Validation of the Random Waypoint Mobility Model Through a Real World Mobility Trace

Andres Rojas, Philip Branch and Grenville Armitage
Centre for Advanced Internet Architectures
Swinburne University of Technology
Melbourne, Australia
Email: {anrojas, pbranch, garmitage}@swin.edu.au

***Abstract*— The Random Waypoint Model (RWP) is a simple mobility model based on random destinations, speeds and pause times. The RWP is one of many mobility models used in simulations of mobile communications networks to model human movement. The RWP is often criticised as not being representative of how humans actually move. Paradoxically, validation of the RWP against real mobility data is seen as being difficult due to the impracticalities of obtaining real mobility data.**

In this paper we consider the RWP as a model of user mobility in networks that cater for a large geographical areas, such as a city. We present results from a real world user movement trace and use these to validate some of the key characteristics of the RWP. The data presented was obtained from one individual's movement around the city of Melbourne, Australia, for a period of two months and included recording the individual's destinations, speed, rest times, and routes of travel.

I. Introduction

The Random Waypoint Model (RWP) is a simple mobility model in which a user is given a random destination to travel to at a randomly chosen speed and, once there, pauses for a random amount of time before the whole process starts again. As with other available mobility models, the RWP is used in simulations of mobile communications systems to investigate the impact of user mobility on the network under test. The RWP is commonly used to model the movements of users in studies of mobile ad-hoc routing and transport protocols [1], [2], [3].

Although the RWP is a good starting point in modeling human movement, it is often criticised as lacking reality [4], [5]. The issue of reality in simulation is an important part of any network performance analysis. Any simulation based on unrealistic assumptions about the way in which users move will have unrealistic results. It is often argued that humans simply do not move in a random fashion such as that represented by the RWP and that the RWP has not been validated against real human movement. Paradoxically, it is also stated that obtaining real life mobility data is a very difficult task, thus making any validation of the RWP or any other model impossible [4], [5], [6].

In this paper, we give details of a real world movement trace and use the results therein to validate some of the RWP's key characteristics. The characteristics of the RWP that we analyze are the distribution of destination states (waypoints) and the distribution of rest time at these destination states. In a typical RWP model, these characteristics are modeled by uniform and exponential distributions respectively.

The real life movement trace used in this study was collected by a simple method that involved manually documenting a daily record for the travel undertaken by one of the authors. This process, which occurred every day for a period of two months, yielded detailed information about the travel method, average speed of travel, destination and route of the trips that were taken throughout the city of Melbourne.

As the data collected pertained to movement of a user over a large geographical area, the validity of the validation results presented should only be applicable to studies of networks that offer mobility over that same scale of geography. A candidate technology is Mobile IP ([7], [8]) which allows nodes to remain reachable while moving around in the Internet. Although the use of the RWP is typical in studies of mobile ad-hoc networks, users in an ad-hoc environment might move with a different pattern of movement than what is presented in this paper.

Although the sample size of the collected data is small, the results and their significance are worthy of discussion by virtue of the fact that there is simply no other validation study in the literature of the RWP within the context explored in this paper.

II. Related Work

Surveys of mobility models, as used in communications network research and simulations are presented in [9] and [6].

In order to model real human mobility, some authors have refined the simple nature of existing models. Bettstetter modified the Random Direction Model (RDM), a model similar to the RWP, to be a better representation of the micro movements that humans exhibit when reaching and leaving destinations, [6]. Bettstetter proposed a more realistic model in which speed and direction values are correlated to previous values, thus producing a smoother representation of movement.

Instead of refining existing simple models, some authors use human characteristics to arrive at a more complex representation of human mobility. Social network theory is used by Musolesi et. al., [5], to generate a mobility model for ad-hoc networks which focused on the mobility of a groups of individuals.

In terms of validation of the RWP, there are cases where having a trace of user mobility, the data is not used to explore whether or not the data verifies the RWP or other simple mobility models. Fitzek et. al., [4], studied the effects of mobility on ad-hoc network performance by using a model derived from multi-player computer games. Although the method of using a virtual trace of movement is novel and produces movement which are human like, an investigation into whether or not simple models, like the RWP, could be verified from this method was not conducted.

