START: Status and Region Aware Taxi Mobility Model for Urban Vehicular Networks

Haiquan Wang
Wenjing Yang
Jingtao Zhang
Jiejie Zhao
School of Software, Beihang University, Beijing, P.R.China
Beijing Key Laboratory of Network Technology, Beijing, P.R.China
Yu Wang*
Department of Computer Science, University of North Carolina at Charlotte, NC, USA

Abstract

Using a realistic mobility model will enhance the validity of simulations. However, the difficulty lies in ~~-~~discovering laws from large amounts of data and applying those rules. Researchers have been working on mobility model extracting from real data set, whereas the taxi behavior differences between ~~different~~ various taxi statuses have been ignored in previous work~~s~~. Based on ~~the~~ experience in daily life, two assump- tions related to the taxi status, one is the behavior of taxi ~~will be~~ is influenced by the statuses and the other ~~one~~ is the macroscopic movement is related with different geographic ~~feathers~~ features in corresponding status, are introduced and estimated by the real data. Based on the two assumptions, a novel taxi mobility model named START is proposed with respect to taxi status. The simulation results illustrate that proposed mobility model has a good reality approximation ~~with reality~~ in trace samples, distribution of nodes and the contact characteristics.

keyword

mobility model, taxi status,region recognition

I. Introduction

In vehicle ad hoc networks (VANETs) [1], realistic mobility model is an important way to improve route planning, control traffic situations, or solve the vehicle-to-vehicle communica- tion problems. However, mobility models ~~will~~ might influence sim- ulation performance, since mobility model defines the nodal mobility pattern including speed and direction. ~~Whereas~~However. large amount~~s~~ of data are difficult to utilized directly. It is necessary to work on realistic mobility models. Some researchers [2], [3] modeled the vehicular mobility, extracting different ~~feathers~~ features from real data sets. Nevertheless, ~~But~~ taxi status is ignored in the previous work~~s~~.

In this work, a St~~T~~atus and Region Aware Taxi mobility model, START, is proposed based on the real taxi GPS data. Two assumptions are introduced in section II. We assume that the taxi behavior and geographic ~~feathers~~ features are related with different status. They are ~~validated~~ verified to be reasonable by the statistical analysis of the data set. START is modeled based on these assumptions. In the macro scope, instead of simply dividing the area into coarse-grain regions, we ~~divide~~ classifies the area into two set of regions according to the passenger loading or droping event density. When a taxi takes a passenger, the current region will be selected from~~in~~ the region set of load-event, meantime~~and~~ the destination region, where the drop-event happens, will be selected from~~in~~ the region set of drop-event. We investigate the relationship between load-event regions and drop-event regions. Paths~~Pathes~~ from the sources to destinations will be found by Dijkstra's algorithm. ~~For microscope~~In detailed view, the *speed* for the two taxi statuses are discussed respectively. Simulations are carried out to compare the similarity of node trace characteristics and contact characteristics. The results show that our mobility model has a good approximation with the real scenario in trace samples, distribution of nodes and the contact characteristics. The rest of our paper is organized as follows: Section II proposes two assumptions which are further validated by statistical results of real data. Section III presents the modeling process. Simulation results are demonstrated in Section IV. Finally, Section V concludes this paper.

II. Assumptions and statistical analysis of Taxi Trace

In this section, we focus on statistical analysis ~~on~~ of the speed and duration characteristics on the data set. Firstly, the data set ~~will be~~ is introduced in section II-A. Then, two assumptions are proposed and validated in the following sections.

A. Trace Dataset: Beijing Taxi Traces

A real-world GPS data set is used in this paper, which was generated by 12, 455 taxis in Beijing, China within 5 days from March 3rd,2011 to March 7th,2011. Each row includes a base station ID, company name, taxi ID (id), timestamp (t), current location (l, including longitude and latitude), speed, event, status, et al. Of all the fields, the taxi ID, time stamp, and current location, status and event are used in this paper. Note that GPS traces from taxis have been used recently for inferring human mobility [4] and modeling city-scale

This research has been partially supported by the US National Science Foundation (NSF) under Grant No. CNS-1319915 and CNS-1343355, the National Natural Science Foundation of China (NSFC) under Grant No.61300173 and No. 61170295, the Project of Aeronautical Science Foundation of China under Grant No.2013ZC51026 and No.2011ZC51024, the Fundamental Research Funds for the Central Universities under Grant No. YWF-12-LXGY-001, and the State Key Laboratory Software Development Environment and Network Information and Computing Center of Beihang University.

