Interactive Visual Exploration of Most Likely Movements

Can Yang and Győző Gidófalvi
Division of Geoinformatics
Deptartment of Urban Planning and Environment
KTH Royal Institution of Technology, Sweden
{cyang,gyozo}@kth.se

Outline

- Introduction
- Problem formulation
- Methodology
- Demonstration
- Conclusions and future work

Introduction

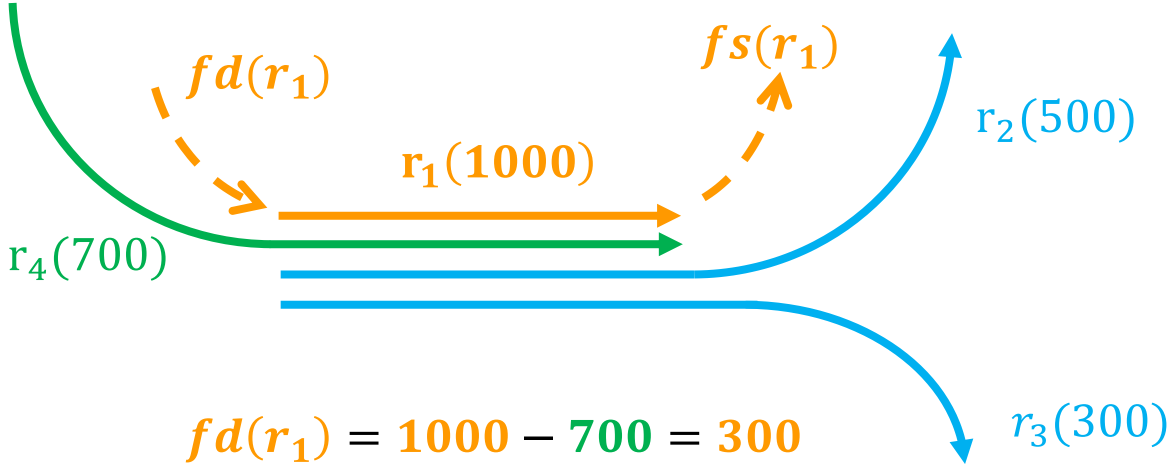
- With location-enabled devices widely adopted, **massive streams of trajectories** have been generated. One way to **compress the data** is to store **frequent patterns**. However, **infrequent movements are lost.**
- In this paper, the proposed method and system
  - aggregates a massive trajectory stream to **limited storage** as time-varying patterns of movements
  - reconstructs from this information the ***k*** **most likely movements** for a selected time period and origin-destination region
  - facilitates querying and explorations of these likely movements using a web based user interface

Introduction – Intuition of Pattern Transition

When only frequent pattern r_i is stored, infrequent movements are lost but some information can be inferred from **free demand** and **free supply** of patterns.

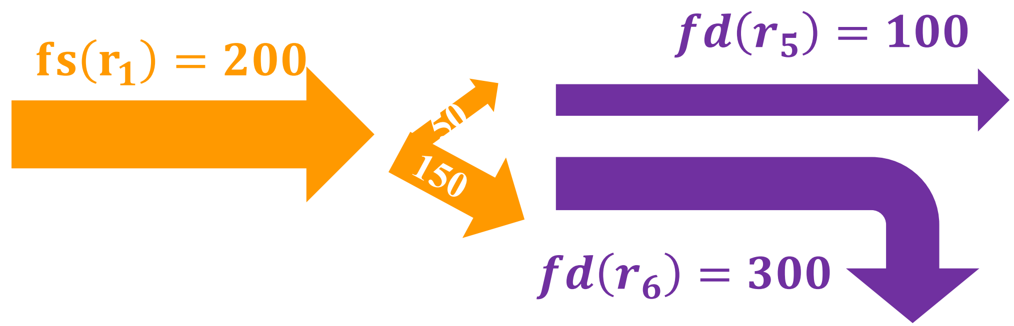
- **free demand** $fd(r_i)$: objects that enter a pattern r_i but not from its **preceding patterns,** $[+, \boldsymbol{r_i}]$
- **free supply** $fs(r_i)$: objects that leave a pattern r_i and do not follow its **succeeding patterns** $[\boldsymbol{r_i}, +]$

Objects in the **free supply** of a pattern can transit to its **connected patterns** proportionally to the free demand of these patterns.