As another example, Tuduce and Gross, [10], collected user mobility data on a university campus WLAN. In addition to producing a framework that can be used to generate mobility data from WLAN traces, the authors also compared their mobility framework with two mobility models, the RWP and the Reference Point Group Model (RPGM), [11], in terms of the impact on the performance of ad hoc routing protocols. They explored the relative speed of nodes, link duration between nodes and the extent of similarity of the velocities of two nodes. However, they did not explore whether or not their mobility traces could validate the RWP.

We also note that many authors, when discussing the difficulty of obtaining real world traces of user mobility, only consider traces of mobility data from networks that are already deployed. As we show in the following section, this is not the only way to obtain a trace of mobility.

III. Methodology

The data presented in this paper was collected using the following rudimentary method. At the end of each day during the collection period, the trips that were undertaken by the subject, one of the authors, were recorded by hand in a log book. Each description of a trip in each daily record consisted of:

- Time at the start of the trip,
- Time at the end of the trip,
- Method of travel,
- Average speed during the trip,
- Destination, and,
- Route taken.

The destinations and route taken were described in terms of Melbourne's street directory's map references [12]. Each map reference was then transcribed into a two dimensional Cartesian coordinate system with x and y coordinates that represent a state. Each state was approximately 300m X 300m in area.

The collected data was obtained during a period of 2 months during October-November, 2004. During this time, there were 58 days where the subject undertook at least one trip and 3 days where there was no travel at all. The methods of travel included walking, riding a bicycle, driving a car, and transportation through public trains and trams.

The best measurement of speed that could be attained with this simple method was that of an average speed during each trip. For most travel methods this was a subjective estimation of how fast each trip was. For trips undertaken while riding a bicycle an average speed reading from a speedometer at the end of the trip was used.

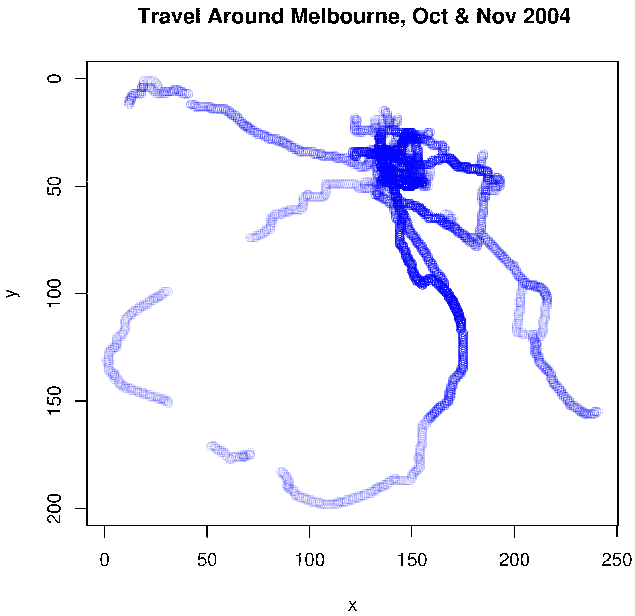


Fig. 1. Trace of movement, Oct-Nov, 2004. States are plotted as transparent points so that a subjective view of how many times each state was traversed can be obtained.

IV. Results

A. Pattern of Movement

The complete trace of movement recorded can be seen in figure 1. The figure shows that movement was predominantly concentrated in the upper central portion of the Cartesian plane. The figure also shows areas where there is a gap in the coverage of the street directory.

Figure 2 shows a 3D plot of the number of times each state was visited throughout the data gathering period. The figure is restricted to show the most visited areas. The most visited state corresponds to the subject's home (approx. at (0.2, 0.8)) in figure 2) with over 100 visits. The second and third most visited states correspond to Melbourne's central train station with over 60 visits (0.3, 0.6) and the subject's place of work with over 40 visits (0.7, 0.4) .

B. Rest Times at Destinations

Figure 3 shows the cumulative rest times at each destination. That is, how much total time was spent resting at each of the destination states. The figure's y-axis is logarithmic as the cumulative rest times of some states are extremely large compared to the destination states that were visited for a short period of time. Figure 3 is restricted to show only the most active areas.

