The Challenges of Accurate Mobility Prediction for Ultra Mobile Users

Jeeyoung Kim
jk2@cise.ufl.edu

Ahmed Helmy
helmy@cise.ufl.edu

Department of Computer and Information Science and Engineering
University of Florida, Gainesville, FL, USA

I. Introduction

Realistic modeling of user mobility is one of the most critical research areas in wireless networks. Currently, several mobility models are proposed based on the analysis of real WLAN traces [1,2,7]. However, the large collection of WLAN usage traces, seem to capture little mobility from the users.

In this paper, we focus on a subset of the wireless users, who use lightweight ultra mobile devices (i.e. VoIP devices). These users leave their devices on most of the time and the devices are light enough to walk-and-talk. Hence, these users show a more mobile (ultra mobile) characteristic than laptop users while connected to the network. We expect many of the users to become ultra mobile in the future with the introduction of the next generation handheld devices. By analyzing these traces we aim to compare behavior of ultra mobile device users to the general WLAN users. This sheds light on the realism of WLAN trace-based models. We also aim to examine the effect of any differences on protocol performance, e.g., prediction protocols.

Particularly, we compare the mobility of VoIP device user traces with whole WLAN traces (as used in previous studies) and also with some samples we have generated based on criteria that distinguish these samples as ultra mobile compared to others, such as prevalence, the number of APs visited and activity range. We compare these different sets of traces using several different predictors such as the Markov O(1), O(2), O(3) and also the LZ predictor. There has been work done with Wi-Fi mobility data using predictors [3] and work done on VoIP users [4] along with related work in the field such as analyzing different characteristics in various WLAN data [5,6,9], most of these works directly based on WLAN traces which can be found under the Mobilib project [10] or the CRAWDAD project [11]. There is even work done on prediction algorithms targeting cellular networks [8] but to the best of our knowledge there has not been any work done using predictors to compare the mobility of users. Our experiments indicate that the number of access points (AP) visited better represent mobility than the actual area range the user has covered and also that the Markov O(2) is the predictor with the highest accuracy among the four predictors and the LZ has the lowest. Surprisingly, all predictors perform quite poorly with VoIP device users compared to general WLAN users, prompting re-visiting of such algorithms for ultra mobile users.

II. Data Sets

The VoIP device data set we use in this work is a subset of the WLAN trace of Dartmouth College [11] and consists of 97 users. Along with the VoIP device data set, we have generated test sets from the same trace in order to validate our findings. There are three test sets used in this work and they are all considered to be ultra mobile users. Sample ap_200 and sample ap_170 are both based on the number of APs visited. Sample ap_200 is a collection of users who have visited 200 APs or more during the length of the trace and sample ap_170 is a collection of users who have visited more than 170 APs but less than 200. Sample range is a collection of users who have covered the largest physical area during the length of the trace. This was done by studying the AP location file and calculating the area range that each user covered. All of these samples were carefully selected from the 3 year long Dartmouth movement trace from 2001 to 2004 [11] and each test set has approximately 100 users each as shown in table 1.

Year	*Labels*	*# of users*	*Characteristics*
2001 - 2004	WLAN	13439	All the users studied in this trace
	VoIP	97	Users that use VoIP devices
	ap_200	112	Users that have visited more than 200 APs
	ap_170	127	Users that have visited more than 170 less than 200 APs
	range	113	Users who have covered the largest physical area range

Table 1 Characteristics of different sample sets extracted from the Dartmouth data trace

III. Mobility Comparison

In our work, we compared the mobility characteristics of WLAN user traces and ultra mobile user traces from several different aspects. The evaluation metrics include prevalence, the number of distinct access points visited and the activity range. The results of the comparison for each of our evaluation metrics is listed and shown as follows.

III.A. Prevalence

Prevalence is one of the mobility metrics proposed in [7], which indicates the time that a user spends at a given AP, as a fraction of the total amount of the time that they spend on the network. Higher prevalence means user spent more time on such an AP, and thus less mobile. Figure 1 and 2 show that VoIP users are more mobile than WLAN users, since the bar is lower. Especially for the most right bar which indicates prevalence higher than 0.95, the WLAN is much higher than VoIP. This means there are larger portion of users in WLAN who spent most of their time on only one AP than that in VoIP.

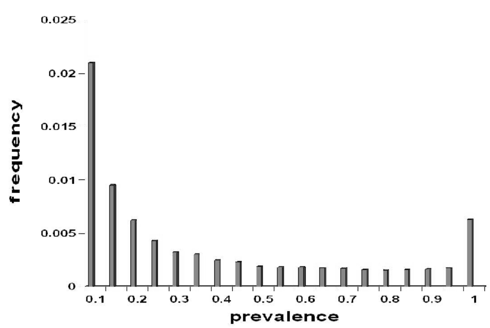

Figure 1 Prevalence of WLAN user trace

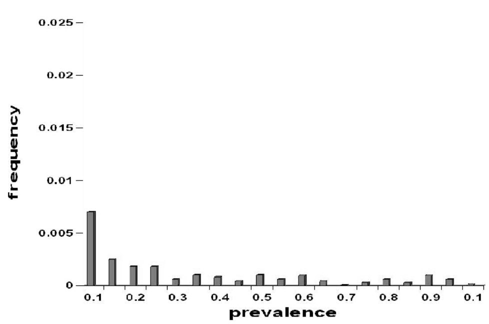

Figure 2 Prevalence of VoIP device user trace

III.B. Number of Distinct Access Points

Figure 3 and 4 shows the number of access points visited distribution CDF by WLAN and VoIP users. This clearly shows VoIP users visited much more access points than WLAN users. The average number of access points that the VoIP users visited is about 4.1 times than that of the WLAN users while the median number of access points that the VoIP users visited is about 7.7 times than that of WLAN.