$$fd(r_1) = 1000 - 700 = 300$$
$$fs(r_1) = 1000 - 500 - 300 = 200$$

Free demand and supply of patterns



Pattern transitions

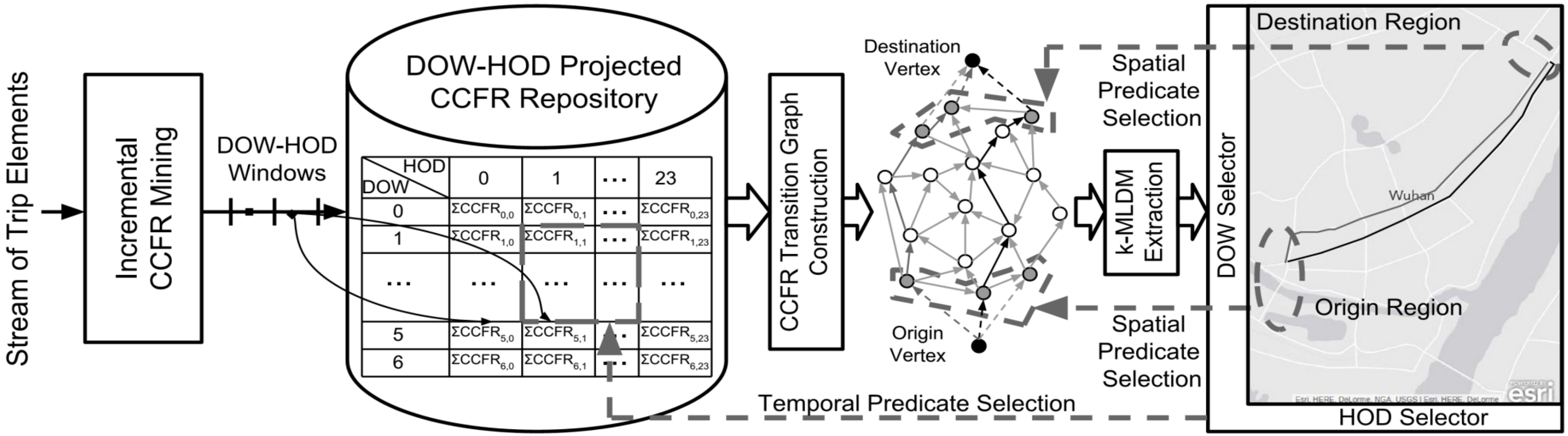
Problem Formulation

The distinct k-Most Likely Movements (MLM) problem is defined as estimating the **distinct *k* most likely movements** of the population given **temporal** predicates such as time periods and **spatial** predicates such as origin and destination.

For instance, what are the likely movements/route choices from the train station to the airport from 8 am to 10 am on Mondays?

Methodology

1. **Extract** and **Store** closed continuous frequent routes / patterns (CCFR) from GPS data
2. **Build** pattern transition graph
3. **Estimate** distinct k-MLMs



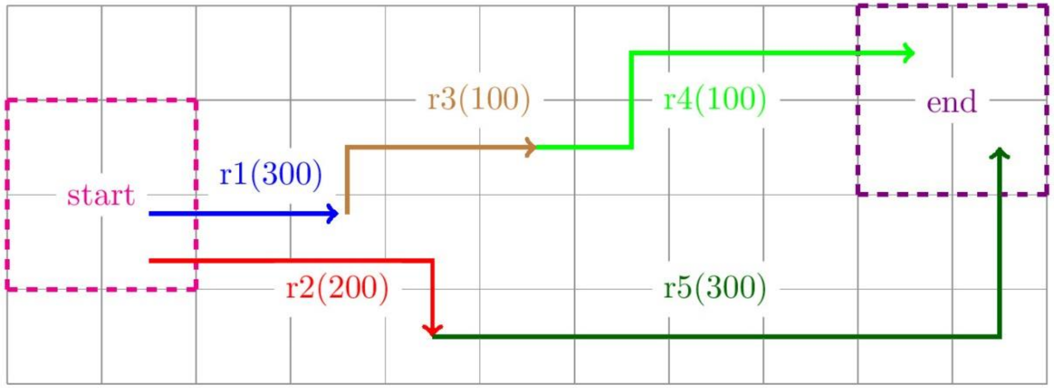
Schematic diagram of methodology

Methodology- CCFR and Pattern Transition Model

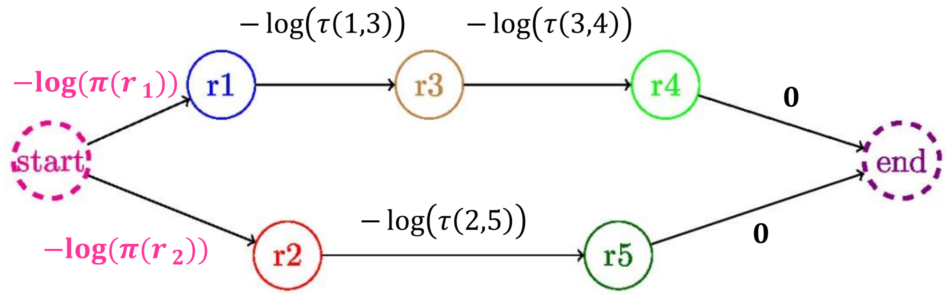
- Info in CCFR
  - Sequence of spatial units traversed + Count of objects
- CCFR movement model
  - At the end of a CCFR an object either probabilistically transit to "connected" CCFRs or stops moving.
- Pattern transition graph
  - 1-1 map of CCFRs to nodes and connections of CCFRs to directed edges
  - weight of edge from r_i to r_j is

