IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 6, 2019, accepted May 21, 2019, date of publication May 28, 2019, date of current version June 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919544

A Large-Scale Urban Vehicular Network Framework for IoT in Smart Cities

HENG LI [1], YONGHE LIU [2], ZHENG QIN [1], HUIGUI RONG [1], AND QIN LIU [1]

[1]College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
[2]Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX 76019, USA

Corresponding author: Heng Li (lih@ hnu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61672221 and Grant 61872133, in part by the China Scholarship Council under Grant 201706135061, in part by the Young Talent Group of National One Thousand Plan of China, and in part by the Science and Technology Key Projects of Hunan Province under Grant 2017GK2272 and Grant 2015 TP1004.

ABSTRACT Gathering extensive public data from the IoT devices scattered throughout a city is a significant challenge in a smart city. In this paper, a novel urban vehicle network we call the location-based urban vehicle network (LUV), is proposed to perform the non-real time data gathering task in smart cities. Different from the most works in-vehicle networks, the empirical research on the real-life 8900 private cars trace data in Changsha, China, compels us to focus on the parked vehicles and the parking places, rather than on the moving cars and the urban roads. The location-based mechanism not only provides more reliable and predictable wireless connections but also dramatically simplifies the system topology. It ensures that the vehicle network deployed on an intricate metropolitan area reaches the desired scale for gathering data.

INDEX TERMS Vehicular network, Internet of Things, scalability, smart city.

I. INTRODUCTION

Nowadays, the smart city has attracted a tremendous amount of interests from government, academia, and industry [1], [2]. In this field, connecting everything in a city and sensing every corner of the city will be within reach if meeting the booming communication demands from the Internet of Things (IoT) [3]. The current estimates of about 20 billion of IoT sensors are connected, and this number will reach 50 billion by 2020 [4]. The Fifth Generation (5G) wireless networks are developing as the network solution of Smart City to connect the billions of IoT devices [5], [6].

A. MOTIVATION

In a smart city, the non-real-time urban IoT data collected over a period are converted to the usable information or knowledge for daily city operation and long-term urban planning, and the real-time IoT data are utilized to monitor the urban environment for rapidly responding the various emergencies [7]. In the scene of gathering non-real-time IoT data, there is no significant difference in performance between the high-speed telecom communication networks (Cable/5G) and the delay tolerant network (DTN), which is usually considered as a temporary emergency alternative for the destructed communication networks by natural disasters or military efforts [8]. That shows the potential demands for the low-cost DTN in smart cities, primarily when the node of DTN can be deployed on the numerous mobile devices in the urban environment. While the 5G networks do the real-time IoT data transmission for emergency monitoring and responding, the DTN perform the daily urban non-real-time IoT data-gathering task for the further urban data mining.

Smartphones and vehicles have been treated as the possible communication and computing resources of a smart city for their characteristics of the copious quantity, the wide-coverage and high-density in the urban area [9], [10]. Considering vehicle can carry larger-size, heavier-weight, and higher-power terminals, the vehicle-based DTN [11] is more suitable as the public non-real-time IoT data gathering platform than the DTN networked by the devices carried by people [12], [13]. By the daily movements of vehicles, the vehicle network can transmit the IoT data distributed in the urban area [14]–[16].

B. CHALLENGE

More connected vehicles signify more resources to be utilized and better network performance when the cars are treated as the potential communication resources and networked for gathering IoT data. Given that there are millions of vehicles in

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu.

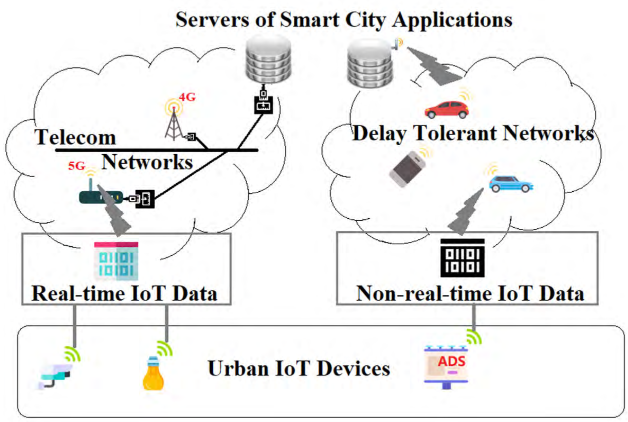


FIGURE 1. Urban IoT data transmission with the telecom networks and the DTNs.

the modern metropolis, the scale of the urban vehicular DTN should reach the same level. Meanwhile, the overall size of traditional VANETs can hardly be scaled up to large-scale.

A decade ago, when the vehicle was just a tool of transport for people and cargo, Vehicular Ad-hoc Networks (VANETs), a type of vehicle-based DTN, were first proposed as a promising approach for future intelligent transportation system (ITS) [17], [18]. Since VANETs are suggested to aid road vehicle driving, VANETs only connects the on-road vehicles. The assigned purpose of VANETs implies that VANETs nodes are always in motion. The task-driven presupposition of VANETs triggers a series of challenges: the network topology is dynamically varying; the contact duration between vehicles is transient; the contacts between cars occur randomly; the wireless channels in the dynamical urban environment are more prone to error [19].

As a result of the contradiction between the highly dynamic network topology and the non-centered network structure, the overhead cost of system collaborations between moving vehicle nodes will grow exponentially in VANETs with the increase of network size. It makes the overall size of traditional VANETs can hardly be scaled up to large-scale. In this case, the more vehicles are networked; the fewer network resources would be available to IoT devices owing to the higher network overhead cost.

Therefore, the primary challenge for the urban vehicle-based DTN for IoT data comes from decreasing the system overhead cost and increasing the system scalability.

C. METHODOLOGY

In this paper, firstly, based on the new purpose of IoT, we simplify the network topology to improve the network scalability. Secondly, we introduce the external network scheduling to reduce the system overhead on the large-scale vehicular DTN node.

To design a large-scale vehicular DTN framework, the properties of urban vehicles the researchers should fully utilize. To comprehensively study the urban vehicles, we analyze 3-month trace data of 8,900 private cars in Changsha, China, collected by a vehicle monitor system. We found that privately-owned vehicles exhibit significant characteristics, such as dominant parking time, highly repetitive moving pattern, and concentrated visited places in urban environments.

Based on these characteristics of privately-owned vehicles we found and the existing urban traffic routing and flow control system, we present a novel type of urban vehicular DTN framework for gathering IoT data, which is called the location-based Urban Vehicle Network (LUV). Instead of relying on the on-road transient connections, LUV focuses on the stationary states of vehicles and centers its network operations on urban places. In the LUV, the city area is divided into numerous places by roads, and the IoT devices of one place submit data to the parked vehicles which belong to the place. While the cars move from one place to another, they will carry the data as well.

