"Space-Time Graph Modeling of Ride Requests Based on Real-World Data"

Abhinav Jauhri[1], Brian Foo[2], Jerome Berclaz[2], Chih Hu[1], Radek Grzeszczuk[2], Vasu Parameswaran[2], John Paul Shen[1]

[1]Carnegie Mellon University, USA; [2]Uber Technologies Inc., USA

February 4, 2017 (AIORSocGood, AAAI)

Outline

1. **Motivation and Approach**
2. **Ride Request Patterns**
3. **Ride Request Graphs (RRG)**
    a. Densification Power Law
    b. Temporal Evolution
4. **Societal Benefits**
    a. Poolability - Observations
    b. Densification and Poolability
5. **Conclusion and Future Work**

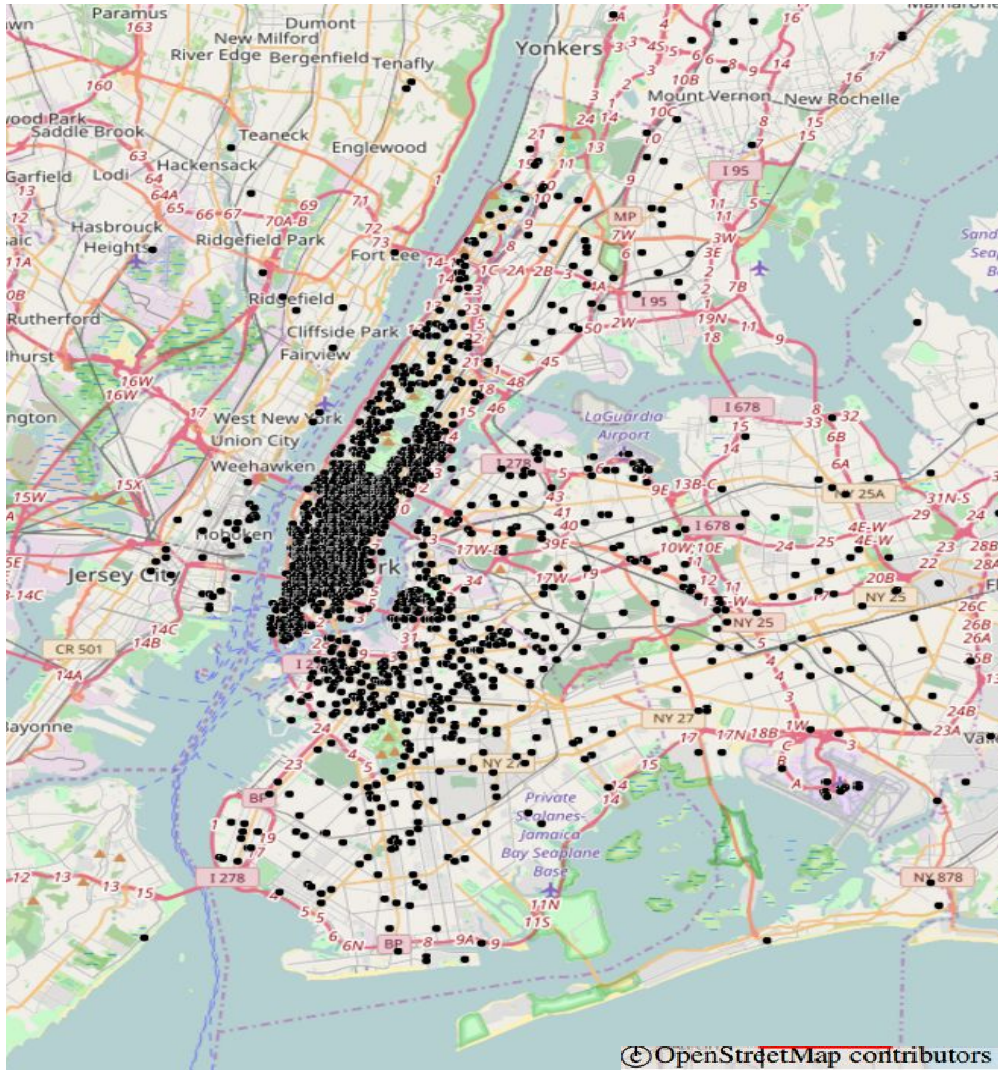
Motivation and Approach

- ➔ Focus on Human Mobility Sensing, Analytics, and Services.
- ➔ Leverage extensive data from a global ride-sharing service.
- ➔ Analyze the temporal and spatial patterns of ride requests.
- ➔ Develop a rigorous graph model based on ride request data.
- ➔ Explore potential usefulness of the ride request graph model.

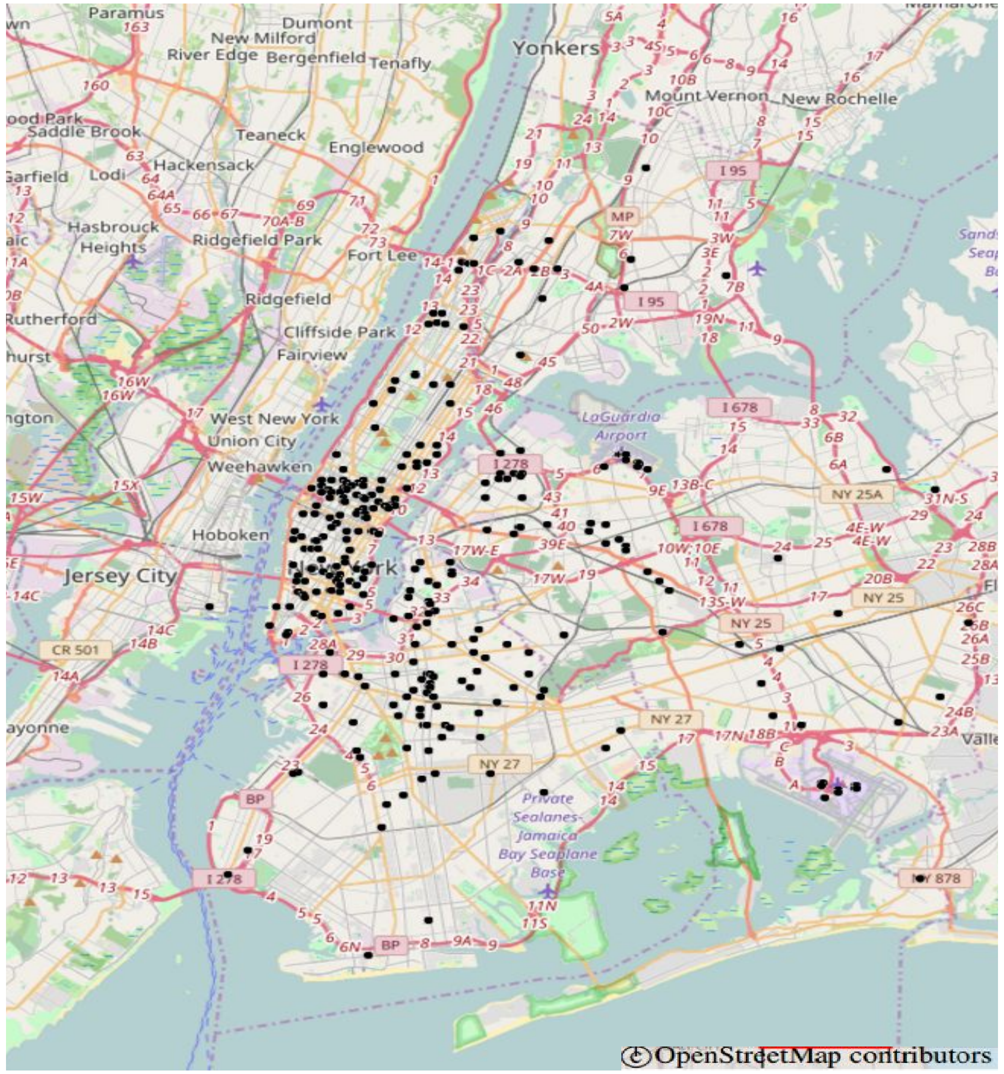
Outline

1. **Motivation and Approach**
2. **Ride Request Patterns**
3. **Ride Request Graphs (RRG)**
   a. Densification Power Law
   b. Temporal Evolution
4. **Societal Benefits**
   a. Poolability - Observations
   b. Densification and Poolability
5. **Conclusion and Future Work**