Only three states had a cumulative rest time larger than 10^5 seconds. The state with the largest cumulative rest time is the subject's home with over $10^{6.3}$ seconds (approx. at (0.2, 0.8)). The second and third largest cumulative rest times

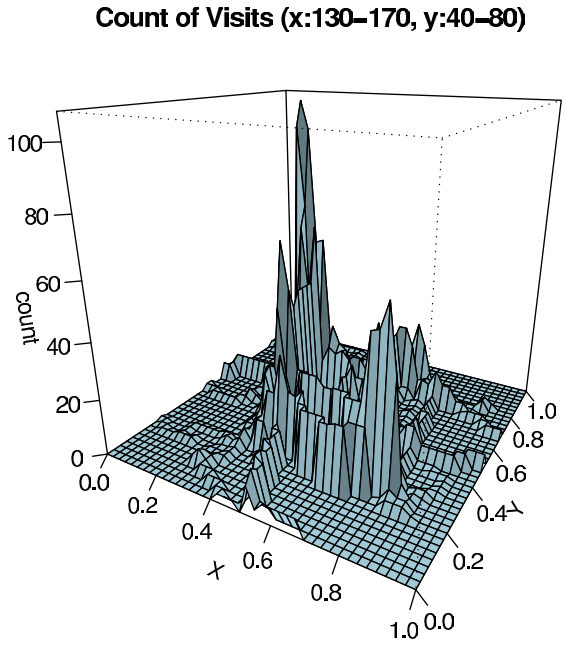


Fig. 2. 3D plot of the number of times each state was visited. Only the most visited part of the whole trace is shown.

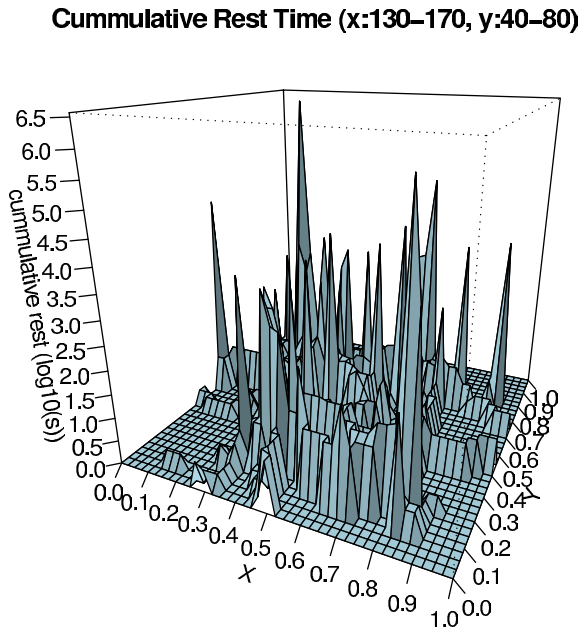


Fig. 3. 3D plot of the cumulative rest times spent at destination states.

corresponded to the subject's work with $10^{5.7}$ seconds (0.8, 0.2) and the subject's in-law's house with $10^{5.2}$ seconds (0.6, 0.9).

C. Distribution of Destinations

A histogram of x-coordinates for all destination states is shown in figure 4. Overwhelmingly, most x-coordinates are those between the value range 140-170.

The y-coordinate of the destination states yields similar results as the results shown here for the x-coordinates, and so, are not shown or discussed further.

Figure 5 shows the cumulative distribution function of the x-coordinates. Superimposed over the data are the cumulative distribution functions of the uniform, normal, and Cauchy distributions. The normal distribution shown has the same mean and standard deviation as the x-coordinate data. The Cauchy distribution shown has the mean of the x-coordinate data as it's location parameter and a shape parameter equal to 3.

From figure 5 it is clear that the x-coordinate data is poorly represented by a uniform distribution. The normal distribution shown is a better match when compared to a uniform distribution, while the Cauchy distribution is an even closer match to the data.

In order to have a clearer comparison between the distribution of the x-coordinate data and the normal and Cauchy distributions, q-q plots are presented in figure 6. In the normal case, figure 6(a), the data shows an S-like pattern with the first and last few points deviating markedly from the line of slope

