The Challenges of Accurate Mobility Prediction for Ultra Mobile Users

Jeeyoung Kim (Student) and Ahmed Helmy (Faculty)
Department of Computer and Information Science and Engineering, University of Florida
Email: {jk2, helmy} @cise.ufl.edu

I. Introduction

Realistic modeling of user mobility is one of the critical research areas in wireless networks. Mobility data based on real human behaviors may give us the opportunity to improve wireless and mobile services for users in many ways. 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* of users.

In this paper, we focus on a subset of the wireless users, who use wireless *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 other heavy device users while connected to the network. By analyzing these traces we aim to compare behavior of ultra mobile VoIP 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.

We generate data samples based on metrics that distinguish these samples as *ultra mobile* compared to other users, such as prevalence, the number of APs visited and activity range. These different sets are then compared using several predictors such as the Markov O(1), O(2), O(3) and 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 Mobilib [10] or CRAWDAD [11]. There is 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 has more to do with 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 Cisco 7920 and Vocera phone users. We also generated 3 test sets from the same trace in order to validate our findings. They are all considered to be ultra mobile users. Test sets ‘ap_200’ and ‘ap_170’ are both based on the number of APs visited. The ‘ap_200’ set is a collection of users who have visited 200 APs or more during the length of the trace and ‘ap_170’ is a collection of users who have visited more than 170 APs but less than 200. The ‘range’ test set is a collection of users who have covered the largest physical area during the length of the trace. All of these test sets 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.

III. Mobility Comparison

In our work, we compared the mobility characteristics of WLAN traces and VoIP traces from several different aspects. The evaluation metrics include prevalence, the number of access points visited, and the activity range. The results of the comparison for our metrics are listed as follows[1].

A. Prevalence

Prevalence is a mobility metric 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 that a user spent more time on a given AP, and thus less mobile. Our studies show that the frequency of WLAN users which have a prevalence higher than 0.95 is 0.6% which is more than 6 times higher than that of the VoIP users which is less than 0.1%. This means there are larger portion of users in WLAN who spent most of their time on only one AP than that in VoIP.

B. Number of Access Points Visited

The average number of access points that the VoIP users visited which is 146 is about 4.1 times than that of the WLAN users which is 36 while the median number of access points that the VoIP users visited which is 131 is about 7.7 times than that of WLAN which is 17.

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.*

About 90% of the WLAN users tend to stay within a limited area (less than 1000000) and 10% covered areas larger than 1000000, while more than 45% of VoIP users covered areas larger than 1000000.

IV. Prediction Comparison

We have run the Markov O(1), O(2), O(3) and LZ predictors for each of the test sets along with the VoIP trace set and the complete 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 1

[1] Further result analysis is omitted for brevity and can be found at http://www.cise.ufl.edu/~jk2/doc/Tech_Report_Kim.pdf

through 5, 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.

When we were first conducting our experiment, we expected that the range of the physical area that each user covered would

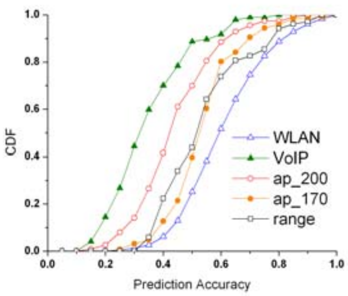

Fig1. Accuracy of Markov O(1)

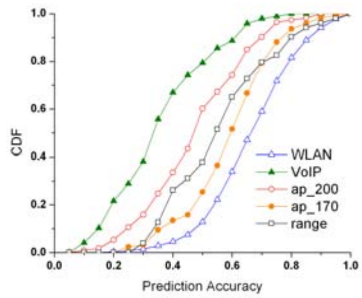

Fig2. Accuracy of Markov O(2)

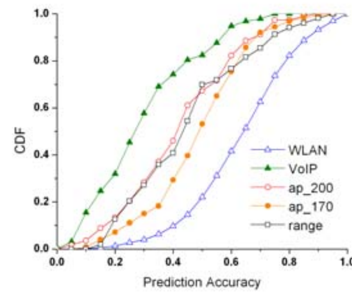

Fig3. Accuracy of Markov O(3)

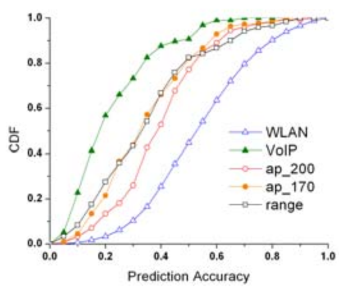

Fig4. Accuracy of LZ

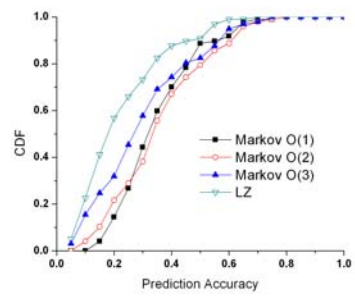

Fig5. Comparison of Predictability in VoIP trace

be a better criterion 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 the 'range' set would return very bad prediction accuracy. Surprisingly, the 'range' set always exhibits performance between of the other two test sets (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 better prediction rates 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.

V. FUTURE WORK

Small yet powerful handheld devices such as the Blackberry or iPhones are increasing drastically. Our findings open the door for improved prediction and modeling of these ultra mobile users. We plan to design a better predictor for "ultra mobile" users, especially the VoIP traces as well as these newly emerging devices. 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 MobileNetworks", 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 maturecampus-wide wireless network", Proceedings of the 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 toImprove 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