    $$-\log(\tau(i,j))$$

    where $\tau(i,j)$ is the transition probability from r_i to r_j based on and adhering free supply and free demand of patterns
  - weight of edge from **start** to r_i is $\mathbf{-\log(\pi(r_i))}$, where $\boldsymbol{\pi(r_i)}$ is the initial probability of a pattern which is the relative free supply of r_i



Pattern topology relationship

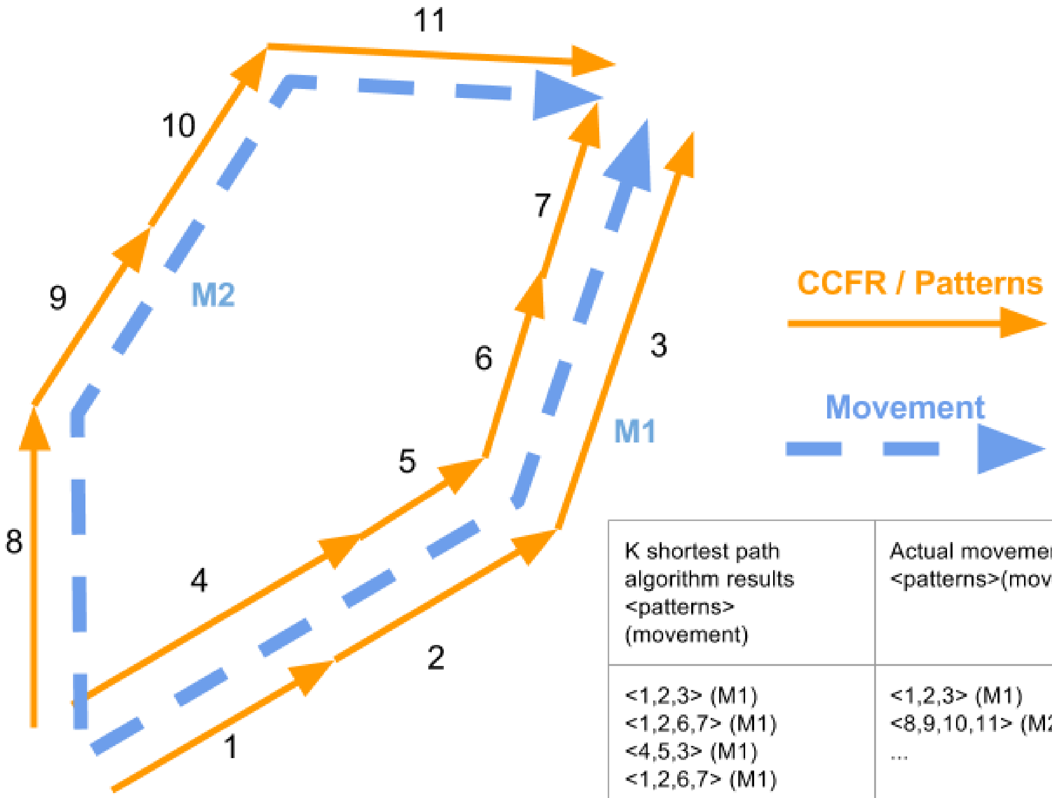


Pattern transition graph

Methodology- Distinct k-MLMs

Problem setting: a **movement** is a sequence of spatial units and can be generated by a large number of sequences of CCFRs.

1. To estimate the likelihood of a movement a dynamic programming approach is used.
2. To extract distinct *k*-MLMs, extract the **current** MLM and **block** the CCFRs that build it and iteratively extract the remaining *k-1* MLMs.



