Synthetic Trace and Simulation Settings for the TVC Model

Wei-jen Hsu*, Thrasyvoulos Spyropoulos†, Konstantinos Psounis‡ and Ahmed Helmy*
*Dept. of Computer and Information Science and Engineering, University of Florida, Gainesville, Florida
†Computer Engineering and Networks Lab, ETH Zurich
‡Dept. of Electrical Engineering, University of Southern California, Los Angeles, California
Email: wjhsu@ufl.edu, spyropoulos@tik.ee.ethz.ch, kpsounis@usc.edu, helmy@ufl.edu

I. Introduction

The purpose of this document is to describe in more details the parameter settings of the time-variant communication model we use in the paper submission [1]. Due to the space limitation, many of these details cannot be fully described in the paper. To make our submission more comprehensive, we list all the relevant parameters here.

This document has two major parts. In section II, we list the parameters we use in section IV of the submitted paper [1] to match the model with various traces. In section III, we describe more details of the models we use in the simulations (section VI) in the submitted paper.

II. Parameters Used in the Synthetic Traces

In this section we describe in detail the parameters we use to construct the synthetic traces with matching characteristics to the realistic traces in section IV of [1].

A. WLAN Traces (the MIT trace)

To generate a synthetic trace with matching mobility characteristics to the MIT WLAN trace, we first divide the 1000-by-1000 unit simulation area into one hundred 100-by-100 unit cells. These 100 cells are the locations we used to tally the mobility characteristics, in terms of the skewed location visiting preferences and the re-appearance probabilities.

For the *complex model*, the STEP1 is to assign some communities from these 100 cells to a given node as its frequently visited locations. Since we do not have a detailed map of the studied environment, each node just picks randomly 15 cells from the 100 available cells as its communities. Note that, concerning matching with the two mobility properties (i.e., skewed location visiting preferences and the re-appearance probabilities), the actual locations of the communities do not make a difference. One can assign different sets of 15 communities to different nodes, according to the popularity of the locations in the actual environment (i.e., the "hot spots" should be chosen by more nodes as one of its communities than less popular locations). We do not consider this option as we do not have the actual information about the target environment.

In STEP3, we first assign the π_j^t value to the j-th community of each node, following the exact value of the location visiting probability to the j-th most popular location in the trace. This leads to the assignment of values in Table I(a). The 16-th community corresponds to the whole simulation field and we allocate the remaining location visiting probability not explicitly captured by the 15 communities to this roaming state (i.e., $\pi_{16}^t = 1 - \sum_{j=1}^{15} \pi_j^t$). When the node chooses this state, it roams around the whole simulation area. Notice that by matching the assignments of π_j^t with the location visiting probabilities and assigning the same values to $\overline{L_j^t}$ and $\overline{D_j^t}$ to all communities in each time period, we naturally get the same location visiting probability distribution as obtained from the trace.

We then assign the values $\overline{L_j^t}$ and $\overline{D_j^t}$ to shape the re-appearance probability, under the chosen on-off patterns of the nodes. We change the value for all communities (expect the community 16, which corresponds to the roaming state) in each time period simultaneously to retain the proper location visiting probabilities. Note that the curves of re-appearance probability represent the probability of an "on" node appearing at the same community after the given time gap, and the peaks appear when the considered points in time are in the same type of time period. Therefore, the peak value on the 7-th day under the weekly schedule is $\sum_{t=1}^{V} \frac{T^t}{\sum_{k=1}^{V} T^k} \sum_{j=1}^{S^t} (P_j^t)^2$, where $P_{on,j}^t$ denotes the probability a node is considered "on". Hence, the fraction of time nodes spend on moving ($\overline{L_j^t}/\overline{v}$) and pause ($\overline{D_j^t}$) can be adjusted (note that in this case, $P_{on,j}^t = \overline{D_j^t}/(\overline{D_j^t} + \overline{L_j^t}/\overline{v})$) to change the peak values in the curve of *periodical re-appearance* property to match with the curves obtained from the trace. Finally, we settle with the values given in Table I(b).