Distribution of pickup locations in New York

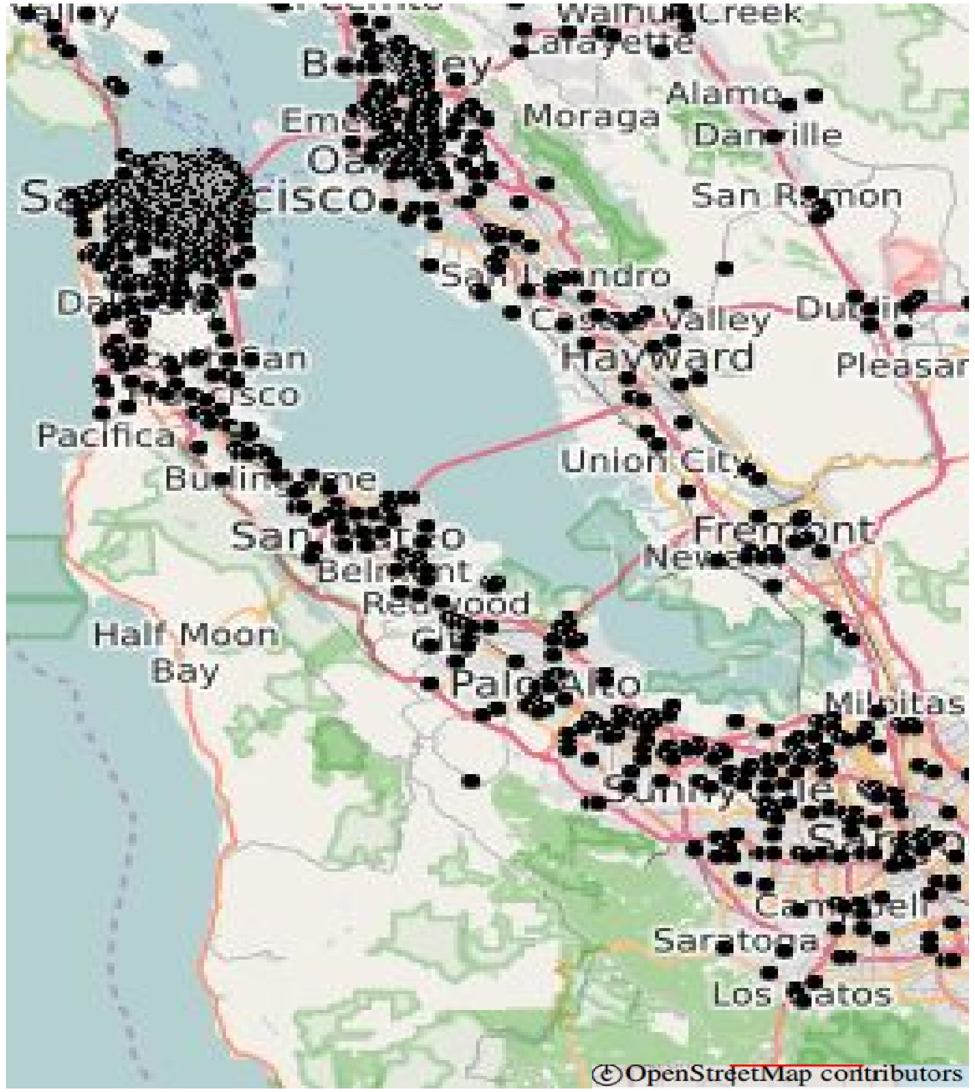


Between 8 & 8:05 pm

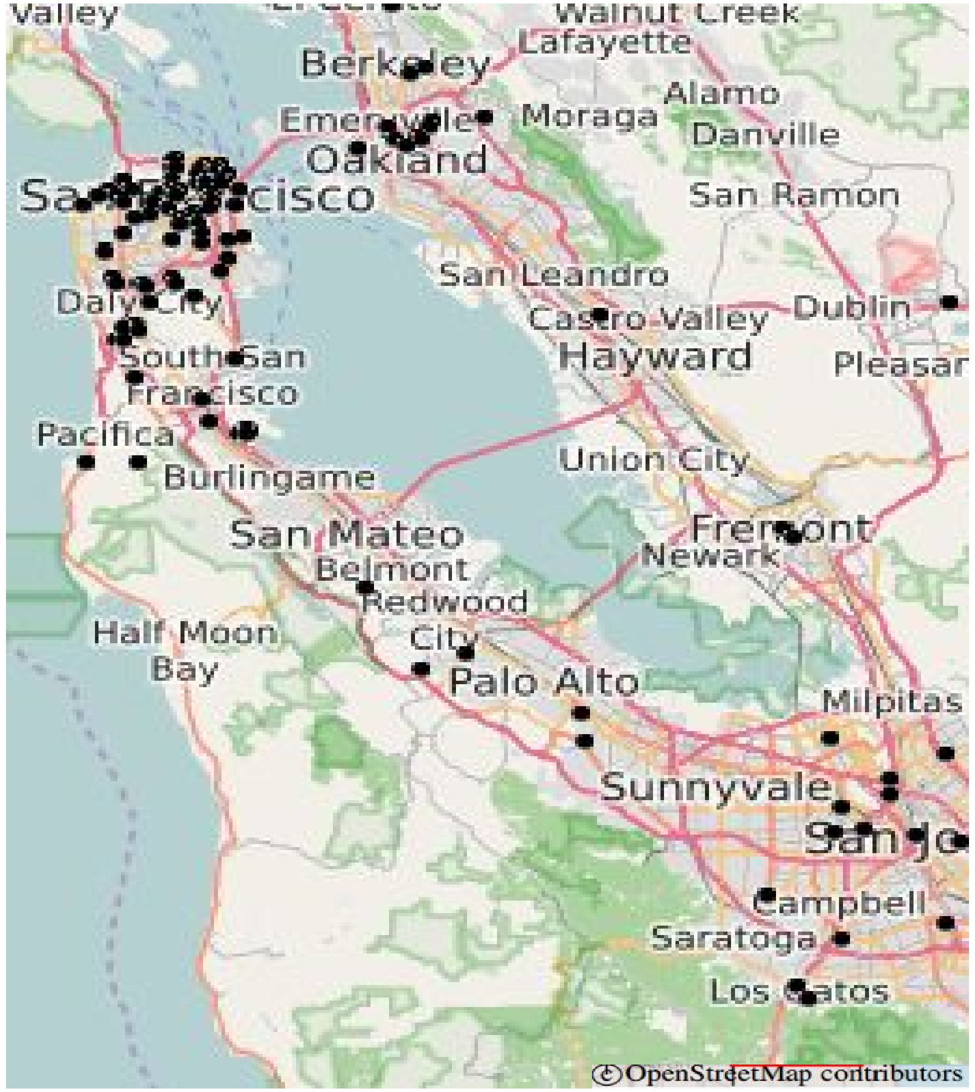


Between 3 & 3:05 am

Distribution of pickup locations in San Francisco

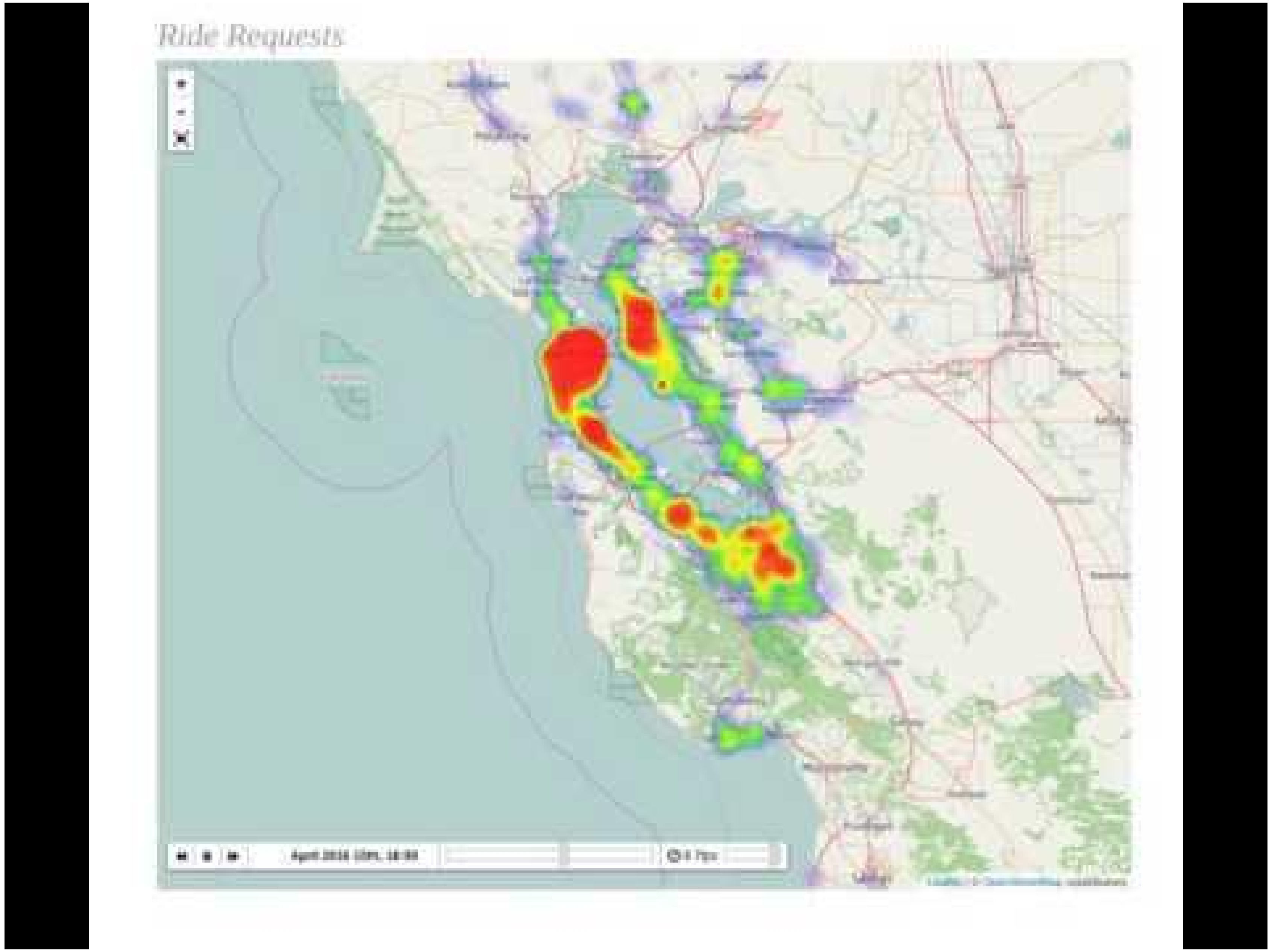


Between 7 & 7:05 pm



Between 5 & 5:05 am

Pickup Locations in San Francisco (one week)



<u>Observation</u>: There is significant variability in the **ride request patterns** from city to city, and across space and time within each city.

Question: Is there a rigorous model that can capture both the spatial and temporal variations of **ride request patterns** in a city?

Outline

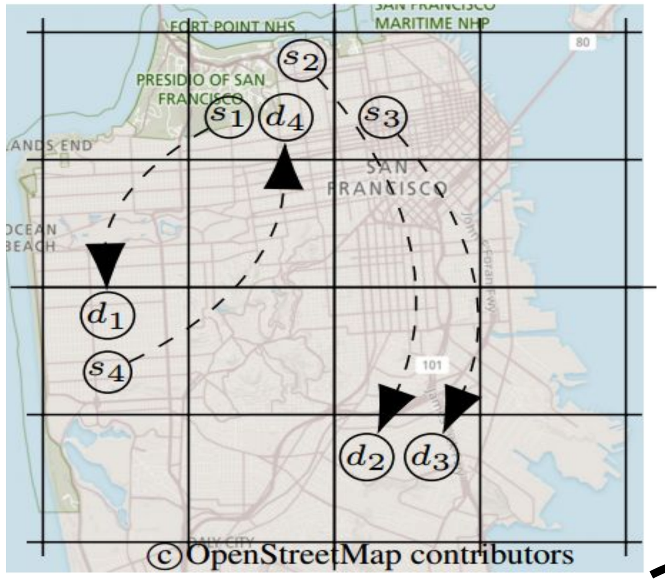
1. **Motivation and Approach**
2. **Ride Request Patterns**
3. **Ride Request Graphs (RRG)**
    a. Densification Power Law
    b. Temporal Evolution
4. **Societal Benefits**
    a. Poolability - Observations
    b. Densification and Poolability
5. **Conclusion and Future Work**

Ride Request Definition

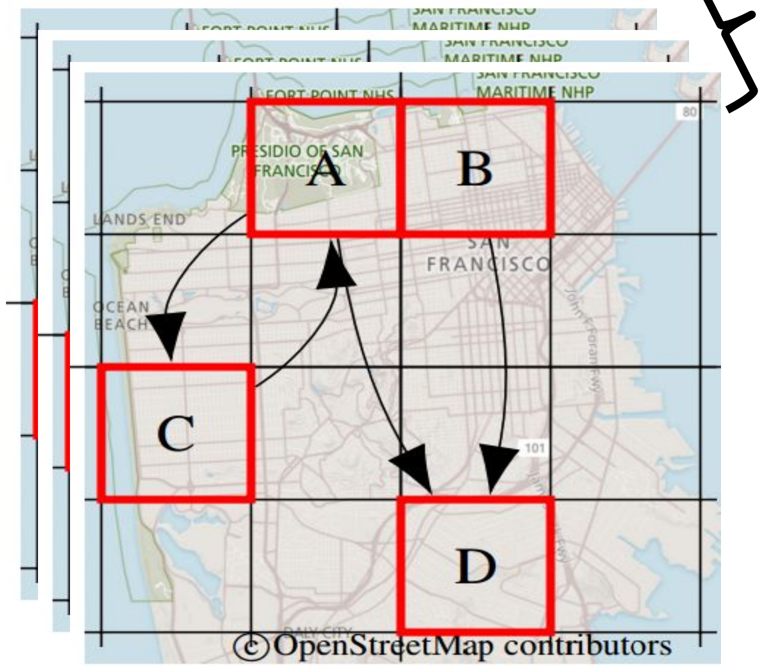
Each ride request is defined by:

1) Pickup location: s = <latitude, longitude>
2) Dropoff location: d = <latitude, longitude>
3) Time of request: t = <timestamp>

Ride Request Graph (RRG)



- Four ride requests distributed spatially over a map at a given short time interval.

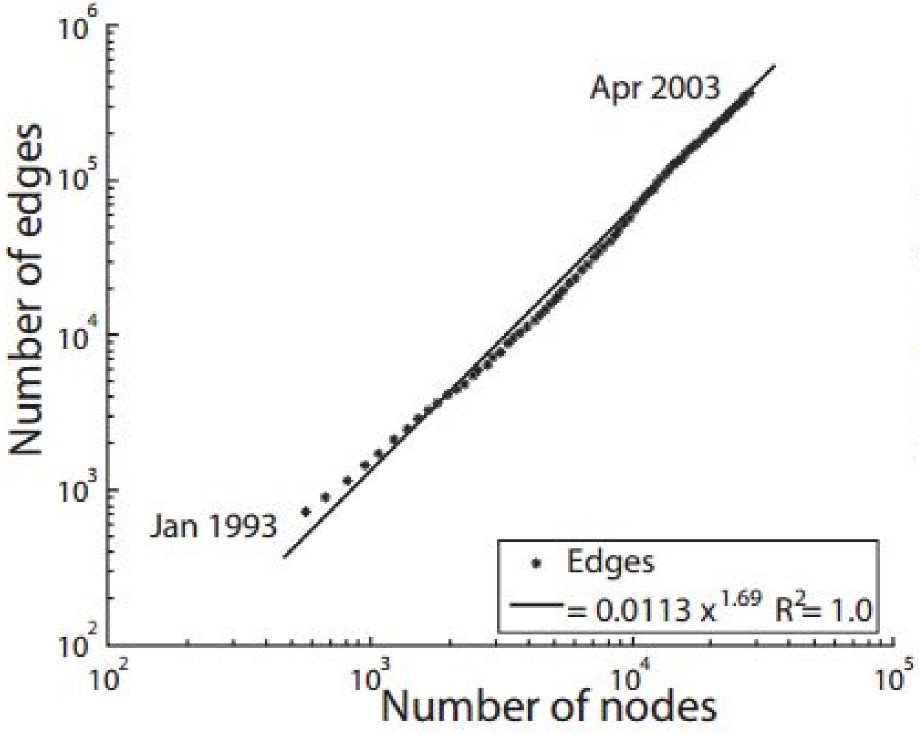


Sequence of RRGs over time intervals

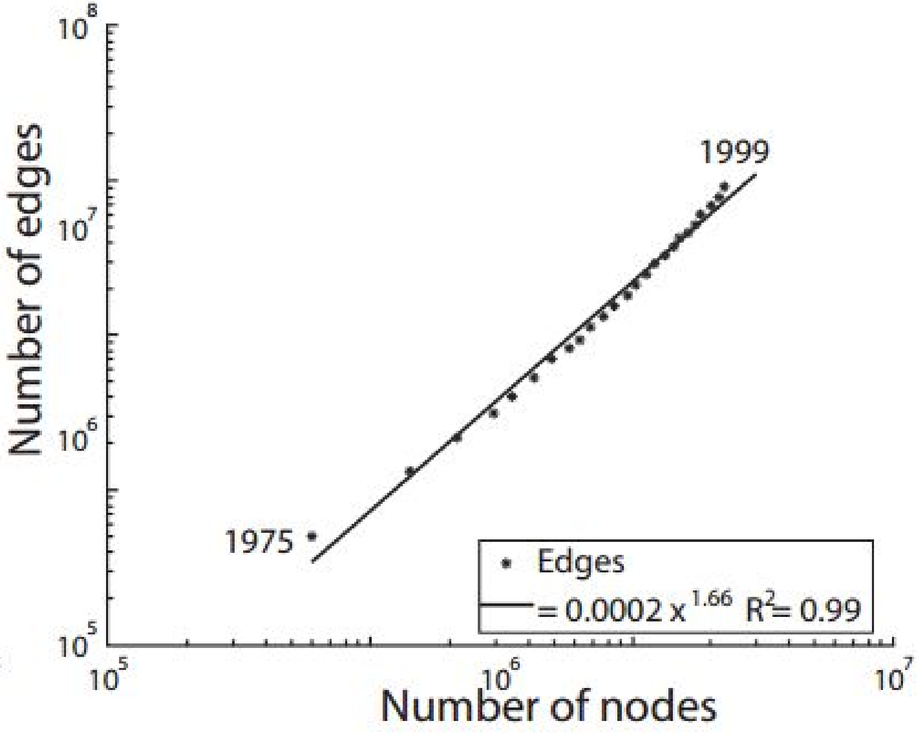
- Corresponding RRG with four nodes and weighted directed edges
  - Nodes correspond to pickup or dropoff locations
  - Edges connect pickup nodes to drop-off nodes

Densification Power Law (DPL)

Time-evolving graph like arXiv citation graph, the Patent citation graph, social network graph, and many others share a common property i.e. DPL.



