Routing in human interaction networks

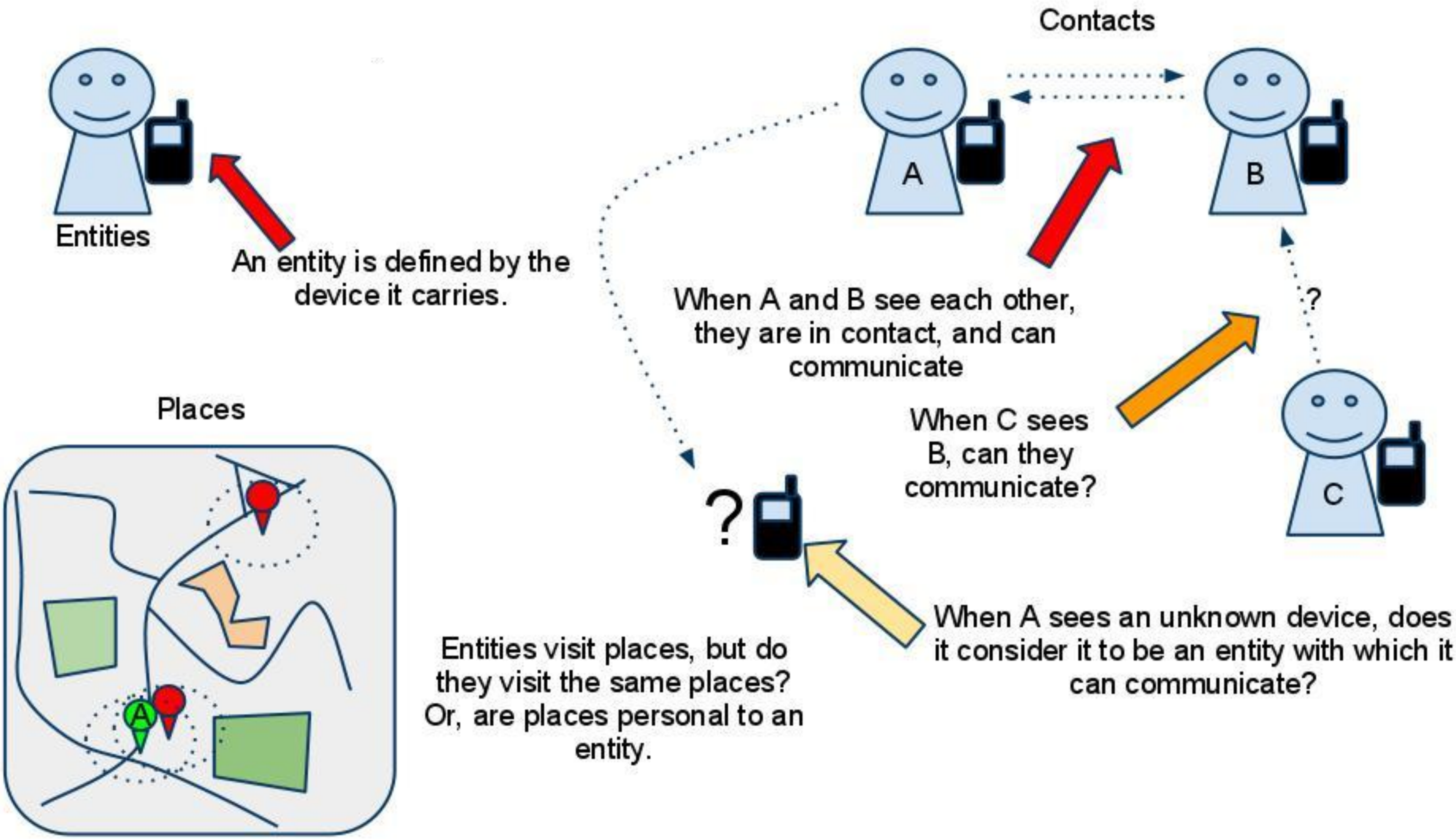
?

- Sending messages *without the need* for infrastructure
  - Low cost, low importance
  - Network specific data
  - Sensor networks?
- Utilising interaction patterns
  - Contacts
  - Locations
- Tolerating delays
  - Sparse network
  - Intermittent connectivity

Outline

- Metrics for routing
- Existing techniques
- Datasets for Simulation
- Persistence Based Routing in detail
  - Results
- Location routing (LBR)
  - Results

Entities, Contacts and Places



Structures

- Visited places graph
  - Node where a person visits
  - Directed edge between places that a person visits
  - Time ordering of
  - Aggregate graph of all visited places

- Encountered individuals graph
  - Node is a person
  - un-directed edge between nodes that have met
  - Aggregrate graph of all node encounters

Metrics for routing

4.1 Encounter age

Rationale	Nodes that have been encountered recently are likely to be encountered again soon.
Properties	– Small values (recent contact) are good indicators of proximity. – As values get larger, they become poorer.
Algorithms	PROPHET [61], Seek/Spray and Focus [79, 81]