批注 [l1]: STRAT??
带格式的: 两端对齐
批注 [l2]: Applying the rules discovered form large amount of data.
批注 [l3]: there are two....句号
批注 [l4]: One is that the behavior...下同
批注 [l5]: 此句主语太长 which are...
批注 [l6]: regard
批注 [l7]: the mobility model we proposed 或者 STRAT
批注 [l8]: to be utilized
批注 [l10]: 先行词不是 paper
批注 [l9]: 不是 STRAT 吗?

traffics [5]. Therefore, we believe that they are suitable to

TABLE I: Explanation of Events and Status

Event	Explanation
0(drop)	A taxi's status change to vacant.
1(load)	A taxi's status change to occupied.
2	Set up defense.
3	Cancel defense.
4	No event happened.
Status	**Explanation**
0(vacant)	A taxi is vacant.
1(occupied)	A taxi is occupied.
2	A taxi is setting up defense.
3	Stop running.

characterize the contact patterns among vehicles in large-scale urban scenario.

Especially, there are five types of events and four types of status~~es. The explanations~~, which are ~~as~~ explained in table I. We only discuss the vacant status and occupied status in this paper. Accordingly we utilize the load-event and drop-event as well.

B. Assumptions

According to the experience in daily life, the following two assumptions are given:

- **Assumption 1:** The behavior of a taxi will change when its status changes. When a taxi is occupied, its destination is certain, and the speed of occupied status will accelerate relatively. ~~In~~ By contrast, when a taxi is vacant, it will slow down or even stop to search for potential passengers along the road. Thus, taxi behavior characteristics, such as speed and the status duration varie~~s~~ ~~consequently~~coherently.
- **Assumption 2:** The movement behavior of taxis associates with geographic ~~feathers~~features. When taxis is occupied, the destination may be ~~tend~~ inclined to be some places, such as the airport. Meanwhile, when taxis are vacant, drives tend to stay around some hot spots where more people want to take a taxi.
  1) The destination selection ~~will be~~ is influenced by different regions.
  2) events occur~~s~~ in different regions un-evenly, passenger drop and load events are distinct.

Therefore, we ~~analysis~~ analyze the speed,duration and passenger load/drop events distribution to estimate our assumptions.

C. Speed analysis

In this section, we investigate the average speed $\overline{speed}$ in each status. For example, taxi i drives in occupied status for a distance d using time t, then the average speed in this status is d/t. From March 3th to 7th, 2011, ~~the~~ we have $\overline{speed_{empty}} = 3.627m/s$, while that for occupied status is $\overline{speed_{occupied}} = 7.083m/s$.

To further investigate the cumulative speed distribution, proportion for every $\overline{speed}$ section is calculated. As shown in ~~fig~~Fig.1. For example, dot(20,0.0245) means 2.45% records fall in the range $[0, 20)km/h$. We also fit the speed to

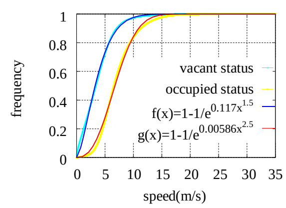


Fig. 1: Speed ~~Distribution~~ distribution for vacant and occupied status~~es~~.

D. Status duration analysis

The duration distribution for each status are shown in ~~fig~~Fig.2. Status duration represents the time length of a taxi staying in a certain status. The red line ~~presents~~ represents the duration time distribution for vacant status, and the green one is for status occupied~~1.~~ A dot of the line means the proportion of the duration. A peak exists in each line, and it is obvious that the peak of the red line is earlier than that of the other line. And the value of duration for status ~~0~~ vacant trends to be ~~smaller~~lower. It ~~accords~~ corresponds to the realistic situation, because drivers tend to shorten the waiting time to raise their income.

