Heterogeneous Community-based mobility model for human opportunistic network

Liang Hu
Technical University of Denmark
2800 LYNGBY, Denmark
liang.hu@ieee.org

Lars Dittmann
Technical University of Denmark
2800 LYNGBY, Denmark
ld@com.dtu.dk

ABSTRACT
Human opportunistic networks can facilitate wireless content dissemination while humans are on the move. In such a network, content is disseminated via nodes relaying and nodes mobility (human mobility). To develop and validate new protocols and services over opportunistic network, it is essential to use real human mobility in the simulation experiment. However, the real mobility traces are limited and their validities are difficult to generalize. We present Heterogeneous Community-based Random Way-Point (HC-RWP) mobility model that can generate synthetic traces that captures important properties of real human mobility: node heterogeneousness, space heterogeneousness, (short term) time heterogeneousness, (long term) time periodicity. These properties are based on intuitive observations of daily human mobility and confirmed by the analysis of real mobility traces. By discrete event simulation, we show HC-RWP captures not only the above observed properties, but also some essential statistical features of real human mobility traces reported in previous studies.

Index Terms—human mobility modeling, Delay-tolerant Network, opportunistic networks, Inter-contact time

1. INTRODUCTION
In recent years, as the new evolution of mobile ad-hoc network, opportunistic network has become an attractive research area for networking small mobile devices carried by human being, vehicles and animals [1]. Opportunistic network is particular useful in challenged environments where the infrastructure network is hard to deploy due to the physical constraints and economic constraints, e.g. disaster-relief, wild-life monitoring and Internet provision for rural areas. As another type of scenario, we focus on wireless content distribution over opportunistic network consist of moving people in urban area. This type of opportunistic network is envision to supplement the traditional cellular networks in terms of extending cellular network coverage and increasing its network capacity, by exploiting node mobility [2]. Within this framework, recently dissemination based routing has attract significant attentions for providing seamless content distribution over opportunistic network such as [3][4]. However, previous studies assume commonly used mobility model such as Random Way-Point (RWP) in a restricted square. Those models are homogeneous mobility model in the sense that: all mobile nodes behave statistically identical to each other (node homogeneousness); each mobile node uniformly picks up a random trip over a given domain without preference (space homogeneousness); their stationary behaviors do not change over time (time homogeneousness). They do provide scenarios that mathematically traceable, yet they are not able to address the complexity of node mobility in real-life settings. In a realistic setting, we believe the mobility of nodes tends to be heterogeneous in the sense that: each node may have very different mobility pattern; In a short-term time scale, each node may visit a number of places very often within a given geographic area than other places outside this area; lastly, node’s repeat the same mobility pattern periodically over long term time scale. In this paper, the notation “node” and “human” are interchangeable.

In principle, real mobility traces could have been more useful in validating new protocols over opportunistic network. However there are several reasons that synthetic model is preferred. Firstly, public available mobility traces are limited and with very low time granularity. Secondly, each trace is specific to its own scenario and thus hard to generalize for all cases. Finally, in some cases, mathematical model of human mobility is needed to analytically study the new protocols and services over opportunistic networking. It also allow us to study the sensitivity of various design parameters.

In this paper, we propose a new synthetic mobility model that can well capture the characteristics of real human mobility: Heterogeneous Community-based Random-Way Point (HC-RWP). HC-RWP well captures heterogeneousness of real-life human mobility: node heterogeneousness, space heterogeneousness and (short term) time heterogeneousness, (long term) time periodicity. In HC-RWP, nodes tend to move and stay locally at set of frequent visited places for the most of the time, while they occasionally roam to other places. Thus, node often meets other nodes that also move and stay within same set of frequently visited places while by chance meet nodes of other areas. We define, for one mobile node, the set of frequent visited places as “home location” and set of less
frequent visited places as “roam location”. Nodes of similar localized mobility patterns are defined as a community, i.e. nodes that have identical home location. Various communities have diverse home locations but may have the same roam location. Nodes of the same community often meet and stay together in their home location, while nodes of different communities less frequently meet in their roam location. Various communities can be, for instance, a group of people that work in the same company (say community A), students that study in the same school (say community B). Home location of A is school canteen, lecture hall, student dorm and sport center. Home location of B is Company restaurant, company building, and company sport centre. Community A and Community B can not meet frequently, as they have very different home location. However, they can meet at Shopping Mall and Train station both of which are common places of their roman locations. Finally, the home location of one node may change periodically over time, e.g. In the evening, home location of A may become Student Dorm, Disco pub and Cinema. In fact, the member of community may also change periodically over time, but we left this for a future work.