By focusing on the stable places and data exchange among parked vehicles and IoT devices, LUV provides more stable connections, more predictable node statuses and simpler network topology which can reduce the system complexity. Moreover, by embedding the external urban traffic routing and flow control system, the system overhead on the nodes does not rise sharply with the growth of the DTN. Hence, LUV can scale well into large-scale networks.

D. CONTRIBUTIONS

In this paper, according to the urban vehicle traffic network and a significant amount of real trace data of privately-owned vehicles, detailed features of the characteristics of the cars are extracted and quantified with mathematic models. A novel vehicle activity model is presented to describe the activity pattern of urban privately-owned vehicles. Based on the activity model, we present a location-based urban vehicular network framework (LUV) for gathering local IoT data.

The remainder of this paper is organized as follows. In Section II, we analyze the urban vehicles and the urban vehicular transportation network. In Section III, we introduce the analysis of real-life monitoring data and the experiment results. In Section IV, we present a daily vehicle activity model. In Section V, we propose the Location-based Urban Vehicular Network framework for IoT. In Section VI, we show the experimental results of LUV over the real-life dataset. We then conclude this paper and outline the directions for future works in Section VII.

II. URBAN VEHICLES & VEHICULAR TRANSPORTATION NETWORK

In this section, we analyze the types of urban vehicles, and the existed urban vehicle transportation network. We aim to find out the factors which bring millions of urban cars out of chaos. By utilizing these factors, we can build a large-scale urban vehicle network.

A. TYPE OF URBAN VEHICLES

Urban vehicles can be characterized into three categories.

TABLE 1. Comparison of the three types of urban vehicles.

Urban Vehicle	Total weight	Type	Random-ness	Relationship of daily trips	Prediction method
Bus	Small	Fixed	Low	Duplication	/
Private vehicle	Dominate	Regular	Middle	Dependence	Bayesian decision
Taxi	Small	Roaming	High	Independence	Markov chain

1) FIXED ROUTE TYPE

Typical examples of this type of vehicles are buses serving as public transportation means. They usually have fixed routes and also set schedules. Certain VANETs have centered design on this type of vehicles owing to regularity and predictability of their movement patterns [20].

2) ROAMING TYPE

Typical examples of this type of vehicles are taxis. A notable feature of this type of vehicles is that their moving routes are seemingly random. Their current path is independent of their past tracks. Often research VANETs models roaming type vehicle with Markov chains.

3) REGULAR TYPE

Often overlooked in the existing study are privately-owned vehicles in urban settings, possibly owing to the difficulty of obtaining real trace data. We term this regular type as these vehicles are the most common in everyday life. It is also more complicated as the states of the vehicles exhibit a combination of certain randomness and repetitiveness. Researchers have used Bayesian decision models for the predicting of the states of the regular type vehicles. Table 1 provides a comparison of the three types of vehicles.

As we discussed above, quantitatively, privately-owned vehicles are typically the dominant type in the most metropolis of the world today. while buses and taxis mainly serve as complementary roles of municipal vehicles. To understand the essential feature of urban vehicles, we prefer to study the characteristics of private cars. To achieve this goal, we present a comprehensive analyzing of privately-owned vehicles in the next section.

B. URBAN VEHICULAR TRANSPORTATION NETWORK

The studying of the existing urban vehicular transportation network helps us understand the rules of the vehicle in an urban environment to design a new type of large-scale urban vehicle network.

In a vehicle trip, a vehicle loads passenger/cargo before leaving its departure place. Then the car moves on the roads and unloads its passenger/cargo after arriving at its destination place. The purpose of a privately-owned vehicle trip is to transfer people/cargo from one place to another. The departure and destination places are fixed before the vehicle starts its tour, while the drivers change the road paths temporarily according to the variable road conditions.

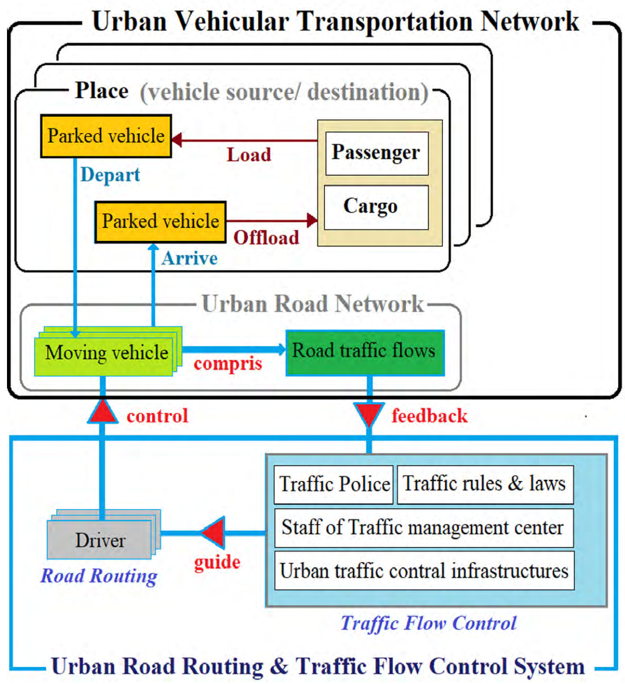


FIGURE 2. Urban vehicular transportation network.

Individually, the drivers usually receive the real-time traffic guidance from the multiple sources and self-execute dynamic path tuning to optimize their trips. Wholly, the traffic police and the staff of the city traffic management center exploit the deployed urban traffic control infrastructures to guarantee the holistic transportation performance against traffic congestion as possible. All of them compose the urban road routing & traffic flow control system, which ensures the efficiency of the urban transportation with millions of vehicles.

If we can embed the urban road routing & flow control system into the urban vehicle network, the network will obtain the capabilities of mega-size network congestion control and scheduling without the need for the massive resources of vehicle nodes.

III. EMPIRICAL DATA ANALYSIS

Vehicle network study from the perspective of private vehicles is rare due to the lack of real-life traces of privately-owned vehicles. In this section, we provide a study of a broad set of traces of privately-owned vehicles. First, we give a brief overview of the monitor dataset source. We next present the primary stationary metrics embedded in the monitor data sets, including average daily parked time, all-day parking ratio, and residential area. Then we carry out a detailed statistical analysis of the pretreated dataset and obtain empirical metric values.