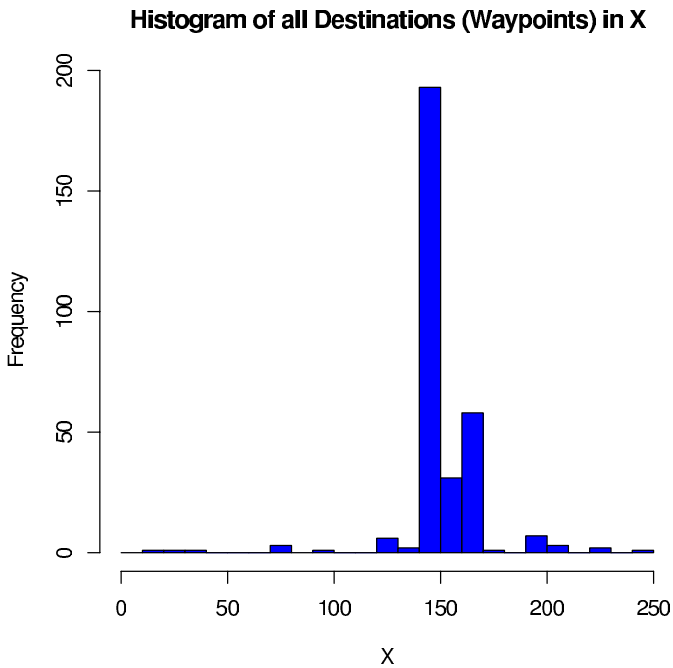
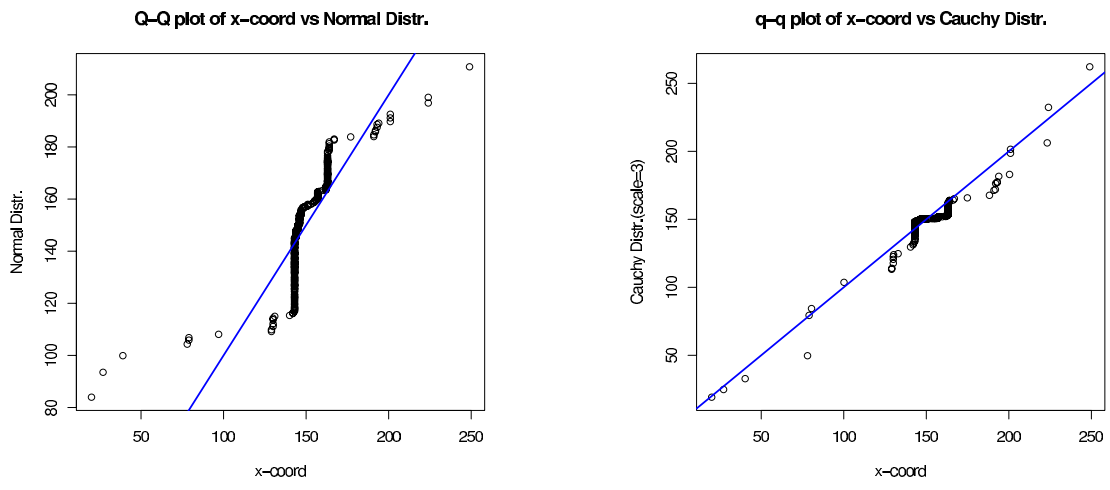


Fig. 4. X-coordinates of destination states.



(a) (b)

Fig. 6. Q-Q plots of the x-coordinate distribution against the Normal (a) and Cauchy (b) Distributions. The Normal distribution has the same mean and standard deviation as the x-coordinate data. The Cauchy distribution shown has it's scale parameter equal to 3. In both figures, a line of slope equal to 1 is also shown.

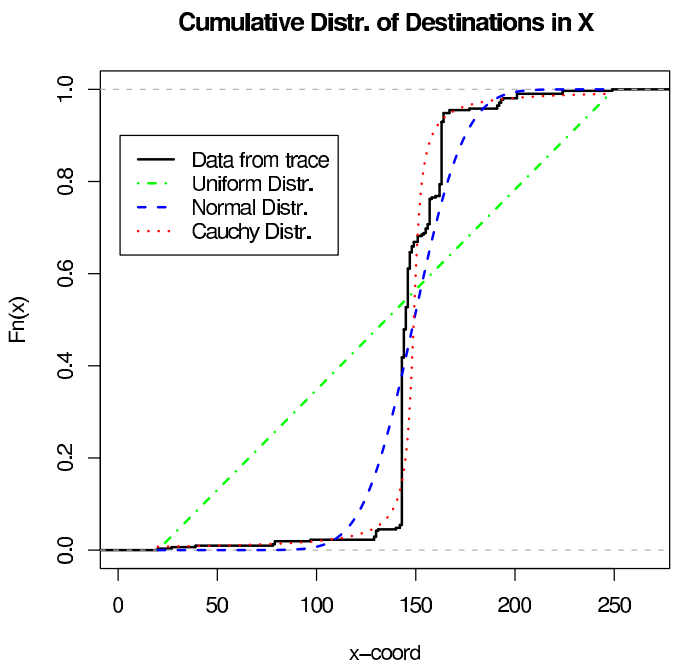


Fig. 5. Cumulative distribution of the x-coordinates of destination states. Uniform, Normal and Cauchy distribution are overlayed for comparison.