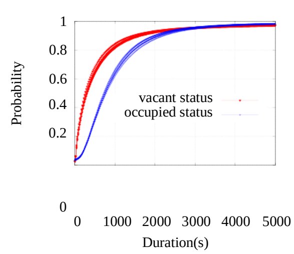


Fig. 2: Status duration distribution.

The statistical results are consistent with the *assumption 1*, that is, the behaviors of taxis are similar within each status while ~~it differs~~differ between the two statuses.

E. Taxi event distribution

To validate $assumption2$, we quantitatively analyze vehicles density ~~in~~ inside one hour. By dividing the whole network into

model the micro scope behavior, which will be shown in section III. Fig. 1 shows that $speed$ distribution differs for each status. For vacant status, the $speed$ gather together at 1 also demonstrates that the speed distribution is with strong regularity for each status.

带格式的: 字体:(中文)宋体,(中文)中文(中国)

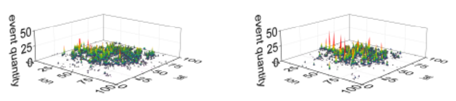
批注 [111]: 不通顺

带格式的: 缩进:左侧: 0.21 厘米,首行缩进: 0.35 厘米,右侧: 1.72 厘米,段落间距段前:3.35 磅, 无项目符号或编号,制表位:不在 0.61 厘米

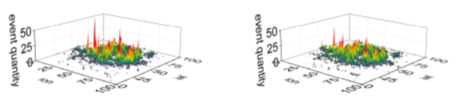
100×100 grid~~s~~, taxi density distributions for loading and droping events are computed in each cell.

~~Figure~~Fig.3 shows the load/drop events distribution in three time sections. In the morning, people begin to get out. so the load-event distribution~~e~~ is more even than that of the drop-event. this phenomenon ~~may be caused by~~is happened because that the load-event spots are mainly at~~ the~~ homes of the citizens, while the drop-event spots tend to gather together at workplaces, railway stations or scenic spots. Besides, in the evening, the dispersion of load-event is of lower degree than that of drop-event, for people ~~coming~~ returning home at that time.

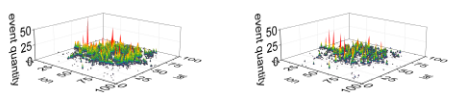
The amount of loading passengers in each cell shows geo- graphic ~~feathers~~features:the distribution is uneven, and the difference between load/drop-event distributions illustrates that the load/drop- event regions are different which supports the *assumption 2*.



(a) $load-event, 7:00-8:00$ (b) $drop-event, 7:00-8:00$



(c) $drop-event, 12:00-13:00$ (d) $load-event, 12:00-13:00$



(e) $load-event, 17:00-18:00$ (f) $drop-event, 17:00-18:00$

Fig. 3: Taxi density for load-event and drop-event in one hour

III. Modeling

Movement model defines the mobility pattern of nodes, which can be represented as a collection of pathes, say $Paths :< p_1, p_2, ..., p_n >$, so ~~dose~~ does the START. A p_i takes two steps to ~~adopt,~~accomplish destination selection and moving process from source location to destination.

Destination selection: In START, to select a destination of a node is closely related to note only its current location but also its current status. Dividing the area into regions by the density of passenger load/drop events or loading passenger events respectively, two transition probability ~~matrixes~~matrices are calculated, one is the probability from passenger drop event regions $\{REGION_{m,drop}\}$ to passenger ~~on~~ load events regions $\{REGION_{n,load}\}$. Note that $\cup\{REGION_{m,drop}\} = \cup\{REGION_{n,load}\} = AREA$. If the status of a taxi changes to vacant, its current location ~~determine~~ is a $REGION_{i,drop}$. Consequently, a destination region in $\{REGION_{n,load}\}$ will be selected by querying the transition matrix from $\{Region_{m,drop}\}$ to $\{Region_{n,load}\}$. Then, START will randomly select a map node in the region as the destination. ~~As to the status of a texi changing to Occupied~~, the destination selection process is similar for the occupied status transition. During this process, the region transition matrix will be utilized according to the current status.