A. MONITORING DATA SET

The monitoring dataset we study is generated by the private vehicle monitoring system (PVMS), which installed on 8900 privately-owned vehicles in Changsha, China. The PVMS provides remote vehicle monitoring service to

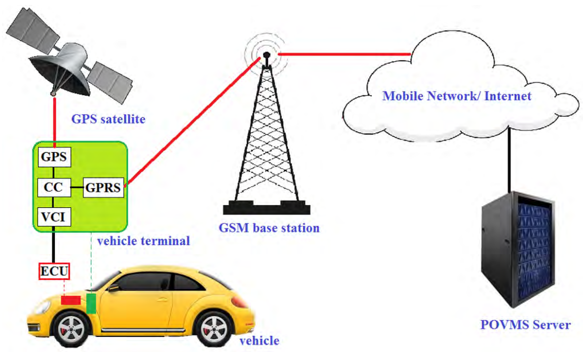


FIGURE 3. System architecture of PVMS.

privately-owned vehicles. The essential functions of PVMS include GPS location tracking, theft alarm, and remote device diagnosis. A vehicle terminal is mounted on the target vehicle for real-time monitoring. It collects real-time vehicle data and sends them to a central monitoring platform through GPRS networks. The system architecture of PVMS is illustrated in Figure 3.

The vehicular terminals reported the monitoring data of vehicles back to the central monitoring platform every 30 seconds. To reduce the total amount of the history monitoring data, the PVMS samples the monitoring data of every vehicle at every 10 minutes. As a result, the history monitor data we study are totaling 42,518,112 records from 8900 vehicles spanning 61 days from March to April of 2013.

During Mar.1-Apr.30, 2013, some privately-owned vehicles joined, some quitted the PVMS, and some had encountered monitoring equipment failure. Therefore, these vehicles' record periods are less than 61 days. To improve the accuracy of data statistics, the data from these vehicles have been excluded from the dataset. After the process of filtering, totaling 37,789,904 records from 7520 privately-owned vehicles in Changsha spanning 61 days were retained.

B. AVERAGE DAILY PARKED TIME

With the statistics of average daily parking time of privately-owned vehicles in Changsha, we can learn the empirical probability of vehicle parking per day. Correspondingly, privately-owned vehicles daily running time is 1.37 hour/day, or 82 min/day. The average daily parked time of privately-owned vehicles in Changsha is about 22.63 hours/day.

The result is highly consistent with previous studies in urban design, which claimed that vehicles are not on the road for 95 percent of the day [21]. It indicates that stationarity is a dominating feature of privately-owned vehicles. In the majority of the time, urban vehicle network nodes are in parked states.

C. ALL-DAY PARKING RATIO

If a vehicle did not move in a day, then the vehicle has an All-day parking state. All-day parking is the most steady and predictable state. It is usually taken as an extreme and rare circumstance, generally ignored in most vehicle network studies. However, our data set indicates that the average daily parking time counts for 95% in a day. A reasonable inference is that the scale of all-day parking states of vehicles in the urban environment more common than we have thought. To study this, we introduce one key metric, All-day parking Ratio (R_ADP), to describe the global privately-owned vehicle All-day parking situation in the urban environment. We define R_ADP as

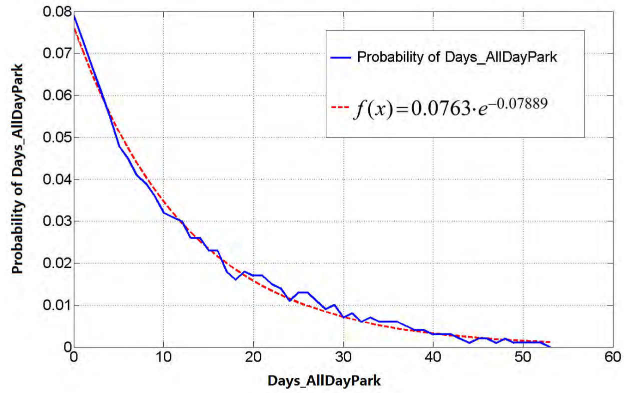


FIGURE 4. PDF of Days_AllDayPark & exponential distribution.

$$R\_ADP = \frac{Days\_All\ Day\ Park}{Days\_statistics} \tag{1}$$

Days_statistics represents the statistical period, and *Days_AllDayPark* represents the number of the all-day parking days in the statistical period. Based on the vehicles' monitoring data statistics, we obtain the distribution of *Days_AllDayPark* in Figure.4. The plots follow a clear exponential distribution.

Let X be a random variable as *Days_AllDayPark* over the 61 days (*Days_statistics*) that has an exponential distribution with the mean $E(X)$ and the variance $VAR(X)$, i.e., $X \sim Exp(\lambda)$.

To identify the exponent constant λ of the exponential distribution of all-day parking day, we apply polynomial regression. The validation of the regression is measured by the coefficient of determination (*R_square*), and the root mean squared error (*RMSE*)

$$R\_square = 1 - \sum_{i=1}^{n}(x_i - \hat{x}_i)^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2}$$

where x_i denotes the sample value with the mean $\hat{x}$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \tag{3}$$

We apply this exercise to the plots in the Figure.4 where the distribution of all-day parking radio is very well approximated ($R\_square = 0.9942$, $RMSE = 0.0016$) by an exponential distribution $Exp(0.07889)$.

$$\Rightarrow E(X) = \frac{1}{\lambda} = 12.68, \quad VAR(X) = \frac{1}{\lambda^2} = 160.68 \tag{4}$$

Let R be a random variable as R_ADP, we have

$$\Rightarrow E(R) = E(\frac{X}{Days\_statistics}) = \frac{1}{Days\_statistics} \cdot E(X) = 0.2079 \quad (5)$$

The expectation of R_ADP in Changsha city is 20.79%. It indicates that on average one-fifth of privately-owned vehicles are in the state of all-day parking. All these All-day parking vehicles can be the transit node (the nodes running DTN routing) in the vehicle network. It indicates that the urban vehicle network has adequate resources for local message storage and forward.

D. RESIDENTIAL AREA

By further analyzing the states of all-day parking, we observe that the locations of all-day parking of vehicles are very predictable and varying little. These mostly are the places of residence of the owners. Being consistent with our common sense, most people only have one residential house. The monitor data shows that the residential area of a vehicle is sole. Additionally, these places have the highest daily visit frequency and the most prolonged parking periods, about 14.6 hours per day. It indicated that the privately-owned vehicle in the urban environment has one fixed place of residence, and it routinely departs from and returns to the residential area with some case of staying at the residential area all day.

It implies that the millions of vehicle nodes in a large-scale urban vehicle network can be naturally divided into the much smaller location-based groups by their places of residence. Moreover, the vehicles nodes addressed in the same place have the similar activity patterns, considering that the residents living in the same community have the same urban living facilities, such as parking lots, schools, and supermarkets.

IV. MODEL ANALYSIS

In this section, we propose our vehicle activity model for the large-scale urban vehicle network.

A. NOTATIONS

We define the following notations.