4.2 Inter-contact time (ICT)

Rationale	In order to minimise delay, route towards nodes that are likely to connect again soon.
Properties	– Aim to minimise delivery delay, rather than explicitly maximise delivery ratio. – Short, bursty connections may lower inter-contact times while total duration of contacts remains low.
Algorithms	MED/MEED [51, 53], RAPID [7], MH* [22]

4.3 Meeting Frequency

Rationale	Nodes that we meet most frequently are likely to be met again soon.
Properties	– Easily counted. – Number of connections does not account for duration or burstiness.
Algorithms	Meets and Visits [17], MaxProp [15]

4.4 Rate of change of connectivity

Rationale	– High rate of change of neighbours imply high mobility. – More mobile nodes are more likely to encounter the destination.
Properties	– Purely a measure of mobility. – High mobility is not necessarily an indicator of a good path to particular destinations.
Algorithms	CAR [63]

4.5 Link duration

Rationale	Nodes which spent a lot of time in contact in the past, are likely to be in contact now or in the future.
Properties	– Explicitly aims to maximise delivery ratio. – Does not count frequency or sequencing of connections.
Algorithms	SEPR [83]

4.6 Network science measures

Rationale	Measuring closeness to a destination via clustering or using betweenness centrality shows how important a node is for quick delivery to destinations.
Properties	– Utilise measures of network properties. – Measure are locally calculable. – Very sensitive to how graph aggregation occurs.
Algorithms	SimBet [24], BubbleRap [48]

PhD Thesis: Williamson, G., Nixon, P. & Dobson, S., 2010. *Routing in Human Contact Netw*

Requirements

- High delivery ratio
- Low delivery cost
- Low delivery latency
- Decentralised*

** ideally*

Existing Techniques

for comparison

Name	Delivery Ratio	Cost
Flooding	Excellent	Very high
Random / Epidemic (1 copy)	OK	High
Prophet	Good	High
Bubble-Degree	OK	Low
Persistance Based Routing (PBR)	Good	Low

Others include: Spray and Focus, Hold and Wait, CAM, MED/MEED, RAPID, MH*, Meets & Visits, MaxProp, SEPR, SimBet, Etc.

Datasets for Simulation

	#Nodes	GPS	WiFi	Bluetooth	Cell	Locations / Places	Contacts / Interactions
SocialSensing	😐	☹	☺	☺	☺	😐	☺
GeoLife	☺	☺	☹	☹	☹	☺	☹
Reality Mining *	☺	☹	☹	☺	☺	☺	☺
Cab Spotting *+	☺	☺	☹	☹	☹	☺	☹
Synthetic + *(per-gen)*	☺	☹	☹	☹	☹	☹	☺

* Cited in literature
+ Not human movement

Others: CenceME, Rollernet, TomStalker, etc.

Example: Persistence Based Routing

simple version

- Route based on amount of time connected to destination node
- Route tables shared amongst nodes and distributed over contact network

7.1 Persistence Based Routing

The Basic Premise:	The more time we spent in contact with a given node in the past, the more likely we are to come into contact with them in the future.
The Routing Question:	"When I come into contact with a node i, is that node more likely to deliver a message to it's destination than I am?"
The Goal:	To maximise the number of successfully delivered messages with preference to reducing delivery time.

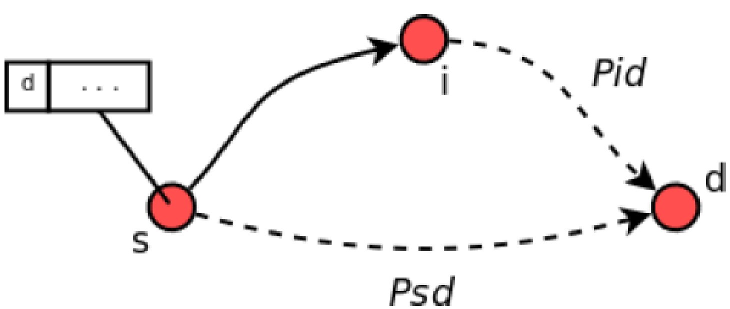
Williamson, G., Nixon, P. & Dobson, S., 2010. *Routing in Human Contact Networks*.