Moving process: When the source location (current location) and destination location is given, next step is to find will find ~~a~~the shortest path from source to the destination, to route on map. The speed of the path ~~should be assigned~~is denoted as the $speed$, which is adopted by the current $speed$ distribution of corresponding status introduced in the following section.

Based on the design above, we model the movement on the speed, duration and region transition ~~matrix~~ matrices respectively.

A. Region transition probability

A travel path of a taxi can be simplified as a multi-hop process, in which a hop indicates ~~an~~ a load/drop event happened. Seeing that, we define a *region transition probability* to figure out the probability of the next hop falling in a ~~definite~~ different region j from the current region i. Particularly, two successive events are different. Likewise, the region i and j are recognized by different metrics, that is, drop or load event distribution. It is more reasonable. For ~~an~~ instance, if the taxi is occupied, the next hop event is the drop one. Hence, choosing a target region from a region set divided by drop event distribution is more logical.

To calculate the region transition probability, the **region recognition process** should be executed in advance.

Firstly, we divide the area into 100×100 grids, and define cells in it as equation 1. Then, we consider region as adjacent cells as equation 2.

$$CELL_{x,y} ::= \{(lon, lat) | x \le \frac{lon}{len_x} < x+1, \; y \le \frac{lat}{len_y} < y+1\} \tag{1}$$

$$REGION_m ::= \{CELL_{x,y} | \exists CELL_{i,j} \in REGION_m \Rightarrow \|x-i\| \le 1, \|y-j\| \le 1\} \tag{2}$$

By clustering cells into regions, two region sets, $\{REGION_{drop}\}$ ... a path. ~~To simplify~~For simplicity, we adopt the Dijkstra algorithm ,which

批注 [112]: Grid 是表格，只有一个

带格式的: 字体:（中文）宋体,（中文）中文(中国)

带格式的: 字体:（中文）宋体,（中文）中文(中国)

带格式的: 段落间距段前: 0 磅, 行距: 固定值 10.6 磅

$_m$ } and $\{REGION_n\}$, can be recognized. The main idea of clustering is to put adjacent cells with event density larger than the event threshold η into the ~~a~~ same region. To avoid the size of the region becomes too large or too small, we set a constant $CLUSTERSIZE$ to restrict a region size, say $\|REGION_i\| \le CLUSTERSIZE$, and only limit the top 200 regions, in which $CELL_{x,y}.events \ge \eta$. After that, the other cells ~~not~~ that do not belong to the top 200 regions, will also be ~~clustered~~ classified into regions, while $\|REGION_j\| \le CLUSTERSIZE$. Consequently, every cell~~s~~ will be ~~clustered~~classified into regions and the size of ~~every~~ each region is not larger than $CLUSTERSIZE$.

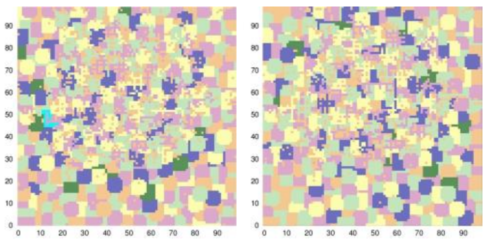
We sort the 100×100 cells by event density in descending order, and begin with the first cell to search its neighbors and ask them whether to join the same region using breadth-first traversal. The region recognition results for load/drop events are shown in ~~figure~~ Figure 4.下同

In figure 4, every colored block represents a region. In addition, the $CLUSTERSIZE = 200$, $\eta = 121$ for ~~on~~ load event and $\eta = 141$ for drop event set by the average event density of the top 5000 cells ordered by its event density. The detailed clustering algorithm is presented in the appendix.