(a) arXiv



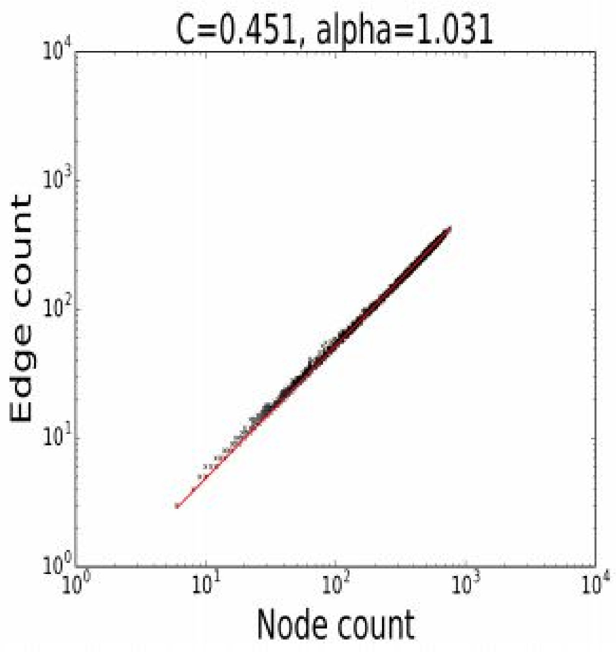
(b) Patents

Image source: Chakrabarti, D., & Faloutsos, C. (2006). Graph mining: Laws, generators, and algorithms. *ACM computing surveys (CSUR)*, *38*(1), 2.

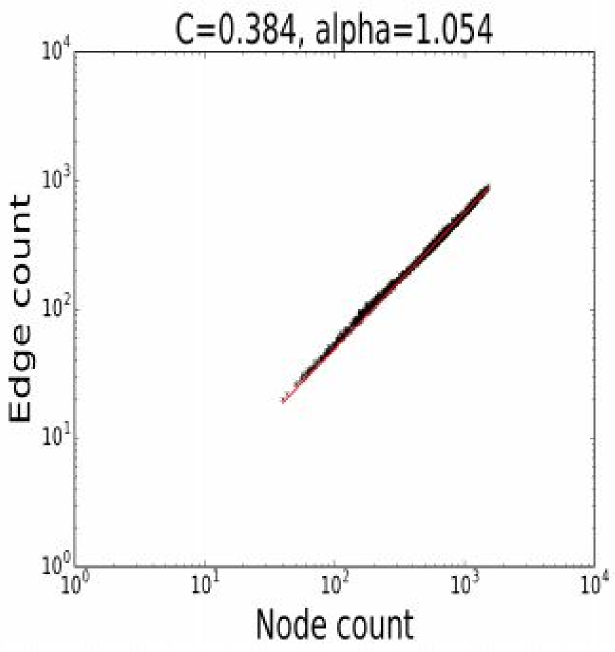
Densification Power Law property of RRG

What is the relation between the number of edges E(t) and number of nodes N(t) for any given time t?

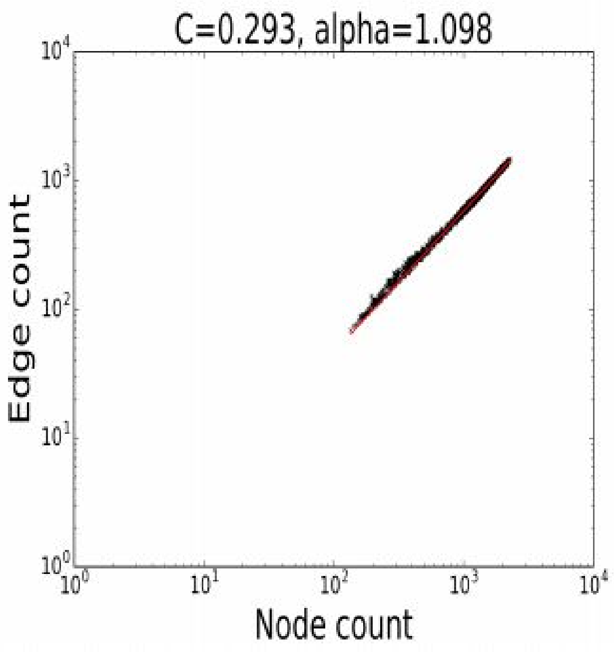
$$E(t) = C * N(t)^{\alpha}$$



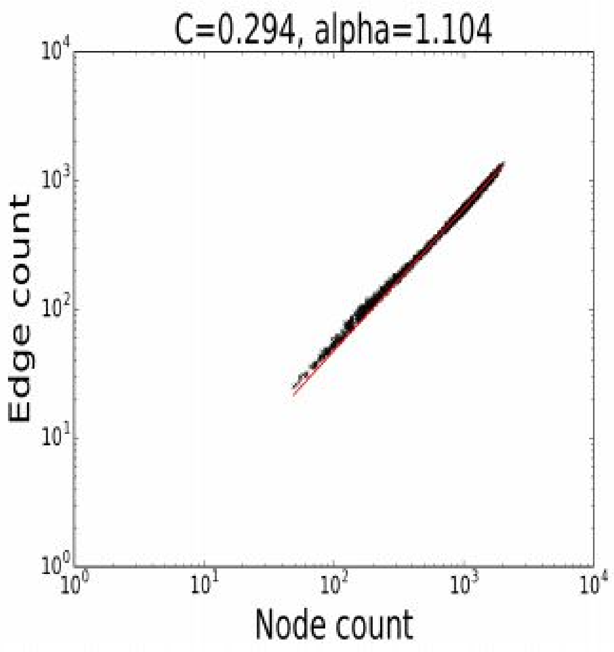
Hyderabad



Paris



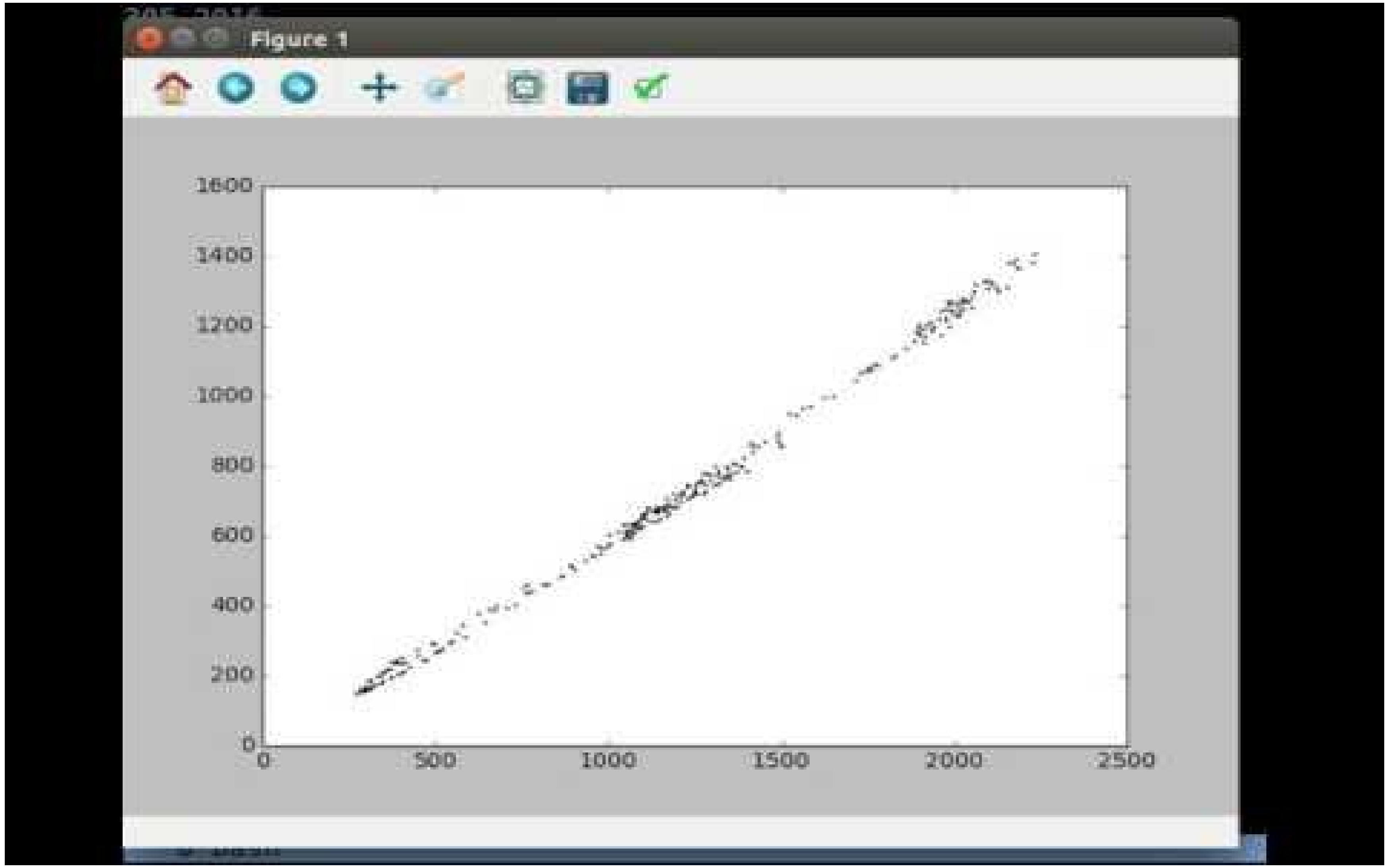
New York



San Francisco

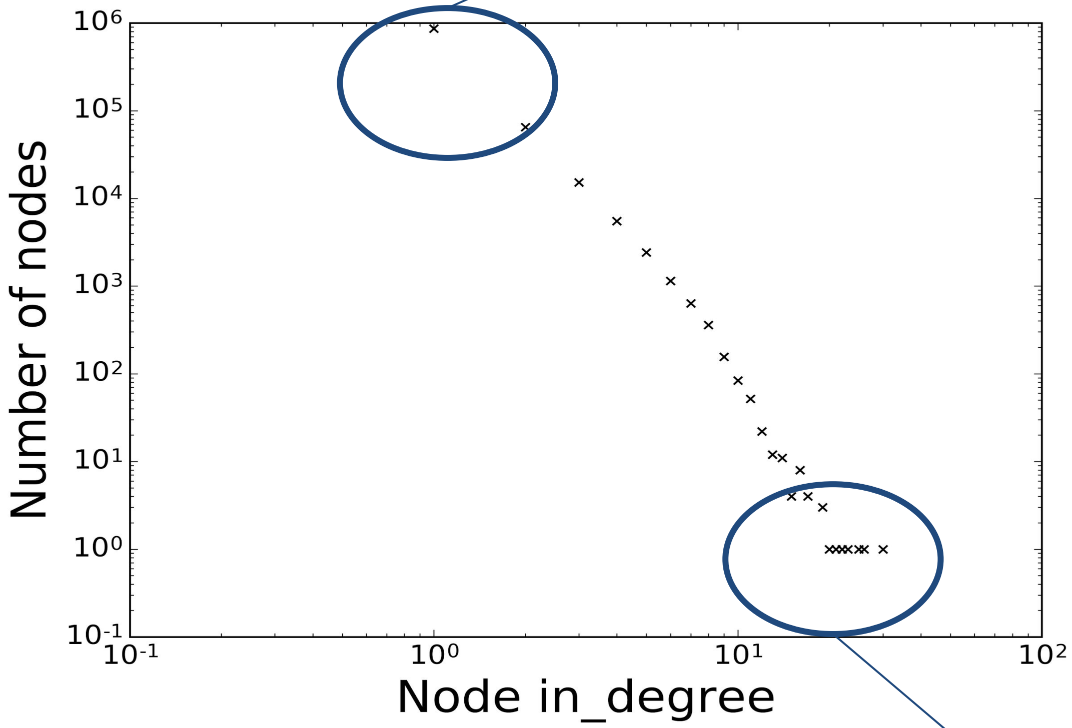
Temporal Evolution of Ride Requests

Important distinction from orthodox graphs: Nodes & edges may not persist across time intervals.

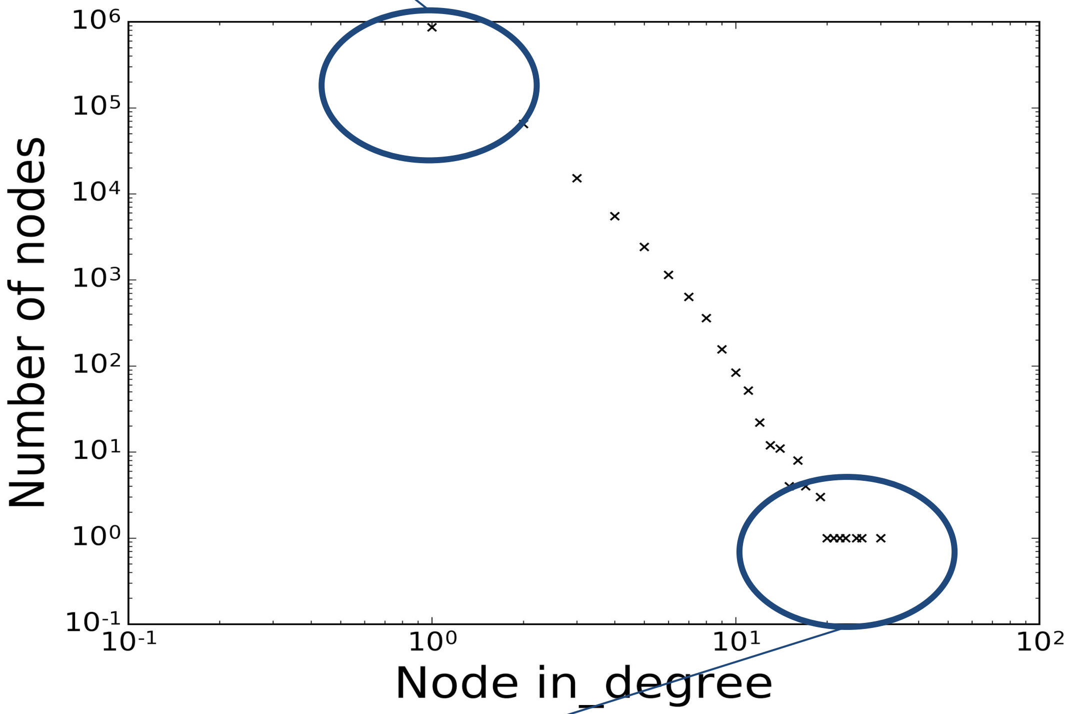


DPL Property => Community Effect

Many nodes with low number of dropoffs (mostly single).