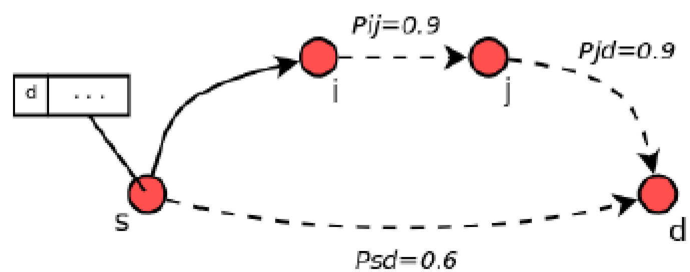
PBR

$$A_{ij}(t) = \begin{cases} 1 & \text{if link } i \rightarrow j \text{ occurs in time-slot } t \\ 0 & \text{otherwise} \end{cases}$$

$$\pi_{ij} = P_{ij} = \frac{\sum_T A_{ij}(t)}{T}$$

if $(\pi_{id} > \pi_{sd})$ forward message to node i
else keep message





$\pi_{sd} = P_{sd} = 0.6.$

$\pi_{id} = P_{ijd} = P_{ij}.P_{jd} = 0.81$

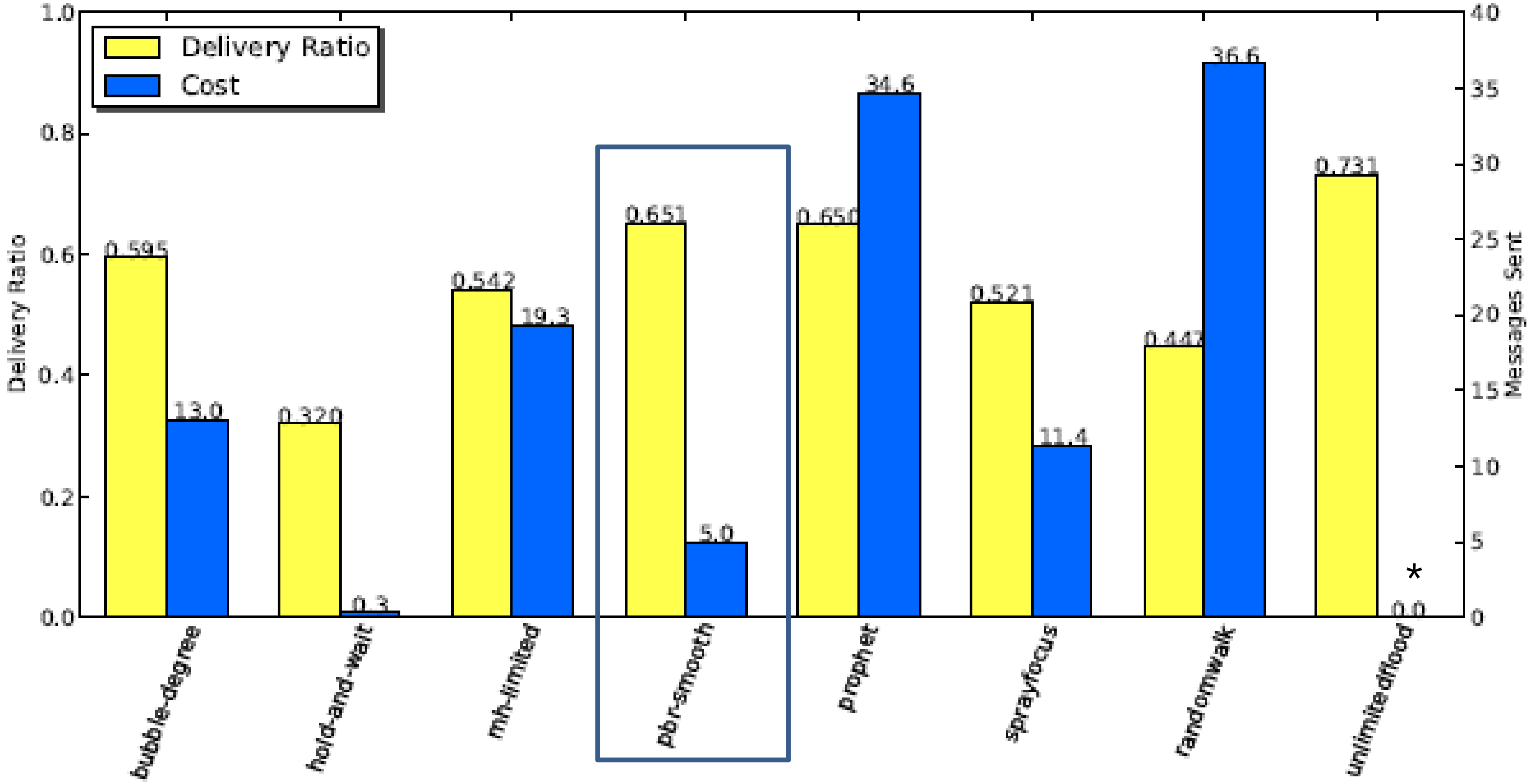
ContactSim

a.k.a LocationSim

- Discreet event simulator
  - Powerful configuration
  - Datasets
  - Protocols
  - Tasks

svn+ssh://kind.ucd.ie/Volumes/Data/svn/software/LocationSim/trunk

PBR Results



Delivery ratio and average cost for each delivery algorithm, using MIT reality mining dataset for October 2004