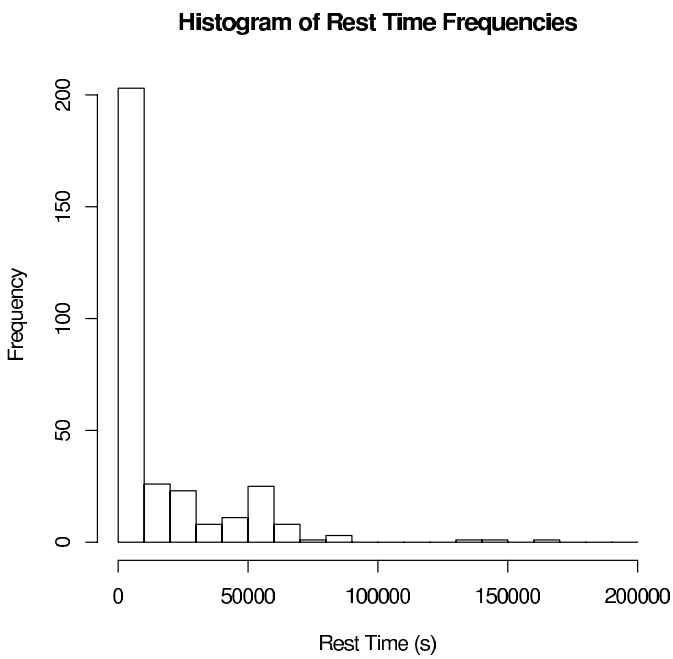


Fig. 7. Histogram showing the rest times at destination states. The majority of rest times are no larger than 10000 seconds.

equal to 1. In the Cauchy distribution's case, figure 6(b), the data fits to the line of slope equal to 1 much closer than the normal distribution.

D. Distribution of Rest Times

A histogram of rest times at destination states is shown in figure 7. The majority of rest times are below 10000 seconds. Most other rest times are between 10000 and 70000 seconds. There are some (less than ten) rest times between 70000 and 170000 seconds.

Figure 8 shows the cumulative distribution function of the rest times at destination states. Superimposed over the data are the cumulative distribution functions of the exponential and Chi-square distributions. The exponential distribution shown has it's rate parameter equal to 1. The Chi-square distribution

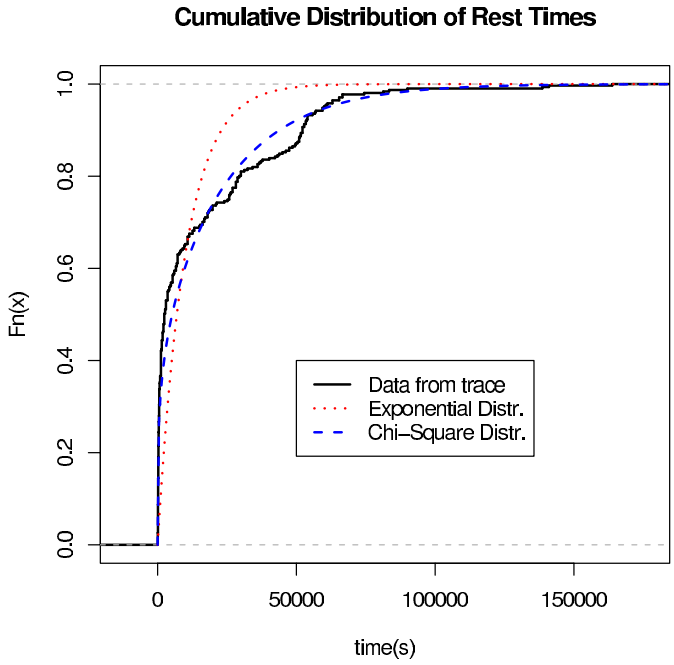


Fig. 8. Cumulative distribution of the rest times at destination states. Exponential and Chi-square distributions are overlayed for comparison. The exponential distribution shown has it's rate parameter equal to 1. The Chi-square distribution shown has it's degree of freedom and non-centrality parameters equal to 0.5 and 1.0 respectively.

shown has it's degree of freedom and non-centrality parameters equal to 0.5 and 1.0 respectively.