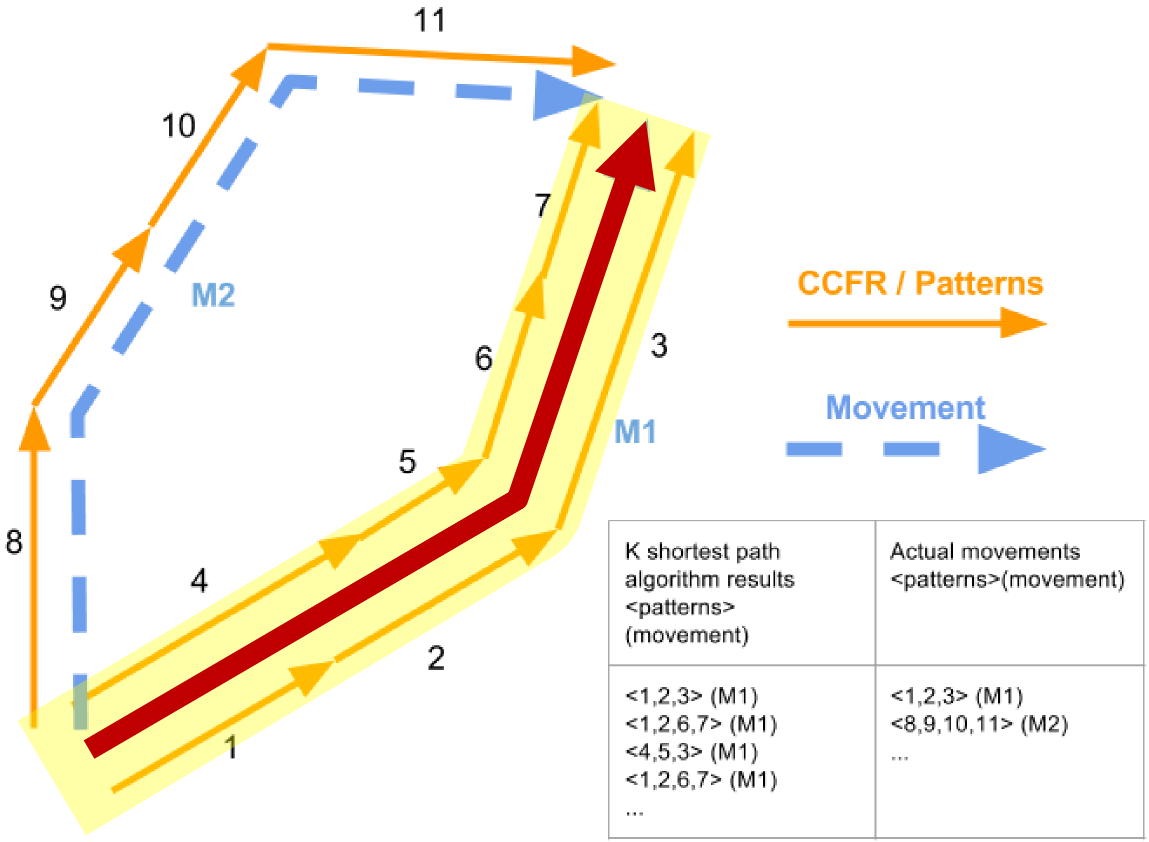
K shortest path algorithm results <patterns> (movement)	Actual movements <patterns>(movement)
<1,2,3> (M1) <1,2,6,7> (M1) <4,5,3> (M1) <1,2,6,7> (M1) ...	<1,2,3> (M1) <8,9,10,11> (M2) ...

Example of Distinct k-MLMs

Methodology- Distinct k-MLMs

Problem setting: a **movement** is a sequence of spatial units and can be generated by a large number of sequences of CCFRs.

1. To estimate the likelihood of a movement a dynamic programming approach is used.
2. To extract distinct *k*-MLMs, extract the **current** MLM and **block** the CCFRs that build it and iteratively extract the remaining *k-1* MLMs.