Using Location

- Calculate ranking of individuals visits to towers
- Route based on rank

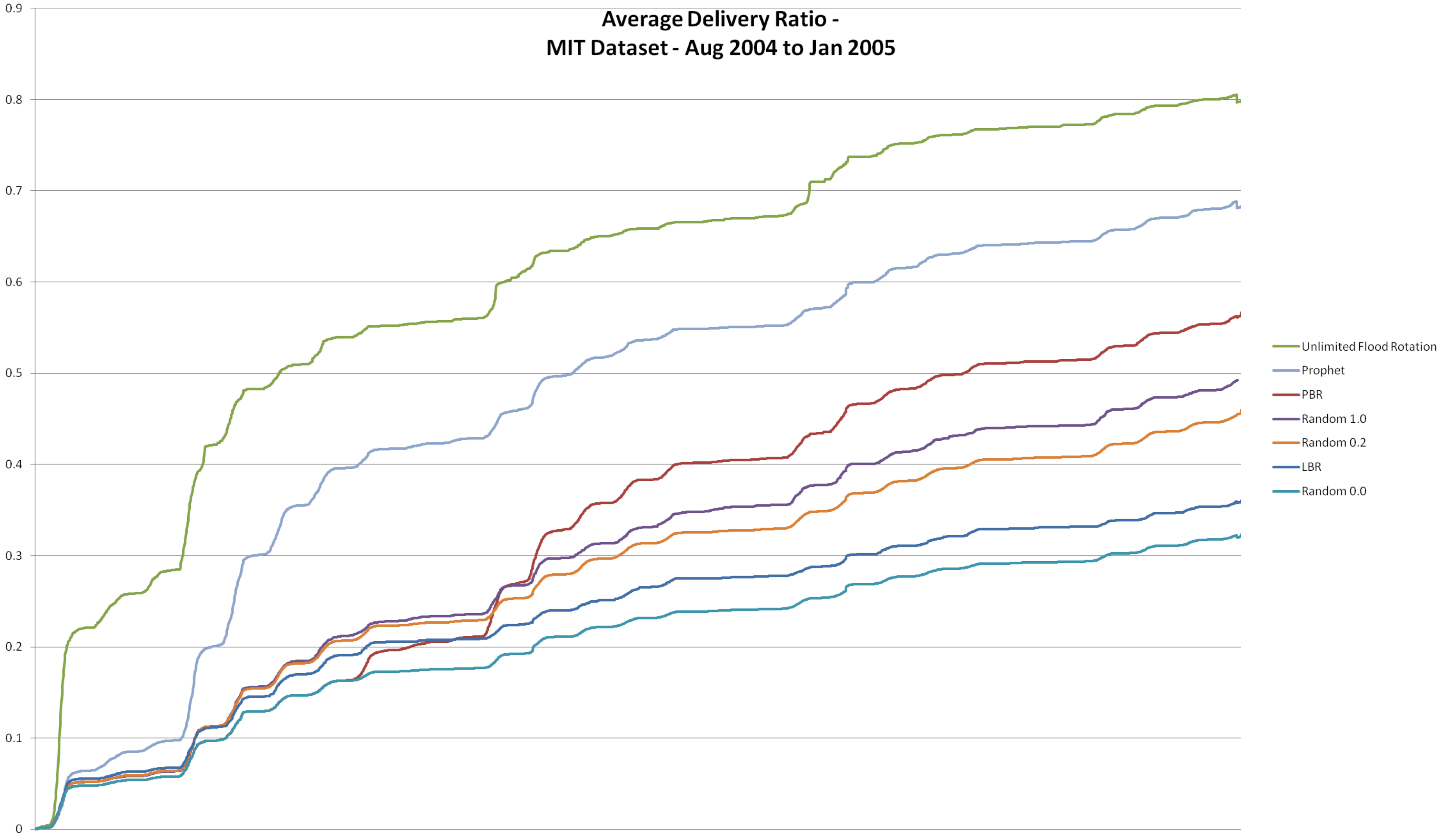
Popularity of user $i = \sum_j \tau_{ij}(p_j + c)$
where τ_{ij} is the time (or number of times) user i has visited the tower j, p_j is the popularity of tower j and c is a parameter which tunes the importance of user mobility

	T1	T2	T3	T4
U1	10	0	0	0
U2	0	6	6	0
...	...	...	...	...
Total Score	100	80	30	20

```
On encountering another node,
for each message do:

if node is destination node
     pass message
else if node has higher rank
     pass message
else
     keep message
```

Simulation Results



28 day time window, at 7 day intervals, (18 runs)

Why does LBR perform so poorly?