Definition 1: The region of Urban is defined as U

Definition 2: The vehicles in the region Re are defined as V_U, the amount of the vehicles in the region is defined as Num_V_U, and the vehicle i is defined as

$$v_i \in V_U, \quad 0 < i \leq Num\_V_U$$

Definition 3: The parked areas which vehicles gathered in the region U are defined as P, P $\in$ U, the amount of the parked areas in the region is defined as Num_P_U, and the parking area j is defined as

$$p_j \in P, \quad 0 < j \leq Num\_P_U$$

Definition 4: The residential place of the vehicle v_i is defined as $RP\_v_i \in P$

Definition 5: The time slots of a day are defined as TS, the amount of the time slots of a day is defined as Num_TS, and the time slot k is defined as

$$ts_k \in TS, \quad 0 < k \leq Num\_TS$$

Definition 6: The parked/active status of the time slot ts_k of the vehicle v_i is defined as $status\_ts_k^i$

$$status\_ts_k^i = \begin{cases} 1, & \textit{vehicle i parked at the time slot k} \\ 0, & \textit{vehicle i with motion at the time slot k} \end{cases} \quad (6)$$

Definition 7: The daily parked time of the vehicle v_i is defined as PT_i

$$PT_i = \sum_{k=1}^{Num\_TS} Status\_ts_k^i \quad (7)$$

Definition 8: The average daily parked time of the vehicles V_U with the time slots is defined as $ADPT_{TS}^V$

$$ADPT_{TS}^V = \frac{1}{Num\_V_U} \sum_{i=1}^{Num\_V_U} \sum_{k=1}^{Num\_TS} Status\_ts_k^i \quad (8)$$

B. TIME UNIT OF TIME SLOT

To observe and describe the daily activities of a vehicle, we divide a day into continuous time slots with a specified time unit. The smaller the time unit is, the more precise the results are. On the other hand, to urban VANETs, a system with millions of nodes, a smaller time unit means massive overhead in data processing. Considering that privately-owned vehicles are not on the road for 95 percent of the time, there is a chance of reducing the system overhead and also maintaining the accuracy of the results by setting coarse-grained time unit.

The dataset we studied was sampled every 10 minutes. The analysis result with 10min time unit shows that the average daily park time is 22.63 hours/day. The result is highly consistent with previous studies in urban design. It shows that 10 minutes time slot have good performance in accuracy.

In this paper, we try to set a coarser time unit than 10 minutes for lower system complexity and overhead. We place 1 hour as an essential time unit and divide a day into 24 slots. With the data set from 7520 privately-owned vehicles spanning 61 days, we obtain the distribution of the total number of daily parked time slots (the time slot unit is 1 hour), as shown in Figure.5.

The average daily parked time average $ADPT_{1h}$ from the 7520 privately-owned vehicles spanning 61 days is

$$ADPT_{1h} = \frac{1}{61 \times 7520} \sum_{d=1}^{61} \sum_{i=1}^{7520} \sum_{k=1}^{24} Status\_ts_k^i = 20.54 \quad (9)$$

Combined with $ADPT_{10m}$ (the time slot unit is 10 minutes) is 22.63 hours/day, the following equation can obtain the coefficient degree of $ADPT_{1h}$.

$$C_{ADPT_{1h}} == 1 - \frac{|ADPT_{10m} - ADPT_{1h}|}{ADPT_{10m}} = 90.76\% \quad (10)$$

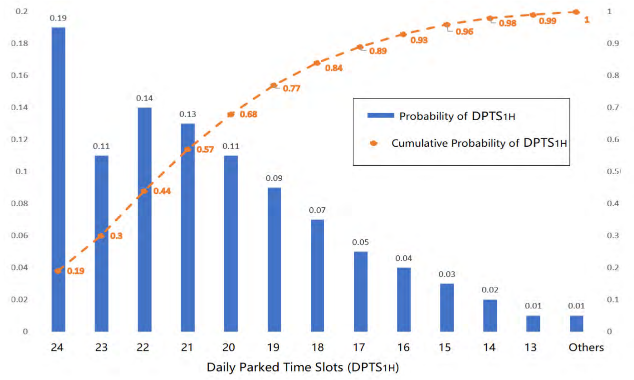


FIGURE 5. Probability distribution of the total number of DPTS_{1H}.

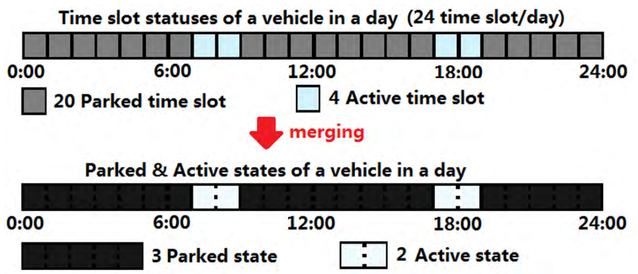


FIGURE 6. Method to obtain the parked states and the active states.

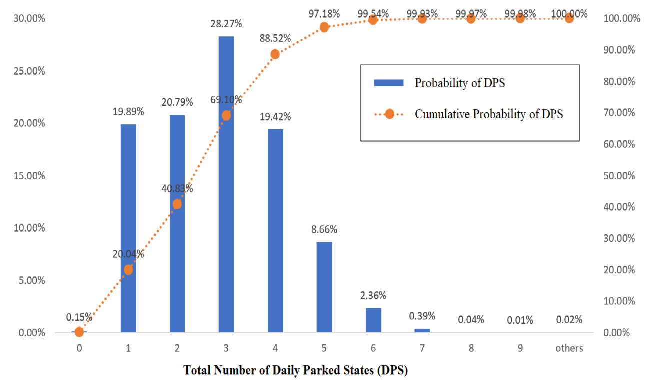


FIGURE 7. Probability distribution of the total number of daily parked states.

The result shows that one hour is an appropriate time unit to observe and describe the daily parked events of vehicles in urban vehicle network.

C. PARKED & ACTIVE STATES

For a privately-owned vehicle, we split a day into some parked states and active states based on the vehicle trips. Hence, by merging the adjacent parked/ active time slots of a day, we obtain the parked and the active states of the vehicle, as shown in Figure.6.

From the data of the 7520 vehicles spanning 61 days, we obtain the probability distribution of the total number of daily parked states (DPS) as shown in Figure.7. The average total number of daily parked states is 2.85.

We know that the All-day parking (none active state, one parked state) ratio is 20%. Now we know that one to three trips/day ratio is 68.58%; four trips/day ratio only is 9%, and more than four trips/day ratio is less than 3%.

TABLE 2. Statistic analysis result of privately-owned vehicles in Changsha.