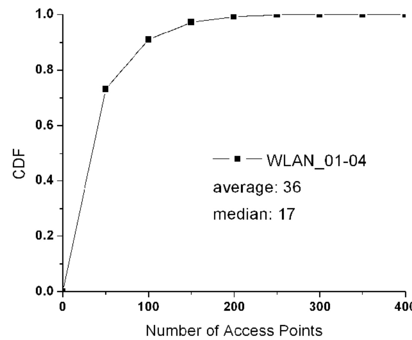

Figure 3 Number of Distinct APs visited per WLAN user

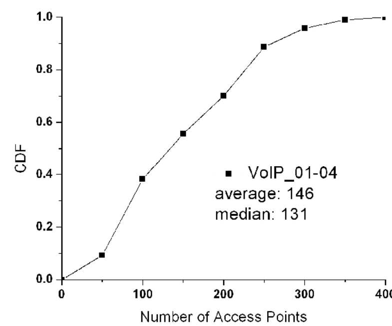

Figure 4 Number of Distinct APs visited per VoIP device user

III.C. Activity Range

Activity range is defined as the smallest square area which can cover all the access points the user has visited in an activity. Figure 5 and 6 shows the activity range distribution for WLAN and VoIP users. The percentage of VoIP users having a larger area of activity range is higher than that of the WLAN users who, most of the time tends to stay in a very limited area.

IV. Prediction Comparison

We have run the Markov O(1), O(2) and O(3) predictors along with the LZ [3] predictor for each of the test sets we have, and also for the VoIP trace set and the whole body of the WLAN trace. We also compared the accuracy of all four predictors with the VoIP trace data to see which one has the best performance. Accuracy is measured as percentage of correct predictions of the next AP to visit. As shown in Figs 7 through 11, the WLAN trace always had the best prediction accuracy for all the predictors with an average of about 60% accuracy. The VoIP trace, by contrast, had the worst prediction accuracy for all of the predictors with an average of approximately 25% accuracy. From these graphs we see that the best accuracy can be no more than 80% for VoIP users, while more than 95% accuracy for WLAN users.

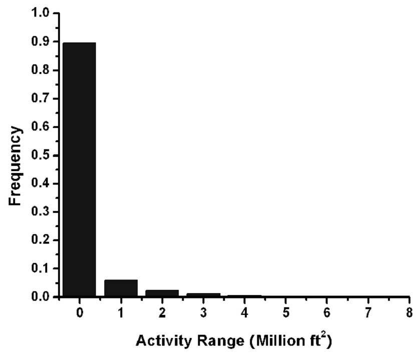

Figure 5 Activity Range Distribution of WLAN users

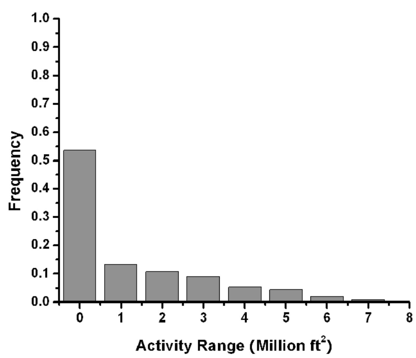

Figure 6 Activity Range Distribution of VoIP device users

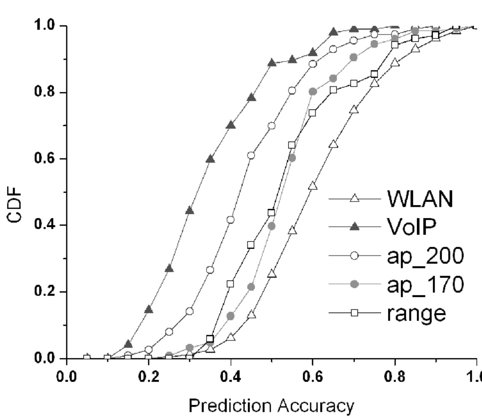

Figure 7 Prediction Accuracy for Markov O(1)

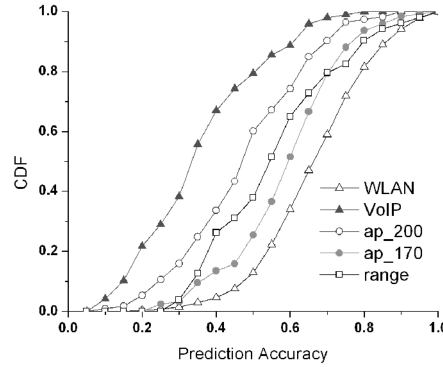

Figure 8 Prediction Accuracy of Markov O(2)