Notice these selected values correspond naturally to what would happen in daily lives: Time period 2, which corresponds to the evening, has less "on" time (recall that the pause time corresponds to "on" time for the WLAN). Time period 1, weekday working hours, has the most relative "on" time. Time period 3, weekend working hours, also sees little "on" time. The epoch lengths are chosen under two criteria: First, the epoch should be long enough, being in the same order as the community size.

TABLE I
PARAMETERS USED IN THE COMPLEX MODEL WITH MATCHING CHARACTERISTICS TO THE MIT TRACE

(a) The probability to visit each community

Community number j	1	2	3	4	5	6	7	8	9	10
$\pi_j^t \quad \forall t$	0.690183	0.175677	0.061342	0.029281	0.016199	0.009591	0.006128	0.003938	0.002469	0.001628
Community number j	11	12	13	14	15	16				
$\pi_j^t \quad \forall t$	0.001084	0.000752	0.000475	0.000326	0.000256	0.000671				

(b) The epoch lengths and average pause times in the communities

$\overline{L_j^1} \quad \forall j \neq 16$	$\overline{L_j^2} \quad \forall j \neq 16$	$\overline{L_j^3} \quad \forall j \neq 16$	$\overline{L_{16}^1}$	$\overline{L_{16}^2}$	$\overline{L_{16}^3}$	$\overline{D_j^1} \quad \forall j$	$\overline{D_j^2} \quad \forall j$	$\overline{D_j^3} \quad \forall j$
50	200	200	1500	2000	2000	150	2	15

TABLE II
PARAMETERS USED IN THE MODEL WITH MATCHING CHARACTERISTICS TO THE VANET TRACE

(a) The probability to visit each community

Community number j	1	2	3	4	5	6	7	8	9	10	
$\pi_j^t \quad \forall t$	0.636117	0.195895	0.066775	0.034782	0.022556	0.010982	0.004363	0.003765	0.003432	0.003000	
Community number j	11	12	13	14	15	16	17	18	19	20	
$\pi_j^t \quad \forall t$	0.002834	0.002331	0.001932	0.001632	0.001384	0.001193	0.001015	0.000871	0.000752	0.000653	
Community number j	21	22	23	24	25	26	27	28	29	30	31
$\pi_j^t \quad \forall t$	0.000563	0.000481	0.000417	0.000362	0.000311	0.000268	0.000226	0.000192	0.000165	0.000139	0.000612

(b) The epoch lengths and average pause times in the communities

$\overline{L_j^1} \quad \forall j \neq 31$	$\overline{L_j^2} \quad \forall j \neq 31$	$\overline{L_j^3} \quad \forall j \neq 31$	$\overline{L_{31}^1}$	$\overline{L_{31}^2}$	$\overline{L_{31}^3}$	$\overline{D_j^1} \quad \forall j$	$\overline{D_j^2} \quad \forall j$	$\overline{D_j^3} \quad \forall j$
800	100	800	1600	1600	1600	60	60	40

Second, the epoch length should give the desired on-off time ratio for the nodes so that the re-appearance probability matches with the actual trace.

B. Vehicle Mobility Traces

The parameter selection process for the TVC model to match with the VANET trace is very similar to the procedure described for the WLAN trace. The only difference is, the node are considered "on" when it is moving, and "off" when it is not moving. The corresponding parameters are listed in Table II. In this case, we have used 31 communities.

C. Human Encounter Traces

To construct a synthetic trace with matching encounter characteristics observed at a conference, in STEP1, we first observe the conference time schedule and the setting of the conference venue. We come up with a rather complex daily time schedule shown in Fig. 1, with five types of time periods. $TP1$ corresponds to the breakfast and lunch sessions where most people dine in a big dining hall. $TP2$ corresponds to the coffee breaks between sessions. $TP3$ corresponds to the technical sessions. $TP4$ corresponds to the early evening part, when people roam around and visit some popular places close by the conference venue. $TP5$ corresponds to the late evening and night, when people stay in their hotel rooms. The time period structure repeats itself daily, for four days (the whole duration of the conference).