Statistic Metric	Empirical value	Unit	Description
ADPT	22.63	hour/day	Average daily parked time
ADPRP	14.6	hour/day	Average daily parked time in the place of residence
ADPTS	20.54	state/day	Average total number of parked time slots per day
ADPS	2.85	state/day	Average total number of parked states per day
ADT	1.85	trip/day	Average amounts of daily trips
AAS	0.73	hour/active state	Average trip time
R_DP_0	20	%	All-day parking radio
R_DP_1	21	%	One trip per day radio
R_DP_2	28	%	Two trips per day radio
R_DP_3	19	%	Three trips per day radio
R_DP_4	9	%	Four trips per day radio

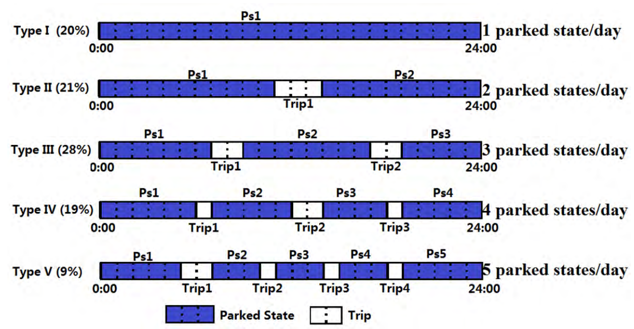


FIGURE 8. The five regular parked patterns accounted for 97%.

The average numbers of daily trips, *ADT*, which denotes the daily vehicular motion frequency, is

$$ADT = \sum_{i=1}^{NumTrip} i \cdot \Pr ob\_Trip_i = 1.85 \qquad (11)$$

Combined with the average privately-owned vehicles daily running time is about 82 min/day, the average time of an active state is 0.73 hours/day from the data. Table.2 summarizes the statistical analysis results of real data from the privately-owned vehicles in Changsha.

D. URBAN PRIVATELY-OWNED VEHICLE ACTIVITY MODEL

Based on the analysis results of normal parked states of vehicles, we obtain five daily parked patterns which account for 97% of the population, as shown in Figure.8.

Combined with the fact that privately-owned vehicle in the urban environment has one fixed place of residence and routinely depart from and return to the place, there are 23 types

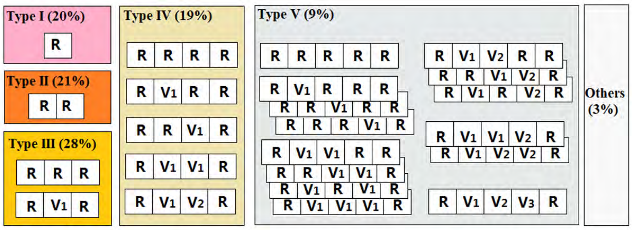


FIGURE 9. The 13 Daily vehicle activity patterns with the parked place.

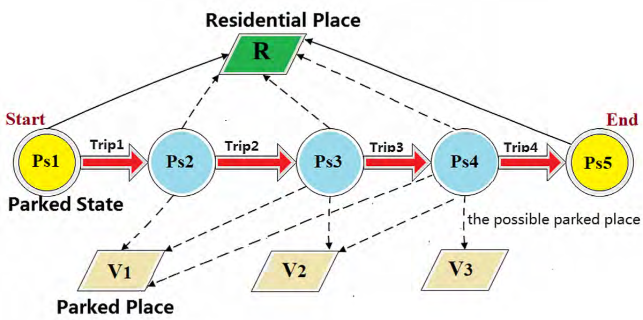


FIGURE 10. Urban privately-owned vehicle activity model.

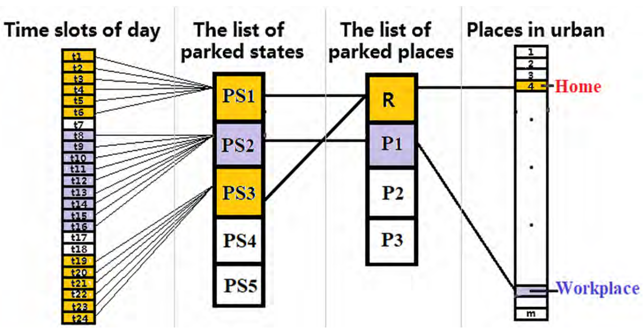


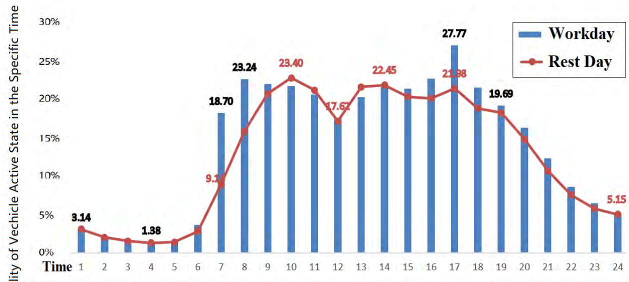
FIGURE 11. The instance of a vehicle's daily activities with the activity model.

of daily activity patterns with the parked places, as shown in Figure.9.

From the above daily activity patterns, the probability of that the number of daily parked places is fewer than five accounted for 97%. Therefore, we propose the activity model of privately-owned vehicles for urban vehicle network based on the privately-owned vehicle activity patterns, as shown in Figure.10.

The activity model has five parked states, three parked places, and one place of residence. It can describe 97% daily activity behaviors of urban privately-owned vehicles in real life. For instance, on a working day, a vehicle departed home at 7:10 am and arrived at the workplace parking lot at 7:50. At 5:30 pm it left the parking lot and returned home at 6:20 pm. The daily parked behaviors of the vehicle belong to the parked type t4, and its activity model description is shown in Figure11.

Based on the different vehicle activity patterns between the workdays and the rest days, as shown in figure 12, there are two activity models for a vehicle, which we define as the workday activity model and the rest-day activity model.



Time	1	2	3	4	5	6	7	8	9	10	11	12
Workday	3. 14%	2. 09%	1. 64%	1. 38%	1. 50%	3. 75%	18. 70%	23. 24%	22. 59%	22. 29%	21. 21%	17. 98%
Rest Day	3. 15%	2. 06%	1. 57%	1. 32%	1. 44%	2. 89%	9. 14%	16. 26%	21. 29%	23. 40%	21. 76%	17. 62%
Time	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Workday	20. 85%	22. 19%	21. 98%	23. 34%	27. 77%	22. 14%	19. 69%	16. 72%	12. 65%	8. 80%	6. 61%	5. 19%
Rest Day	22. 18%	22. 45%	20. 89%	20. 66%	21. 98%	19. 32%	18. 72%	15. 19%	10. 96%	7. 77%	5. 92%	5. 15%

FIGURE 12. Distributions of daily activities of vehicles in workday & rest day.