The paper is organized as follows: in section 2, we review the related work in real human mobility measurement and modelling. In section 3, we describe the general HC-RWP model and provide a simplified version and its implementation. In section 4, we provide extensive simulation results of HC-RWP model with two purposes: to demonstrate how it captures properties of real human mobility? What are statistic distributions of the contact time and inter-contact time compared with real mobility trace? Finally, we conclude the paper and present future work in section 5.

2. RELATE WORK

The initial inspiration of our work comes from the Restrict Random Way-Point model (R-RWP) presented in [7]. However, their model only captures certain space heterogeneousness, but not node heterogeneousness, (short term) time heterogeneousness, and (long term) time periodicity.

Inter-contact time and contact time are typical performance metrics for characterizing nodes mobility in mobile opportunistic network. Inter-contact time is the time interval between successive contacts of a specific node pair. Contact time is the time interval that a specific two nodes stay connected before they move apart from the radio range. Inter-contact time corresponds to how often two nodes meet to send each other message, while contact time corresponds to how much data two specific nodes can exchange during each contact. In previous studies, Inter-contact time and contact time distribution are employed to characterize the various real mobility traces or synthetic models.

There are several different opinions on the distribution of inter-contact time and contact time of real mobility traces. An early study of real human mobility is presented in [9], where they observed the inter-contact time is well approximated by a power-law over the range [10 minutes, 1day]. Their observation is confirmed using eight distinct experiment sets. In [10], author presents that the inter-contact time distribution of 90% contacts of mobile bus nodes approximately follows an exponential distribution. For a wide range of mobility trace, Karagiannis et al [8] show the inter-contact times are only power-law distributed up to 12 hours, and have an exponential cut-off after that. A possible course for this observation is the daily periodicity people have.

Han Cai et al. [11] show that simple random mobility models on boundless area can produce a power-law distribution of inter-contact times. They also show the exponential cut-off effect is in many cases a side-effect of bounded area. We believe even if simple random mobility model on boundless areas can produce power-law, it does not necessary show the general properties of real human mobility, as the human mobility is in fact most likely within a bounded area. The assumption of boundless area is not realistic.

Author [12] proposes a social network based mobility model. This model is based on the idea that node prefers to move to areas with higher social attractiveness. The social attractiveness is defined as the number of friends in a specific square. Friends can change periodically depends on the time of the day, for instance node meets colleagues as friends in the day and meet their family as friend instead in the evening. The paper does not show the inter-contact time distribution behavior for more than roughly one third of a day. Also, the model does not capture the essential properties such as node and space heterogeneousness.

In [13], a community-based random walk model is presented. Community is defined as a set of frequent visited places. In a concentration period, node visit home community more often than other places. In normal period, nodes pick up community uniformly with equal probability. In contrast, our work assumes node has a list of frequent visited places and a list of less frequent listed places. Then, we define community as node with similar mobility patterns which are determined by the set of most visited places. In other words, our community is node centric, rather than the physical place centric. Moreover, in [13] authors do not show the inter-contact time and contact-time distribution and their comparison to real mobility trace.

3. Heterogeneous Community-based Random-Way-Point Model

In this section, we firstly present several key properties of human mobility, based on intuitive observations of real human mobility and analysis of real mobility traces. Then we describe the HC-RWP model in details and show how the model captures the properties of real human mobility.

The intuition of real human mobility is that: node visits a few locations very frequently while only occasionally visit other locations. We refer this property as space heterogeneousness. Besides, different nodes may have very different mobility pattern i.e. nodes have different most frequently visited places. We refer this property as node heterogeneousness. The third property is that human mobility tends to show (short-term) time heterogeneousness. The set of frequently visited places could be different at different periods of the day. For example, in day time,
office lady more often stays at her office, while in the
evening time she more often stays at home with her family.
Lastly, human mobility pattern are repetitive every one or
multiple days, e.g. with the high probability, she re-visits
the same set of places regularly. This is also called (long-
term) time periodicity. Besides the intuitions of real human
mobility, the real trace analysis [5] [6] indeed confirms the
above mentioned properties. By studying the real user
traces, they found that that node only visit few WLAN APs
in campus areas. They also show nodes mobility while
using the network is very low and one node only meets a
small portion of all other nodes in the area. Finally, they
also show the repetitive patterns of node movement with a
period of one day and heterogeneity among nodes.

