Behavioral Modeling in Networked Mobile Societies

Funded by:

Gautam Thakur*, Ahmed Helmy* and Wei-Jen Hsu§

*CISE, University of Florida, Gainesville

§Cisco Systems Inc.

{gsthakur, helmy}@ufl.edu , wehsu@cisco.com

Wireless Mobile Networking Lab ***http://nile.cise.ufl.edu***

Research Areas

- Mobility Modeling
- Analysis and Impact of Behavioral aware routing protocols
- Network Theory
- Measurements
- Vehicular Networks (Just started)

Behavior Characteristics

1. Similarity
2. Community detection
   1. Profile-Cast
   2. Bubble Rap

1. Encounter Statistics
2. Epidemic Routing

Structural Dynamics

Routing Performance

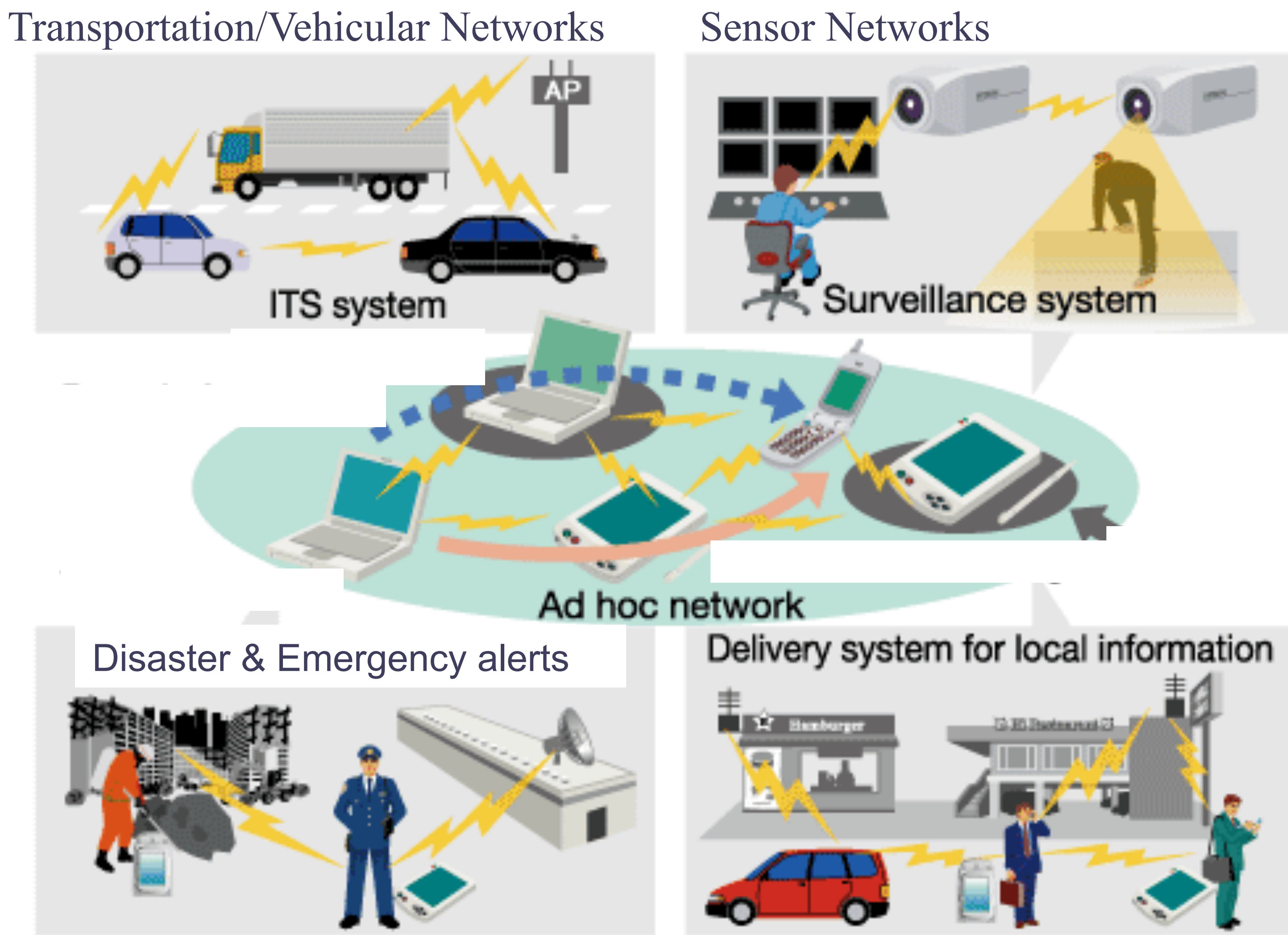
Topics

- The Missing Link
- Analysis of Encounter Statistics
- The Behavioral Mobility Framework
- SHIELD

Outline

- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Networked Mobile Societies Everywhere, Anytime



Mobile Ad hoc, Sensor and Delay Tolerant Networks

Emerging Behavior-Aware Services

- Tight coupling between users, devices
  - Devices can infer user preferences, behavior.
  - Capabilities: comm, comp, storage, sensing.

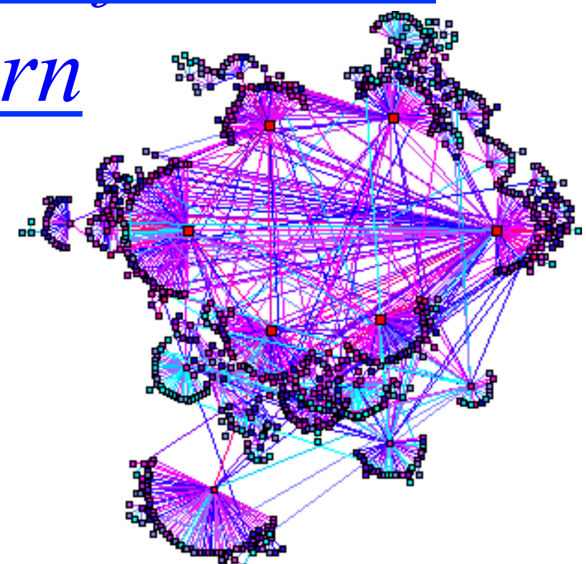
- New generation of behavior-aware protocols
  - Behavior: mobility, interest, trust, friendship,…
  - Apps: interest-cast, participatory sensing, crowd sourcing, mobile social nets, alert systems, …

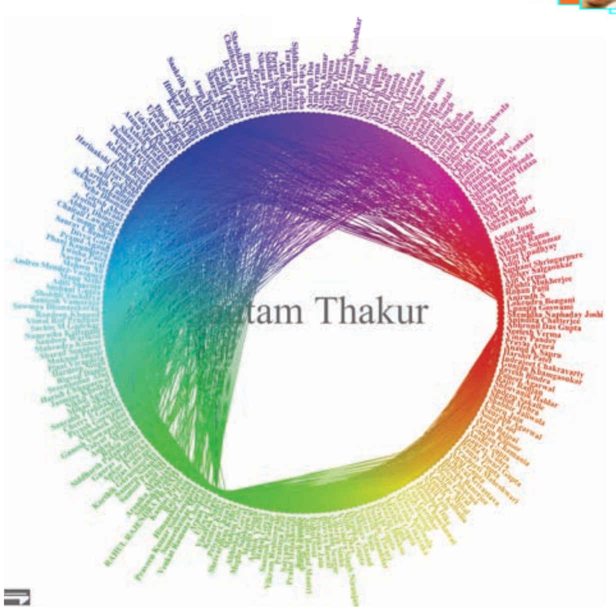
New paradigms of communication?

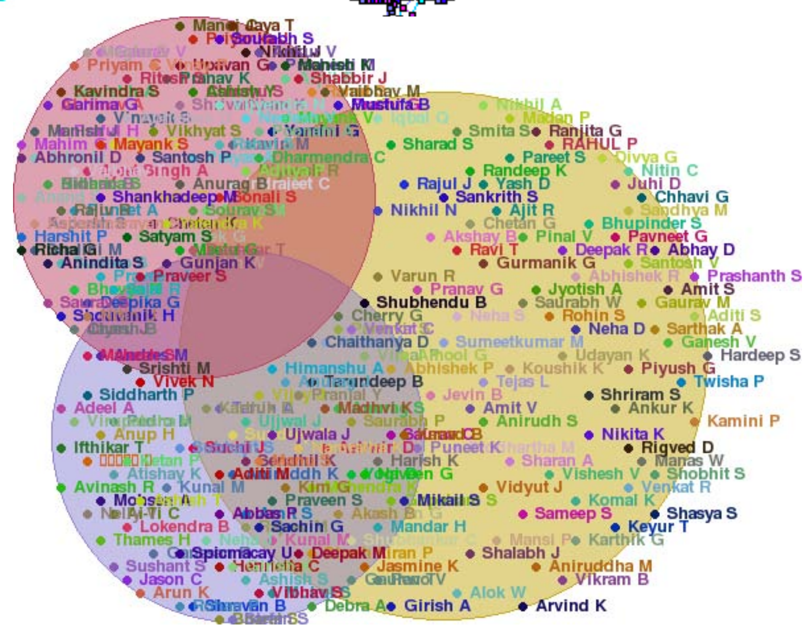
Similarity: *A Characteristic of Human Behavioral Pattern*

- Interest based likings attached to locations.
- Synchronization in schedules etc.
- Websites and Twitter followers.
- Affiliation to groups.

Similarity: *A Characteristic of Human Behavioral Pattern*







Outline

- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Similarity: Aid to DTN Paradigm

- Future social networks targeting mobile users' behavioral profiles and preferences.

- DTN using behavioral profiles and Social Structures
  - Profile-Cast [1], Bubble-rap [2] etc.

- Efficient use of Similarity in DTNs
  - interest-based target message forwarding, localization.
  - encounter- based routing, mobile resource discovery
  - Establishing trust [Poster].

1. Hsu, W., Dutta, D. and Helmy, A. CSI: A Paradigm forBehavior-oriented Delivery Services in Mobile Human Networks, *IEEE/ACM Transactions on Networking (submitted), 2009*.
2. Hui, P., Crowcroft, J. and Yoneki, E. *Bubble-rap: social-basedforwarding in delay tolerant networks*. *ACM MobiHoc, 2008*.

Mobility based Similarity

- Similarity – Defined by Mobility Preferences.
- Users interest in form of on-line usage captured through WLAN measurements.
- *May not reflect social ties, but does reflect mobility-related behavior affecting connectivity and network topology dynamics in a DTN.*

I knew we had a lot in common. I'm crazy too!