By embedding the calendar into the vehicle system, the vehicle can distinguish whether the current day is a working day or a rest day. Therefore, the urban vehicles generate the two activity models, depending on their history traces records of the working days and ones of the rest days respectively. Then the vehicles adaptively choose the right model in their daily operations.

V. URBAN VEHICULE NETWORK FRAMEWORK

In this section, we propose our vehicle activity model for the Motivated by the characteristics of privately-owned vehicles in urban environment discussed above, we present a location-based urban vehicle network, which we term LUV, for the urban IoT data transmission.

A. LOCATION-BASED MECHANISM

The unique place of residence of a vehicle implies that it can address the vehicle node in the vehicle network. When the vehicle network needs to transfer IoT data to a specified location, it can easily find the target carrier by comparing the place of residence of the candidate vehicle and the data destination.

In the local-based network, the vehicles provide the datalinks between their place of residences and their visiting places. The distributed IoT data can be transported from one place to another place by the movements of the urban vehicles. We termed this vehicle network as the Location-based urban vehicle network (LUV).

LUV is quite unlike the traditional road-based vehicle network. It only considers the departure places and the destination places of urban vehicles. Since every vehicle has its road routing mechanism, it can autonomously select a reasonable path to reach its destination. Taking advantage of the factor, LUV obtains the urban road routing capability by following the movement of vehicles.

The Location-based mechanism simplifies the vehicle network topology remarkably considering the complexity of the urban road network. Moreover, the mechanism drastically reduces the network dynamics by removing the random factor about the temporary road choosing of the vehicles.

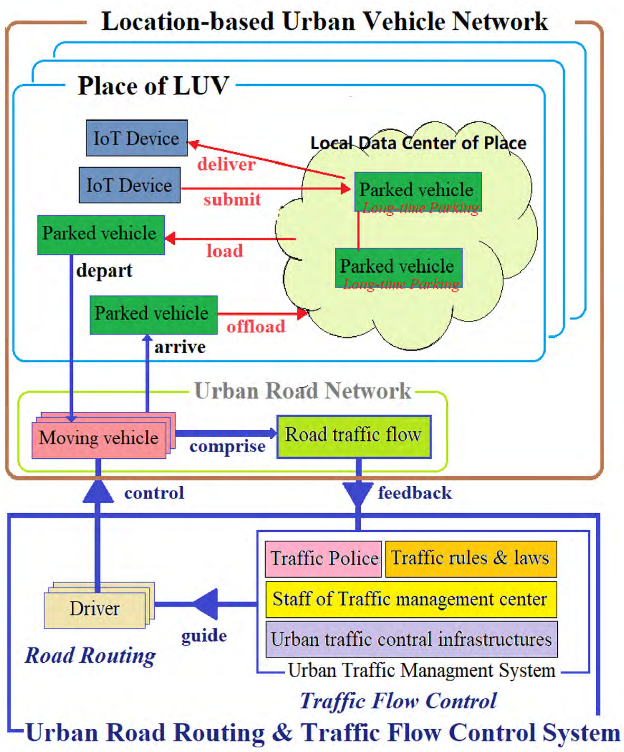


FIGURE 13. Network framework of LUV.

Hence, the Location-based mechanism offers the possibility for the urban vehicle network to accommodate millions of self-originating vehicle nodes in the sophisticated city environment.

B. NETWORK FRAMEWORK

By following the existed people/cargo transmission pattern of vehicles, LUV is parasitic on the existed urban vehicle transportation network to obtain the mega-size network traffic control and path routing in the complex urban environment.

If we consider the IoT data of an urban place as a particular type of cargo and follow the existed people/cargo transmission pattern of vehicles, the non-central urban vehicle network can integrate the urban vehicle transportation network to improve its network capabilities and scalability. Hence, we present a novel location-based network framework for the large-scale urban vehicle network, as shown in Figure 13. The parked vehicles exchange and route IoT data in the local data center which is deployed on the long-time parking vehicles in the area and transfer data on the roads under the effect of the existing urban traffic routing and flow control system.

The primary network operations are composed of three components: message exchange at a place between IoT devices and the local data center which is deployed on the parking vehicles, message routing (relay-place) at a place by the local data center, and message forwarding from place to place by the moving vehicles.

C. LOCAL DATA CENTER OF PLACE

Based on our previous analysis of the data traces, we find that almost one-fifth of private vehicles are in the state of all-day parked and more vehicles are in the states with long-time parking. It shows the practicability of deploying the distributed local data center of place on the all-day parked vehicles and the long-time parked vehicles in the area.

To describe the deployment scheme of place data center, we introduce two types of nodes in place as follows.

Local node: a parked node at its place of the parked list of its activity model.

Foreign node: a parked node not in its place of the parked list of its activity model.

We devise the following approach for deployment of the local data center among the local nodes in steps.

a: OBTAINING THE DEMANDS OF THE DATA CENTER IN THE PLACE

By the registration information from the IoT devices deployed in the place, LUV can obtain the IoT data cache demands of the area for the local data center. Depending on the data needs of the IoT devices and the IoT data cache capacity of a vehicle node, the number of required vehicle nodes in the local data center can be obtained.

b: CHOOSING DATA CENTER NODES

The probability of remaining in parked states (Prob_RPS) during a time slot can be inferred from the local node's activity model which is generated and updated by its history trace records. We merely randomly choose these local nodes as data centers when their Prob_RPS satisfy the predefined threshold Thr_RPS, till the amount of data center nodes fulfill the demands. For there are plenty of all-day parked vehicles and long-time parked vehicles in reality, LUV can generally find enough candidates for the local data centers from the local nodes in the places of urban.

c: RETIRING DATA CENTER NODES

Prob_RPS of a vehicle node varies with time. As a part of the local data center, when its Prob_RPS is lower than the predefined threshold Thr_RPS, the data center node stops caching IoT data and sends saved cache data to other data center nodes. Then the node into a carrier mode and load the IoT data from the local data center, and carry the IoT data to the destination place with its trip.

D. PLACE OF LUV

LUV obtains places from the urban street blocks. Street blocks, as the results of the ongoing urban planning process, is the fundamental units of the urban traffic network. Generally, the urban traffic network is designed and operated to meet the traffic demands of every street block in the city. Since LUV has the same transmission pattern with the urban traffic network, the places of LUV come from the urban street blocks in reality when the digital map of the city identify its boundaries.

Based on its historical traces, each vehicle can determine its place of residence and the daily visiting place list. According to the activity model of a vehicle, the places where the two parked states belong to compose a place-place pair. After all parked vehicles in the same area share their place-place pairs in the local data center, the local data center obtains the vehicle table and the place routing table. Moreover, according to the vehicles owners' will, the parking time in each parked place from the activity model can be shared with the local data center. The related system files of the local data center are copied in each of the local nodes.