In HC-RWP, to model the space heterogeneousness, for
each node, we define the home location as a set of most
visited places and roam location as a set of less visited
places. For simplicity, we model home location and roam
location of one node from set of discrete places into a
continuous area which covers those places. Thus, home
location of one node is an area that covers its most frequent
visited places for given time interval [t1, t2]:

\[ H_i(t_1, t_2) \], for node i

roaming location of one node is an area that covers its less
frequent visited places at time interval [t1, t2]:

\[ R_i(t_1, t_2) \], for node i

Different node i have its own home location and roam
location which captures node heterogeneousness. Further
more, the home location and roam location of one node are
updated on different time interval [t1, t2], which captures
(short-term) time heterogeneousness and. Finally, the
updates of home location and roam location repeat
periodically over a period T e.g. one or multiple days,
which captures the (long term) time periodicity.

To give a clear presentation, we present a simplified
version of HC-RWP. We classify the set of nodes that have
the same home location and roaming location (thus
identical mobility pattern) as one community. Assume the
number of node is N, the number of communities is X, and
set of nodes of community i is \( C_i \), the following holds:

\[ \sum_{i=1}^{X} |C_i| = N \], where \( |A| \) denotes the cardinality of finite
set A. Node movement is modeled into two states: “home
state and “roam” state. In “home” state, nodes of
community i move or stay within area home location. In
“roam” state, nodes of community i move or stay within
roam location. Nodes travel between “home” and “roaming” states which can be characterized by a two-state
Markov Chain model showed in Markov transition diagram
in figure 1 (a). The details of node movement are as follows.

As shown in figure 1, we denote the \( \pi_h^i \) as the
probability that node is in a “home” state and \( \pi_r^i \) as the
probability that node is in a “roam” state.

From elementary Markov chain theory, for node in
community i, we get the following:

\[ \pi_h^i = \frac{p_h^i}{p_h^i + p_r^i} \text{ and } \pi_r^i = \frac{p_r^i}{p_h^i + p_r^i} \]

We also defined two terms “home trip” and “roam trip”

- Home trip is a random way-point movement towards a
  point in home location, i.e. either a random way-point
  movement within home location, or a random-way
  point movement from roaming location to home
  location. To be specific, node picks up a point
  uniformly sampled from home location area and
  moves towards it with a constant moving speed. Upon
  reaching it, pause for a constant duration.

- Roaming trip is a random way-point movement
  towards a point in roam location, i.e. a random-way
  point movement inside the roaming location or from
  home location to roaming location. To be specific,
  node picks up a point uniformly sampled from roam
  location area and moves towards it with a constant
  moving speed. Upon reaching it, pause for a constant
duration.

We assume the period T is one day (excluding the sleep
time in the night) which is divided into two periods: day
time period \( T_1 \), evening time period \( T_2 \). In principle, it can
be divided into more than two time intervals. We assume
the global area \( M \) is a large square consist of \( K \) small
squares (grids) \( m_j, j=0, 1, 2...K \), the following holds:

\[ M = m_1 \cup m_2 \cup m_3... \cup m_K \]

For the period \( T_1 \), nodes of community i is pre-assigned one
grid out of \( K \) grids as the home location, e.g. drawn from a
probability distribution function. Nodes of community i is
also pre-assigned one grid as roam location, e.g. drawn
from another probability distribution function. For the
period \( T_2 \), we follow the same instruction as in \( T_1 \).

Without loss of generality, we describe an algorithm
that implements Waypoints Selection function of HC-RWP
for community i. All other communities follow the same
instructions. The algorithm is shown in fig1 (b).
INITIALIZATION:
- Assignment of home and roam location for community i:
  \[ H_i(T_1), R_i(T_1), H_i(T_2), R_i(T_2) \]
- Locate initialized positions of nodes of community i such that node position distribution corresponds to the time-stationary distribution of HC-RWP model, employing sampling algorithm of Perfect Simulation [7].