Similarity Analysis Approach

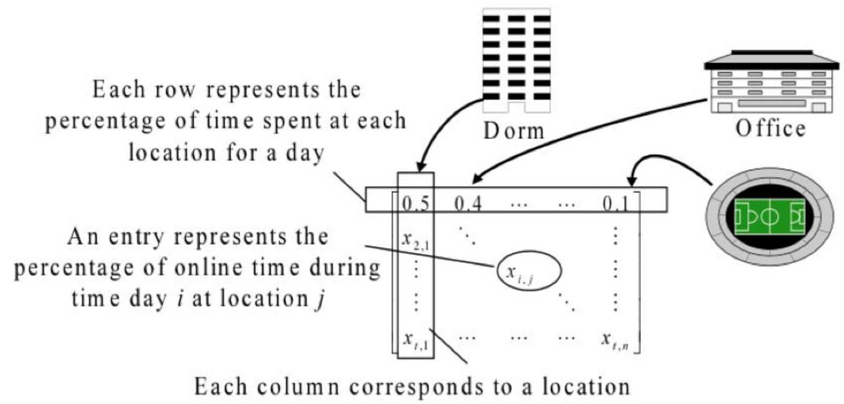
- Introduce and capture mobility-based behavior and Similarity
  - Behavioral representation
    - On-line Association Matrix.
    - Singular Value Decomposition (SVD).
  - Calculating Similarity
    - Quantitatively compare the behavioral trends.
    - Weighted cosine product of Eigen behaviors.
- Investigate similarity in real measurements
  - Distribution, Stability and Clustering.
- Analyze Similarity in existing mobility models
  - Random and Community based Models.

Outline

- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Capturing Spatio-temporal Mobility Preferences

- To build spatio-temporal profile of Mobile user
  - *We use longitudinal wireless activity sessions.*
- *Association Matrix:*
  - Has the location of association and session time information for that user.



Row → on-line time (e.g., a day)

Column → Location (e.g building, AP)

x_{ij} → Percentage of time spent at a location

$$A = \begin{bmatrix} 0 & 0 & 0.5 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.4 & 0.1 & 0 \\ 0 & 0 & 0 & 0.6 & 0.4 \\ 0.2 & 0.3 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0 & 0 & 0 & 0.9 \end{bmatrix}$$

Calculating Similarity *[Sim(X,Y)]*

- Singular Value Decomposition of Association Matrix to capture ***Dominant Behavior***

$$A = U \cdot S \cdot V^T$$

- Calculate weights for Eigen Vectors of Association Matrices

$$w_i = \sum_{i=1}^{k} s_i^2 \Big/ \sum_{i=1}^{rank(A)} s_i^2$$

- Quantitative measure of Similarity for user ***X & Y***

$$Sim(X,Y) = \sum_{i=1}^{rank(X)} \sum_{j=1}^{rank(Y)} w_{x_i} w_{y_j} \left| x_i \cdot y_j \right|$$

Outline

- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Mobility Models under Study

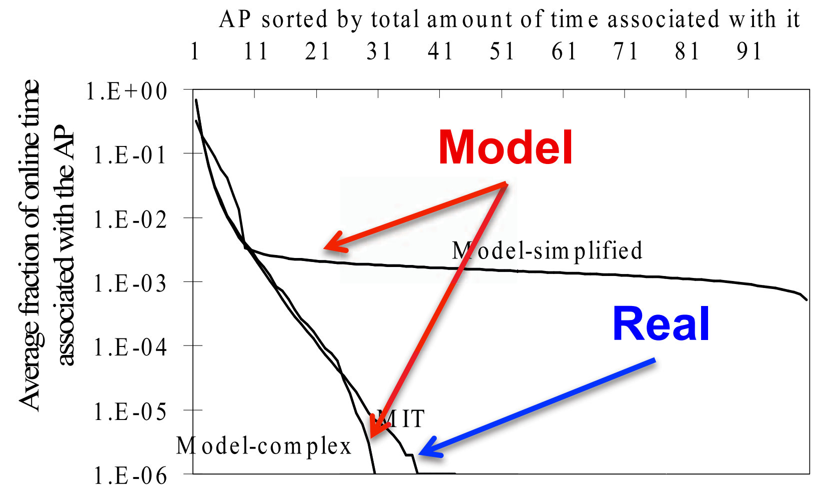
- Random Direction Model [1]
  - This model is more stable as compared to other random models and provides quantitatively even distribution of nodes in the simulation area.
  - to investigate the effect of random movements on DTN performance.
  - to baseline for comparisons.

1. Royer, E. M., Y, P. M. M. and Y, L. E. M. *An analysis of the optimum node density for ad hoc mobile networks*. *IEEE ICC, 2001, 857-861*.

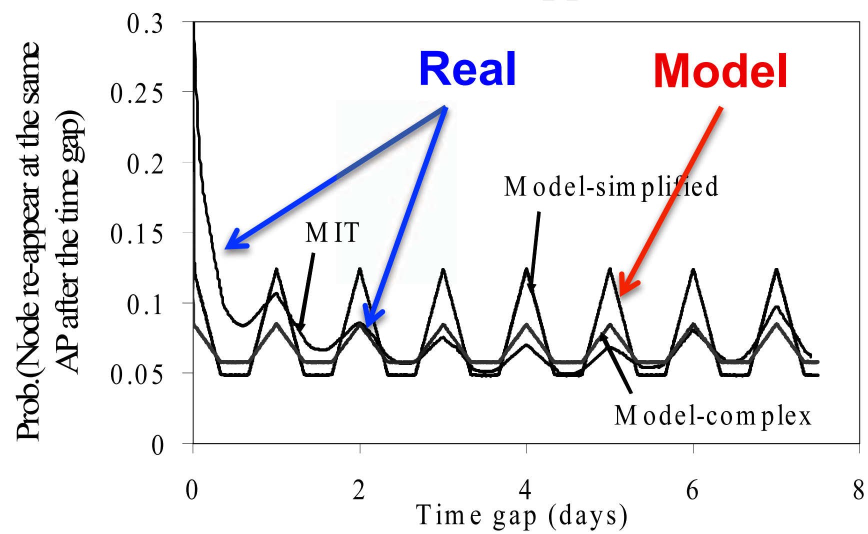
Mobility Models under Study

- Time Variant Community Model (TVC)[1]
  - two realistic mobility features.
  - skewed location visiting preferences.
  - periodic re-appearances.
- Include other models
  - Community Model[2]

Location Visiting Preferences.



Periodic Re-appearances



1. Hsu, W., Spyropoulos, T., Psounis, K. and Helmy, A. *Modeling spatial and temporal dependencies of user mobility in wireless mobile networks*. IEEE/ACM *ToN, pp. 1564-1577, 2009.*
2. M.MusolesiandC.Mascolo,"*A community based mobility model for ad hoc network research*," *(REALMAN), May 2006.*

Outline

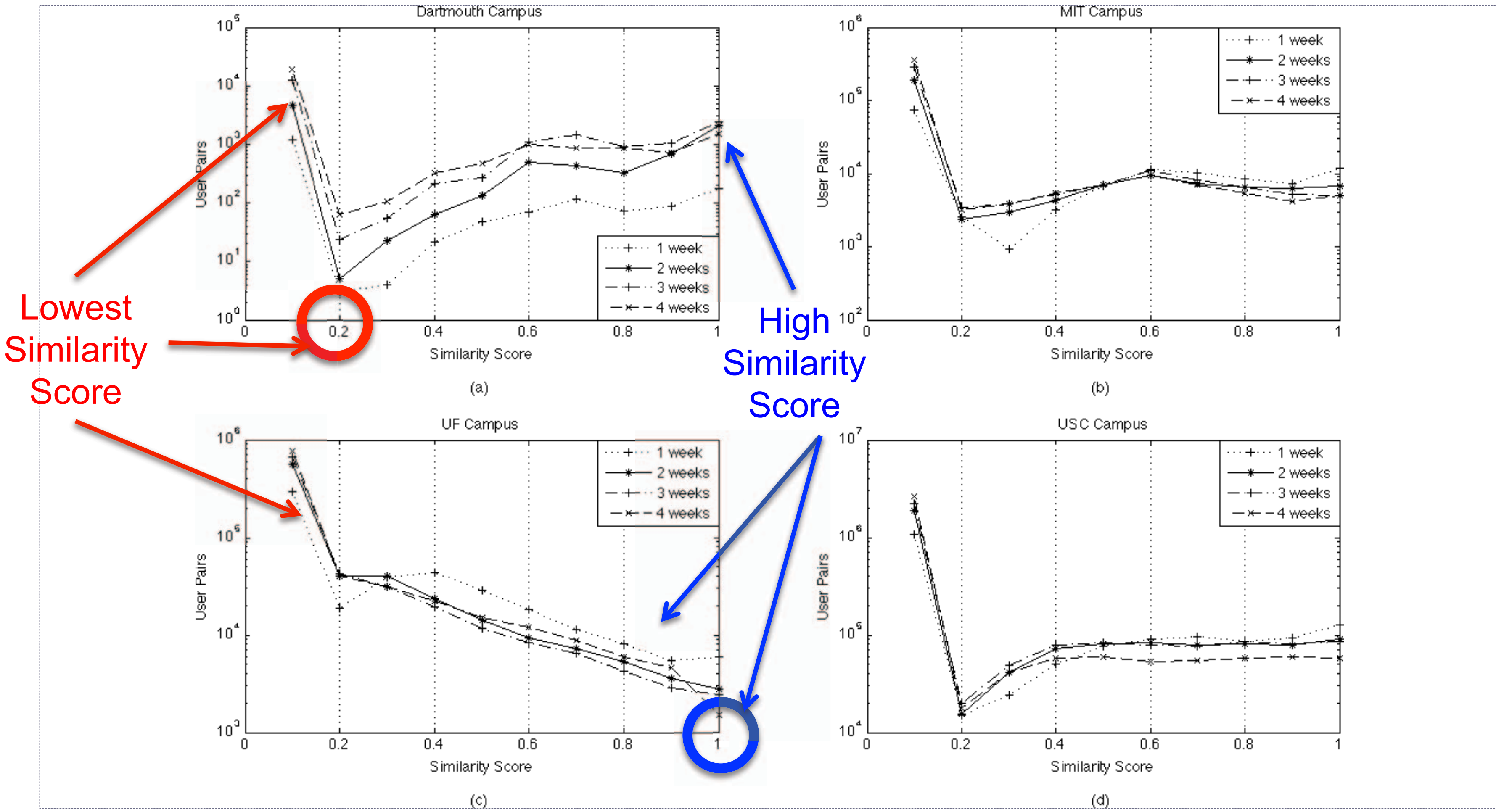
- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Mobility Data Sets (Real Traces)

- Extensive repository of anonymous WLAN session traces.
- Collected from four University campuses.
- Sampled for one month duration.