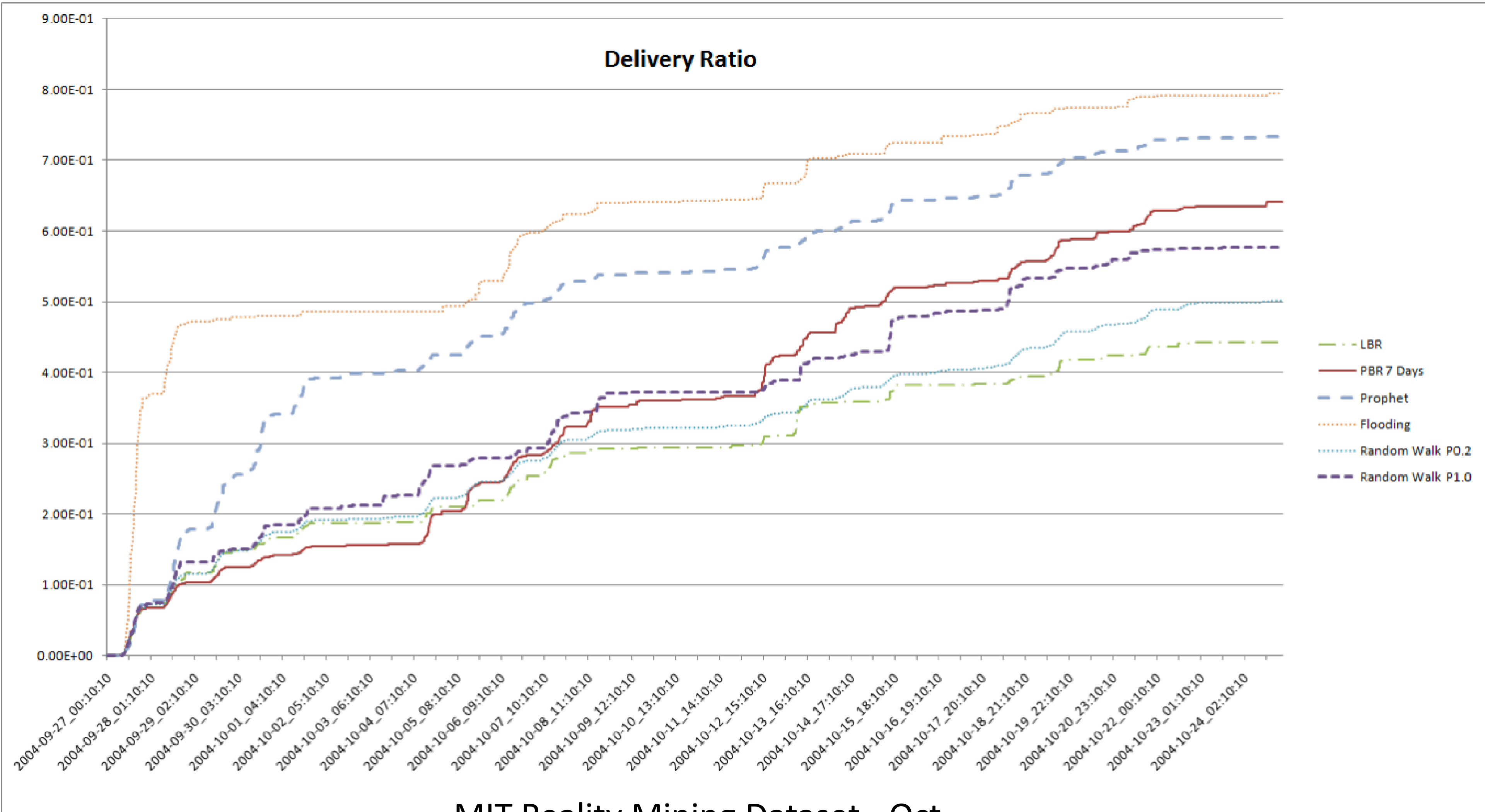
- Different communities visit different places?
- Highly mobile individuals do not see everyone?
- Different carriers?

Improvements to LBR?

