

Visualization and Representation of Mobile Networks

Jeeyoung Kim¹, Wei-jen Hsu², Sungwook Moon¹, Udayan Kumar¹, Ahmed Helmy¹

Department of Computer and Information Science and Engineering, University of Florida

Email: ¹{jk2, smoon, ukumar, helmy}@cise.ufl.edu, ²{wjhsu}@ufl.edu

I. INTRODUCTION

In recent years, the popularity of portable network capable devices has increased drastically. Devices such as laptop computers, PDAs, or smart phones have become commodities for a big part of the society, and the users integrate these devices with many different activities in lives. Such an unprecedented rapid adoption of new technology calls for the need of understanding the usage of these devices in realistic scenarios. This need is partially fulfilled by extensive collections of user behavior traces (e.g., see trace archives at [1], [2]) and detailed analysis of user behaviors based on the traces (e.g., [3], [4], [5]). The researchers seek to understand user behaviors based on a multitude of activities, such as time-location information (i.e., mobility), network usage, device types, etc.

As the number of users and the usage of existing users both increase significantly, the analysis of user behavior data faces several challenges. First, the sheer amount of data makes processing of the data a heavy task. More importantly, among the many potential aspects of investigation, the researcher has to quickly identify important and interesting hidden trends for further study. To tackle these challenges, it is very helpful to *visualize* a selected part of the large-scale trace data in the same geographic context (e.g., overlay the users on top of the campus map) from which the data is collected.

The visualization tool we demonstrate addresses the above challenges. Leveraging MySQL database and the animation capability provided by Google Earth [8], we provide a framework for the users of the tool to quickly animate a desired part of the data. This animation framework has several important features. (1) By using the MySQL database, it allows the maximum *flexibility*. Multiple data sets (those available from [1], [2]) can be inserted to the database under a properly designed schema to retain most of the information from the raw data format. The users, on the other hand, use various query criteria to obtain only the entries that meet the interest for the current investigation. (2) By converting the selected part of the raw data into the KML syntax and leveraging Google Earth, we provide a light-weight, easy to share animation file. (3) The visualization of trace data combined with the temporal and geographic context allows the researcher to effectively grasp an existing trend as the first step to validate ideas that worth further investigation.

In section II we further describes the methodology we use for the visualization framework. In section III we show several examples of visualizing our earlier research results. We discuss future work and conclude in section IV.

II. METHODOLOGY

In our study of traces from multiple sources[1], [2], we found that most of the mobile network's traces have a similar structure. They contain node ID (can be MAC ID or its anonymized version), association start time, association end time and the name of the Access Point(AP). They

frequently contain other information as well. For user visualization, we need to have at least these values for each session and information about its geographical context. We hence combine the network trace with the coordinates of each of the APs appearing in the traces. We put this information in a tabular structure in a Database. Now we allow a user to query this database for specific patterns but the results are not directly returned to the user. Instead, the results are passed into a KML (Keyhole Markup Language) file generator which puts the fields of the output of the database in the KML format in a file. This file can be visualized using Google Earth. The framework is illustrated in Fig. 1.

We have tried to keep our method flexible and generic. It is not tied to a particular trace source or a specific trace attribute; database allows arranging traces in a similar schema. Thus a researcher can analyze many different traces with similar queries. Often researchers want to analyze the same data in various representations (such as node level, group level or cumulative representation), and we allow this facility in our implementation. Using database makes annotation to existing trace entries straight-forward (by adding new columns or tables in the database), thus providing a flexible way to combine new information obtained from additional processing with existing data entries. Choice of a freely available cross platform tool like Google Earth enables easy sharing of animation files among the researchers. Demo video, codes, and more details are available at [9].

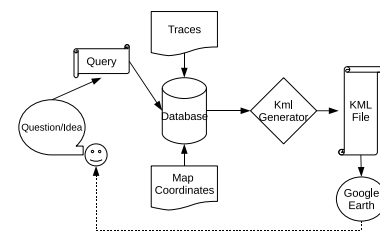


Fig. 1. Visualization framework.

III. VISUALIZATION EXAMPLES

Here, we provide examples of research problems where the visualization tool can help. These include clustering nodes, different node behaviors by gender, hot spots such as the most populated area and path of each node.

A. User Clustering

It has been shown users on university campuses can be classified based on their mobility preferences [4]. The mobility preferences of a user can be summarized in a matrix representation and singular value decomposition can be applied to obtain its important components, and the similarity of these components are applied to classify users. In Fig. 2 we display a snapshot of the animation of the users on campus, where users from different group are displayed with different colors. This animation allows us to see clearly the distinction of mobility among groups over time and space, visually validating the classification of users.

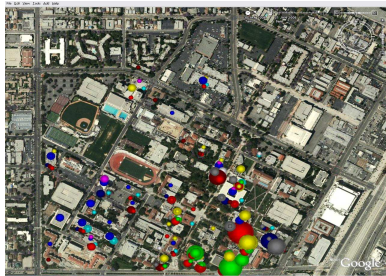


Fig. 2. Snapshot of user clustering animation. Users in different groups are shown with different colors.