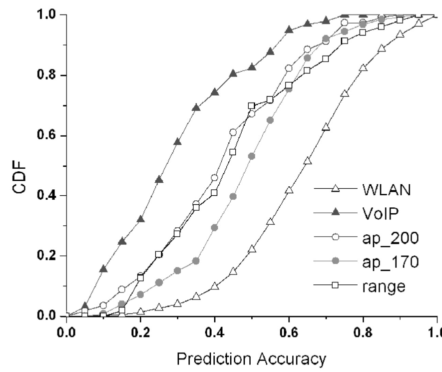

Figure 9 Prediction Accuracy of Markov O(3)

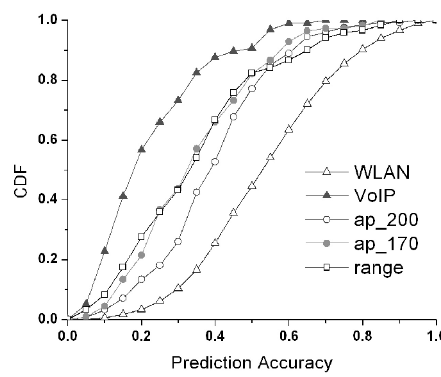

Figure 10 Prediction Accuracy of LZ

When we were first conducting our experiment, we expected that the range of the physical area that each user covered would be better criteria to measure mobility than the number of APs visited since we consider a person to be more mobile when that person covers more ground. Hence, we expected that sample range would return very bad prediction accuracy. Surprisingly, sample range always exhibits performance between the other two samples (ap_200

& ap_170), which indicates that the users that covered larger areas physically most likely have visited an average of 200 APs during their lifetime.

To explain this result intuitively, the sample of users which had visited less APs have a better prediction rate than that of the user who have visited more APs. The difference of the prediction accuracy between the two samples is always around 10% near the median.

As for the comparison of the predictors on the VoIP data set, the LZ predictor showed the worst prediction rate and the Markov O(2) showed the best prediction accuracy by a very minimal difference from the Markov O(1). Markov O(3) did not show a good prediction and these results indicate that a larger data structure and higher complexity does not help in making better predictions. However, the four predictors that are used in this work do not provide good prediction for the VoIP data set.

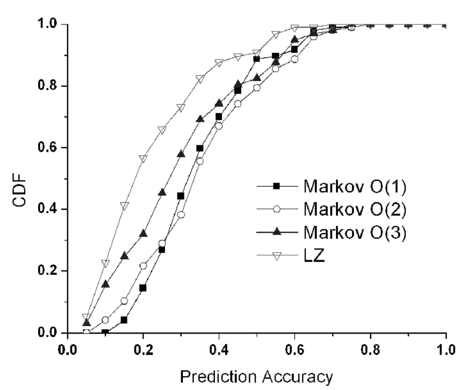

Figure 11 Comparison of Predictability on VoIP device users

V. Future Work

Our findings open the door for improved prediction and modeling of ultra mobile users. We plan to design a better predictor for “ultra mobile” users, especially the VoIP traces. Our plan includes investigating domain specific knowledge, regressions, schedules and repetitive or preferential user behavior. The success rate should also be taken in to consideration since depending on the granularity of the success rate the prediction accuracy may be highly affected. We shall also examine the adequacy of WLAN trace based mobility models for ultra mobile and VoIP users, that are likely to increase in the future.

References

[1] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, “Modeling Time-variant User Mobility in Wireless Mobile Networks”, Proceedings of IEEE INFOCOM 2007.

[2] C. Tuduce, T. Gross, “A mobility model based on WLAN traces and its validation”, Proceedings of INFOCOM 2005: Miami, FL, USA 664-674

[3] L. Song, D. Kotz, R. Jain and X. He, “Evaluating location predictors with extensive Wi-Fi mobility data”, Proceedings of IEEE INFOCOM 2004.

[4] M. Kim, D. Kotz, and S. Kim, “Extracting a mobility model from real user traces”, Proceedings of IEEE INFOCOM 2006.

[5] M. Balazinska and P. Castro, "Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network" in International Conference on Mobile Systems, Applications, and Services, May 2003

[6] T. Henderson, D. Kotz, I. Abyzov, “The changing usage of a mature campus-wide wireless network”, Proceedings of MOBICOM 2004: 187-201

[7] W. Hsu and A. Helmy, “On Modeling User Associations in Wireless LAN Traces on University Campuses” The Second International Workshop on Wireless Network Measurement (WiNMee 2006), Boston MA, Apr. 2006.

[8] H. Zang and J. C. Bolot, “Mining Call and Mobility Data to Improve Paging Efficiency in Cellular Networks”, Proceedings of the MOBICOM 2007: 123-134

[9] F. Chinchilla, M. Lindsey, M. Papadopouli, “Analysis of wireless information locality and association patterns in a campus”, Proceedings of the IEEE INFOCOM 2004, Hongkong, China

[10] http://nile.cise.ufl.edu/MobiLib/

[11] http://crawdad.cs.dartmouth.edu