63 records/trip

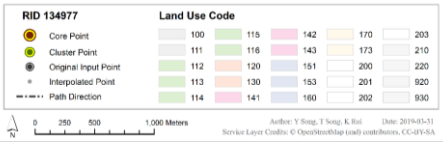
$MinT = 300\ sec$ (signal loss/device off)
$SampT = 15\ sec$ (most frequent interval)
$Eps = 800\ m$ (80% quantile, 3-4 blocks)
$MinPtCnt = MinT/SampT = 20$

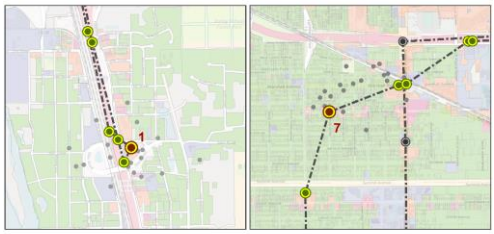
HOURCAR Study Case

Results – Type of Stops

Residential	
112	Seasonal or Vacation
113	Single Family Detached
114	Single Family Attached
115	Multifamily
116	Manufactured Housing Parks
Non-residential	
120	Retail and Other Commercial
130	Office
151	Industrial and Utility
160	Institutional
Mixed Use	
141	Mixed Use Residential
142	Mixed Use Industrial
143	Mixed Use Commercial
920	Outside TC Metro Area *

Recreation	
170	Park, Recreational or Preserve
173	Golf Course
Agriculture	
100	Agricultural
101	Extractive
111	Farmstead
Transportation	
200	Local Road *
201	Major Highway
202	Major Railway
203	Airport
Others	
210	Undeveloped
220	Open Water
930	Outside MN *





(a) one stop accessed by highway (b) one stop accessed by local roads

(c) two consecutive stops in one cluster (d) two separate stops in two clusters

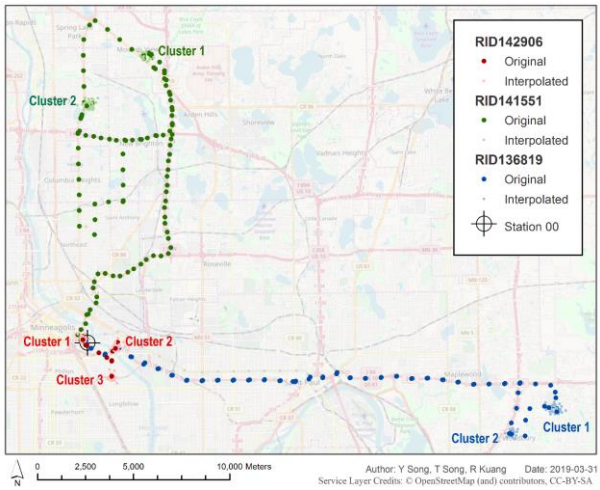
HOURCAR Study Case

Results

– Stops along one trajectory

- ✓ actual travel behaviors
- ✓ trip scheduling

SID	Res. ID	Cls. Seq.	(Point ID: LU Code-Seconds)
0	136819	1	(3226:**200**-2542) (3353:**200**-2593)
0	136819	2	(3539:**120**-4375)
0	141551	1	(75307:**120**-565)
0	141551	2	(75325:**120**-629) (75328:**120**-321)
0	142906	1	(117198:**200**-4534)
0	142906	2	(117358:**200**-2307)
0	142906	3	(117569:**200**-758)

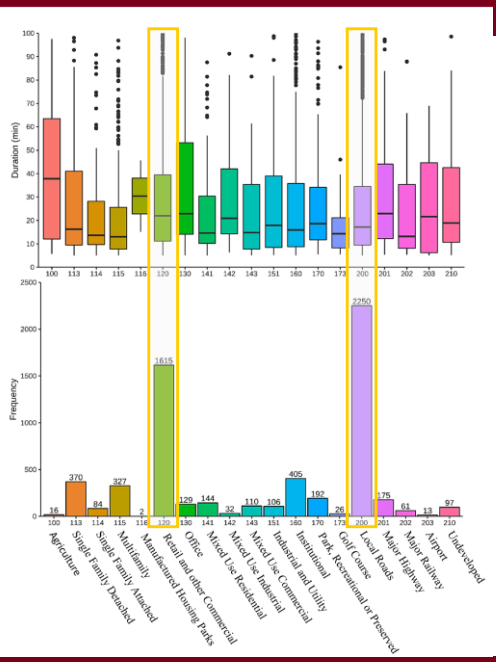


HOURCAR Study Case

Results

– Stops along all trajectories

- ✓ **Stop types**
  (retail & other commercial; local roads)
- ✓ **Stop durations**
  (retail & other commercial: 10 min to 40 min)

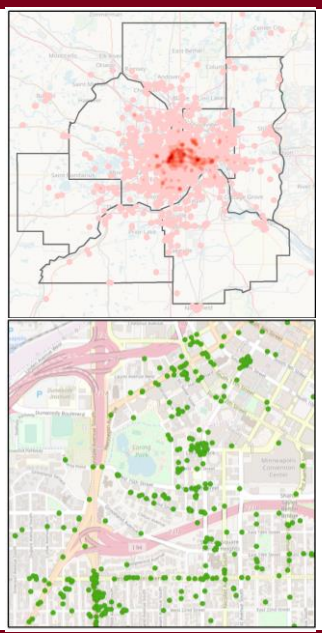


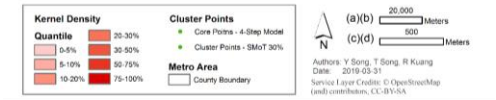
HOURCAR Study Case

Results

– Stops along all trajectories

- ✓ Stop types
  (retail & other commercial; local roads)
- ✓ Stop durations
  (retail & other commercial: 10 min to 40 min)
- ✓ **Stop locations**
  (downtown Minneapolis)
  (Loring park? Congestion?)





Conclusion

- **From movement trajectory data To useful information (stop & moves)**
- **4-Step method to handle irregular time intervals between two consecutive points**
  1) parameter setting – data-driven / domain knowledge
  2) space-time interpolation – truncated Brownian bridges
  3) density-based spatial clustering – modified C-DBSCAN
  4) geographic context – general / specific
- **HOURCAR vehicle tracking data**
  - ✓ Actual trip (stops & routes) of all users
  - ✓ Extract and Describe stop episodes along each reserved trip
  - ✓ analyze stop type and duration for a single stop, a single trajectory and all trajectories

Ying Song

yingsong@umn.edu

Geography, Environment and Society
University of Minnesota–Twin Cities

Acknowledgement

- The material in this paper is based on research project supported by Digital Technology Imitative Seed Grants (DTI) from the **Digital Technology Center (DTC)** at University of Minnesota entitled "Advancing knowledge of car-sharing behaviors by learning with high-resolution GPS data".
- The data is provided by **HOURCAR** carsharing services in Twin Cities, MN.