Campus	# Users	# AP	Time Period
Dartmouth	1500	756	Fall 2007
MIT	1366	173	Fall 2006
Univ. of Florida	3000	1161	Fall 2008
USC	3000	128	Fall 2007

Similarity Distribution in Real Traces



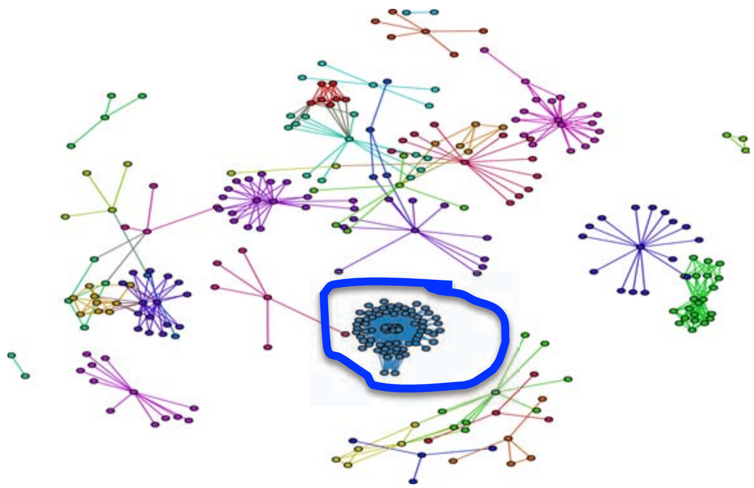
All the four time intervals show near consistent user pair counts for a particular similarity score. Lowest similarity score (0.0 - 0.1) shows that users have very different spatio-temporal preferences. A fraction of the user pairs are also very similar with (0.9 - 1.0) similarity score.

Detection of Mobile Societies

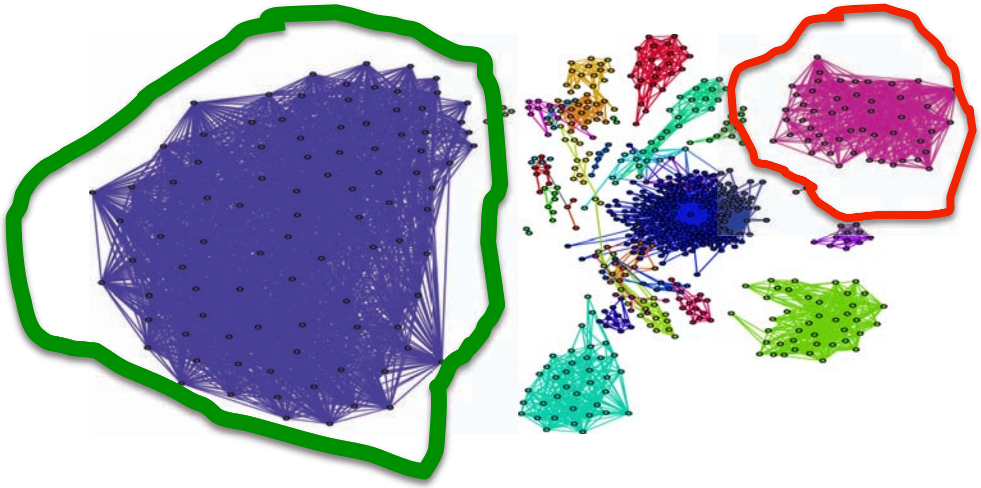
- Based on Similarity score among mobile users' pairs.

- Modularity:
  - Centrality-index driven method to discover clusters.
  - To understand the underlying structure of mobile societies, *the similarity distribution is not sufficient.*

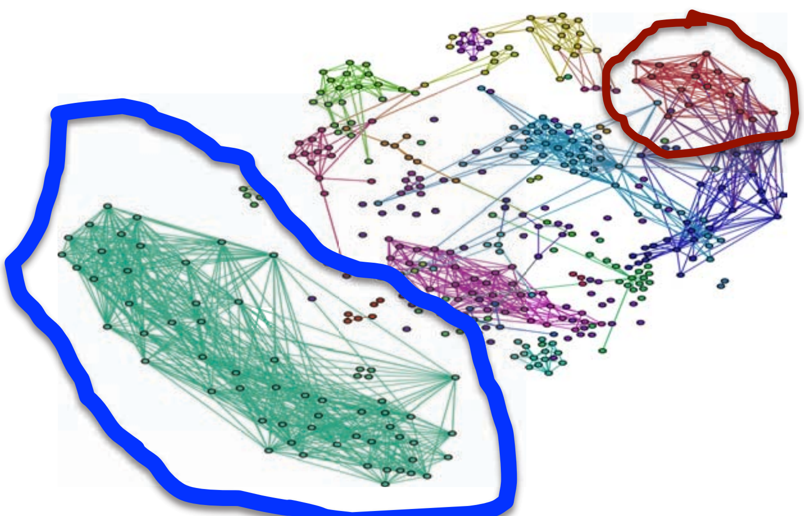
Clustering Results for Real Traces



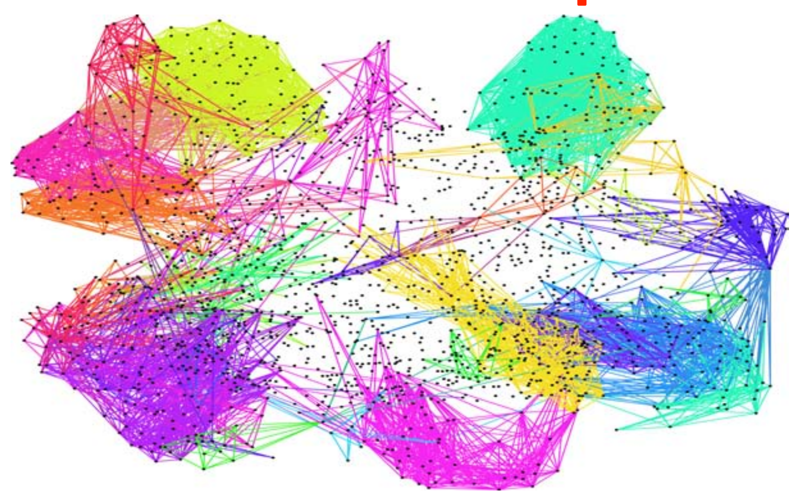
Dartmouth-Campus



MIT-Campus



UF-Campus



USC-Campus

Network Analysis of Real Traces

Dataset	Clustering Coefficient		Average Path Length		Modularity	
	Orig	*Rand*	*Orig*	*Rand*	*Orig*	*Rand*
Dartmouth	0.89	0.05	0.10	2.47	0.63	0.2
MIT	0.92	0.05	0.40	2.12	0.79	0.14
UF	0.78	0.051	0.30	2.605	0.67	0.24
USC	0.91	.05	0.19	2.0	0.46	0.11
**Orig = Original Dataset Graph*			**Rand = Random Graph*			

Large

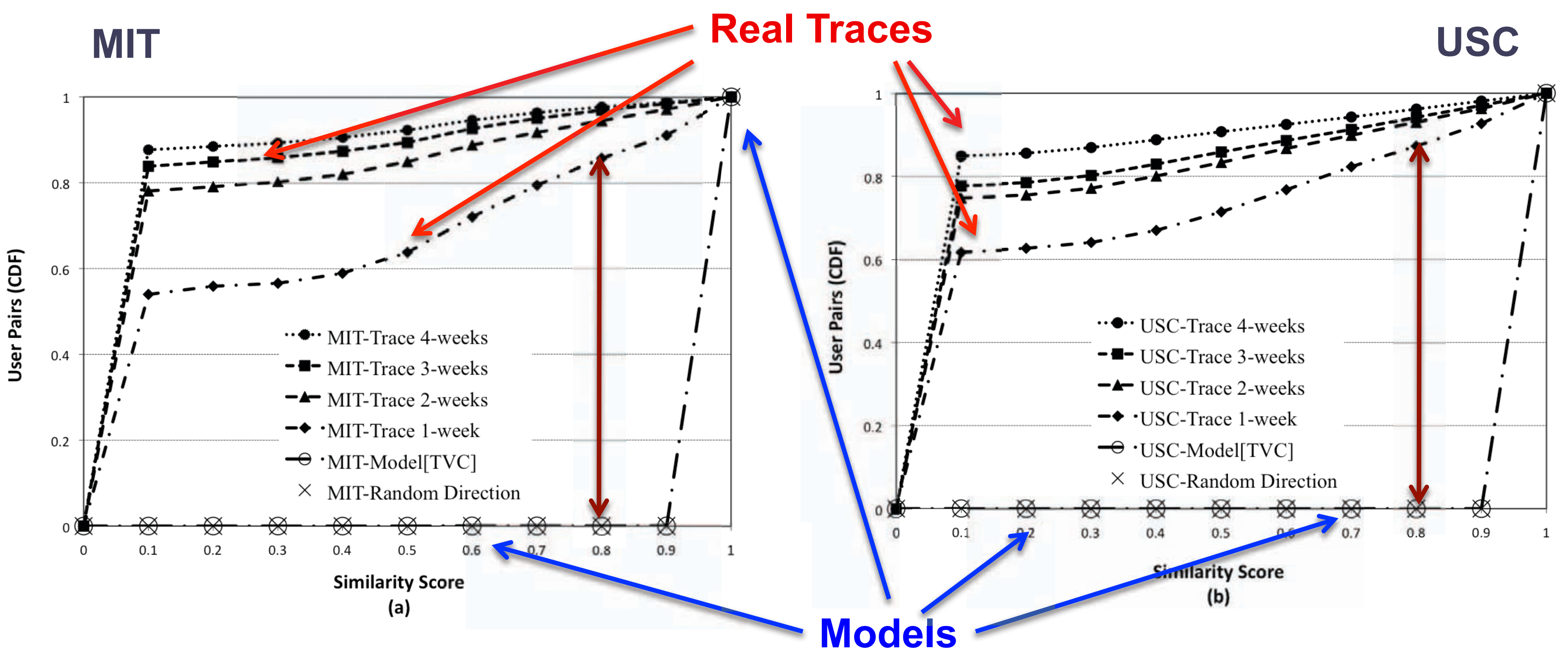
Small

- The comparative values in the table clearly show that mobile societies exhibit **small world characteristics**. On-going work, further research needed.

Evaluation – Mobility Model

- TVC and Random Direction Model are unable to accurately capture the richness of similarity distribution.
- Both models consider homogenous population
- For all values of similarity score except **0.9-1.0**, the models yield no user pairs.