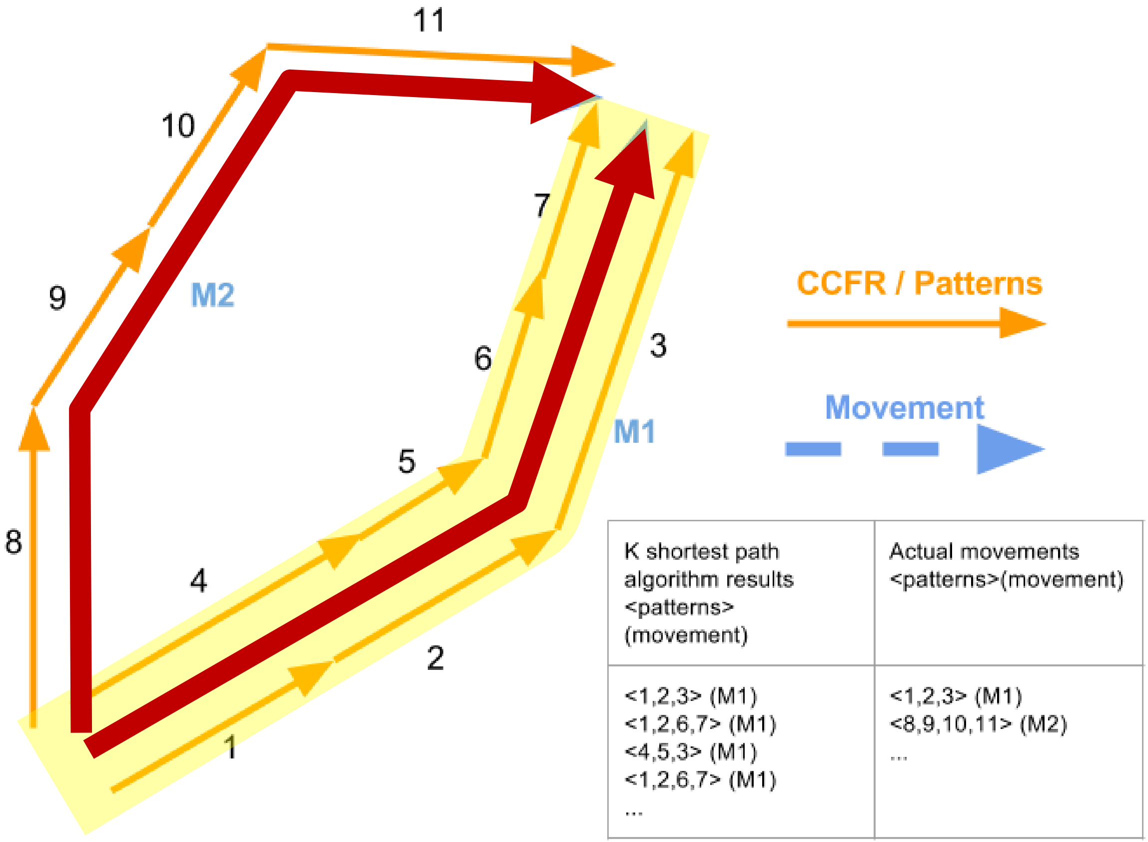
K shortest path algorithm results <patterns> (movement)	Actual movements <patterns>(movement)
<1,2,3> (M1) <1,2,6,7> (M1) <4,5,3> (M1) <1,2,6,7> (M1) ...	<1,2,3> (M1) <8,9,10,11> (M2) ...

Example of Distinct k-MLMs

Methodology- Distinct k-MLMs

Problem setting: a **movement** is a sequence of spatial units and can be generated by a large number of sequences of CCFRs.

1. To estimate the likelihood of a movement a dynamic programming approach is used.
2. To extract distinct *k*-MLMs, extract the **current** MLM and **block** the CCFRs that build it and iteratively extract the **remaining *k-1* MLMs**.



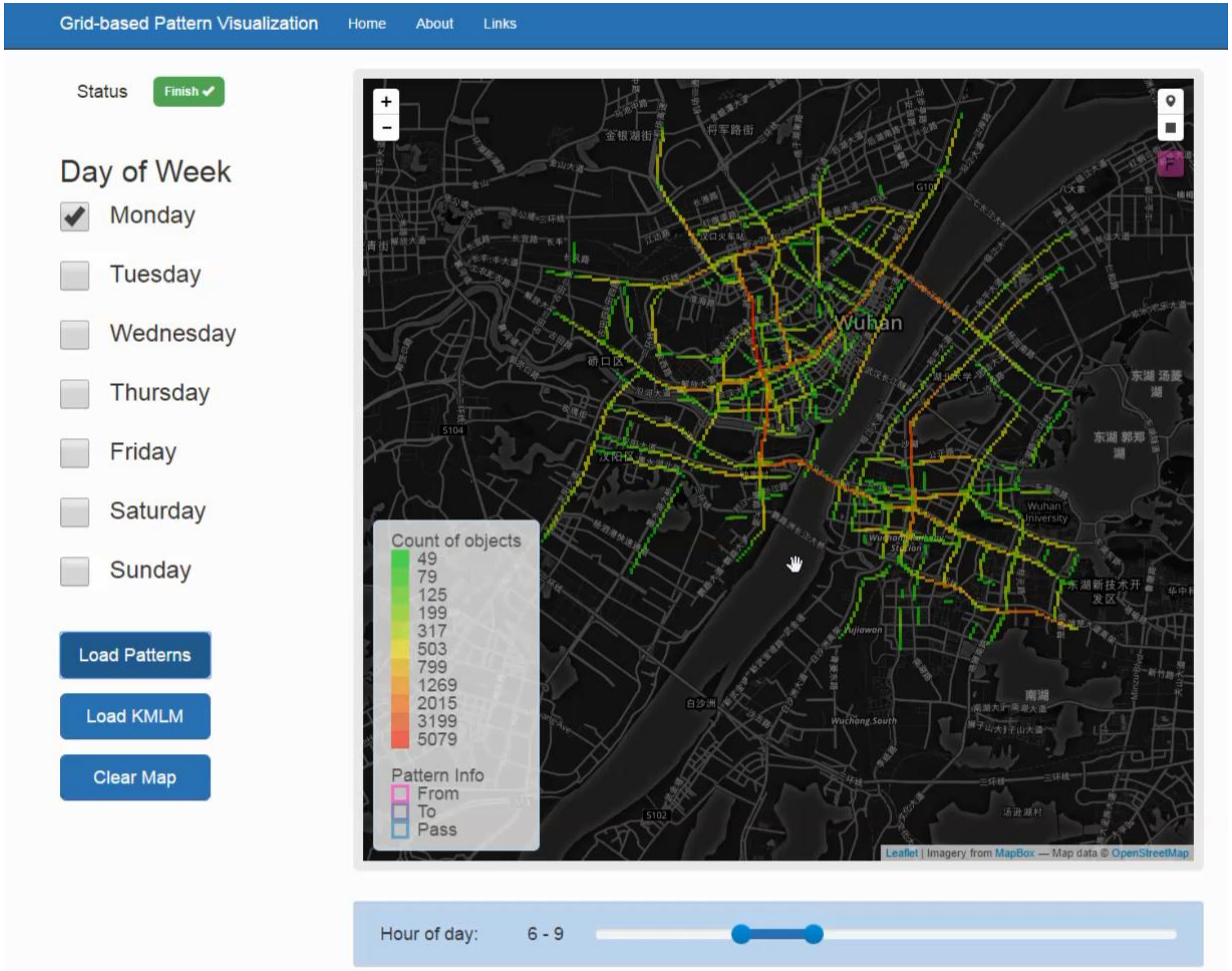
K shortest path algorithm results <patterns> (movement)	Actual movements <patterns>(movement)
<1,2,3> (M1) <1,2,6,7> (M1) <4,5,3> (M1) <1,2,6,7> (M1) ...	<1,2,3> (M1) <8,9,10,11> (M2) ...