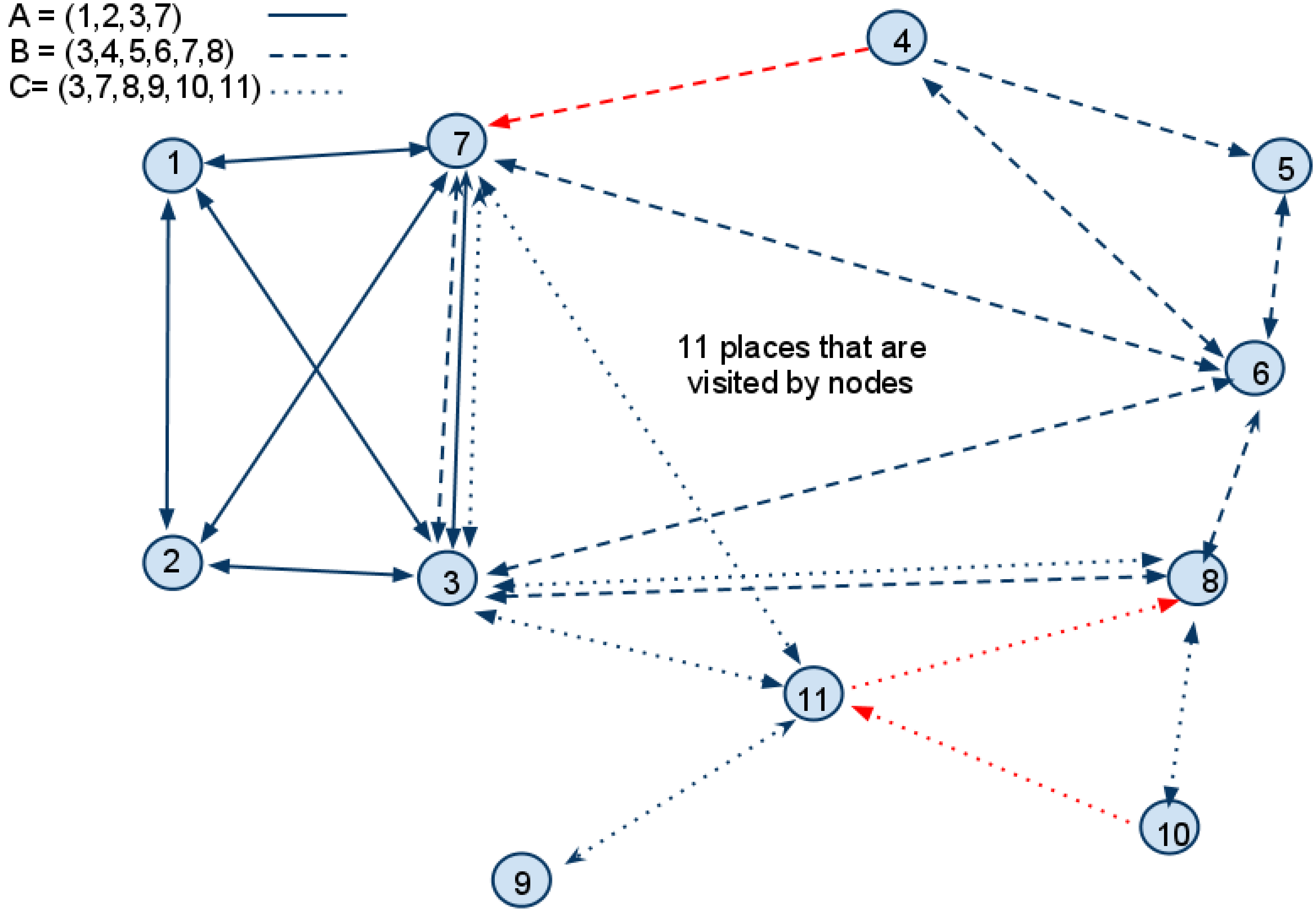
- Consider only popular locations
- Time ordering of visits for future location prediction
- Detect communities based on places visited
- Combine with PBR

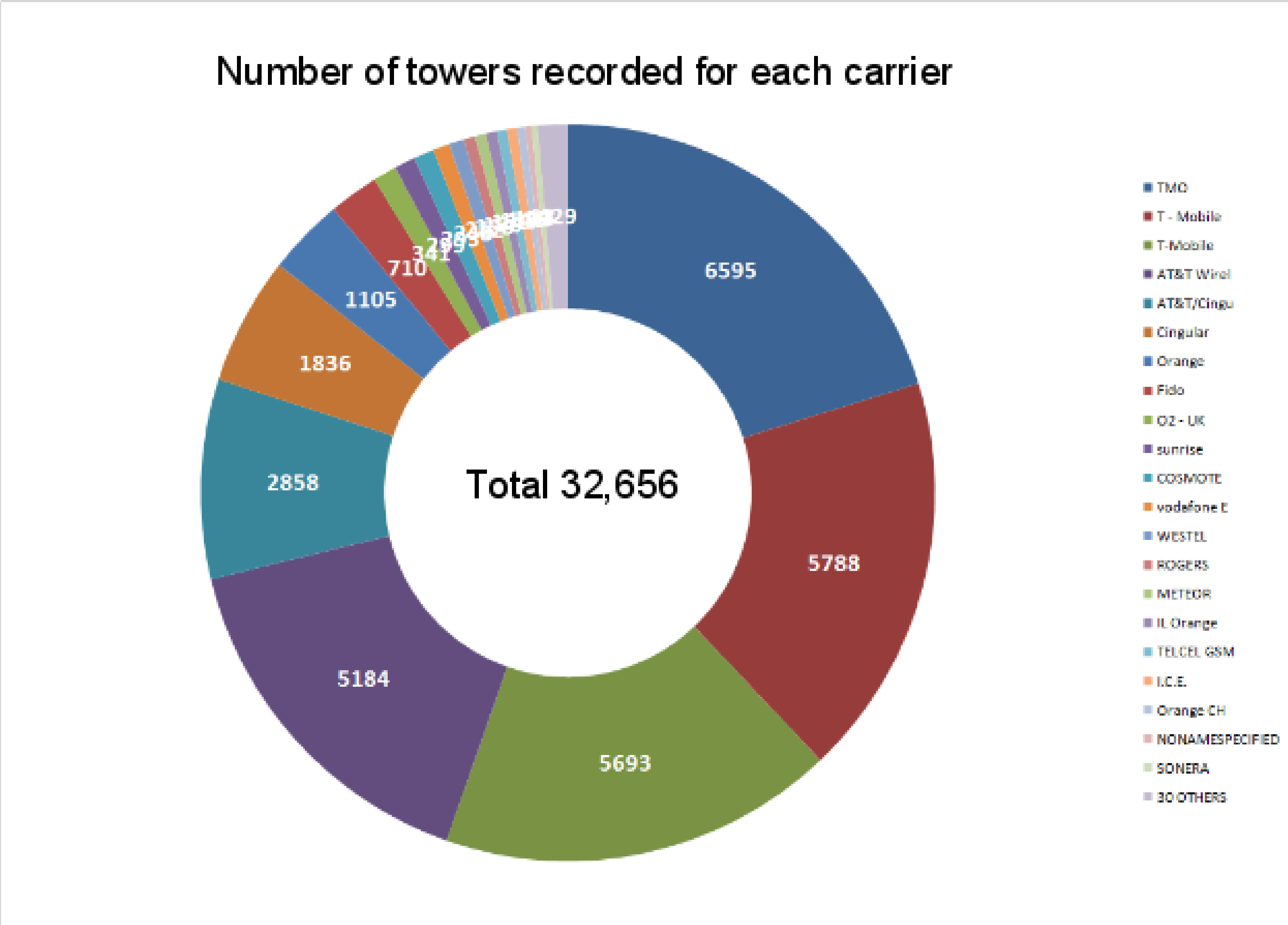
Questions/Ideas?

