*Computer and Information Science and Engineering Department*
*UNIVERSITY OF FLORIDA*

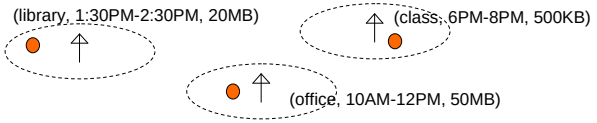
More info: http://nile.cise.ufl.edu/MobiLib/

# Visualization and Representation of Mobile Network Users

Udayan Kumar, Sungwook Moon, Wei-jen Hsu, Ahmed Helmy ukumar@cise.ufl.edu, smoon@cise.ufl.edu, wjhsu@ufl.edu, helmy@cise.ufl.edu

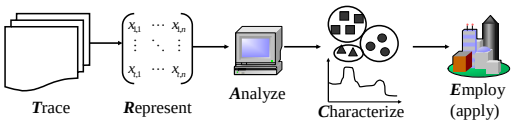
## 1. Introduction

- Large-scale deployments of wireless networks (e.g., on university or corporate campuses) provide rich sets of user activity logs[1][2].

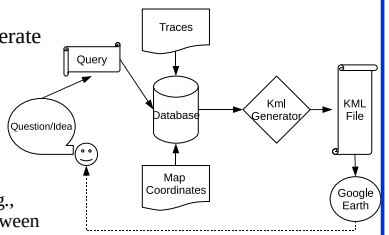



- **Visualization of such data sets helps to**
  - **Display the dynamics of users in its geographical context**
  - **Visually verify the discovered trends in the user behavior**
  - **Show-case various concepts**

## 2. Framework



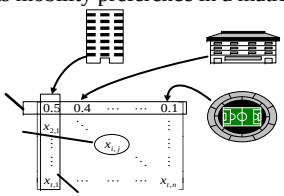

- In the *TRACE* framework we seek solid understanding of user behaviors based on empirical traces first, and then propose applications – bottom-up approach.
- In the demo we show an automatic process to generate KML-based animation on Google Earth [3].
  - Store WLAN traces and additional pre-processed information in a database
  - Specify queries to find out matched user (e.g., Who visits the library between 2PM-5PM?)



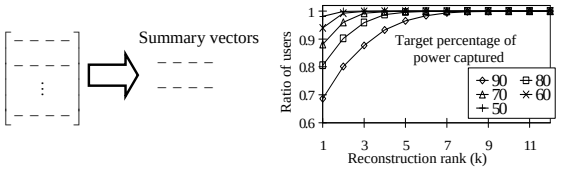

  - The returned entries from the database are parsed into KML format for Google Earth display.

## 3. Application: User Clustering [4]

- **How can we distinguish users with different behaviors from their activities logged in the WLAN trace?**
  - We use "summarized mobility history" as an example
- (STEP1) Each user represents its mobility preference in a matrix form




- (STEP2) The mobility preference matrix is summarized by singular value decomposition (SVD)




  - Reduce the computation overhead to calculate the similarity of users based on their "eigen-behavior vectors"

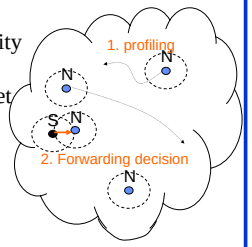
$$Sim(U,V) = \sum_{\forall i,j} w_i w_j \left| u_i \cdot v_j \right|$$

- (STEP3) Hierarchical clustering is used to classify users into ~hundreds of groups with distinct mobility trends

| Significance score of top eigen-behavior for | USC | Dartmouth |
|---|---|---|
| Its own group | 0.779 | 0.727 |
| Other groups | 0.005 | 0.004 |

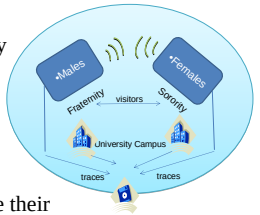
## 4. Application: Profile-cast [5]

- **A new communication paradigm in which the receivers are identified by their behavioral patterns, not the network IDs.**
- (STEP1) Each user represents its mobility preference and summarizes.
- (STEP2) The sender specifies the "target profile (TP)". Message follows nodes with increasing similarity to the TP.
- (STEP3) A group-spread phase follows after the message reaches a node that is highly similar to TP.




## 5. Application: Gender -Based grouping of WLAN users[6]

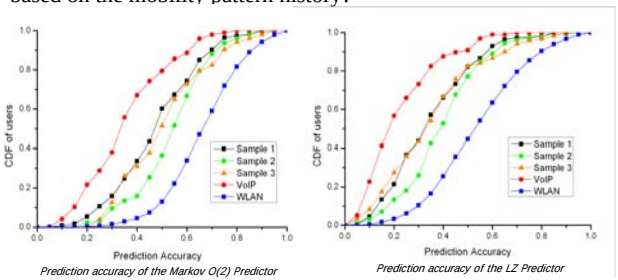
- **A method to classify users into gender based groups for mobile student communities.**
- People associating to AP in fraternity and sorority are classified into male and female after removing visitors using statistical methods.[6]




- The visualization framework enables us to display these users to show-case their mobility characteristics and also to look at other statistical data like manufacturer preference combined with gender information and mobility in a meaningful way.

## 6. Future Application: Mobility Prediction [7]

- **Are users predictable in terms of mobility? If so, how predictable are they (we)?**
- Mobile device users can be tracked with respect to their mobility patterns, and these patterns tend to be repetitive.
- Can we actually predict the location where the user will go next, based on the mobility pattern history?




*Prediction accuracy of the Markov O(2) Predictor* *Prediction accuracy of the LZ Predictor*

- Our previous findings indicate that WLAN users are more predictable than on-the-go VoIP users. How true is this?
- **By visualization we could estimate how close or far off our prediction really is.**

## 7. Future Application: Regular and irregular appearances

- **Regular appearance can help in deciding the messenger nodes in terms of probability to deliver data bundles.**
- Analyze the regular/irregular encounter patterns of the potential messenger nodes with the destination nodes.
- Only the nodes with regular encounters are candidates with calculated probability.

## 8. Future Application: Privacy and K-Anonymity

- **Ability to trace users, gives us an opportunity to understand human behavior, however it also leads to the challenging question about user's privacy**
- We want to apply/modify traditional techniques of data anonymization and see how they hold up in the mobile world.

[1] MobiLib: USC WLAN trace and pointers to many WLAN trace archives available at http://nile.usc.edu/MobiLib.
[2] CRAWDAD: A Community Resource for Archiving Wireless Data At Dartmouth. http://crawdad.cs.dartmouth.edu/index.php.
[3] Google Earth. Download from http://earth.google.com/
[4] W. Hsu, D. Dutta, and A. Helmy, "Extended abstract: Mining behavioral groups in large wireless LANs," in Proceedings of MOBICOM 2007. Longer version of technical report available at http://arxiv.org/abs/cs/0606002
[5] W. Hsu, D. Dutta, and A. Helmy, "Profile-Cast: Behavior-Aware Mobile Networking," in Proceedings of IEEE WCNC, Las Vegas, NV, Mar. 2008.
[6] U. Kumar, N. Yadav and A. Helmy, "Gender-based Grouping of Mobile Student Societies", in MODUS workshop, St. Louis, MO, April 2008 (colocated with IPSN 2008)
[7] J. Kim, Y. Du, M. Chen and A. Helmy, "Comparing Mobility and Predictability of VoIP and WLAN Traces", in CRAWDAD workshop, Montreal QC, Canada, Sep. 2007

This worked is supported by NSF CAREER Award 0134650