ALGORITHM:
Waypoints Selection (simulation_time, node of community i):
If ((simulation_time \mod T) < T_i) { 
  For each node of community i, select next movement:
  If node is in "home" state, the next movement is a home trip with probability \( 1 - p_{hi} \), or is a roaming trip with probability \( p_{hi} \).
  If node is in "roam" state, the next movement is a roaming trip with probability \( 1 - p_{ri} \), or a home trip with probability \( p_{ri} \).
}
If ((simulation_time \mod T) = T_i) { 
  For each node of community i:
  Re-set the home location to \( H_i(T_2) \);
  Re-set the roam location to \( R_i(T_2) \);
}
If (T_i \mod (simulation_time) < (T_i + T_2)) { 
  For each node of community i, select next movement:
  If node i is in “home” state, the next trip is a home trip with probability \( 1 - p_{h} \), or a roaming trip with probability \( p_{h} \).
  If node i is in “roam” state, the next trip is a roaming one with probability \( 1 - p_{r} \), or a home one with probability \( p_{r} \).
}
If ((simulation_time \mod T) = (T_i + T_2)) { 
  For each node of community i:
  Re-set the home location to \( H_i(T_1) \);
  Re-set the roam location to \( R_i(T_1) \);
}

Figure 1 (b): Algorithm for Waypoints Selection of HC-RWP for community i

4. SIMULATION AND VALIDATION
In this section, by discrete event simulation, we firstly show HC-RWP model well captures the observed properties of real human mobility. Then we validate the statistic features of HC-RWP model by comparing the collected real mobility trace.

We implement HC-RWP in our own simulator in C language [4]. The simulator is based on a simple communication model: two nodes can communicate with a nominal bit-rate if their geometric distance is smaller than a threshold value. This geometric distance is set 40 meters (the radio range of 802.11b). We consider a simple case of HC-RWP model. We assume 100 mobile nodes are equally grouped into four communities \( c_1, c_2, c_3, c_4 \). We assume the global area M is a large square with diameter [1500 m, 1500 m] consist of four small squares (grids), m1, m2, m3 and m4 and five intermediary areas, as shown in figure 2. Each of the grids is [500 m, 500 m] size. These four grids are physically separated by intermediary areas, but nodes can pass by those areas to reach any grids. For the preliminary study, the simulated time is set to 16 hours which corresponds to one day time and one evening period \((T_1 + T_2) = 8\) hours \(+8\) hours). During both \( T_1 \) and \( T_2 \), the home location and roam location of community i are pre-determined before simulation and summarized in the table 1:

<table>
<thead>
<tr>
<th>Community</th>
<th>Home location</th>
<th>Roam location</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_1 )</td>
<td>m1 m2</td>
<td>m2 m1</td>
</tr>
<tr>
<td>( c_2 )</td>
<td>m2 m3</td>
<td>m3 m2</td>
</tr>
<tr>
<td>( c_3 )</td>
<td>m3 m4</td>
<td>m4 m3</td>
</tr>
<tr>
<td>( c_4 )</td>
<td>m4 m4</td>
<td>m4 m4</td>
</tr>
</tbody>
</table>

As in the table, for simplicity, we assume the home location of \( T_1 \) is pre-assigned with roam location of \( T_1 \), while roam location of \( T_1 \) is assigned with home location of \( T_2 \).

Figure 2: HC-RWP model with four grids and four communities

Furthermore, we assume the transition probability between “home” and “roam” states in fig 1 are the same for all communities and are defined specifically in various scenarios. The first step is to validate the observed properties of real human mobility: node heterogeneity, space heterogeneity, (short term) time heterogeneous. We divide the whole simulation areas into 36 equal size grids. Each grid is covered by one of the 36 virtual Access Point (AP). Each AP keeps track of the time duration that nodes stay within its coverage area (aggregate over all nodes). In other words, we keep records of aggregate fraction of time over all nodes that stay within the each of 36 sub-areas of the whole square. The mapping between sub-squares and AP index is presented in table 2:

<table>
<thead>
<tr>
<th>Square</th>
<th>AP Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>1, 2, 7, 8</td>
</tr>
<tr>
<td>m2</td>
<td>25, 26, 31, 32</td>
</tr>
<tr>
<td>m3</td>
<td>29, 30, 35, 36</td>
</tr>
<tr>
<td>m4</td>
<td>5, 6, 11, 23</td>
</tr>
</tbody>
</table>

In figure 3(a) (b), we show HC-RWP can capture several properties of real human mobility: space heterogeneousness, node heterogeneity, time heterogeneousness. In fig 3(a), y-axis shows the aggregate...
time duration that nodes stay within the coverage area of each AP at period T1, while the x-axis shows the AP index. It is clear from fig 3(a) that, for all the four communities, nodes visit some AP coverage areas of home location much more often than other AP coverage areas, which captures space heterogeneousness. Also, nodes of different community have different set of frequent visit areas or home location, e.g. nodes of c1 mostly visit AP 1, 2, 7, 8, while nodes of c2 mostly visit AP 29,30,35,36, which captures node heterogeneousness. Finally, fig 3(b) shows the aggregate time duration that nodes stay within the coverage area of each AP at period T2. We observe that each of the community exchange its home location and roam location, compare to the case of period T1. For instance, nodes of C1 mostly visit AP 25,26,31,32, while they only occasionally visit AP 1, 2, 7, 8, which is the opposite case of period T1. In this way, the HC-RWP captures time heterogeneousness of real human mobility, i.e. nodes have time-variant home location and roam location. Of course, dividing one day into two periods T1 and T2 is a low granularity approximation of time-variant real mobility pattern. A more accurate model could be developed by dividing one day into multiple periods which is larger than two.