CDF of Similarity



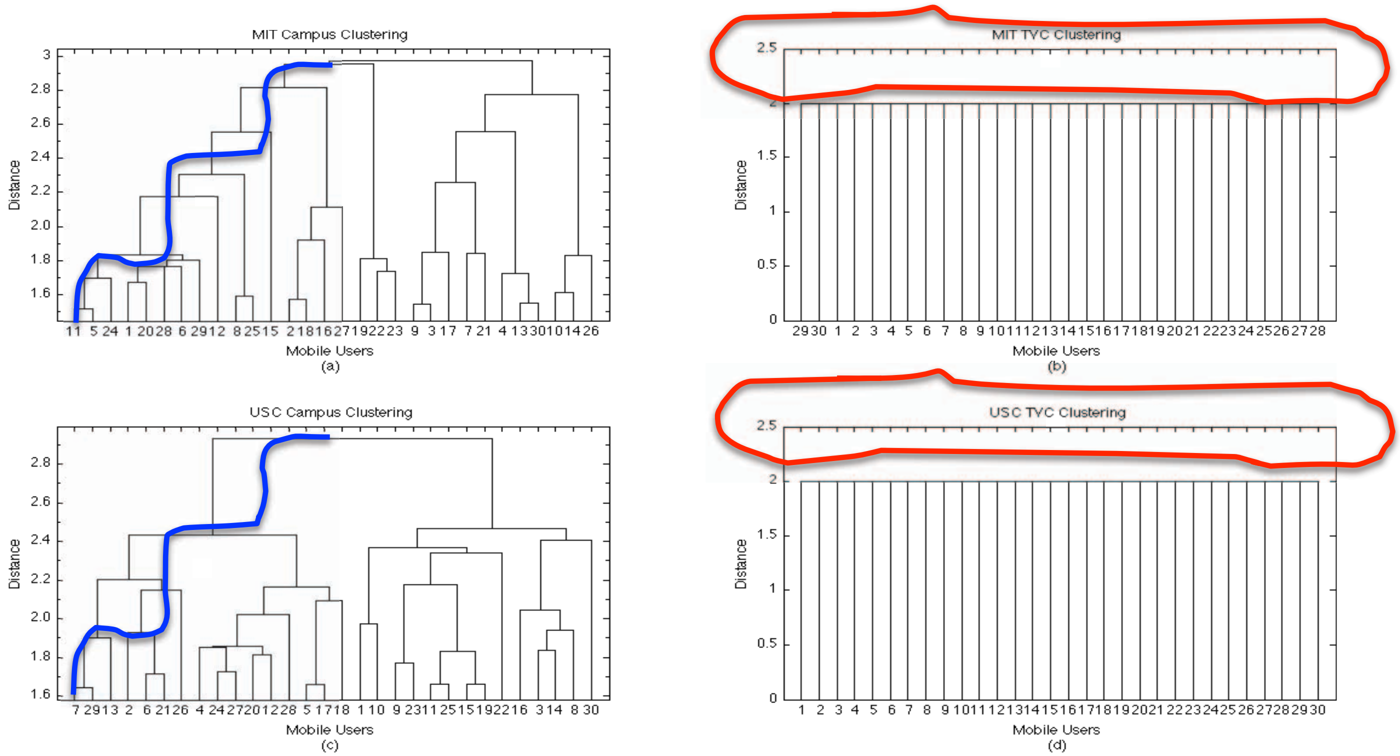
Cumulative distribution function of distances for the similarity score of mobile users. Real trace curves show a conformance with user pairs for different values of similarity score, while TVC and Random Direction Model has all users pairs in the 0.9 score range.

Hierarchical Clustering

Real Traces | Model

MIT

USC



Dendrograms giving visual representation of *two-dimensional hierarchical clustering*.
Real traces (Figure a & c) show an incremental built-up of components based on the similarity score strength between mobile user.
Models (Figure b & d), output only one cluster containing all mobile users. Invariably, models treats all mobile users to have same preferences.

Outline

- Introduction
- Similarity in DTNs
- Representation of Similarity
- Mobility Models
- Data Sets and Result Analysis
- Conclusion

Conclusion

- Similarity as a cornerstone
  - behavior-aware protocols and services designed for DTNs.
- Representation of Similarity
  - Proposed SVD-based weighted-cosine similarity index to quantitatively compare mobility profiles.
- Presence of Similarity
  - Clusters in *real measurements*.
  - Models lack richness of clusters of similar behavior.

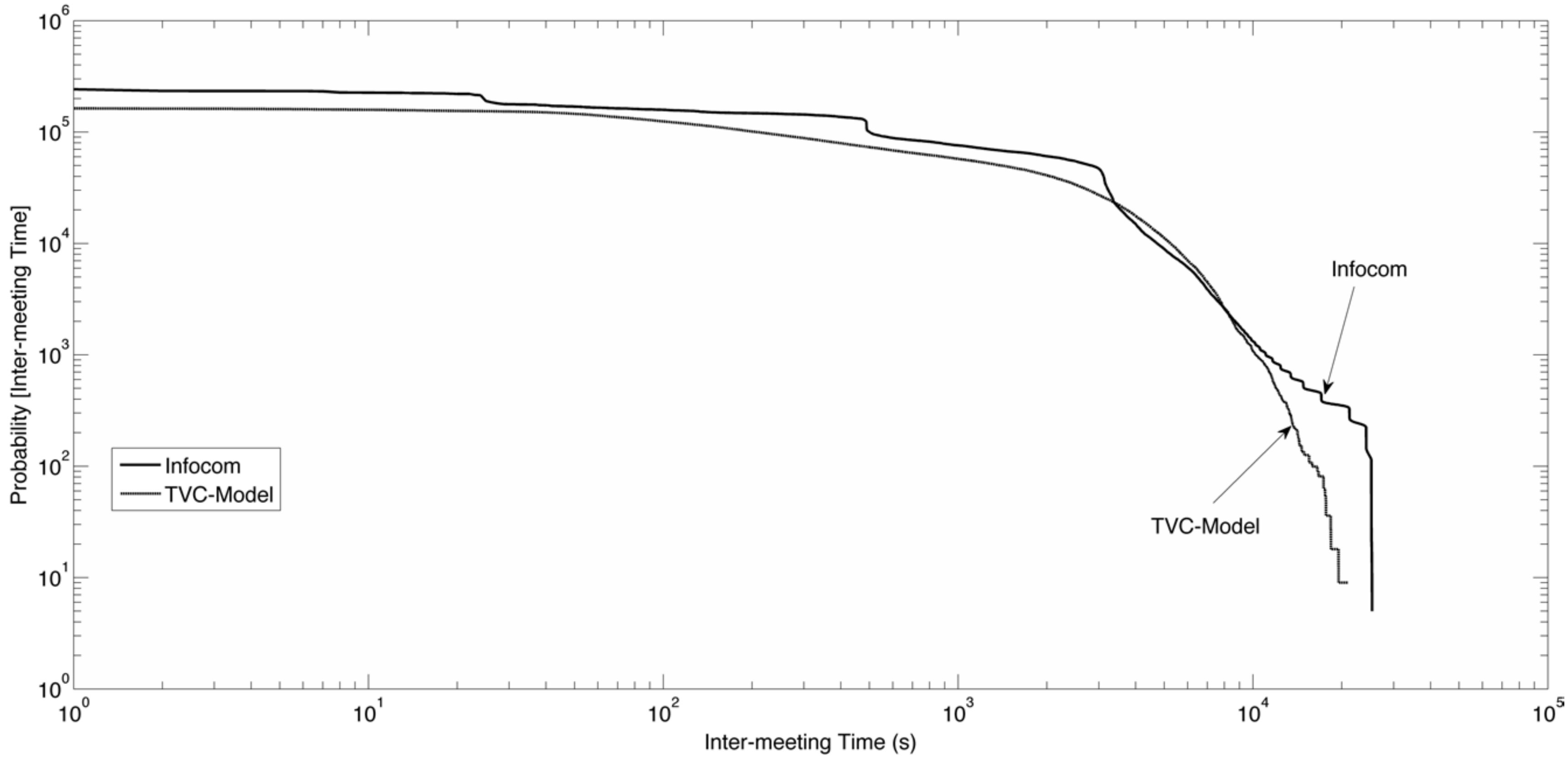
Conclusion

- Questions to address:
  - How can we accurately model *Similarity* in Mobility Models?
  - Is *Similarity* sufficient to produce accurate simulations for DTN protocols and services?
    - » Our On-going work ….

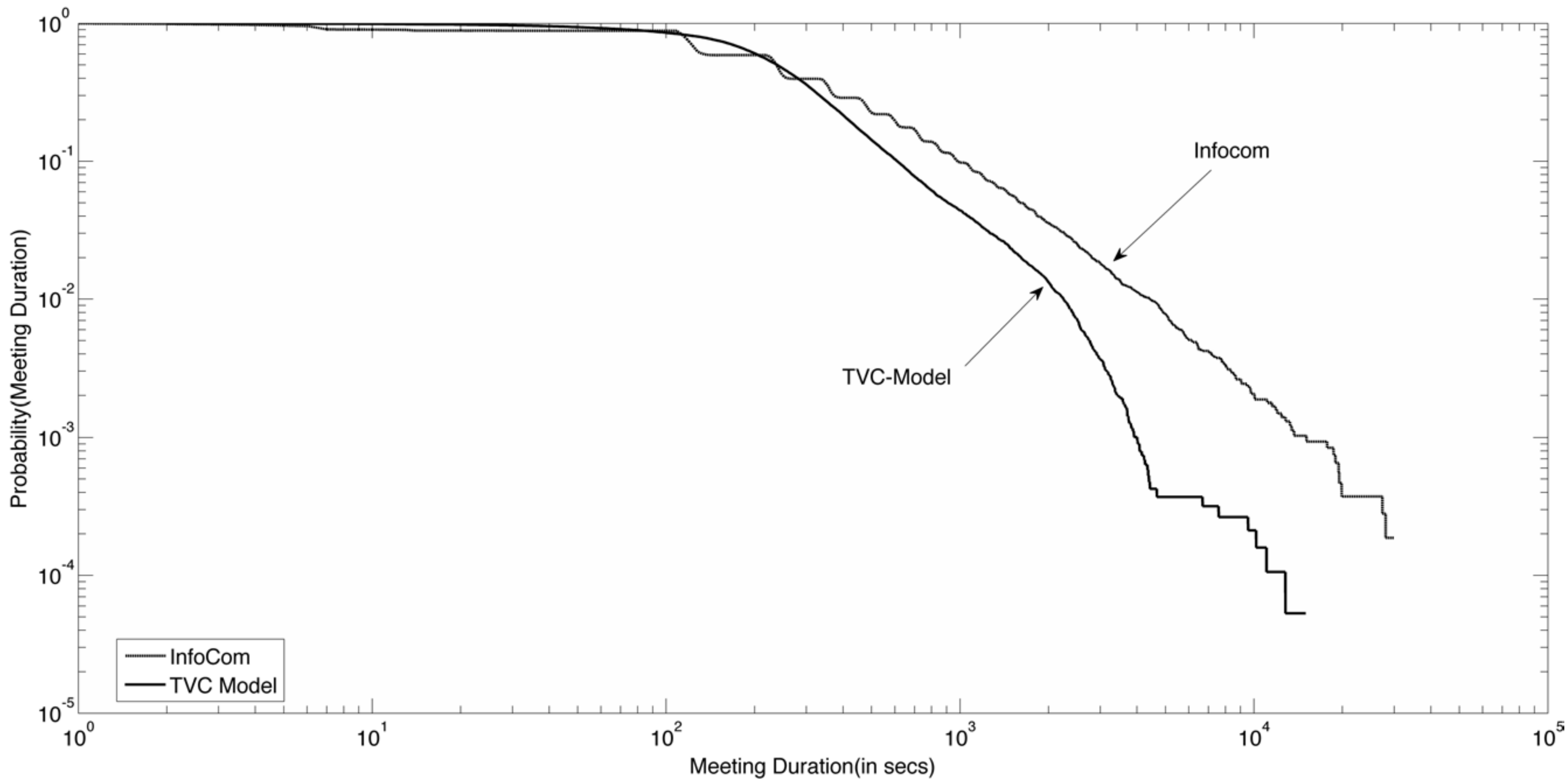
Encounter Statistics

- Inter-meeting Time
  - Time between two meets
- Meeting Duration
  - Duration of meeting of two users.