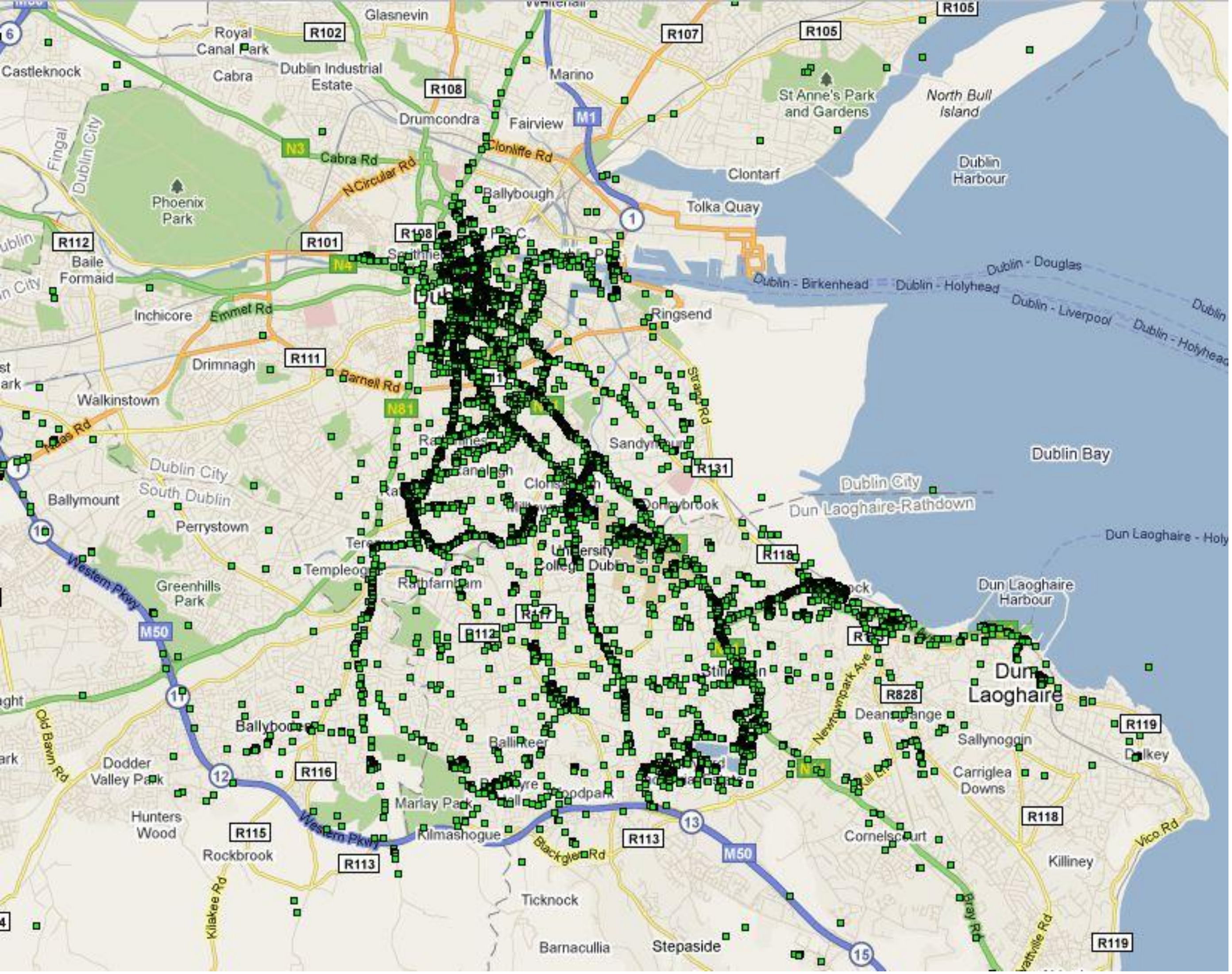
Simulation Results

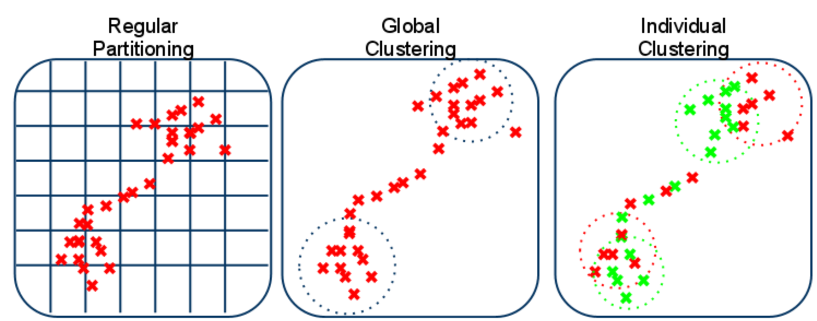


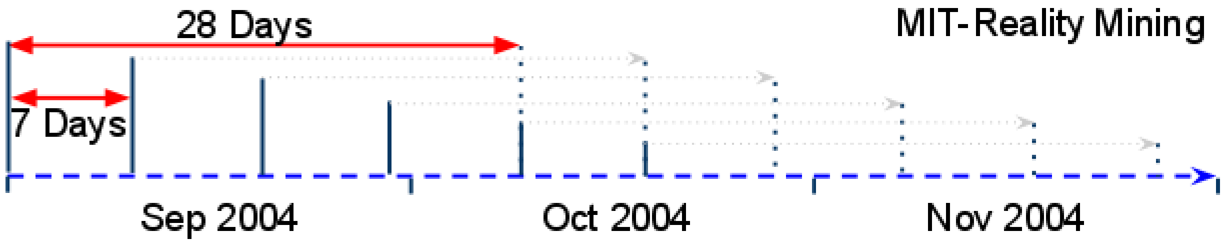
MIT Reality Mining Dataset - Oct


A = (1,2,3,7)
B = (3,4,5,6,7,8)
C= (3,7,8,9,10,11)
11 places that are visited by nodes
1
2
3
4
5
6
7
8
9
10
11


Number of towers recorded for each carrier
6595
5788
5693
5184
2858
1836
1105
710
341
Total 32,656
TMO
T - Mobile
T-Mobile
AT&T Wirel
AT&T/Cingu
Cingular
Orange
Fido
O2 - UK
sunrise
COSMOTE
vodafone E
WESTEL
ROGERS
METEOR
IL Orange
TELCEL GSM
I.C.E.
Orange CH
NONAMESPECIFIED
SONERA
30 OTHERS


Glasnevin
Royal Canal Park
Castleknock
Cabra
Dublin Industrial Estate
Marino
Drumcondra
Fairview
St Anne's Park and Gardens
North Bull Island
Phoenix Park
Cabra Rd
N.Circular Rd
Clonliffe Rd
Ballybough
Clontarf
Dublin Harbour
Tolka Quay
Baile Formaid
Inchicore
Emmet Rd
Ringsend
Dublin - Birkenhead
Dublin - Holyhead
Dublin - Douglas
Dublin - Liverpool
Drimnagh
Barnell Rd
Walkinstown
Sandymount
Dublin Bay
Ballymount
Perrystown
Dublin City
Dun Laoghaire-Rathdown
Donnybrook
Dun Laoghaire - Holy
Templeogue
Greenhills Park
Rathfarnham
Dun Laoghaire Harbour
Western Pkwy
Dun Laoghaire
Deansgrange
Sallynoggin
Carriglea Downs
Dalkey
Dodder Valley Park
Ballinteer
Marlay Park
Hunters Wood
Kilmashogue
Cornelscourt
Killiney
Vico Rd
Rockbrook
Ticknock
Barnacullia
Stepaside
R102
R107
R105
R108
R112
R101
R111
R131
R118
R828
R119
R116
R115
R113
M50
N3
N4
N81
M1


Regular Partitioning
Global Clustering
Individual Clustering


28 Days
MIT-Reality Mining
7 Days
Sep 2004
Oct 2004
Nov 2004