New York



San Francisco

Very few nodes with high number of dropoffs.

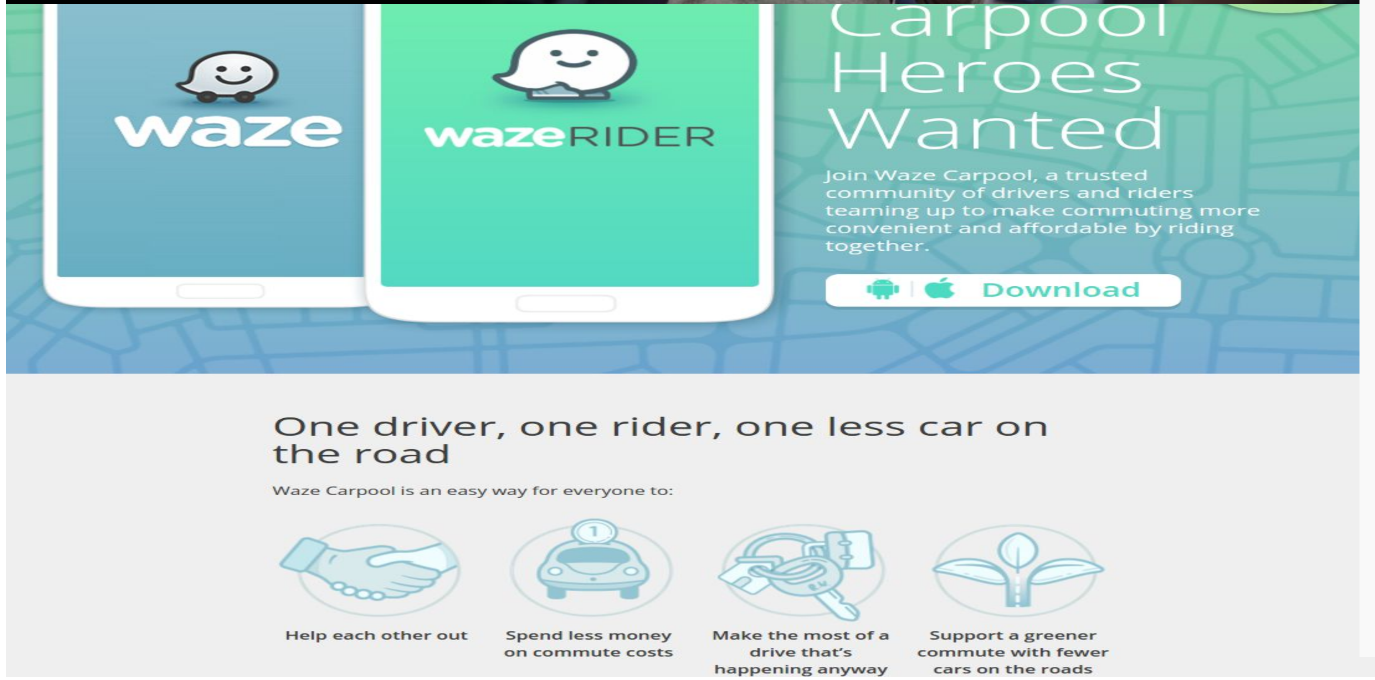
Similar plots for out degree.

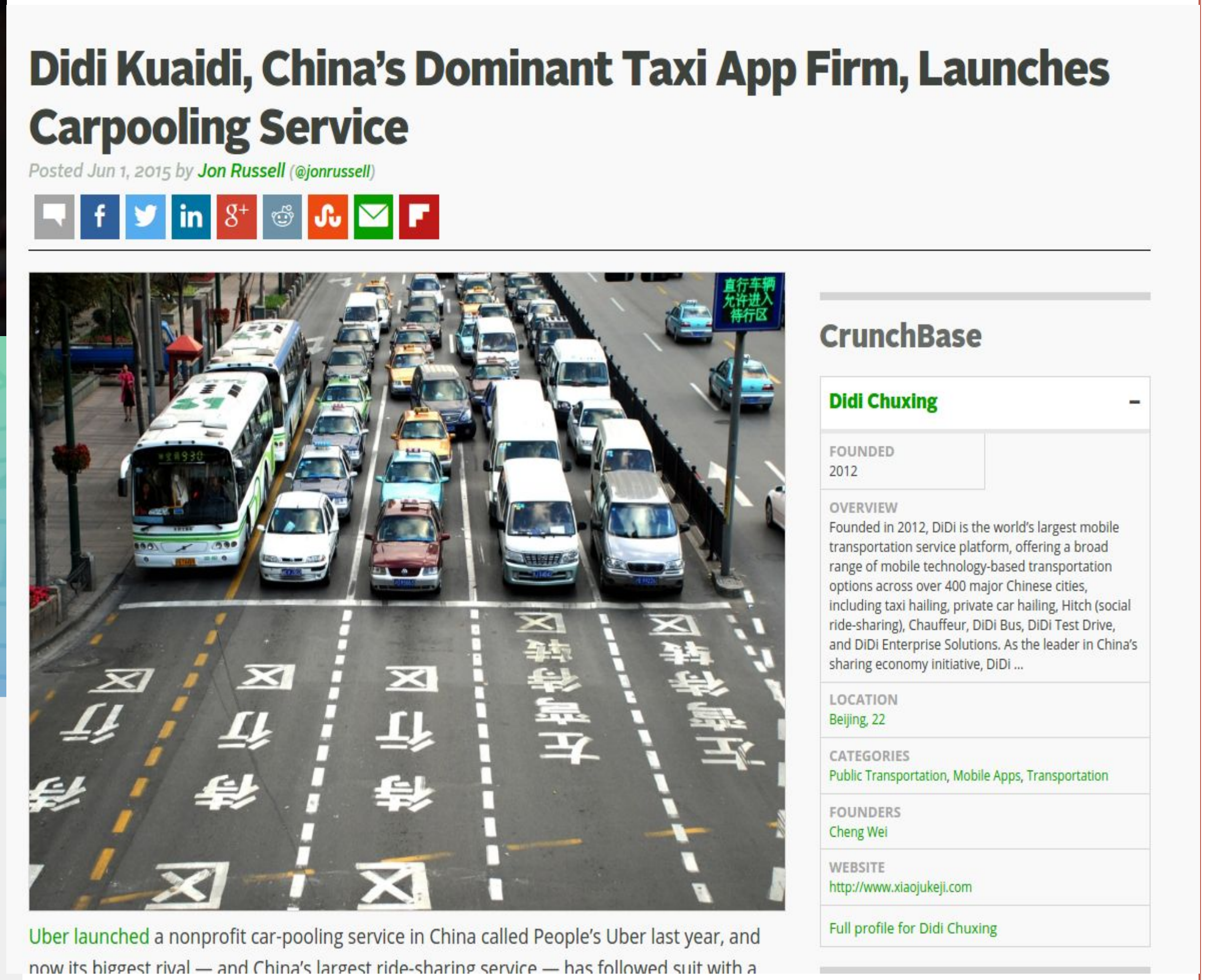
Question: How can such modeling help identify and interpret potential societal benefits?

Outline

1. **Motivation and Approach**
2. **Ride Request Patterns**
3. **Ride Request Graphs (RRG)**
   a. Densification Power Law
   b. Temporal Evolution
4. **Societal Benefits**
   a. Poolability - Observations
   b. Densification and Poolability
5. **Conclusion and Future Work**

All major ride-sharing services are introducing **"ride pooling"** to reduce number of vehicles needed on the roads and to increase the overall efficiency.



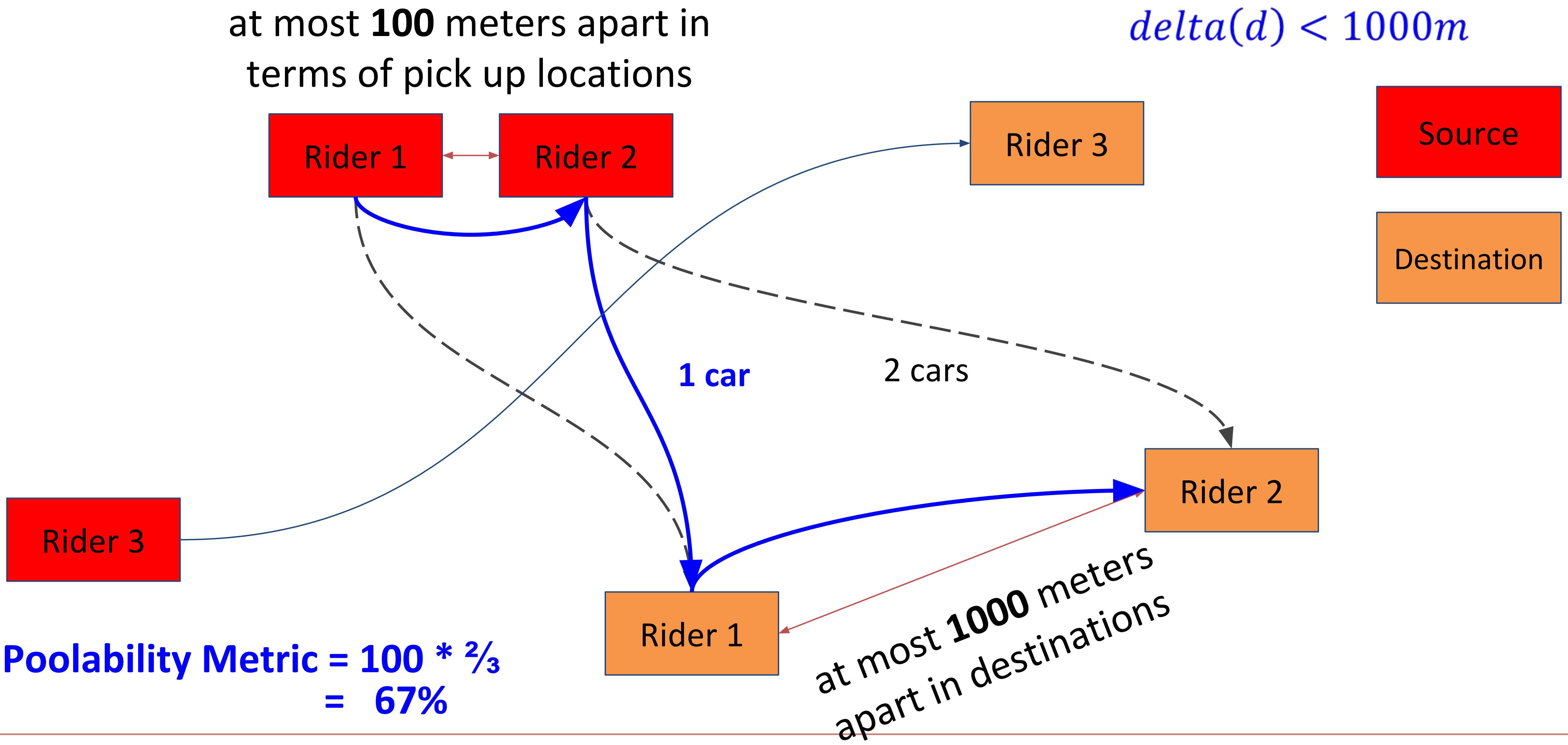


What is "Ride Poolability"?