Encounter Statistics



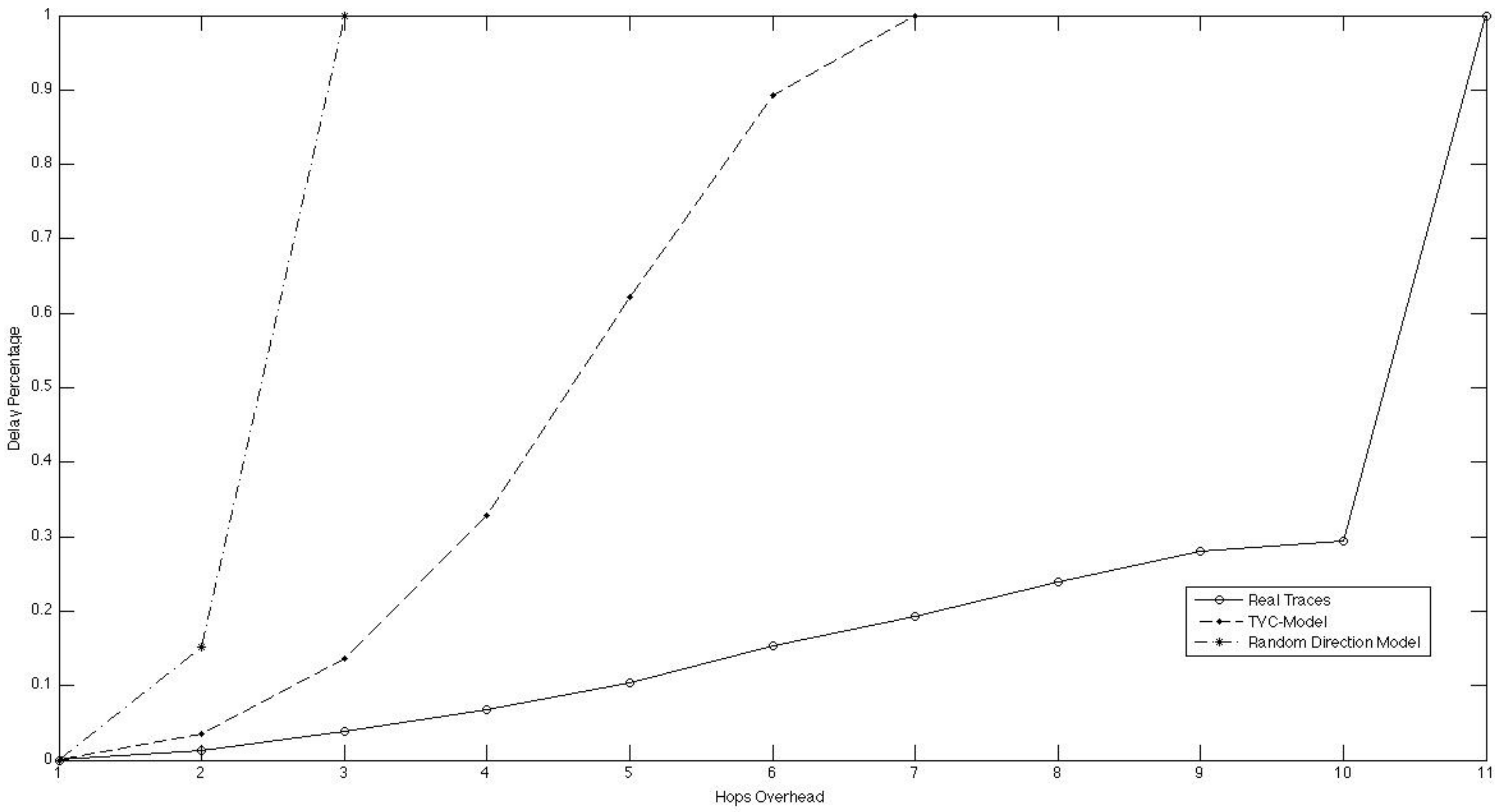
Encounter Statistics



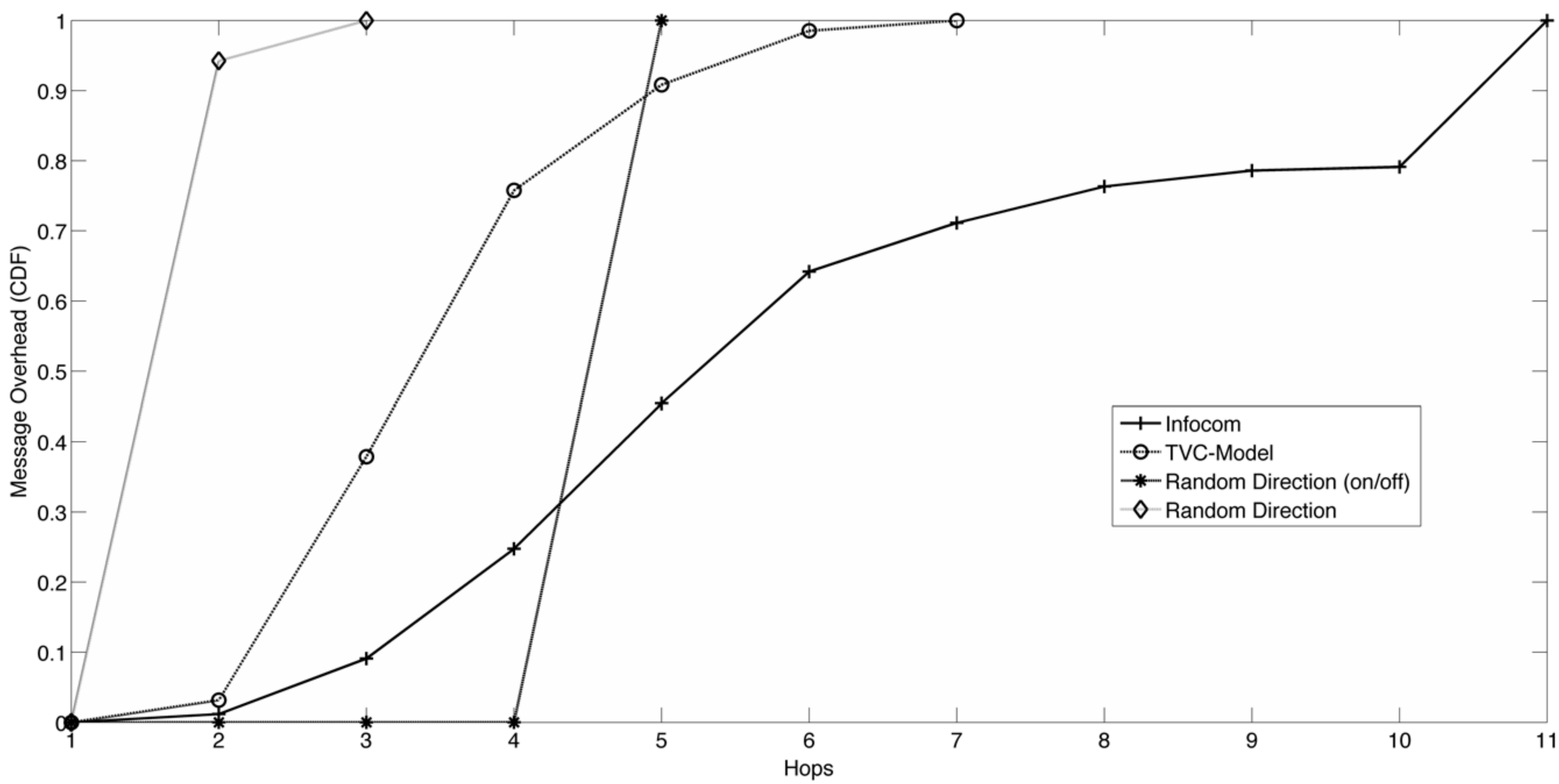
Routing Performance

- Message Overhead
  - Number of messages generated
- Delay
  - Number of Hops it took for messages to reach
- Time
  - Time taken for messages to reach all nodes