The exponential distribution shown seems to fit the first part of the data well, though does not fit well in the middle of the data and towards the end. The Chi-square distribution fits the data well.

In order to have a clearer comparison between the distribution of the rest time data and the exponential and Chi-square distributions, q-q plots are presented in figure 9. In the exponential case, figure 9(a), after the first part of the data having a somewhat close fit to the line with slope equal to 1, the data deviates and is consistently lower than the exponential distribution data. In the Chi-square case, figure 9(b), apart from a section of data around the 50000 second mark, the majority of the data shows a close fit to the line with slope equal to 1.

V. Discussion

In the classic use of the RWP, the uniform distribution is used as a basis for choosing the destination to travel to next. From figure 5 we can see that the uniform distribution is a poor representation of the data gathered. Two distributions, which are explored with q-q plots in figure 6, the normal and Cauchy distributions, offer better approximations to the data. The Cauchy distribution, figure 6(b) shows a very close approximation. The Cauchy distribution for destinations allows a longer tail than the normal distribution. Destinations that are a long way from the center of the distribution are thus more probable. Other long tailed distributions, such as the Pareto distribution, might also show close approximation to the data.

The data for x-coordinates of destination states showed that there was a bias towards traveling to destinations to the east of the subject's home state. This can be seen in figure 4 where most destinations are to the right of the classification in the histogram with the highest amount of destination states (140-150). The movement patterns of other individuals might be more symmetric around their primary state (home state), though this would depend on the relative positions of their major destinations.

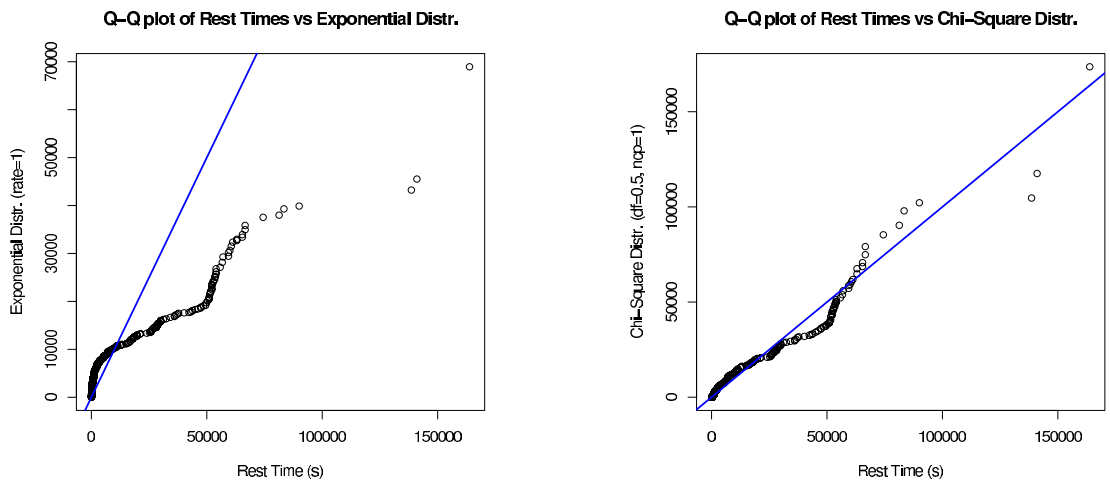
The exponential distribution is commonly used as a basis for choosing the amount of time to rest at destinations for the RWP. Figures 8 and 9 shows that the exponential distribution is a fair approximation to the data when compared against a uniform distribution, for example. The Chi-square distribution is a better match to the data compared to the exponential, especially when compared to the middle and tail of the data.

The use of the RWP with these distributions for destinations and rest times should be put into context. As the data collected pertained to movement over a large geographic area, we can recommend that these distributions be used only in simulations of networks that offer support for mobility over the same geographic scale. Mobile IP ([7], [8]) is a technology that promises to provide this support for IP networks, allowing nodes to be reachable and sessions to be stable whilst the user is mobile.

The collected data demonstrated that the RWP is not realistic when it comes to the route taken to travel from destination to destination. In the RWP, this route is a straight line to the next destination from the current destination. The data, as shown somewhat in figure 1, shows that the route taken from destination to destination is not a straight line and seems to be much more complex in nature. We do not examine the disposition of these routes in this paper.