Each Ride Request: <*t,s,d*>

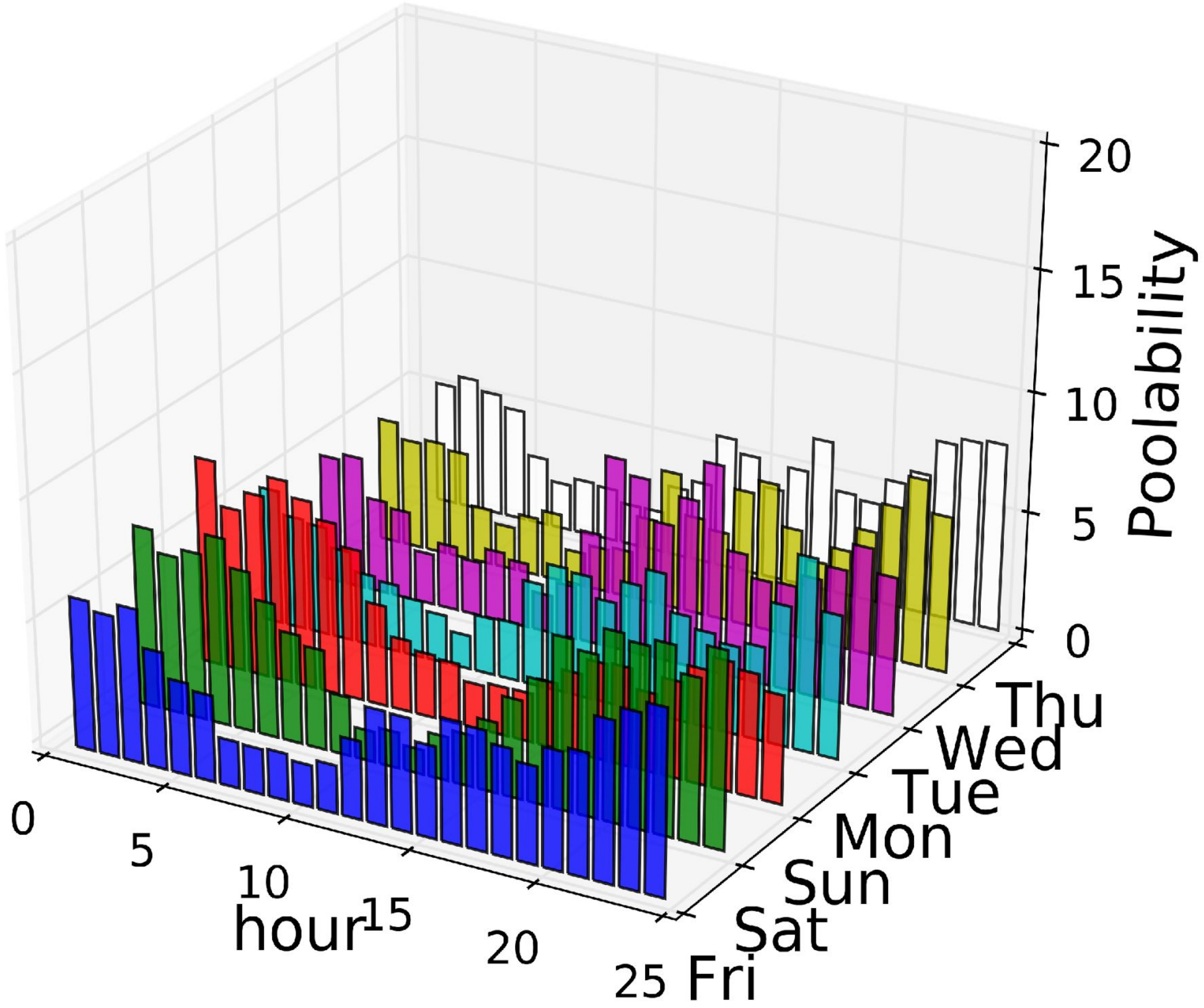
Proximity Constraints for Pooling:

$$delta(t) < 5min$$
$$delta(s) < 100m$$
$$delta(d) < 1000m$$

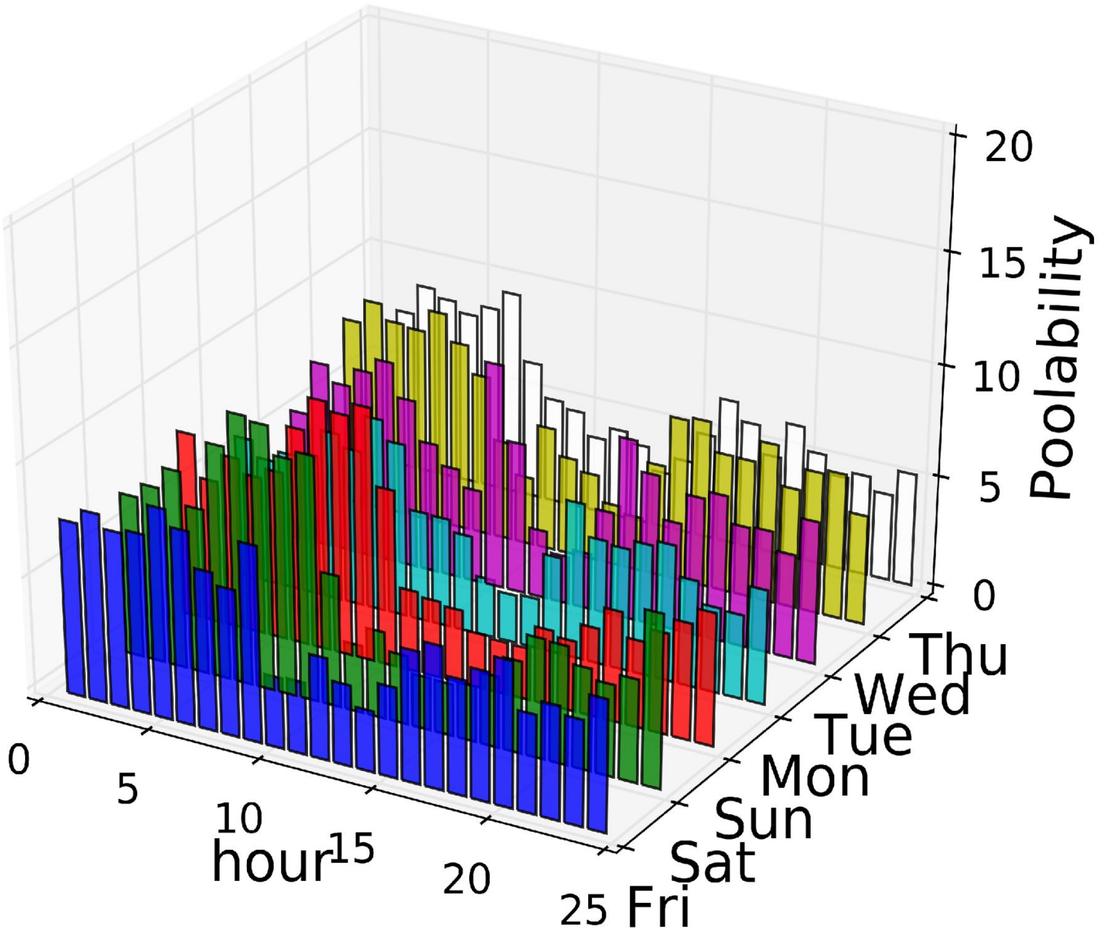


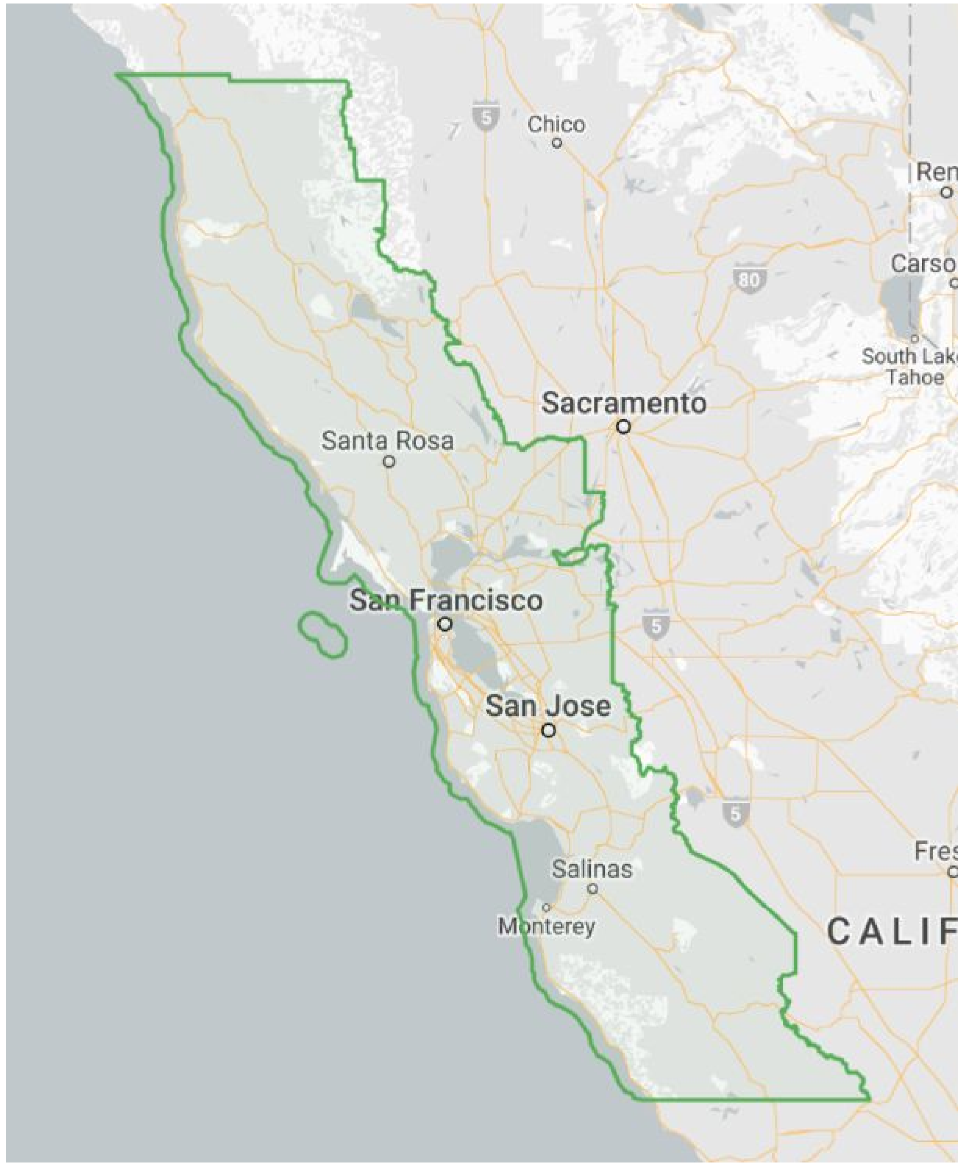
Poolability Metric = 100 * ⅔
= 67%

Ride Poolability Profile for New York



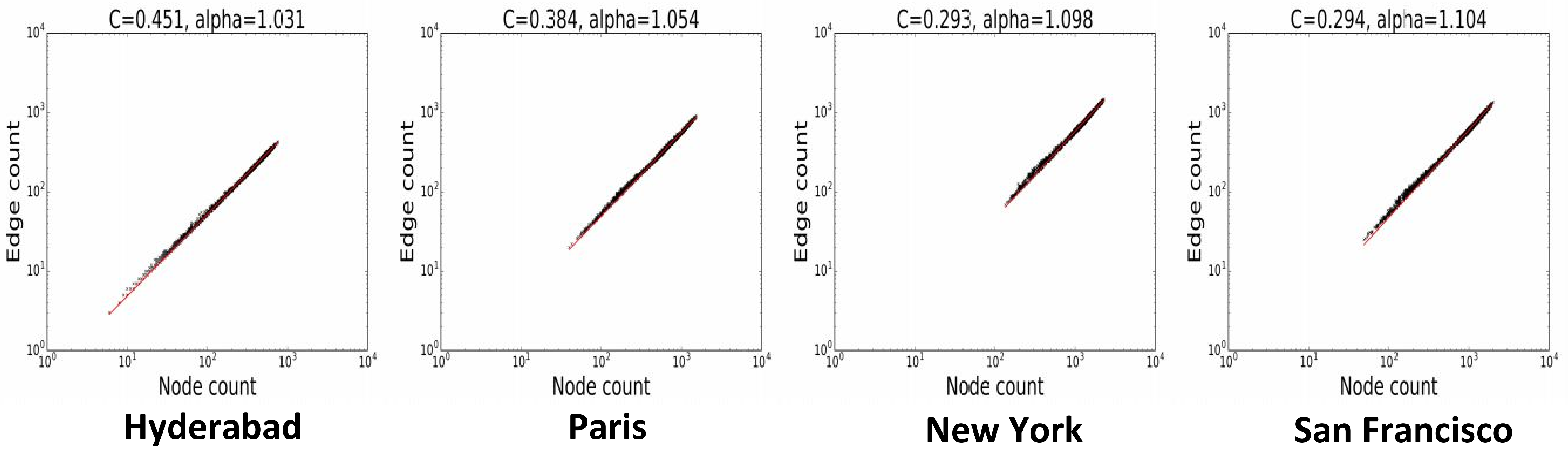
Ride Poolability Profile for San Francisco





Question: What is the correlation, if any, between the **ride request patterns** of a city and the **ride poolability profiles** of that city?

DPL and Ride Poolability



Alpha is correlated with ride poolability due to the **"human community"** effect.