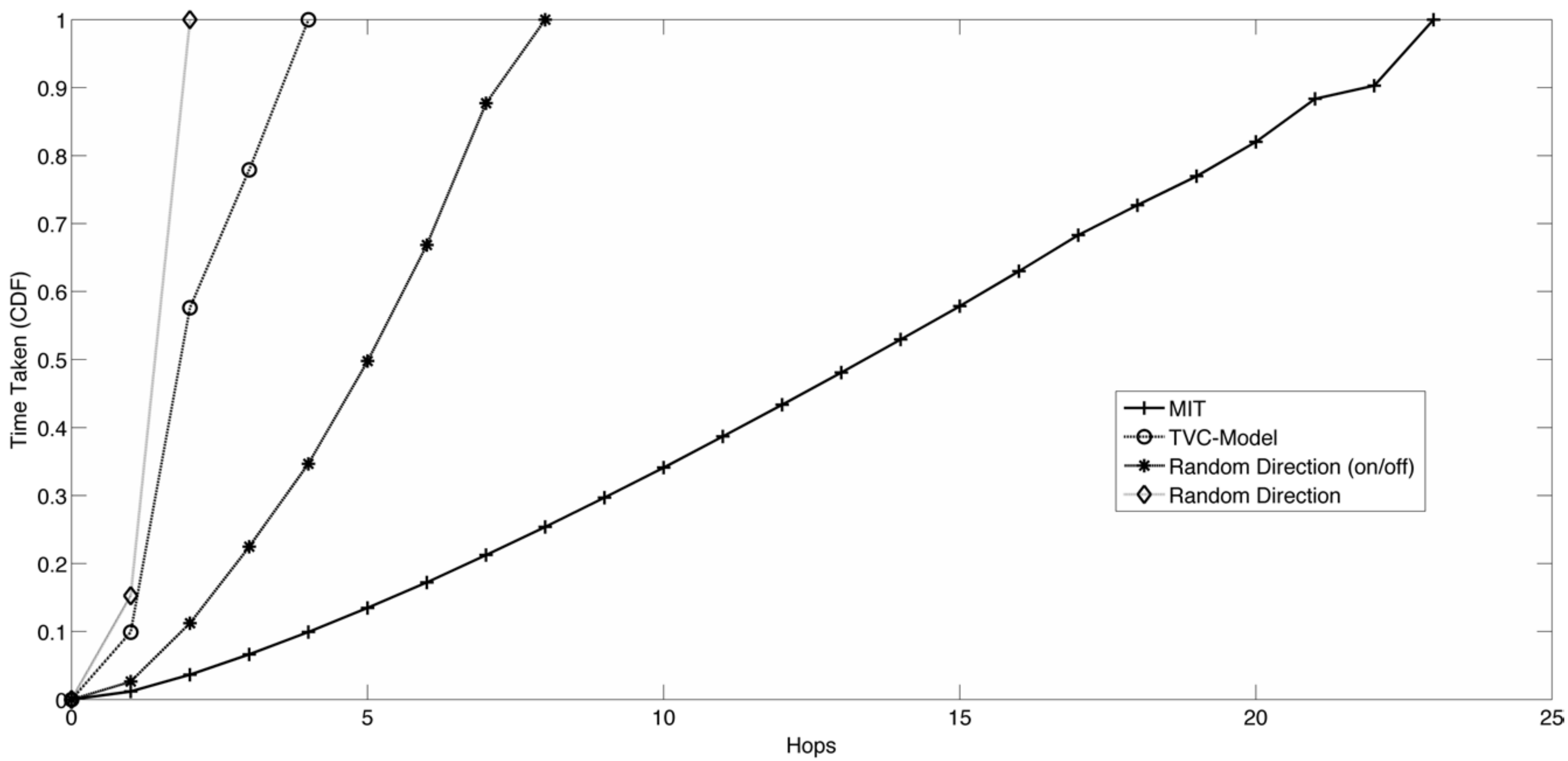
Delay


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Delay Percentage
1
2
3
4
5
6
7
8
9
10
11
Hops Overhead
Real Traces
TVC-Model
Random Direction Model

Message Overhead



Time Taken



Time Taken (CDF)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
Hops
MIT
TVC-Model
Random Direction (on/off)
Random Direction

Conclusion

- spatio-temporal preferences and encounter statistics are inadequate
- Models completely miss out on the performance criteria
- fundamental characteristics that drive the dynamics in challenged networks
- we should maintain the current characteristic but also look out for structural semblance and topological realisms between simulation and similarity

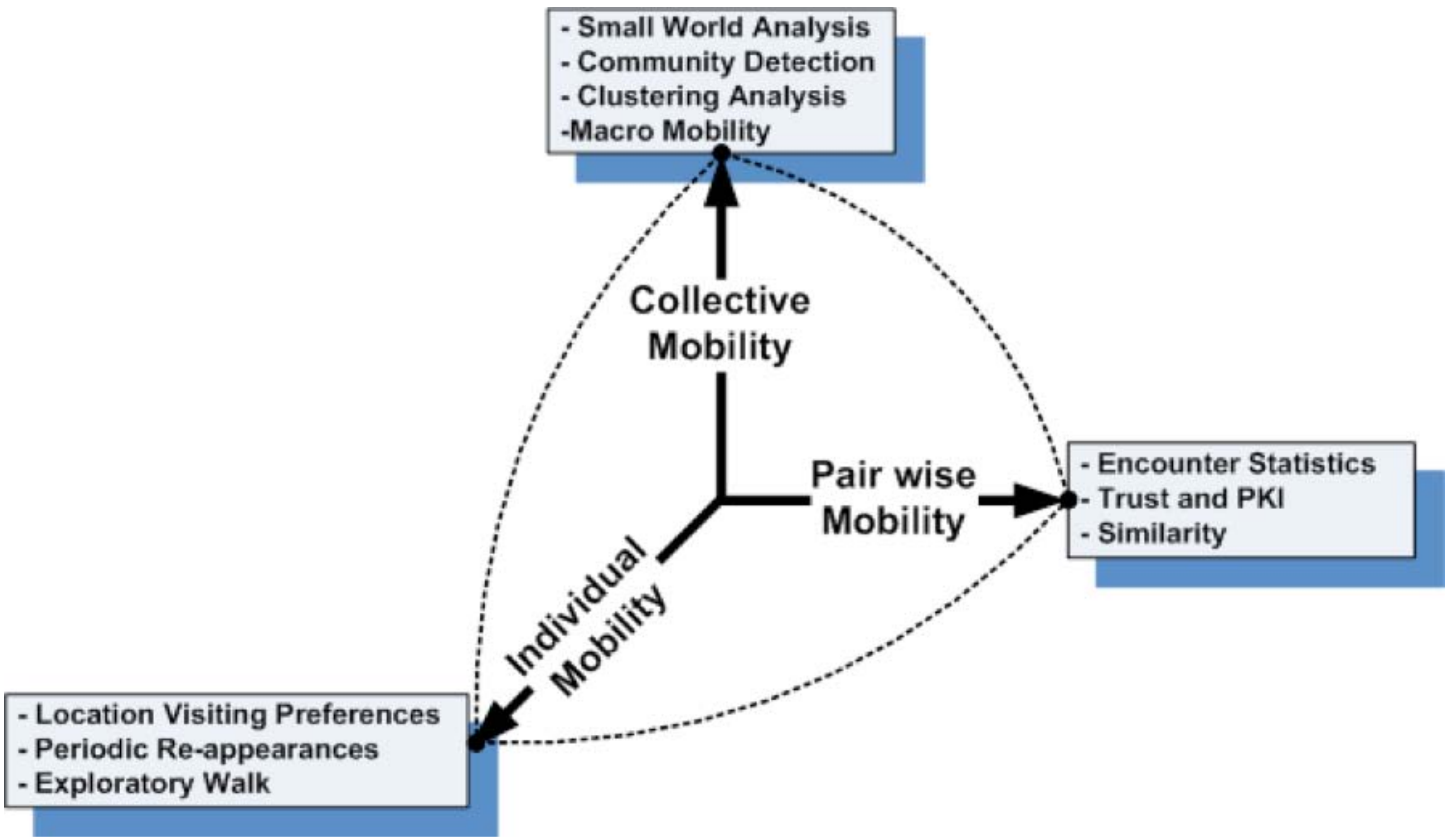
Framework for Behavioral Modeling in Networked Mobile Societies

- Studies have shown – Behavior impact Mobility.

- In turn, Mobility impacts the performance in routing protocols

=> Characterization of Human Behavior is very important.

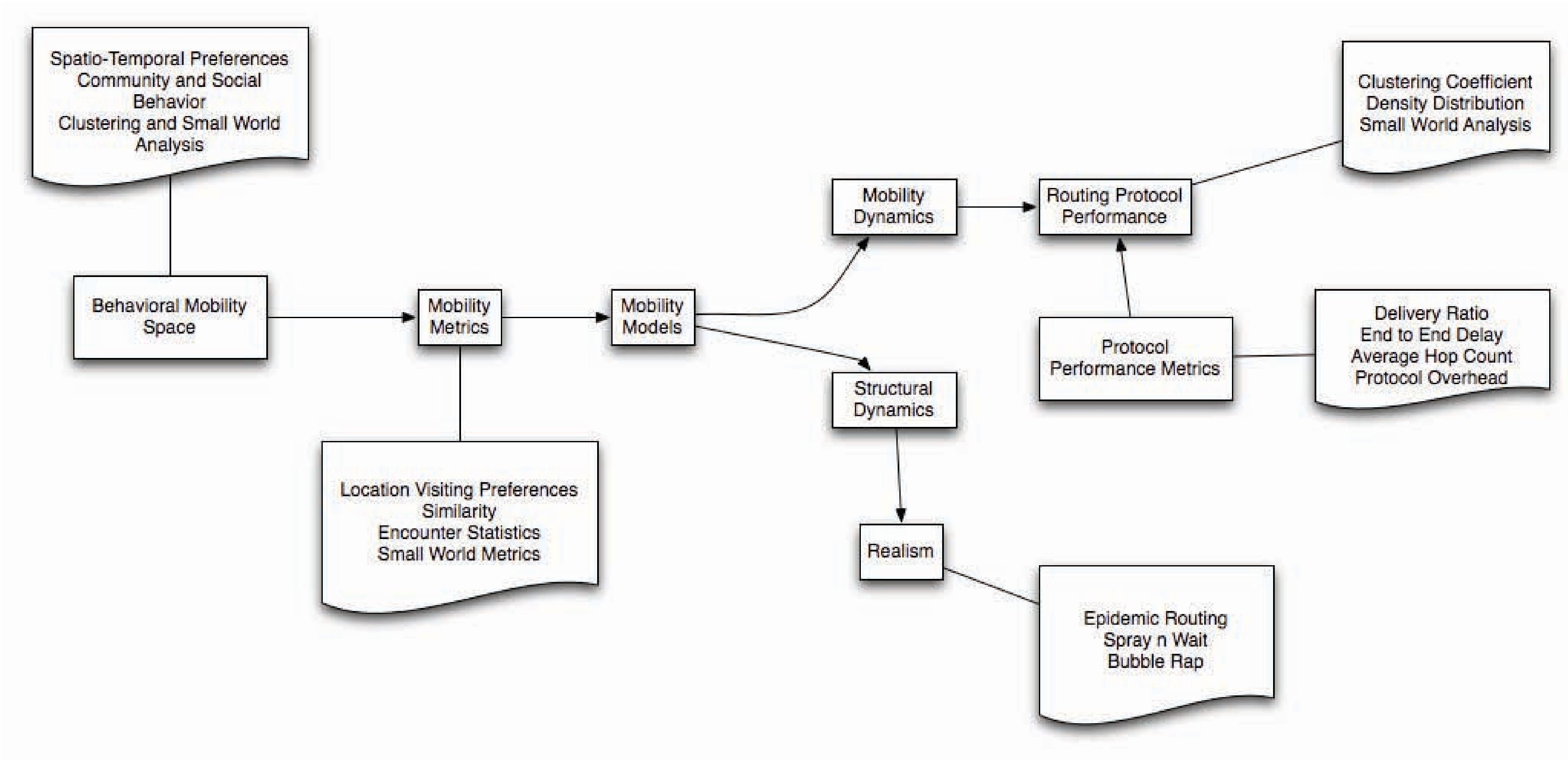
Behavioral Dimensions



Dimensions

- Individual
  - Spatio-temporal preferences
  - Levy Walk, Random Models
- Pair-wise
  - Encounter Statistics
  - Similarity
- Collective
  - Community Detection, Clustering
  - TVC, SWIM etc.

Framework



Spatio-Temporal Preferences
Community and Social Behavior
Clustering and Small World Analysis
Behavioral Mobility Space
Mobility Metrics
Location Visiting Preferences
Similarity
Encounter Statistics
Small World Metrics
Mobility Models
Mobility Dynamics
Structural Dynamics
Realism
Epidemic Routing
Spray n Wait
Bubble Rap
Routing Protocol Performance
Protocol Performance Metrics
Clustering Coefficient
Density Distribution
Small World Analysis
Delivery Ratio
End to End Delay
Average Hop Count
Protocol Overhead

Thank You !