The calculate process of the region transition probability: After clustering cells into regions, the transition probability from $REGION_i^{load/drop}$ to $REGION_j^{drop/load}$,

带格式的：字体：（中文）宋体，（中文）中文(中国)

批注 [113]：有 cluster sth. Into sth.这个用法吗？之前也多次出现



(a) passenger drop event regions (b) passenger load event regions

Fig. 4: Region recognition

donated as $p^{load\to drop/drop\to load}_{i\to j}$. To substantiate, the calculate process of $p^{load\to drop}_{i\to j}$ will be introduced in detail. The records in $REGION_i^{load}$ can be ~~re~~acquired from the data set easily. the record amount is ~~donated~~denoted as $\|RECORDS_i^{load}\| = \|\{record|record.location \in REGION_i^{load} \cap record.event = load\}\|$. For $record \in RECORD_i^{load}$, the next event and location can be easily ~~required~~acquired. Therefore, the record can be associated with the its next hop information to $(taxiid, location_{current}, event, event_{next}, location_{next})$. The $RECORDS_{i\to j}^{load\to drop} = \{record|event = load \cap event_{next} = drop \cap location_{current} \in REGION_i^{load} \cap location_{next} \in REGION_j^{drop}\}$ will be obtained.

$$p_{i\to j}^{load\to drop} = \frac{\|RECORDS_{i\to j}^{load\to drop}\|}{\|RECORDS_i^{load}\|} \tag{3}$$

$$P^{load\to drop} = (p_{i\to j}^{load\to drop})_{m\times n} \tag{4}$$

$$P^{drop\to load} = (p_{i\to j}^{drop\to load})_{n\times m} \tag{5}$$

B. Parameter estimation of $\overline{speed}$ distribution

In this section, we ~~modeling~~ modeled the speed distribution. ~~we~~ We fit the cumulative status average speed distribution to get the cumulative probability distribution function, and then take a derivative with it to obtain the speed probability distribution.

From figure. 1, the $\overline{speed}$ distribution shows exponential law. Given that, we ~~set~~ define the function form as follows:

$$\begin{cases} f(x) = 1 - 1/exp(a_1 x^{1.5}) \\ g(x) = 1 - 1/exp(a_2 x^{2.5}) \end{cases} \tag{6}$$

The fit formulas are given as formula~~s~~ 6. $f(x)$ is the function form for the $speed$ distribution of vacant status, and ~~the other one~~ $g(x)$ is for that of occupied status.

TABLE II: The parameter and the rms of residuals of fitting curves

Categories	rms of residuals	
$f(x)$	$a_1 = 0.117$	0.0113159

The rms of residuals for each fit are as ~~table~~ Table II. The smaller rms of residuals means better fitting. In the table II, the values are all less than 0.02, showing good similarity.

IV. Model Verification

In this section, START mobility model is validated on the aspects of node distribution and contact characteristics. All mobility models are implemented on Opportunistic Networking Environment (ONE)[6].

Shortest Path (SP) mobility model based on the map in Beijing is ~~an other~~a model for comparison, which is implemented by ONE. It also moves along the map roads ~~by~~ based on Dijkstra algorithm. The RWP model is another comparison model, because it is proved to be an efficient model modeling the nodal movement in VANETs. ~~But~~ However, it takes no consideration of the node statuses and geographical distribution.

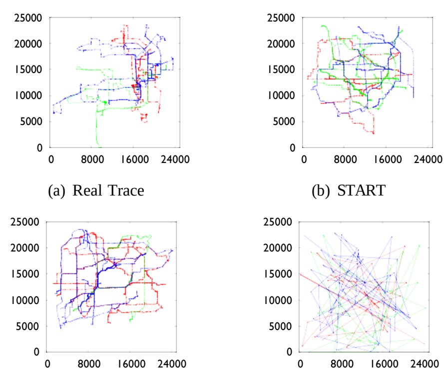
批注 [114]: Model 不能 move

The START, SP and RWP mobility model are compared with the real trace. In simulations, Node number is set as 4000 and scenario ~~in area~~size is $24445 \times 23584m^2$ (a sub-map of the whole area), including fourth ring roads in Beijing. The simulation time is three hours and the warm up time for reports is one hour, so that the nodal movement and position will not be affected by its initial position. The communication range is $200m$.

A. Traces and distribution of nodes

Trace samples and their snapshots are demonstrated in this section, shown ~~as~~ in fig. 5 and fig.6.