City	Alpha	Mean Poolability
Hyderabad	1.031	2.23
Paris	1.054	2.39
New York	1.098	4.48
San Francisco	1.104	5.48

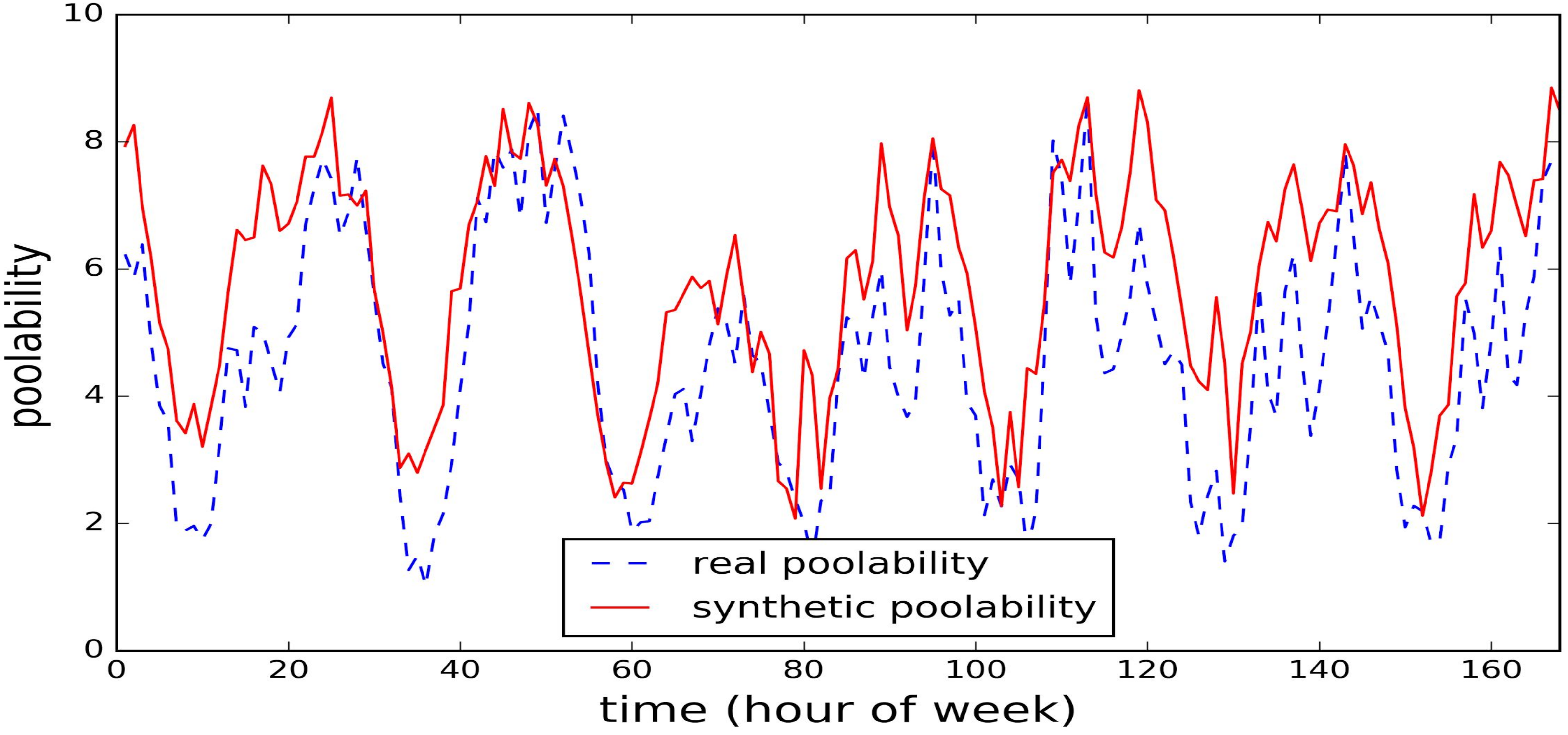
Outline

1. **Motivation and Approach**
2. **Ride Request Patterns**
3. **Ride Request Graphs (RRG)**
   a. Densification Power Law
   b. Temporal Evolution
4. **Societal Benefits**
   a. Poolability - Observations
   b. Densification and Poolability
5. **Conclusion and Future Work**

Conclusions

1. We study the evolution of ride requests over space and time based on extensive real world data.
2. We introduce the Ride Requests Graph model that captures the temporal and spatial variations of ride requests.
3. We discover:
   a. Densification Power Law applies to Ride Request Graphs.
   b. Poolability metric is correlated with the exponent of DPL.
4. We propose an algorithm for generating synthetic RRGs that exhibit similar DPL metrics as RRGs extracted from real data.

New York Poolability Results



Promising avenues for future research. Some research questions:

1. If the ride pooling proximity constraints, both temporal and spatial, can be relaxed, is it possible to significantly improve ride poolability?
2. Can the pattern of temporal and spatial variation of ride poolability be leveraged to create intelligent predictive ride pooling algorithms?
3. Can we rigorously characterize the relationship between ride poolability and ride request graph densification power law factor?
4. Is it possible to use the space-time graph model extracted from historical data to perform real-time traffic congestion prediction and alleviation, as well as real-time accurate travel time prediction?

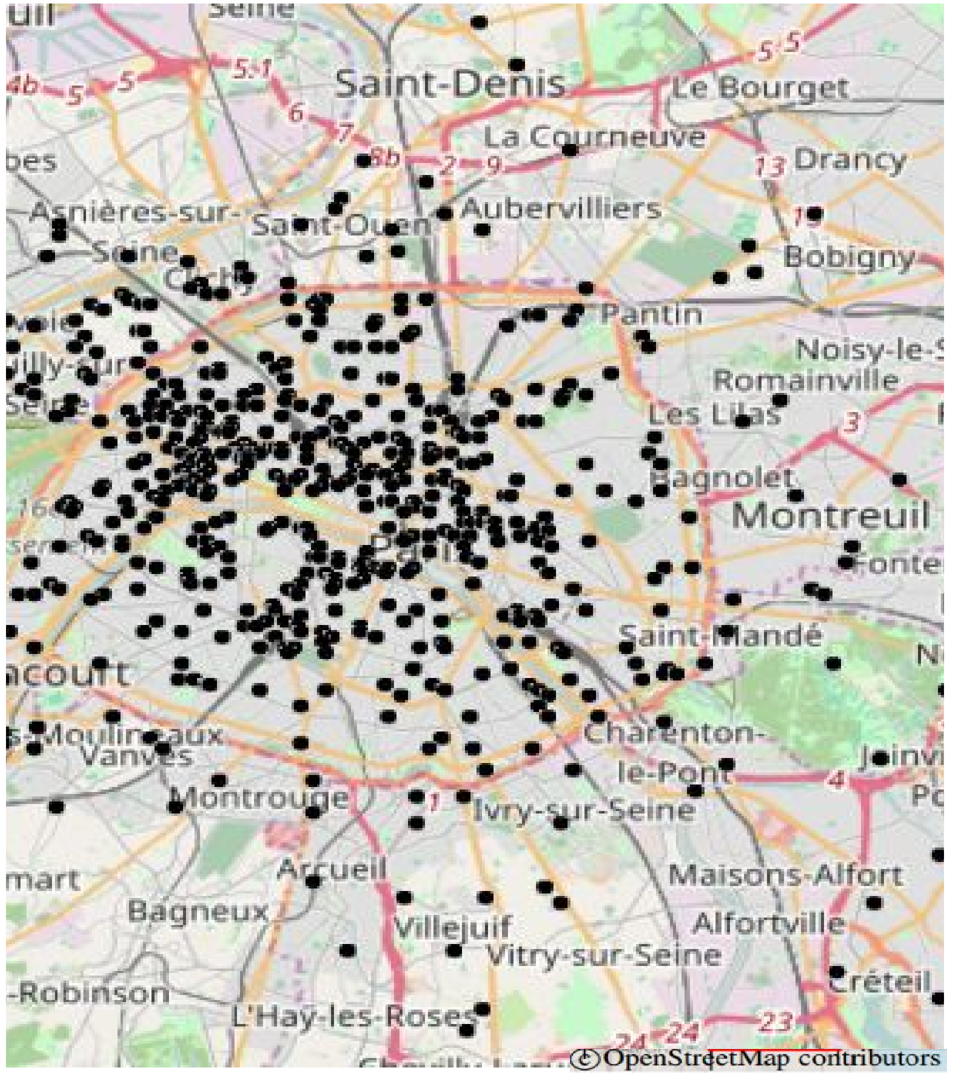
Questions?

ajauhri@cmu.edu

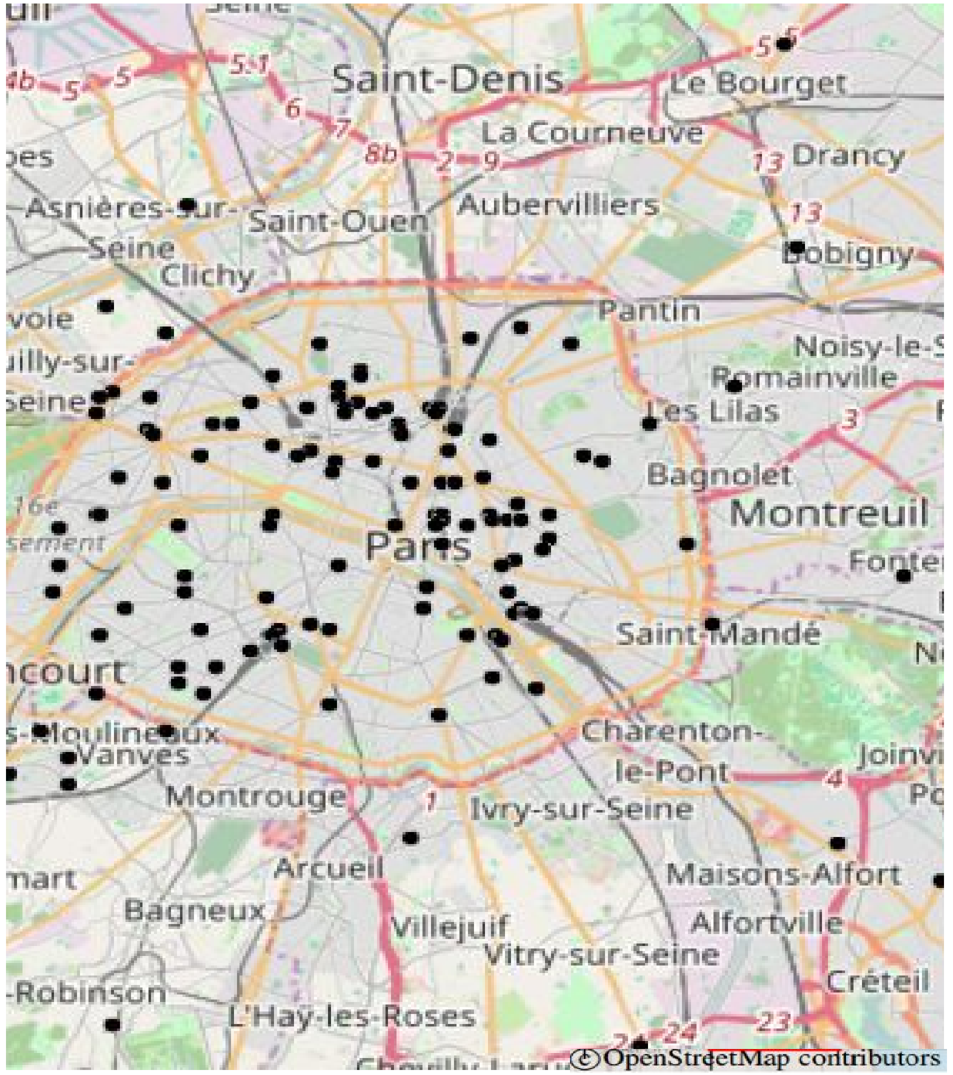
Paper: http://ece.cmu.edu/~ajauhri/publications.html