B. Gender Difference

In [5] attempt has been made to group users biased on gender. In the technique, people associating to AP in fraternity and sorority are classified into male and female. After the classification, gender gap is studied in terms of preferred manufacturer, major or average on-line time. The visualization framework enables us not only to display these users to show-case their mobility characteristics but also to look at other statistical data like manufacturer preference combined with gender information and mobility in a meaningful way. Such a capability raises the issue of protecting user privacy when detailed traces are available for long time. We are also working on techniques to study k-anonymity for wireless network traces.

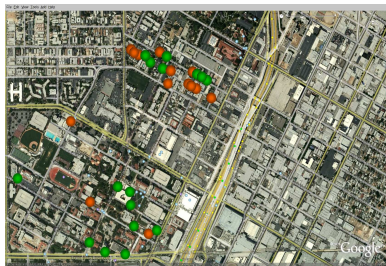


Fig. 3. Snapshot of gender difference animation. Orange nodes are males, green nodes are females.

C. Profile-cast

In future mobile networks, it is sometimes desirable to provide a service where the messages are delivered to nodes that match with certain target properties. This service is useful for targeted advertisement or resource discovery (when a given *property* of nodes is considered as a resource to search for within the general population). In mobile networks, it is difficult to maintain a mapping or directory service between network IDs of the nodes and their properties. We design the profile-cast service [6] to achieve message delivery to a *target profile* without first resolving the property to network identity mapping. Using the animation techniques presented in the demo, we show the progress of message forwarding in a realistic environment. Starting from the sender, the message is forwarded only to another node if it is more similar than the current message holder to the *target profile*. The message moves towards the intended receivers following the increasing similarity metrics, and eventually spreads among the intended receiver group. An snapshot for this animation is shown in Fig. 4.

D. Hot Spots and Predicted Paths

Other events of interest for trace visualization can be the population change in each location and the flow of users among locations. Polygon representation can show the dynamics of population changes in various locations, as shown in Fig. 5. It is also possible to display popular paths of user movements, to highlight user mobility across

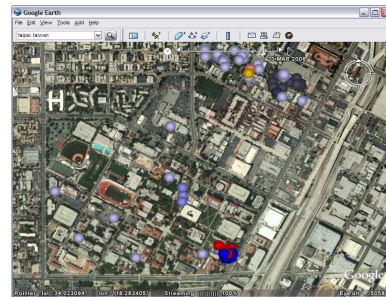


Fig. 4. Snapshot of profile-cast animation. The dark blue nodes are the ones with high similarity to the target profile (i.e., the intended receivers). Dissimilar nodes are in lighter colors. The sender is the yellow node, and all nodes that have received the message turn red. We can see that the message mainly reaches the similar users to the target profile, cutting down unnecessary transmissions.

the campus. This can be used to visualize and represent the mobility prediction [7] of users by allowing us to see different degrees of successful prediction. By creating a path between the predicted locations and the actual locations a user has traveled, we can determine the degree of separation between the two paths. A threshold will be used to enable us to quantitatively evaluate the success of a prediction.

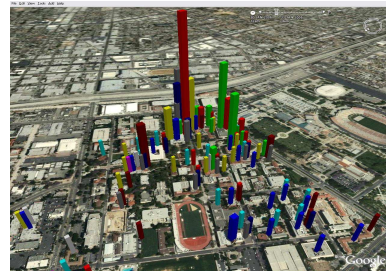


Fig. 5. Using polygon bars to visualize the most-visited hot spots on campus. The height of the bar is the number of current visitors.

IV. FUTURE WORK AND CONCLUSION

The visualization methodology we have developed provides flexibility for many scientific investigations into network traces. We show-case the applicability of the visualization framework to several earlier research results based on various data sets and ways to represent users. Google Earth and the KML representation provide a convenient cross-platform vehicle to share animations with geographic context.

In the future, we propose to build a web-based application with access to several traces. With various query criteria, the user can filter the trace and automatically generate the KML-based animation to help understanding the user behaviors. We will also investigate into more primitives for representing the data.

REFERENCES

- [1] MobiLib: <http://nile.usc.edu/MobiLib>.
- [2] CRAWDAD: <http://crawdad.cs.dartmouth.edu/index.php>.
- [3] T. Henderson, D. Kotz and I. Abyzov, "The Changing Usage of a Mature Campus-wide Wireless Network," ACM MobiCom, Sep. '04.
- [4] W. Hsu, D. Dutta, and A. Helmy, "Mining Behavioral Groups in Large Wireless LANs", ACM MobiCom, Sep. '07.
- [5] U. Kumar, N. Yadav, and A. Helmy, "Gender-based Feature Analysis in Campus-wide WLANs," ACM MC²R Jan. '08
- [6] W. Hsu, D. Dutta, and A. Helmy, "Profile-Cast: Behavior-Aware Mobile Networking," IEEE WCNC, Apr. '08.
- [7] J. Kim, Y. Du, M. Chen and A. Helmy, "Comparing Mobility and Predictability of VoIP and WLAN Traces", CRAWDAD workshop, ACM MobiCom, Sep. '07.
- [8] Google Earth. Download from <http://earth.google.com/>
- [9] More details for the poster at <http://nile.cise.ufl.edu/pages/demo-amp-presentations.php>