Example of Distinct k-MLMs

Demonstration

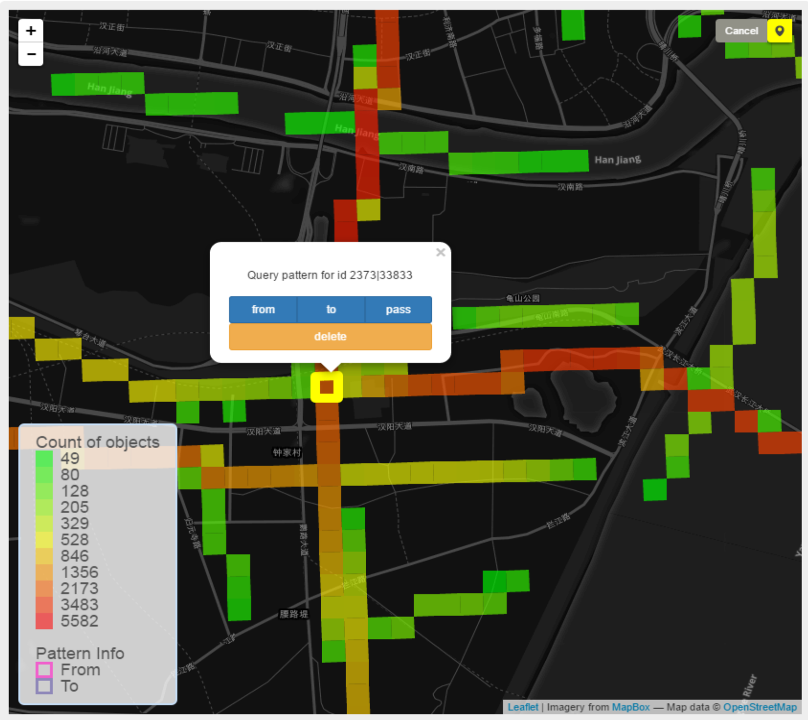
- Implementation
  - Server
    - NodeJS
  - Client
    - Leaflet

- Data
  - Totally 2.26 million trajectories collected from 11000 taxis over a 6 day period in Wuhan, China.