Single vehicle node of LUV only needs to load the subnetwork files of the places where belong to its activity model. When visiting a place where not in the parked place list, as a foreign node of the place, the node has not necessary loaded the local subnetwork information and does not join in the local data center. The foreign nodes share their parked place lists with the local data center there and exchange the local IoT data with the local data center.

Place file (*P_List*) includes the LUV places information. It comes from the system initialization process of place identification. Every vehicle node copies the Place file. With the Place file and their current GPS position, urban vehicles can detect the LUV place where they are parking. Datalink file (*L_List*) includes the information about the data links between places (place-place pairs) of LUV. It comes from the parked place lists of vehicles. From the Datalink file, the place-place pairs of vehicles are stored in an incidence matrix, PV. The adjacency matrix of the matrixes PV, NL, used to describe the number of data links between any two places. Matrix NL is used to calculate the shortest path between any two places, and the matrix PM stores the results. By binarization of the element values of the matrix PM, a place connection matrix, PC, is generated to indicate the place-place connections in LUV. When element PCij is 1, LUV provides data servers between the place i and the place j; otherwise, the data servers of LUV are not available between the two places.

E. IoT DATA SUBMISSION & DELIVERY

In a place of LUV, the IoT devices exchange data with the local data center by building the wireless connection with the nearby local nodes. The local data center delivers the IoT data to the vehicles which are ready to leave, and receive the local data from the vehicles arrived, as shown in figure.14.

The wireless connection range of the vehicle is longer than the one of the IoT devices. Although the communication between the vehicle and the IoT device is limited by the communication range of the IoT devices, the distance of the V2V communication can reach farther. Such an extended communication range make it easy for building the vehicle-based local data center to cover the entire area.

The local data center connects the other local nodes and the foreign nodes in the area for data delivery and submission.

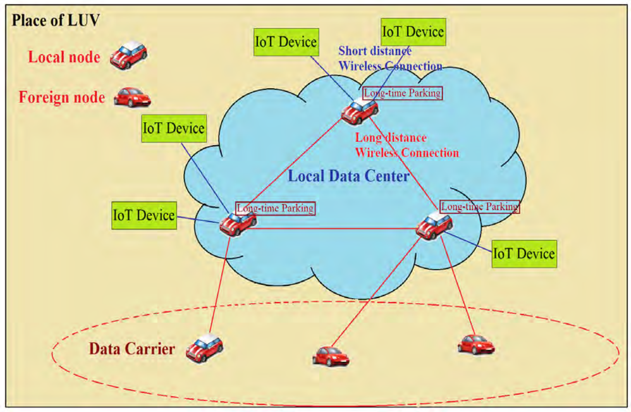


FIGURE 14. **The wireless connections of IoT-vehicle and V-V in the LUV place.**

It needs to choose a proper vehicle node as data-carrier among the connected vehicles when it prepares to deliver IoT data to the destination place. We devise the following approach for the local data center choosing the carrier for the IoT data to the destination place. We present it in the following steps.

a) *The vehicle nodes of the local data center build the wireless connections with the nearby parked vehicles.*
b) *The local data center sends to the connected vehicles the requests for their visiting probabilities of the data destination place. Then the vehicles send back their responses which contain the calculation results based on their history traces.*
c) *The local data center selects the vehicle which has the highest visiting probability as the candidate node.*
d) *If the likelihood of the candidate is lower than the threshold, the local data center wait for a new vehicle to submit data. Otherwise, it sends the IoT data to the candidate.*
e) *When the data carrier arrives at the data destination, it uploads the data to the local data center of the destination.*
f) *If the local data center has not found a suitable data-carrying vehicle during the current waiting time cycle, it decides whether to into another waiting time cycle or switch in the relay mode.*
g) *In relay mode, depending on the place routing table, which can be obtained from the data links matrix NL, a suitable relay place is chosen. Then the data is sent to a parked vehicle which will visit the relay place. When the vehicle carrying the data arrives at the relay place, it submits the data to the local data center, and the center tries to find other vehicles to send the data to the destination place.*

F. ENERGY CONSUMPTIONS

Nowadays, most of the vehicles have not to be equipped with high-power batteries yet. The vehicular terminals have to be turned off after the vehicles stop. However, according to the electrification trend of the automotive industry, it can

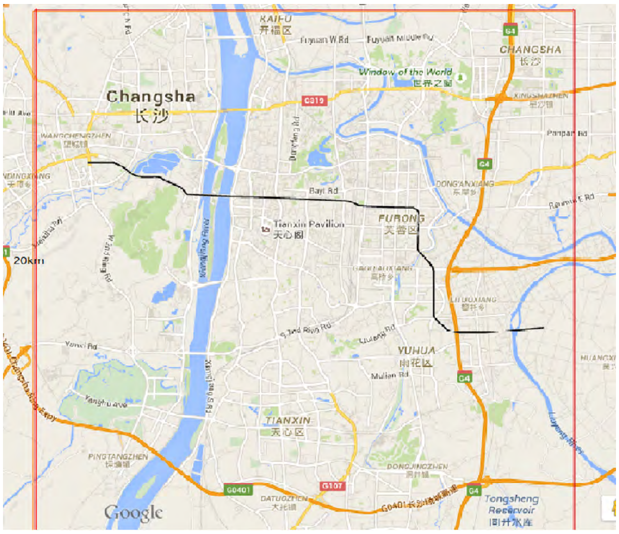


FIGURE 15. Targeted LUV area in Changsha, China.

be expected that the next generation of vehicles will have the ability to power vehicle-mounted terminals around the clock. Therefore, during its parking, by peer to peer wireless connections, the vehicles can collect data from the IoT devices, submit data to the servers, and interact with the other parked vehicles. Although the energy scarcity of the IoT devices and sensors limits their wireless communication range, the high density of parked vehicles in the urban area ensures that the IoT devices can be connected to the nearby vehicles within their short wireless communication range.

G. INCENTIVES FOR CAR OWNERS

Given the fact that the car owners pay for the power spending of the vehicular terminals or other indirect costs, the incentives for them to participate in the urban vehicular network is necessary.

Considering that urban vehicular networks for collecting public IOT data are non-profit networks served smart cities, municipalities should provide public resources for the incentives. The system records the total amount of historical IOT data transmitted by each vehicle. Similar to airline mileage conversion points, assign points to the vehicles according to their total amounts of data transmission. When the vehicles have earned enough points, they gain priority access to specific public resources in the city by redeeming their points. For instance, the car owners can obtain one-time free parking in a particular area with its points. In the case, it sounds be attractive to car owners due to the lack of parking resources in modern cities.