Backup Slides

Distribution of pickup locations in Paris

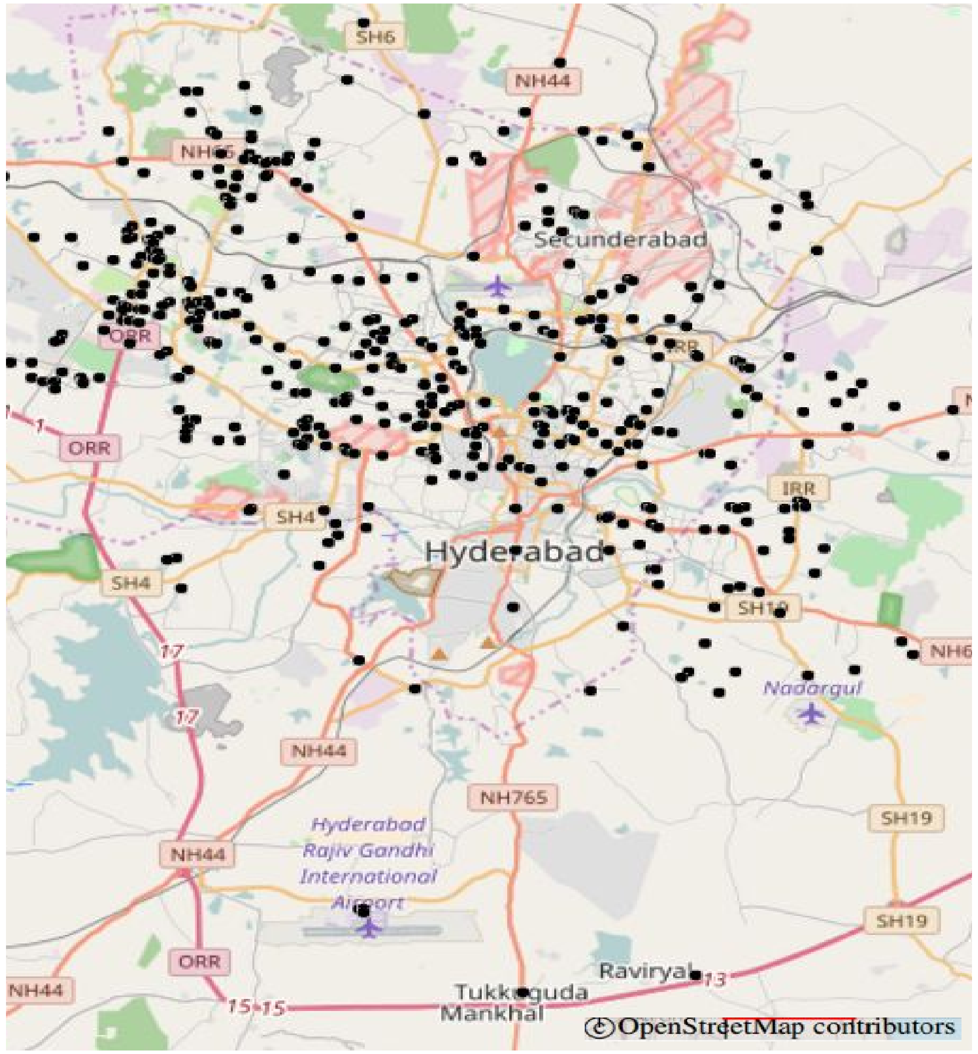


Between 2 & 2:05 pm

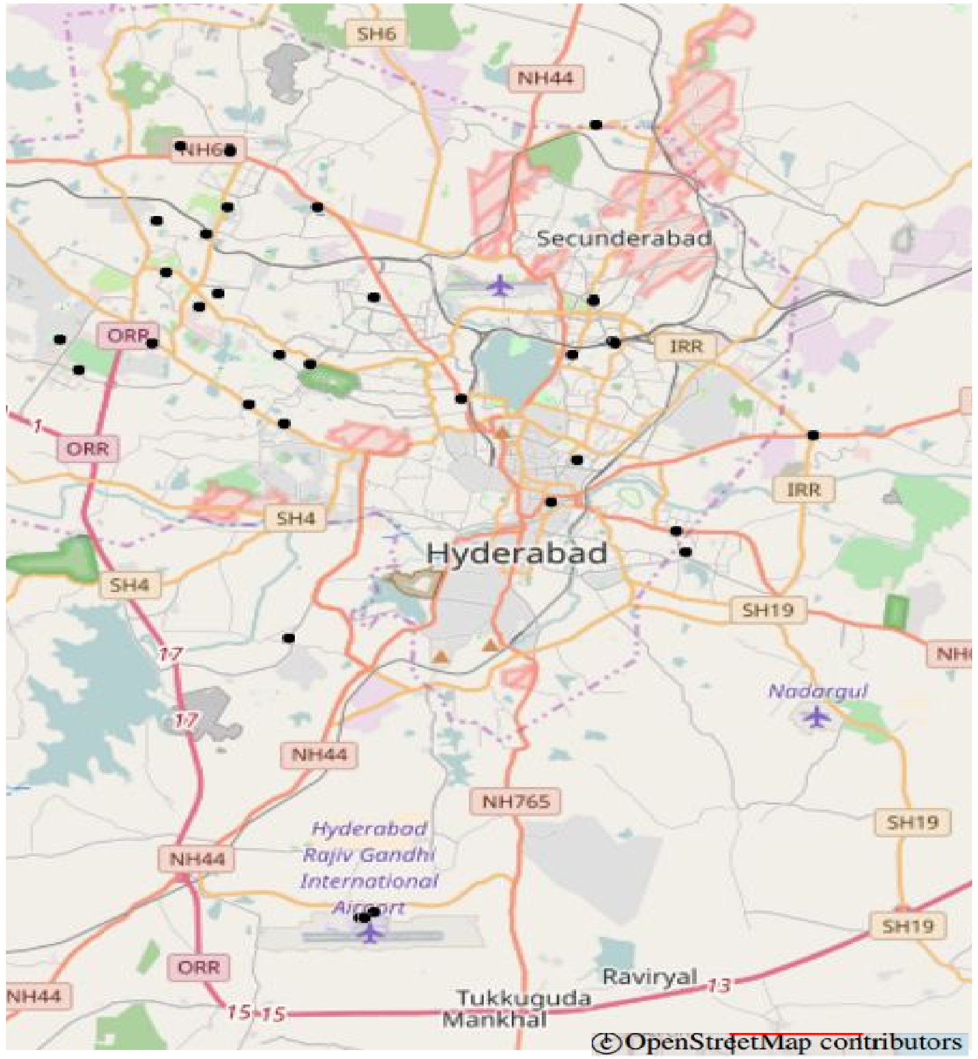


Between 4 & 4:05 am

Distribution of pickup locations in Hyderabad

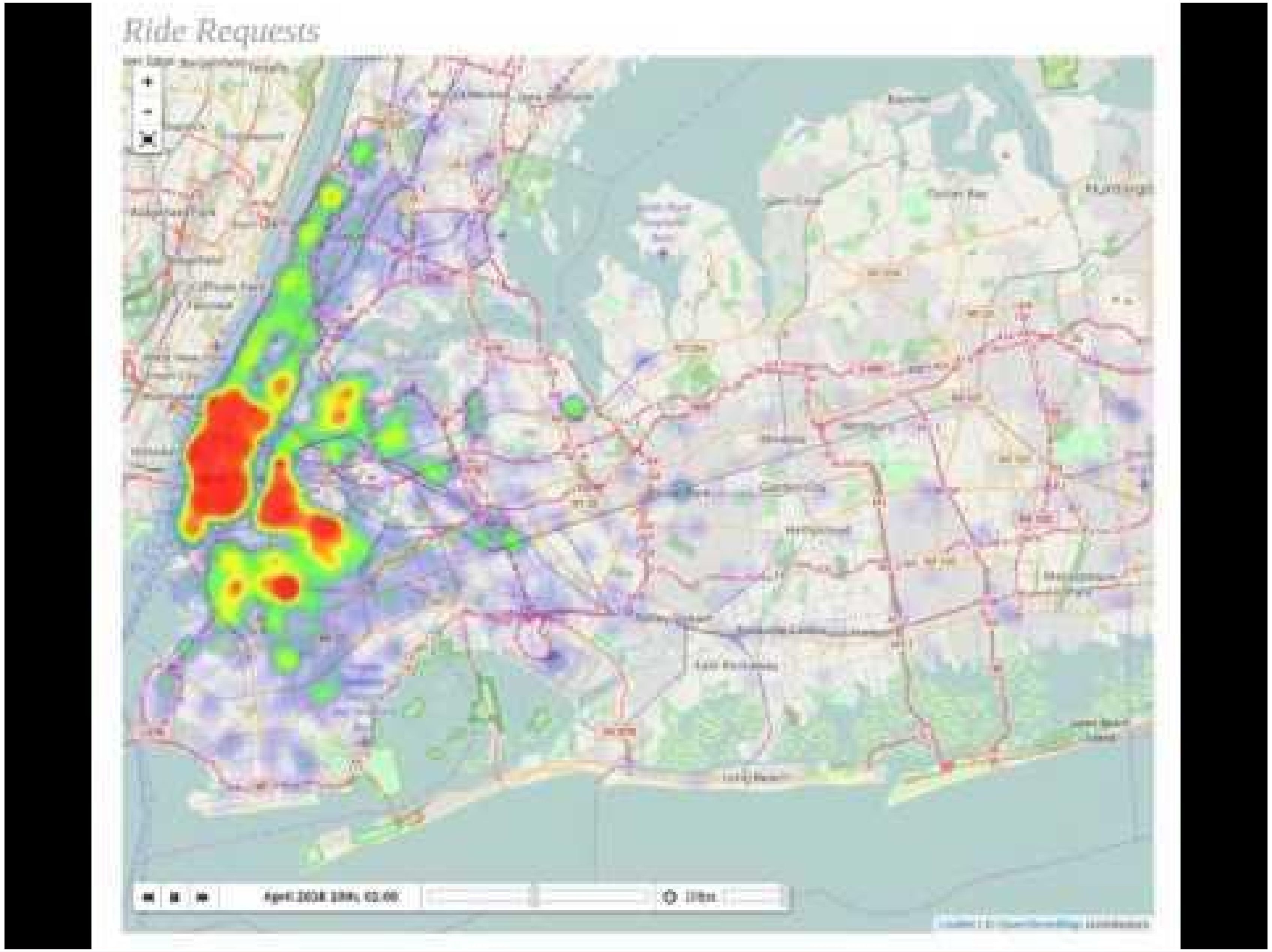


Between 11 & 11:05 am



Between 5 & 5:05 am

Pickup Locations in New York (one week)



Question: Is it possible to auto generate RRGs that exhibit the same DPL metrics as those RRGs extracted from actual real data?

Synthetic RRG Model - Intuition

How to emulate human travel patterns?

1. Find densely populated regions in a city, and mark them as pickup points.
2. Connect pickup or dropoff points with high probability from densely populated regions of a city, and with low probability from sparsely populated regions.
3. From time to time visit pickup or dropoff points which have already been visited before.

We used OSM’s public data on node density to distinguish between dense and sparse regions of a city

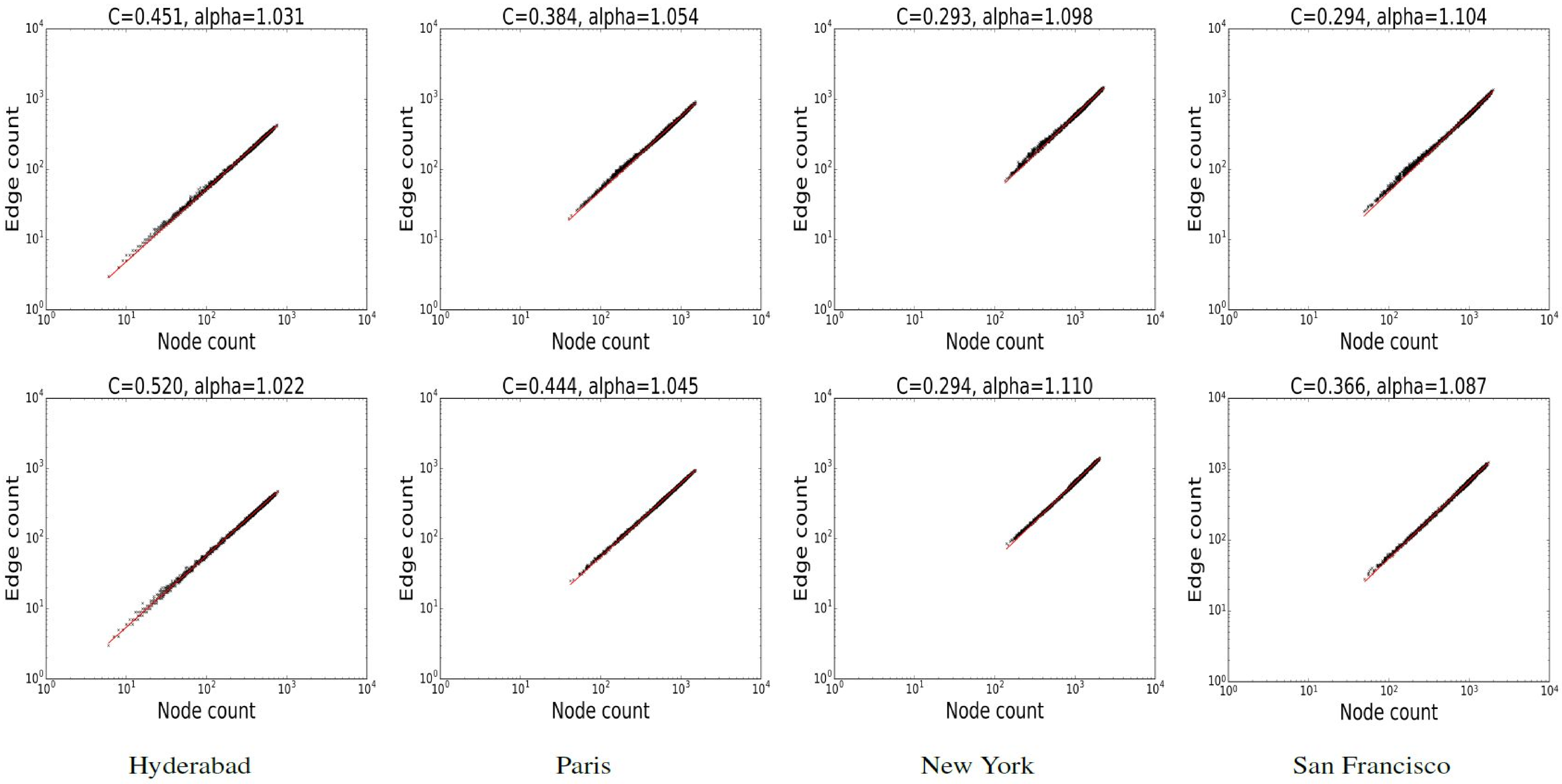


Figure 4: DPL plot from real data (top row) and synthetic data (bottom row) for four cities. The red line is the least square fit of the form $y = Cx^{\alpha}$, where y and x are number of edges and nodes respectively. $R^2 \approx 1.00$ for all of them.