Contact us:

Gautam S. Thakur
gsthakur@cise.ufl.edu

Dr. Ahmed Helmy
helmy@cise.ufl.edu

Dr. Wei-Jen Hsu
wehsu@cisco.com

Questions and Discussion

On-going Work - Analysis

- Network theoretic approach to understand similarity distribution
  - Small world analysis
  - Following known distribution, like Powerlaw
  - Dynamics and Stability over a large period
  - Similarity Percolation.
  - Trust generation and network resiliency.

On-going Work – Development

- Developing a mobility that
  - Showcase similarity pattern observed in realistic scenarios
  - Showcase Encounter Statistics, inter-contact time and duration of meeting
  - Showcase statistically equivalent network performance, viz. Delay, Reachability and Overhead.
- DTN Test-bed to reflect automated scenarios generation and performance analysis.

SHIELD

- an on-campus emergency rescue and alert management service
- based on proximity-enabled trust and cooperation
- relies on nearby localized responses sent using Bluetooth and/or WiFi
- augmenting the traditional notion of centralized emergency services

SHIELD – Basic Idea

- trust from mobile user encounters
- context aware service localization and signaling of historical crime log statistics
- preemptive response in averting the possibility of incident

SHIELD – Trust Generation

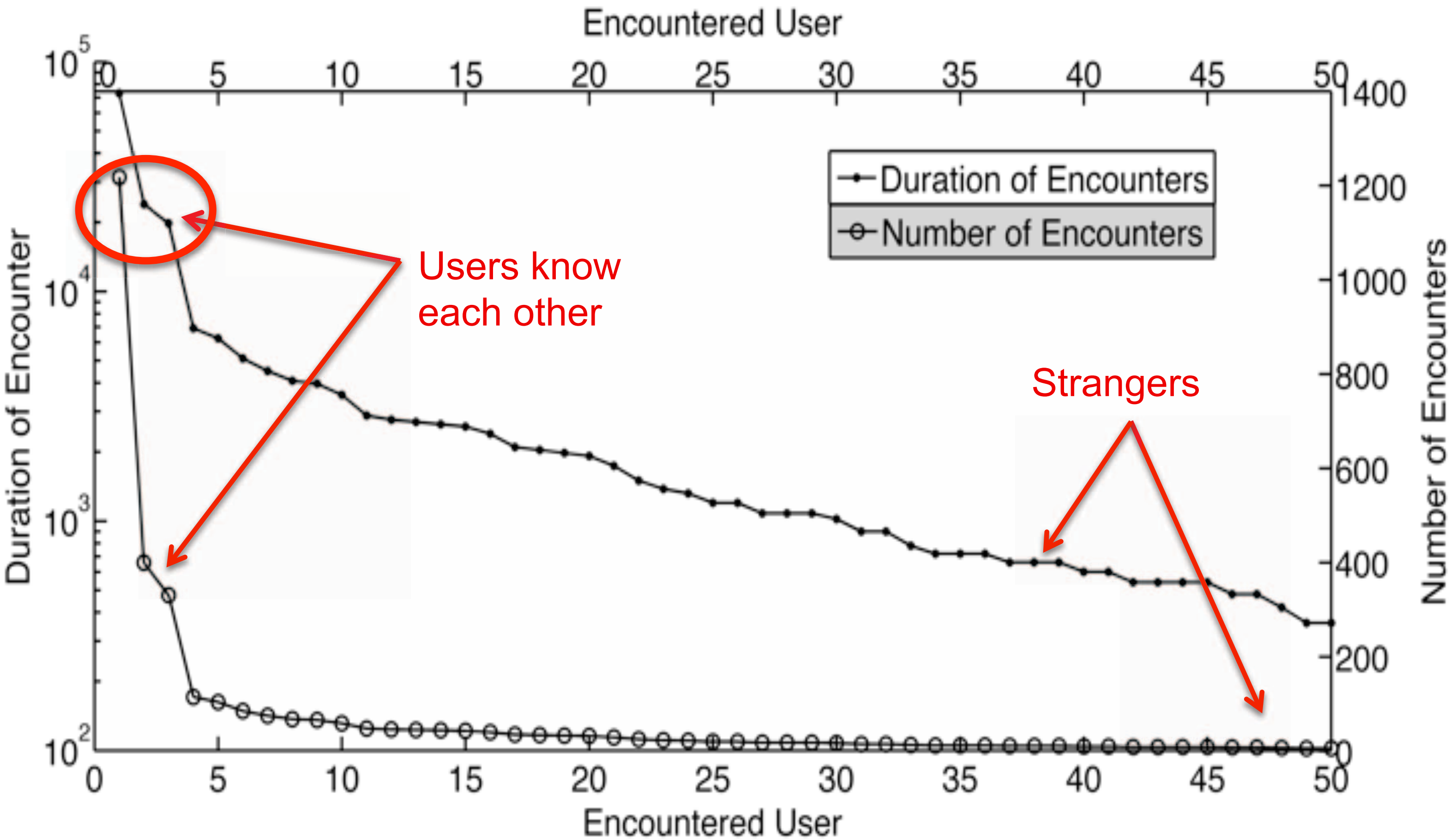
1. Number of Bluetooth encounters

$$f(i,j) = \sum_{i,j=1}^{n} \delta(i,j)$$

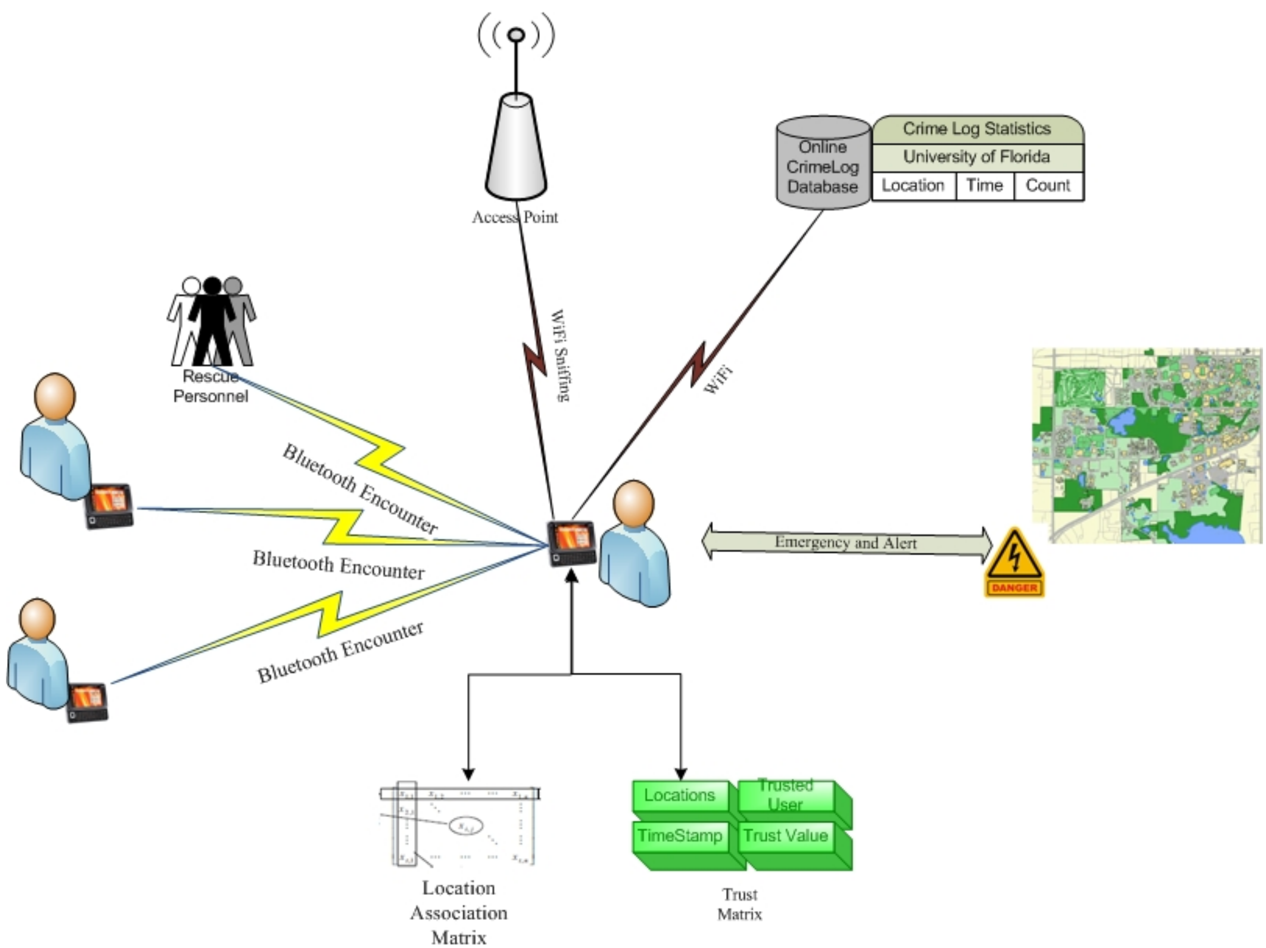
2. Duration of Bluetooth encounters

$$D(i,j) = \sum_{i,j=1}^{n} d(\delta(i,j))$$

Encounter Trace Analysis



SHIELD

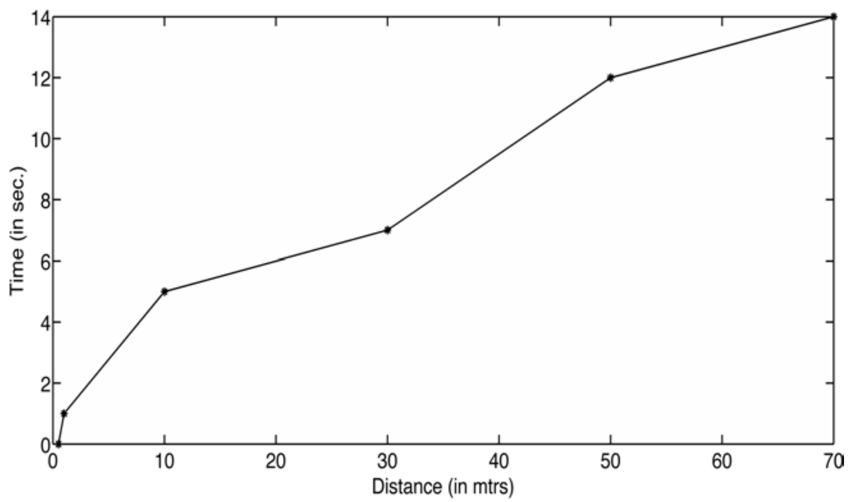


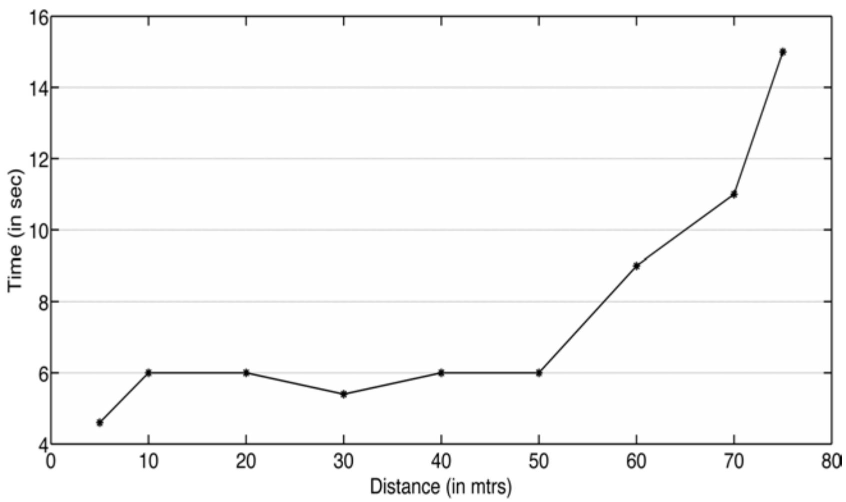
Access Point
Online CrimeLog Database
Crime Log Statistics
University of Florida
Location
Time
Count
WiFi Sniffing
WiFi
Rescue Personnel
Bluetooth Encounter
Bluetooth Encounter
Bluetooth Encounter
Emergency and Alert
DANGER
Locations
Trusted User
TimeStamp
Trust Value
Location Association Matrix
Trust Matrix