In fig 3 (c), we show HC-RWP captures the time periodicity of real human mobility pattern. Here, we assume the simulation time is 32 hours and the period T is 8 hours consist of T1=4 hours and T2=4 hours. According to the algorithm in fig 1(b) and table 1, each community update their home location and roam location every 4 hours, while the transition probability does not change. The y-axis is the aggregate time duration over all nodes (per hour) that stay within the coverage area of AP index 1 during the simulation time 32 hours. The unit of y-axis is second. It is obvious that aggregate time duration (per hour) that nodes stay within coverage area of AP index 1 is periodical with peak value roughly every four hours. The same observations remain if the set of T, T1 and T2 are chosen other values.

To validate HC-RWP model, we investigate CCDF of inter-contact time between mobile nodes under the impact of various transition probabilities $p_i^h$ and $p_i^l$ defined in section 3. In figs 4(a)(b), we show that CCDF of the inter-contact time on log-log and line-log scales. The simulation time is 32 hours and the period T is 8 hours consist of T1=4 hours and T2=4 hours. The simulation parameters are as follows in the table 3:

<table>
<thead>
<tr>
<th>Moving Speed</th>
<th>Pause time</th>
<th>$(p_i^h, p_i^l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 m/s</td>
<td>100, 600 s</td>
<td>(0.9,0.1)</td>
</tr>
<tr>
<td></td>
<td>1200 s</td>
<td>(0.6,0.4)</td>
</tr>
</tbody>
</table>

Firstly, fig 4(a) (b) show the CCDF of inter-contact time (under all parameters) approximately follows exponential distribution, which is in line with the analysis of real mobility traces presented in [10] and [8]. Secondly, we observe that, for given $P(r) = 0.1$, $P(h) = 0.9$, the pause time does impact the inter-contact time distribution. In particular, larger pause time (e.g. 1200s) incurs larger inter-contact time on average than small pause time (e.g. 100 s and 600 s). This trend is nature, as pausing nodes produce longer contact durations but less frequent associations. Secondly, for a given pause time 600 second, fig 3 (a) (b)
show transition probability \((P(r) = 0.4, P(h) = 0.6)\) on average gives larger inter-contact time than \((P(r) = 0.1, P(h) = 0.9)\). This is because node tends to move around a larger area with transition probability \((P(r) = 0.4, P(h) = 0.6)\), while node moves more locally with transition probability \((P(r) = 0.4, P(h) = 0.6)\). Nodes that move around a larger area tend to less frequently meet each other, compared to the case of moving within a smaller area.

According to the above analysis, we claim that HC-RWP parameters approximately follows exponential distribution under all scale. For most nodes contacts, the CCDF of contact time divided one day into more than two time periods, each of which have different mobility pattern. Also, we plan to tune system parameters of HC-RWP such that it can well match statistic features of all real mobility trace. Finally, as the current real mobility trace is rather limited, the full validation and tuning of our model is possible upon the availability of large-scale, high time granularity real mobility traces.

5. CONCLUSION AND FUTURE WORK

We present a new synthetic mobility model HC-RWP for mobile opportunistic networking research. By discrete event simulation, we show it captures four properties of real human mobility: node heterogeneousness, space heterogeneousness and (short term) time heterogeneousness, (long term) time periodicity. Those four properties are observed according to daily intuitions of real human movement and confirmed by the measurements of real traces. Besides, in terms of inter-contact time and contact time distribution, we show HC-RWP do provide synthetic traces that have the same statistic features as real mobility traces.

As the future work, we intend to extend our model to capture higher granularity time-variant node mobility, e.g. divided one day into more than two time periods, each of which have different mobility pattern. Also, we plan to tune system parameters of HC-RWP such that it can well match statistic features of all real mobility trace. Finally, as the current real mobility trace is rather limited, the full validation and tuning of our model is possible upon the availability of large-scale, high time granularity real mobility traces.

6. REFERENCES