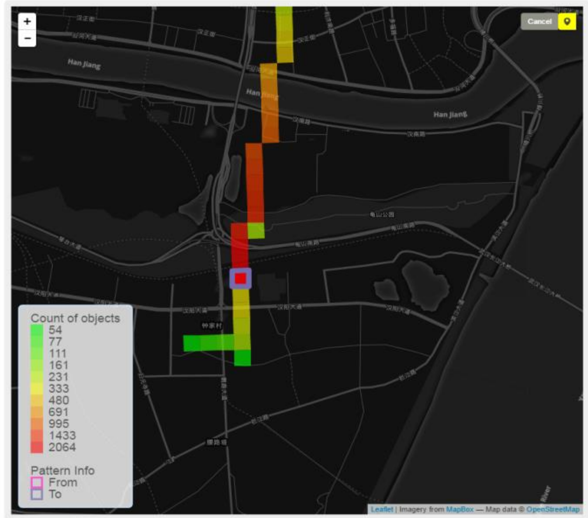


Screenshot of User Interface

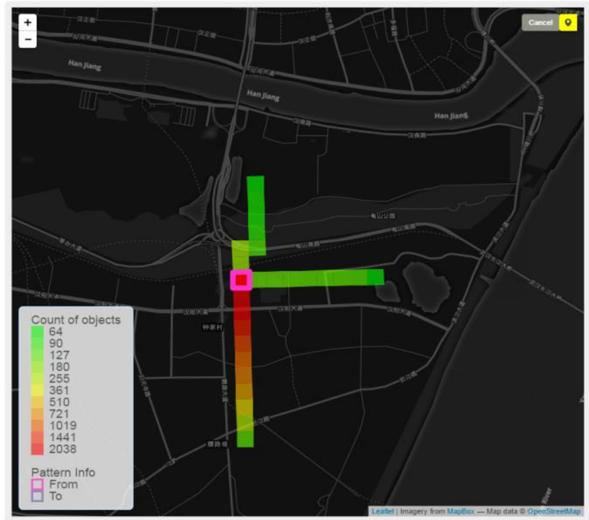
Demonstration- Pattern Exploration



Interactive query of patterns
from/to/pass a grid

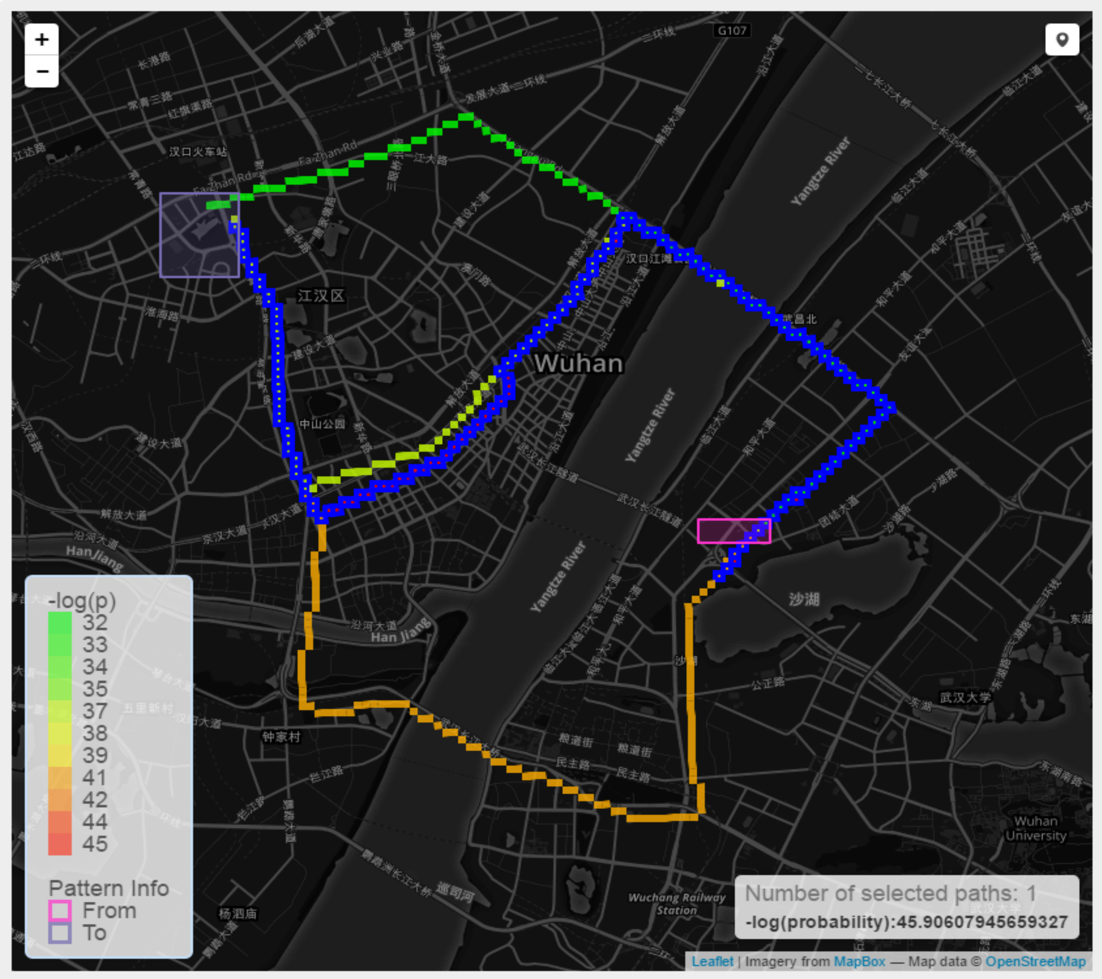


Patterns starting from specific grid

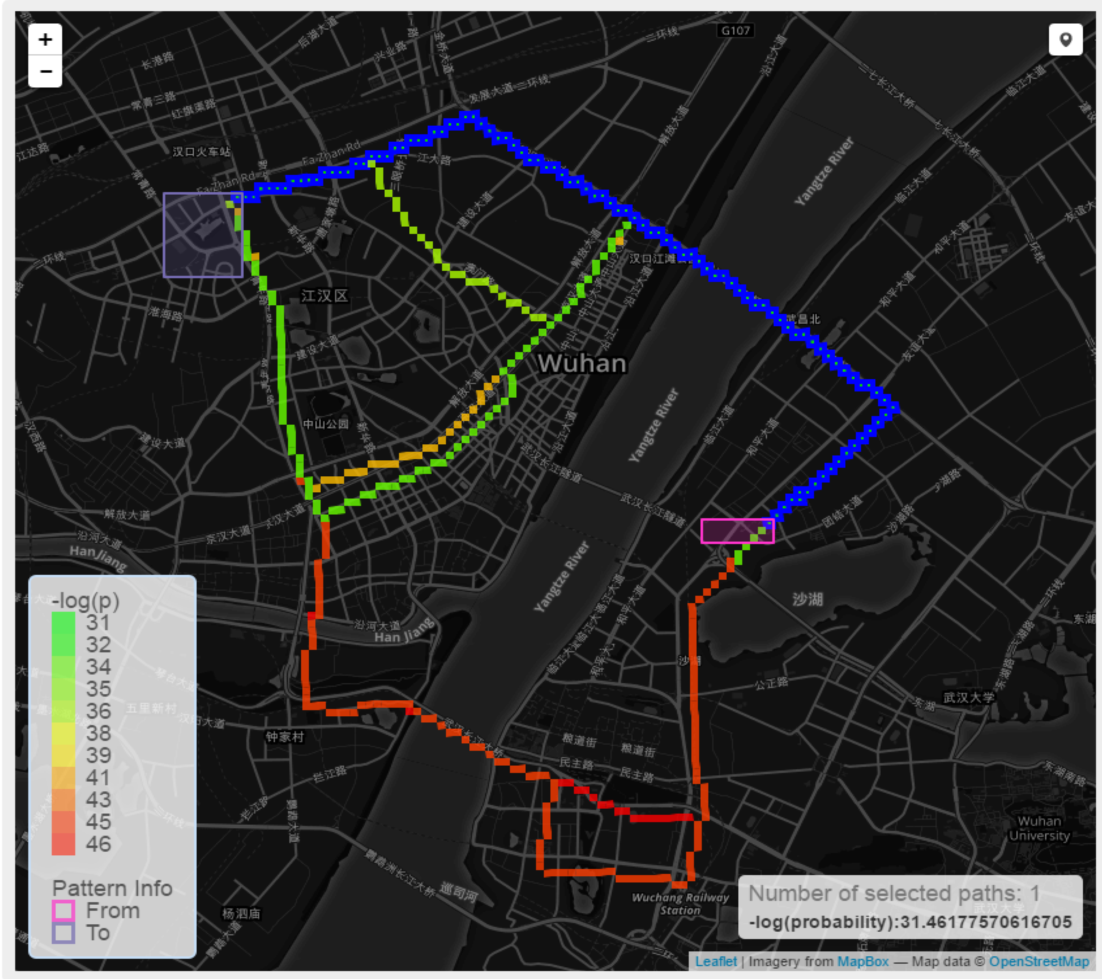


Patterns ending at specific grid

Demonstration – Distinct KMLM



(a) 4 distinct movements between 2 regions in the morning 06:00 – 09:00



(b) 6 distinct movements between 2 regions in the afternoon 16:00 – 19:00

Time-varying Distinct K-MLM generated from the model (the blue path is the movement highlighted by user)

Conclusions and Future Work

- Conclusions
  - The paper proposed a method that in an effective manner extracts complex, time varying movement patterns from a stream of moving object trajectories, regenerates likely movements based on these patterns, and facilitates the visual querying and explorations of these likely movements using a simple map interface.
- Future work:
  - **Alternative models** considering topological relationship between CCFRs
  - **Empirical validation** of the model
  - Extend the model to **other types of spatial units**

Thank you for your attention!
Q/A?