We consider two nodes for each run of our simulation, referred to as node a and b below. We set communities for these two nodes differently in each time period – there are shared communities representing commonly visited parts of the conference venue, and "private" communities that are visited only by one of the nodes. These "private" communities serve as the "hotel rooms" for node a and b, respectively, and when they go back to these rooms, they cannot meet with the other node. This captures the behavior of conference attendees skipping part of the conference. The detailed setup of the communities for each time period for node a and b is illustrated in Fig. 2. In $TP1$, the community in the center ($Comm_1^1$) represents the dining hall, and the community at the corner of the simulation field ($Comm_2^1$) represents the "hotel room". In $TP2$, the same community construction applies. In $TP3$, we use six small-sized communities ($Comm_1^3$ through $Comm_6^3$) to capture the conference rooms for technical sessions, as people sitting in the same session are highly likely to meet. There is an additional community surrounding the conference rooms ($Comm_7^3$), capturing the phenomenon people sometimes roam around the conference venue without actually going to one of the sessions. The "hotel room" is also there to capture people skipping part of the conference. In $TP4$, the nodes are free to roam about the whole simulation area ($Comm_3^4$), but there are two additional communities

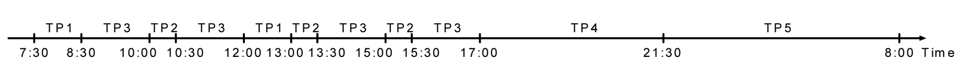

Fig. 1. Illustration of the time schedule we use for producing a conference environment.

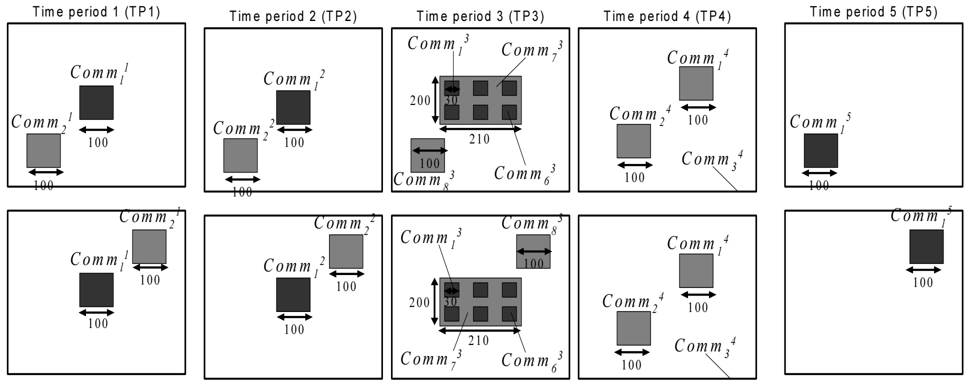

Fig. 2. Illustration of the community setup we use for producing a conference environment, when considering two attendee nodes, a and b. Top row: communities for node a, bottom row: communities for node b. The dimensions of the communities are given in the illustration, and the size of the simulation field is 1000-by-1000.

TABLE III
PARAMETERS USED IN THE MODEL WITH MATCHING CHARACTERISTICS TO THE INFOCOM-CAMBRIDGE TRACE

(j,t)	$(1,1)$	$(2,1)$	$(1,2)$	$(2,2)$	$(x,3)$ where $x=1,...,6$	$(7,3)$	$(8,3)$	$(1,4)$	$(2,4)$	$(3,4)$	$(1,5)$
π_j^t	0.9	0.1	0.9	0.1	0.1	0.2	0.2	0.2	0.2	0.6	1.0
$\overline{L_j^t}$	150	20	150	20	50	600	20	200	200	1000	10
$\overline{D_j^t}$	1800	1800	600	600	2400	100	2700	10000	10000	300	900

($Comm_1^4$ and $Comm_2^4$) to capture the popular meeting places (e.g., the hotel lobby or close-by restaurants) after the conference. In $TP5$, nodes go back to their respective rooms. The parameters (i.e., π_j^t, $\overline{L_j^t}$ and $\overline{D_j^t}$) for each community are listed in Table III.