In terms of implementation, the normal distribution for destinations and the exponential distribution for rest times are commonly implemented and readily available, at least in the often used ns-2 network simulator [13]. The distributions with better approximations, Cauchy for destinations and Chi-square for rest times, though less well known are still freely available [14].

Unfortunately, the simple method chosen for data collection did not allow a detailed study into the dynamics of the speed characteristics of the data or comparison against the uniform distribution used in the RWP to model speed. However, one comment that can be made is that this component of each trip is probably the least deterministic in nature since a subject is limited in the destinations he/she might travel to, and limited in the route selections to get to those destinations, but, the speed at which this travel is undertaken is influenced by other people, traffic signals, tiredness (in the case of walking, cycling), speed restrictions, etc. It would be unrealistic to model speed as



(a) (b)

Fig. 9. Q-Q plots of the rest time distribution against the Exponential (a) and Chi-Square (b) Distributions. The exponential distribution has it's rate parameter equal to 1. The Chi-square distribution has it's degree of freedom and non-centrality parameters equal to 0.5 and 1 respectively. In both figures, a line of slope equal to 1 is also shown.

being constant for the duration of each trip.

VI. Conclusion

Mobility models are used in simulations of mobile communications networks to study the effect of user mobility on different network characteristics. One of the commonly used mobility models is the Random Waypoint Model (RWP) which is based on random destinations, speeds, and rest times. The RWP is criticised as being an unrealistic representation of how humans move.

In this paper we use the results from a real world movement trace to validate some of the characteristics of the RWP model. The real world movement trace consisted of the movements of one individual throughout the city of Melbourne.

In the RWP, the uniform distribution is used to represent the distribution of destinations that are visited. This was found to be a poor representation according to the data. The normal distribution was found to be more representative, whilst the Cauchy distribution was an even closer fit to the trace.

In the RWP, the exponential distribution is used to represent the distribution of rest times when a user is at a destination. This was found to be a fair representation according to the data found in the trace. The Chi-square distribution was an even closer fit.

With these small changes to the distributions used in the RWP, we have shown that the RWP can still be used as a good model of movement for mobility over large geographic areas - such as a city..

Although the data for this study was limited, at the time of writing, there is no other published validation study into the applicability, or otherwise, of the RWP for the modeling of movement in such a context. The implication of this work is that there now exists real evidence that in simulations of networks where users are mobile throughout large metropolitan areas, the RWP is valid with the aforementioned distributions.

References

[1] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Mobile Computing and Networking*, 1998, pp. 85–97.

[2] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *IEEE Workshop on Mobile Computer Systems and Applications*, 1999.

[3] G. Holland and N. H. Vaidya, "Analysis of TCP performance over mobile ad hoc networks," in *Proceedings of IEEE/ACM MOBICOM '99*, August 1999, pp. 219–230.

[4] F. Fitzek, L. Badia, M. Zorzi, G. Schulte, P. Seeling, and T. Henderson, "Mobility and Stability Evaluation in Wireless Multi-Hop Networks Using Multi-Player Games," in *ACM NetGames 2003*, May 2003.

[5] M. Musolesi, S. Hailes, and C. Mascolo, "An ad hoc mobility model founded on social network theory," in *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, October 2004.

[6] C. Bettstetter, "Mobility modeling in wireless networks: Categorization, smooth movement, and border effects," *ACM Mobile Computing and Communications Review*, vol. 5, no. 3, pp. 55–67, July 2001.

[7] C. Perkins, "IP Mobility Support," RFC2002, Oct. 1996.

[8] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6," RFC 3775, June 2004.

[9] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," *Wireless Communication and Mobile Computing (WCMC)*, vol. 2, no. 5, pp. 483–502, 2002.

[10] C. Tuduce and T. Gross, "A Mobility Model Based on WLAN Traces and its Validation," in *Proceedings of IEEE International Conference on Computer Communications*, March 2005.

[11] X. Hong, M. Gerla, G. Pei, and C. Chiang, "A group mobility model for ad hoc wireless networks," 1999. [Online]. Available: citeseer.ist.psu.edu/hong99group.html

[12] Ausway Publishing, "Melway: Greater Melbourne Street Directory," 1999. [Online]. Available: http://www.ausway.com/ausway.html

[13] (2005) Ns (network simulator). [Online]. Available: http://www.isi.edu/nsnam/ns/

[14] Galassi et al, *GNU Scientific Library Reference Manual (2nd Ed)*, 2003.