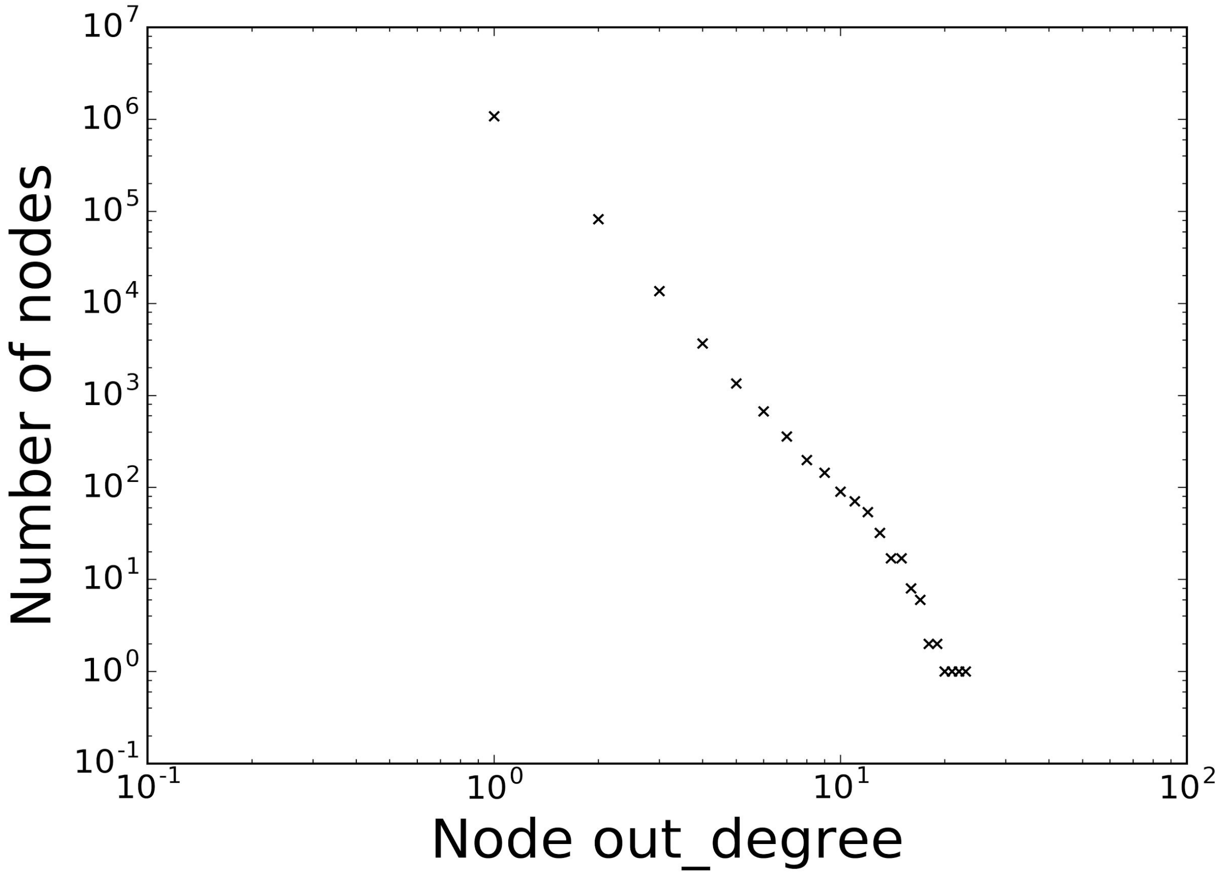
Auto generation property of DPL graphs

- → Graphs exhibiting DPL can be automatically generated
- → Auto generation is based on a small number of parameters

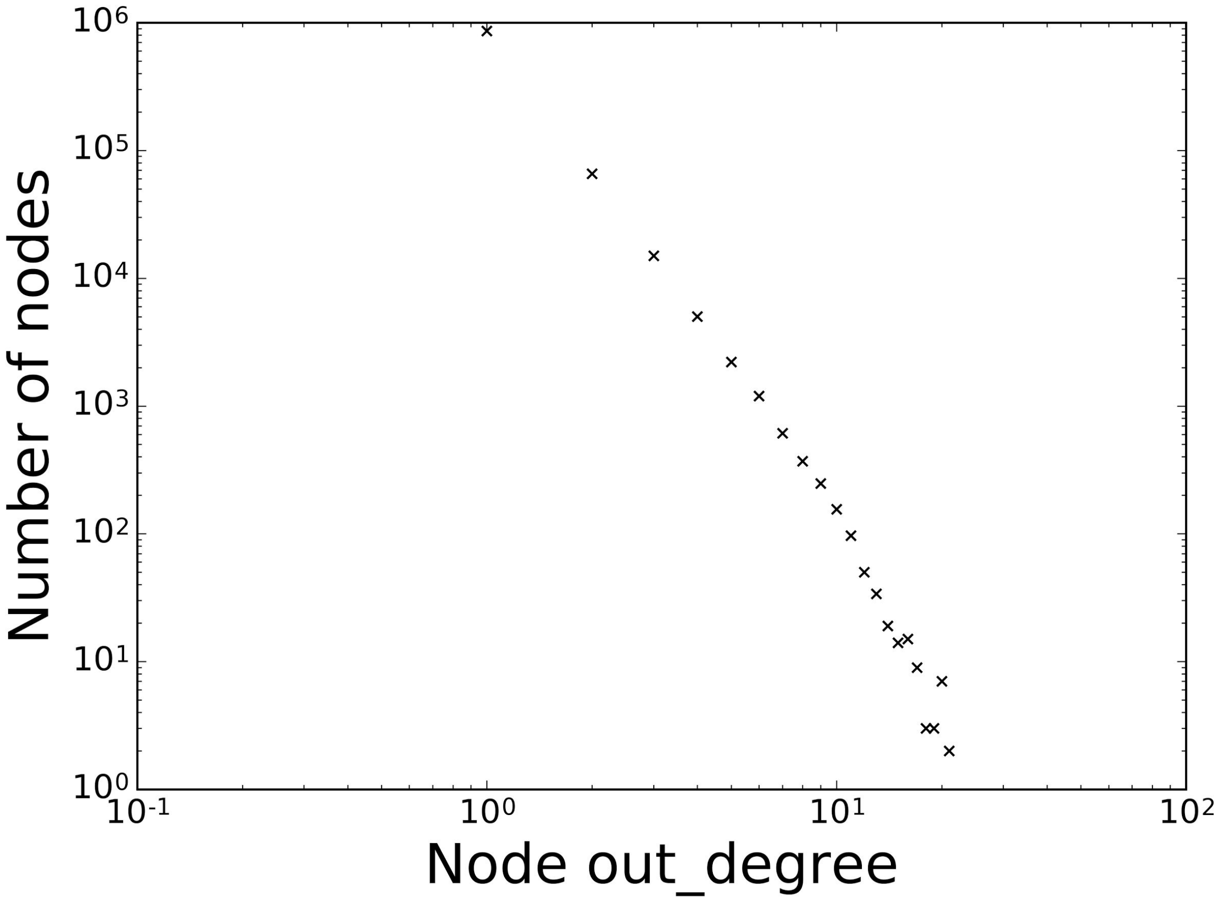
Desiderata for auto generation of RRGs

- → Generated graphs should exhibit very similar attributes of interest with graphs extracted from real data
- → Should exhibit the same **human community** effects
- → Should be efficient and scalable

DPL property - Community Effect (2)

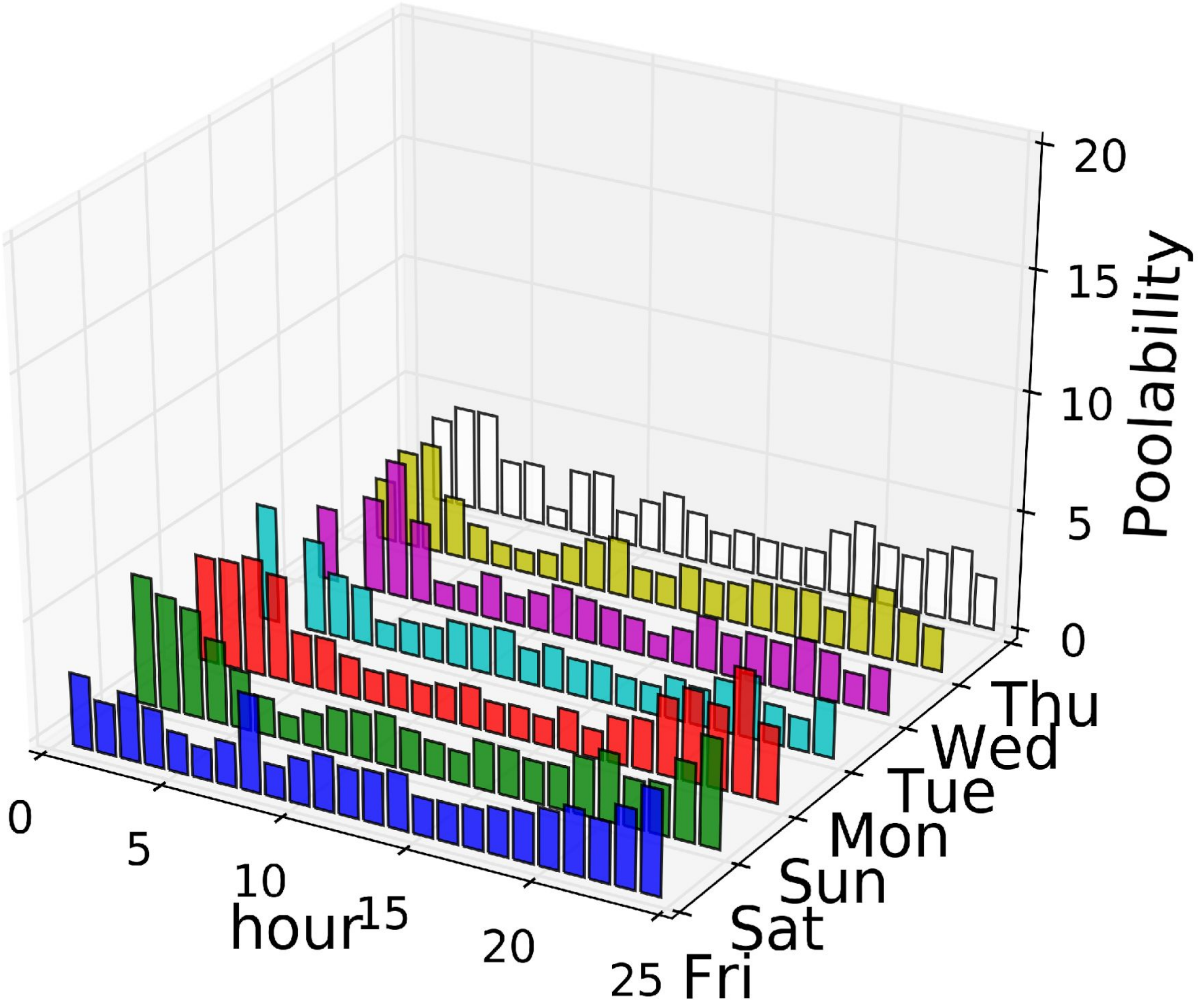


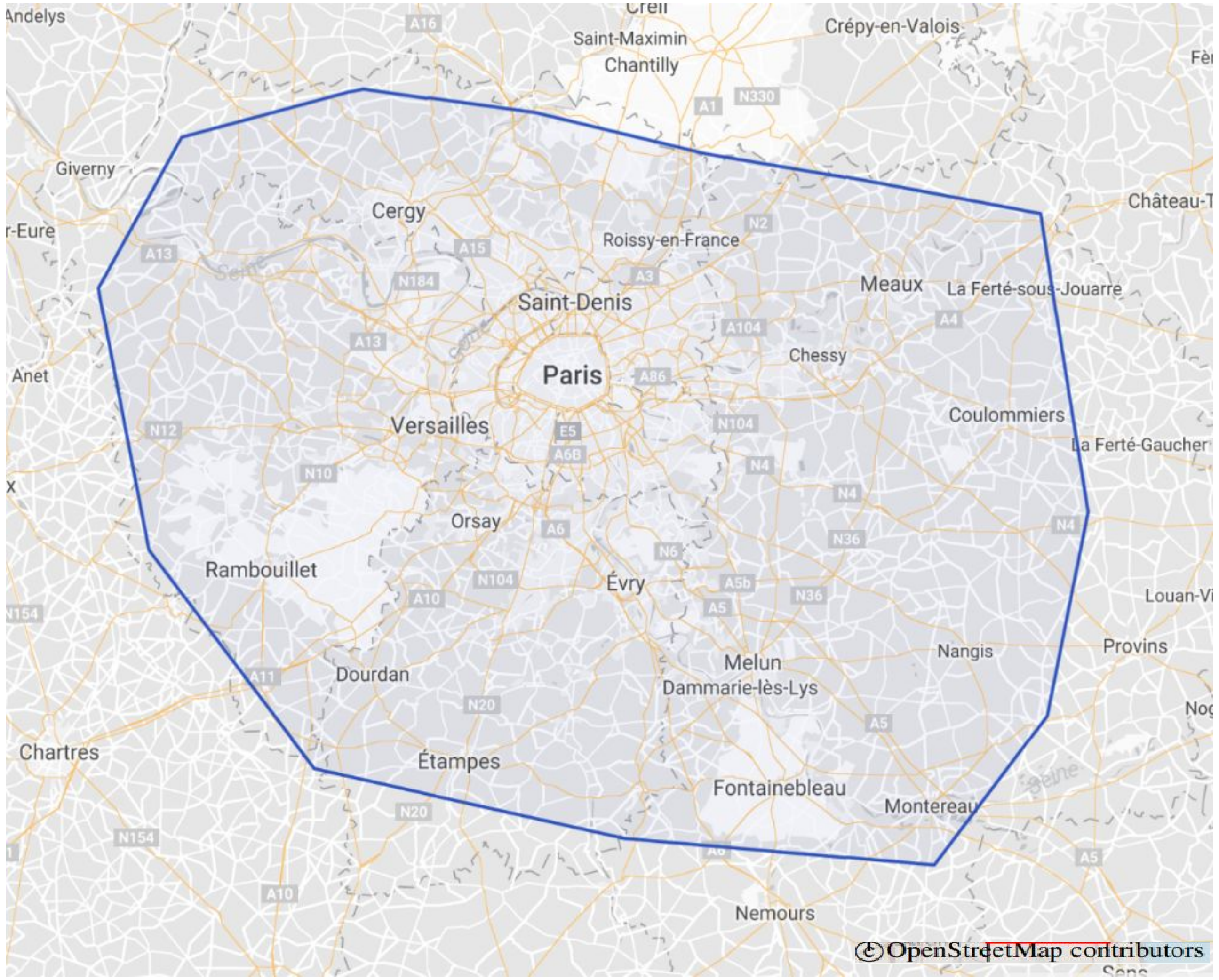
New York



San Francisco

Ride Poolability Profile for Paris





Ride Poolability Profile for Hyderabad