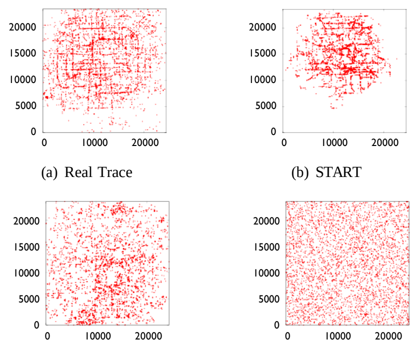
Fig.5 shows the trace in one day. The trace of the real data and START only cover some parts of the area, while the trace of SP and RWP almost go through the whole area. Because SP and RWP ~~will~~ select ~~a~~ destinations randomly, while START takes the associations between current regions and destinations into consideration ~~which satisfies the~~based on movement rules of taxis.



(a) Real Trace (b) START (c) SP (d) RWP

Fig. 5: Trace samples

From fig.6, Real trace START and SP mobility model exhibit the road structure. However, the node distribution of RWP is much uniform. As to START, the destination section process decides that it tends to select a destination in the regions with higher load/drop event probability. Therefore, with the decline of the randomness, the snapshot of START becomes much clearer and more centralized on the main roads.



(a) Real Trace (b) START (c) SP (d) RWP

Fig. 6: Trace snapshots

B. Contacts characteristics

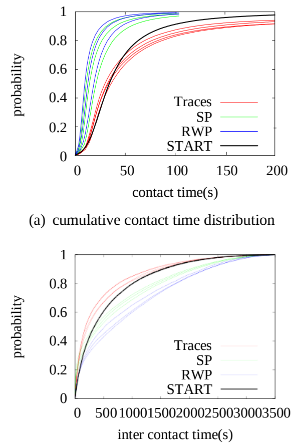
The contact time and inter-contact time are evaluated as the indicators to validate the similarity. The speed range of RWP and SP need to be configured. To ensure the accuracy, we choose three speed ranges: $[0, 44.4]m/s$,$[0, 33.3]m/s$ and $[0, 22.2]m/s$, that is, the upper bounds of speed are 80,120,160 km/h.

Fig. 7. (a)(b) is the contact time and inter-contact distribution, which shows the probability of the contact or inter-contact time smaller than certain time length. To substantiate, a point $(200, 0.5)$ in these figures means the probability is 0.5 when contact or inter-contact time is shorter than $200s$. In addition, a point (x, y) in 8.(a) means the sum of contact time is y when the simulation time is x. In this figure, the three green lines of SP and the three blue lines of RWP are ~~coincide~~ overlapped with each other. To recognize the differences, we set y axis as log-scale in fig.8.(b). The curves of the total contact time vs. time present~~s~~ a ~~liner~~ linear law. For SP and RWP, the differences of speed ranges show little influence on these curves. Clearly, the rank of the contact characteristic similarity with the real data is $START > SP > RWP$.

To conclude, by comparing the node distribution and contact characteristics, the START mobility model performs good similarity with the real data.The evaluation results conform to our expectations. Because $START$ takes use of speed and geographic ~~feathers~~ features related ~~with~~ to status, ~~While~~ however, SP employs the map information and RWP is a random model taking use of no realistic data.

V. Conclusion

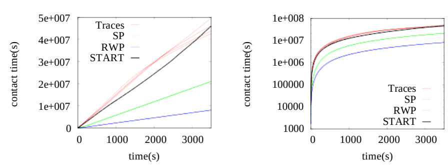
Since the mobility model is important for mobile network, a novel mobility model START based on real GPS data is



(a) cumulative contact time distribution

(b) cumulative inter contact time distribution

Fig. 7: Contact time, inter contact time distribution



(a) time vs. total contact time (b) time vs. total contact time in ~~logscale~~ log scale

Fig. 8: time vs. total contact time

proposed. By assuming the taxi behavior is related with its statuses and geographic ~~feathers~~ features, statistical experiments are conducted to verify those assumptions using the real trace data. Further, its parameter–average speed of each status ~~are~~ is estimated respectively, and the region transition probability is calculated. In this case, macroscopic movement, a node moves switch between load-event regions and drop-event regions, and microscopic movement(speed for each status) can be defined. Finally, the START is implemented on ONE simulator and estimate it by comparing with the real trace, RWP and ~~Shortest Path~~ SP mobility Model. Comparing the node distribution and contact ~~feathers~~ features, START shows better performance. Simulation results demonstrate that START ~~has~~ gives a good approximation ~~with~~ of reality.