III. Parameters Used in the Simulations

We summarize the parameters for the tested scenarios in our simulation cases for theory validation (in section VI of [1]) in Table IV. Table IV(a) lists the parameters we use for a simplified model (two time periods with two communities in each time period, where one of the communities is the whole simulation field). For more complex models, we try out the setup of tiered communities (as shown in Fig. 3(a)) and multiple randomly placed communities (as shown in Fig. 3(b)). In the tiered communities layout, a randomly chosen point in the simulation field serves as the *center* of the communities, and multiple tiers of communities with different sizes share the same center. This construction is suggested by a common observation from our daily lives: People visit the vicinity area of locations that bear importance to them more often than roam far away. When we assign the tiered community structure, it naturally makes sense to have the node visit the outer tiers less frequently than the inner tiers, although this is not required for the theoretical derivation. In the simulations, we use two alternative time periods with a two-tier local community in each time period, and the parameters are listed in Table IV(b). In the multiple randomly placed communities layout, multiple communities are instantiated randomly to show that our theory is not limited to a single community. We use two time periods with two randomly placed communities each for this scenario. Other than the difference in community setup and sizes, we again use the parameters in Table IV(b) for this case. Our discrete-time simulator is written in C++.

References

[1] W. Hsu, T. Spyropoulos, K. Psounis, A. Helmy, "Modeling Spatial and Temporal Dependencies of User Mobility in Wireless Mobile Networks," In submission.

(a) Concentric multiple-tier communities setting.

(b) Multiple randomly placed communities setting.

Fig. 3. Illustration of the community setup for the generic cases of time-variant community model.

TABLE IV
PARAMETERS FOR THE SCENARIOS IN THE SIMULATION

Common parameters: For simplicity, we use the same movement speed for all node: $v_{max} = 15$ and $v_{min} = 5$ in all scenarios. In all cases we use two time periods and they are named as time period 1 and 2 for consistency. In the simple model we use a single local community (with subscript l) in each time period. For the generic model, we test with two different configurations: (1) A two-tier community in each time period, as illustrated in Fig. 3 (a). In this scenario the inner tier community and the outer tier community has edge length C_{l1} and C_{l2}, respectively. (2) Two randomly placed communities in each time period, as illustrated in Fig. 3 (b). In this scenario the communities both have edge length C_{l1}, but the parameters correspond to the two communities are different (i.e., correspond to subscript $l1$ and $l2$ in the table). In all cases, there is also a roaming state (with subscript r) in which the node moves about the whole simulation area (i.e. the whole simulation area is a community).

(a) The simple model.

Model name	Description	N	C_l^1	C_l^2	$D_{max,l}$	$D_{max,r}$	$\overline{L_l}$	$\overline{L_r}$	π_l^1	π_r^1	π_l^2	π_r^2	T^1	T^2
Model 1	Match with the MIT trace	1000	100	100	100	50	80	520	0.714	0.286	0.8	0.2	5760	2880
Model 2	Highly attractive communities	1000	200	50	100	200	52	520	0.667	0.333	0.889	0.111	3000	2000
Model 3	Not attractive communities	1000	100	100	50	200	80	800	0.5	0.5	0.667	0.333	2000	1000
Model 4	Large-size communities	1000	200	250	50	100	200	800	0.7	0.3	0.889	0.111	2000	1000

(b) The generic model.

Model name	N	C_{l1}^1	C_{l2}^1	C_{l1}^2	C_{l2}^2	$D_{max,l1}^1$	$D_{max,l2}^1$	$D_{max,r}^2$	$D_{max,l1}^2$	$D_{max,l2}^2$	$D_{max,r}^2$
Model 5	1000	100	300	100	300	25	15	1	30	25	3
Model 6	1000	150	450	150	450	50	20	15	30	20	30
Model 7	1200	160	480	160	480	50	20	15	30	20	30

Model name	$\overline{L_{l1}^1}$	$\overline{L_{l2}^1}$	$\overline{L_r^1}$	$\overline{L_{l1}^2}$	$\overline{L_{l2}^2}$	$\overline{L_r^2}$	π_{l1}^1	π_{l2}^1	π_r^1	π_{l1}^2	π_{l2}^2	π_r^2	T^1	T^2
Model 5	300	500	1000	200	300	1000	0.6	0.3	0.1	0.85	0.1	0.05	5760	2880
Model 6	100	300	1000	200	500	1000	0.5	0.35	0.15	0.7	0.2	0.1	5760	2880
Model 7	140	600	1500	200	500	1600	0.8	0.15	0.05	0.7	0.2	0.1	5760	2880