Test Bed Implementation

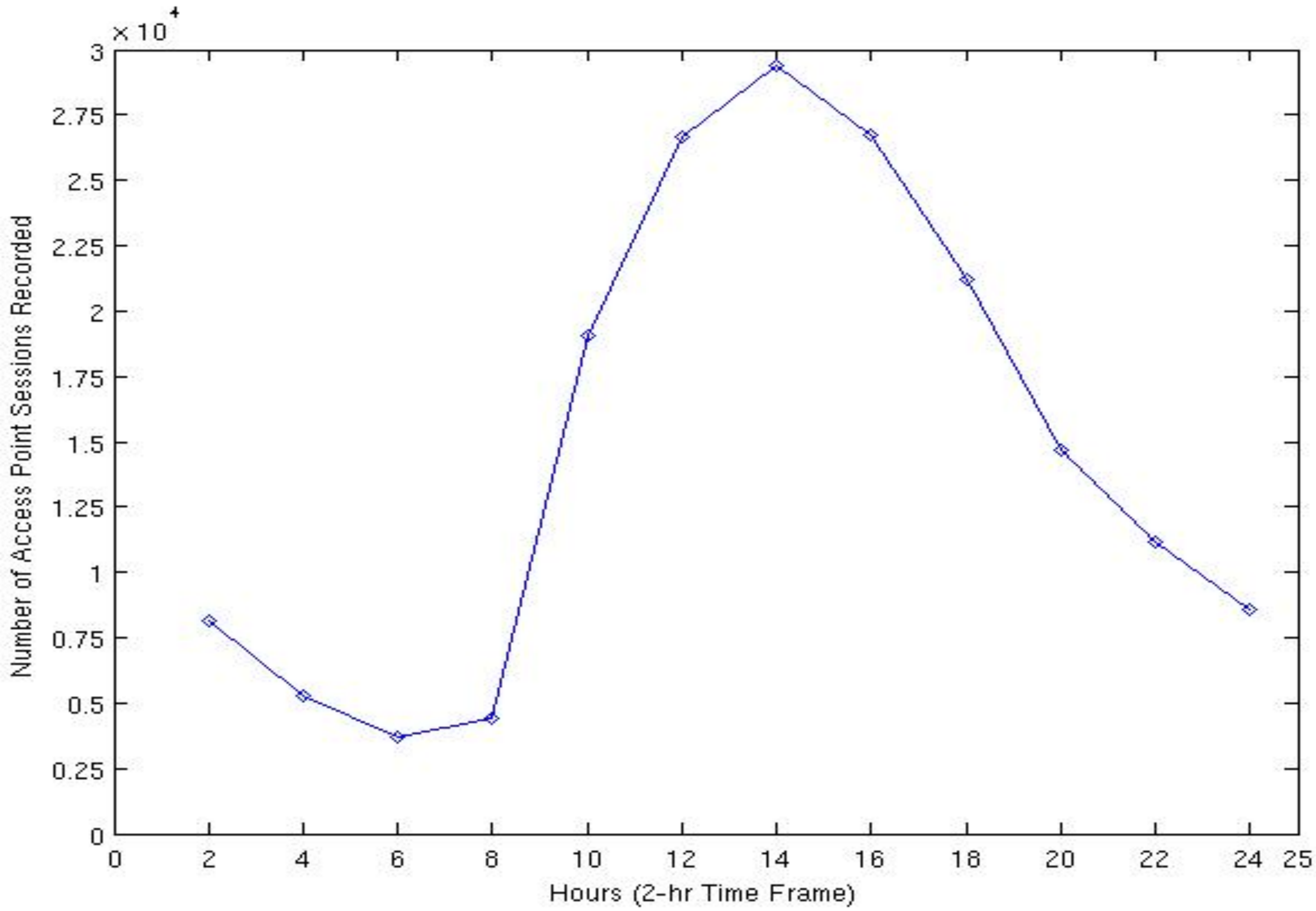
Bluetooth Scanning time that varies with distance

Connection & Transfer time for a message on Bluetooth



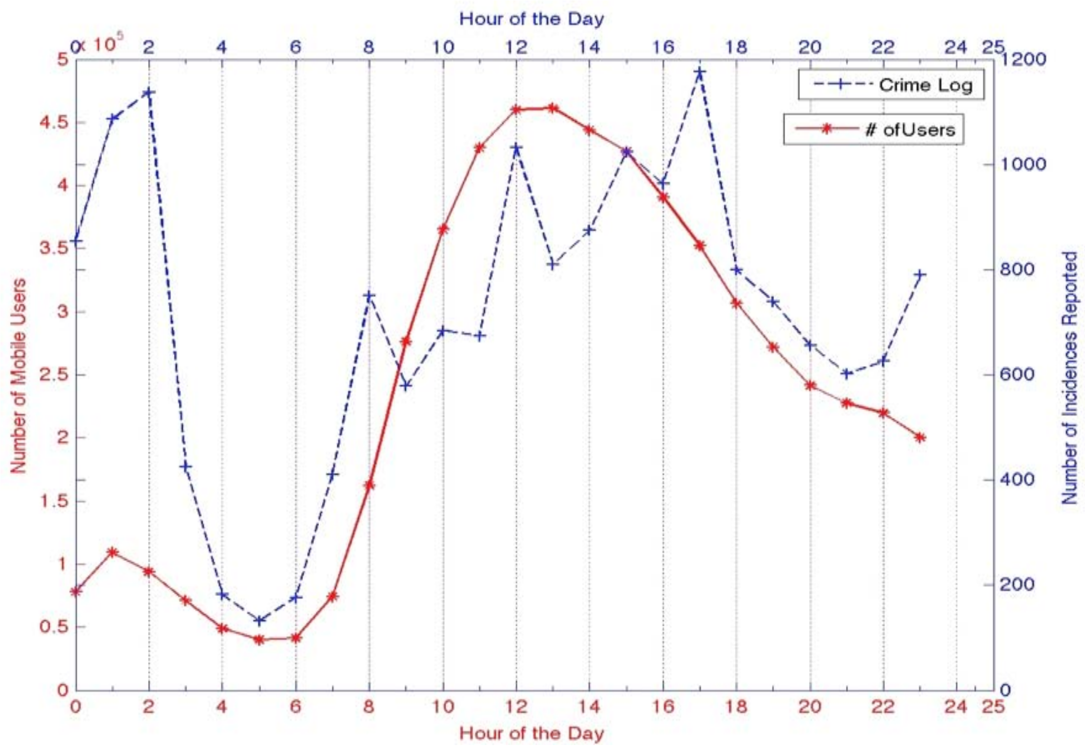


SHIELD – Login Patterns

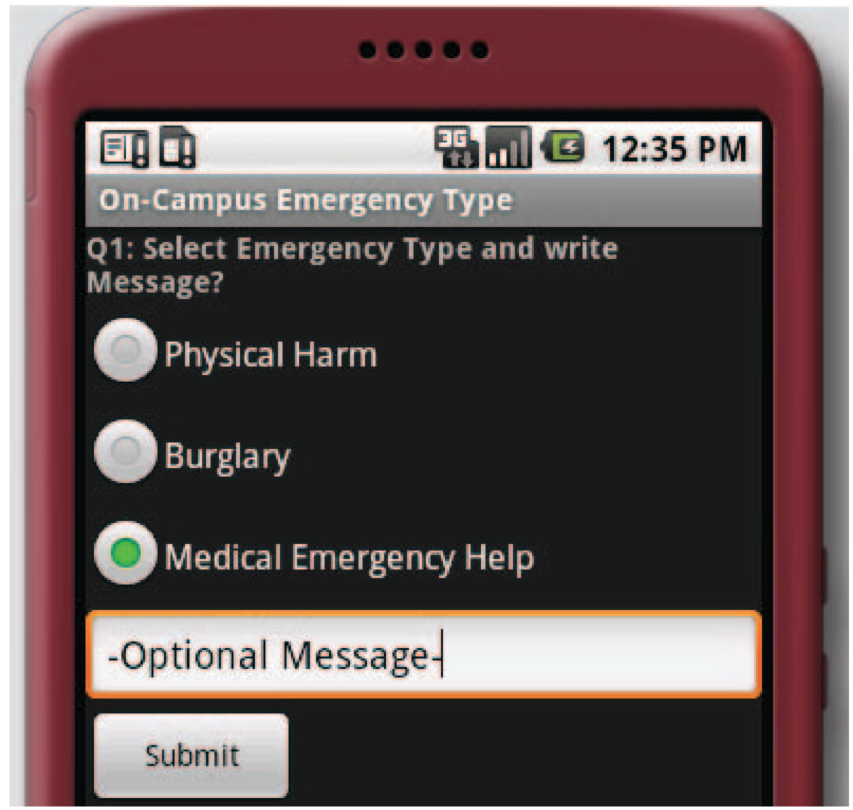


Crime Statistics and Mobile Users

- There is a positive correlation (~55%) between the incidences and the number of active mobile users.
  - *Thus, these incidences can be very well averted given proper preparedness exists for the mobile users.*



Application Prototype



Conclusion

- Utilize handheld devices in emergency rescue and alert scenarios
- SHIELD to establish spatio-temporal trust and cooperation for use in localized emergency alerts.
- Our statistical analysis reveals a positive correlation (55%) between on-campus crime incidences and density distribution of users.