References

[1] S. Yousefi, M. Mousavi, and M. Fathy, "Vehicular ad hoc networks (vanets): Challenges and perspectives," in *ITS Telecommunications Proceedings, 2006 6th International Conference on*, June 2006, pp. 761–766.

[2] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces." in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, vol. 6, 2006, pp. 1–13.

批注 [115]: 没查到这个用法

批注 [116]: 语法不通顺

批注 [117]: 语法不通顺

[3] H. Huang, Y. Zhu, X. Li, M. Li, and M.-Y. Wu, "Meta: A mobility model of metropolitan taxis extracted from gps traces," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, 2010, pp. 1–6.

[4] R. Ganti, M. Srivatsa, A. Ranganathan, and J. Han, "Inferring human mobility patterns from taxicab location traces," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '13. New York, NY, USA: ACM, 2013, pp. 459–468.

[5] J. Aslam, S. Lim, X. Pan, and D. Rus, "City-scale traffic estimation from a roving sensor network," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '12. New York, NY, USA: ACM, 2012, pp. 141–154.

[6] A. Keraen, J. Ott, and T. Karkkainen, "The ONE simulator for DTN protocol evaluation," in *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 55.

APPENDIX

Algorithm 1 Clustering

Require: $Cells = \{CELL\}$
  η ##events threshold
  $CLUSTERSCALE$
  $REGIONSEED = 200$
  $ClusterQueue = \emptyset$
  $UsedCells = \emptyset$
  Sort $Cells$ by events DESC
  for $CELL_{x,y} \in Cells$ **do**
    if $CELL_{x,y} \notin UsedCells$ **then**
      $CELL_{x,y}.region = REGIONSEED$
      $size = 1$
      $CLUSTERSEED = CLUSTERSEED - 1$
      $ClusterQueue.enqueue(CELL_{x,y})$
      $UsedCells.add(CELL_{x,y})$
      while $ClusterQueue \neq \emptyset$ **do**
        $CELL_{x,y} = ClusterQueue.dequeue()$
        if $REGIONSEED \geq 0$ **and** $CELL_{x,y}.events \geq \eta$ **then**
          $enqueueNeighbor(CELL_{x-1,y})$
          $enqueueNeighbor(CELL_{x-1,y-1})$
          $enqueueNeighbor(CELL_{x-1,y+1})$
          $enqueueNeighbor(CELL_{x+1,y})$
          $enqueueNeighbor(CELL_{x+1,y-1})$
          $enqueueNeighbor(CELL_{x+1,y+1})$
          $enqueueNeighbor(CELL_{x,y-1})$
          $enqueueNeighbor(CELL_{x,y+1})$
        else
          $enqueueNeighborOthers(CELL_{x-1,y})$
          $enqueueNeighborOthers(CELL_{x-1,y+1})$
          $enqueueNeighborOthers(CELL_{x-1,y-1})$
          $enqueueNeighborOthers(CELL_{x+1,y})$
          $enqueueNeighborOthers(CELL_{x+1,y-1})$
          $enqueueNeighborOthers(CELL_{x+1,y+1})$
          $enqueueNeighborOthers(CELL_{x,y-1})$
          $enqueueNeighborOthers(CELL_{x,y+1})$
        end if
      end while
    end if
  end for

Algorithm 2 $enqueueNeighbor(CELL_{x,y})$

  if $CELL_{x,y}.events \geq \eta$ **and** $size < CLUSTERSCALE$ **and** $CELL_{x,y} \notin UsedCells$ **then**
    $ClusterQueue.enqueue(CELL_{x,y})$
    $CELL_{x,y}.region = REGIONSEED$
    $UsedCells.add(CELL_{x,y})$
    $size = size + 1$
  end if

Algorithm 3 $enqueueNeighborOthers(CELL_{x,y})$

  if $size < CLUSTERSCALE$ **and** $CELL_{x,y} \notin UsedCells$ **then**
    $ClusterQueue.enqueue(CELL_{x,y})$
    $CELL_{x,y}.region = REGIONSEED$
    $UsedCells.add(CELL_{x,y})$
    $size = size + 1$
  end if