VI. NETWORK EVALUATION

In this section, the set of real-life data on privately-owned vehicles is utilized for the network evaluation. To reduce the amount of calculation, we set the urban area scope of Changsha from N28.07 to N28.27, E112.90 to E113.10, which is shown in Figure 15.

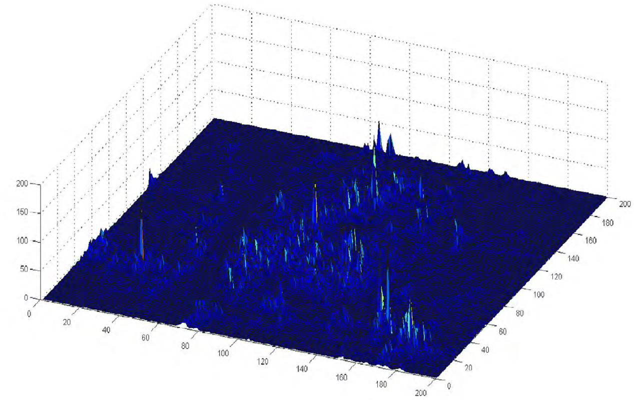


FIGURE 16. LUV place distribution with Thr_density = 0 (Changsha, China).

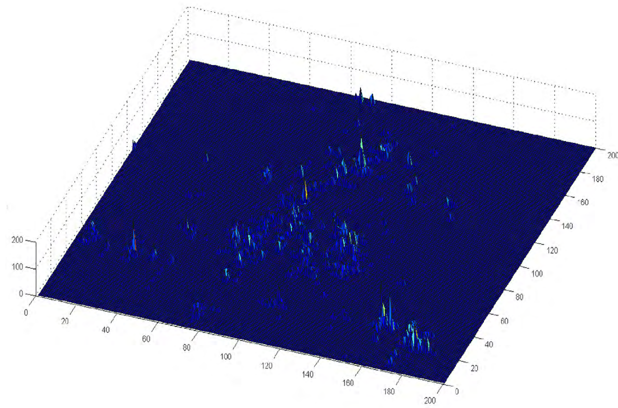


FIGURE 17. LUV place distribution with Thr_density = 19 (Changsha, China).

A. PLACE IDENTIFICATION IN LUV

Based on our approach for place identification in LUV [25], we first divide the area, as shown in Figure.16, which is 20km by 20km into 40,000 units, each with an area of 100m by 100m. The total residence time of the vehicle in each unit is accumulated over the entire observing period and stored in a matrix, MS. The row of the matrix MS is vehicle ID, and column is unit ID. The total number of vehicles which belongs to the area is 5190. As the number of units is 40000, the size of matrix MS is 5190 by 40000. A vehicle's total residence time in each unit is stored in the elements of the matrix.

If we merely set the *Thr_time* (the predefined threshold of residing time in a unit) to 1, the matrix becomes sparse, and the non-null rate is about 0.43% (89870 non-empty elements). If we set *Thr_density* (the vehicle density threshold of unit) to 19, we identify 873 LUV places which account for 2.18% of the total area. Figure.16 is the intermediate result where parameter *Thr_density* has not been applied yet. Figure.17 illustrates the distribution of the 873 LUV places.

LUV allows messages exchange within places to ensure relative stable and reliable connection among vehicles. At the same time, the places where the vehicles will visit next

	177	239	223	403
269	358	531	339	
	275	691	330	
		488		
		203		

FIGURE 18. LUV places(4km*4km) and the local vehicles.

Place:	1	2	3	4	5	6	7	8	9	10	11	12	13
1		6	1	2	3	5	4	2	1	2	0	3	0
2	8		8	5	0	8	10	3	2	3	2	2	0
3	1	7		8	1	3	7	6	2	5	2	3	0
4	3	8	20		3	7	7	4	2	2	5	6	2
5	4	1	3	2		14	5	0	3	8	0	2	0
6	13	9	3	5	7		18	2	9	9	3	6	2
7	10	19	14	6	7	28		35	5	21	12	9	4
8	3	5	9	5	0	3	16		3	7	14	4	1
9	2	1	4	2	3	14	2	3		18	2	9	4
10	6	7	10	7	6	16	26	10	23		41	31	5
11	1	2	3	4	1	1	8	9	3	24		8	1
12	5	5	6	3	6	7	13	6	11	16	12		17
13	1	0	0	1	2	1	2	0	2	8	1	11	

FIGURE 19. Data links of LUV places.

provide a likely route for message delivery. By using simple threshold methods, we can identify concentrated vehicle places in an urban area that can effectively serve as places in LUV. Given the multiple tunable parameters, the method can effectively control the size of the place, the number of vehicles, and the required frequency or residence time in a place.

B. DATA LINKS OF PLACES IN LUV

Considering the vehicle nodes which from real data set is much less than the real quantity of vehicles in Changsha, China, we zoom in the place area to acquire enough local nodes in one place to study the connectivity of places in LUV over the real-life data. We divide the area into 25 units, each with an area of 4km by 4km. By our threshold-based approach for identifying the places, we identified 13 LUV places, and 4526 local vehicles belong to these places. The result as below:

The residential place and the visited place of one vehicle build a data link between the two places. Therefore, we obtain the data links between these places as shown in figure.19.

The total number of data links between places fits logarithmic normal distribution, as shown in Figure.20.

The connectivity rate in LUV is 96.15%. It indicates that urban vehicles can provide data service between places and the network performance depends on the amounts of vehicle nodes. As the network scale increases, the performance of

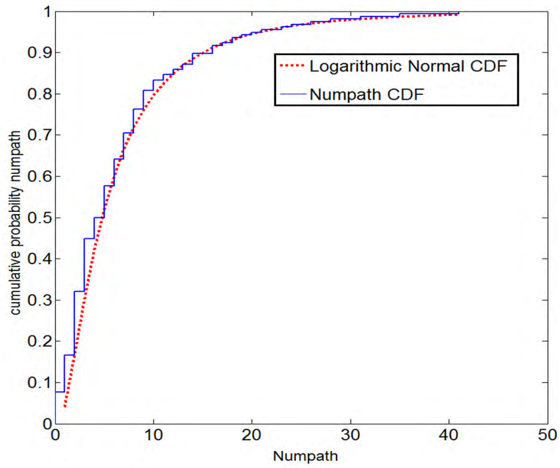


FIGURE 20. The total number of data links between places CDF.

LUV will improve as well. Therefore, within LUV, the millions of urban vehicles can provide low-cost local data access services to hundreds of millions of IoT devices in the city.

VII. CONCLUSION

With the rapid development of Smart City related technologies such as IoT, big data, and wireless communication, the demands for the IOT data in smart cities have snowballed. In addition to the 5G network, smart cities can also deploy the vehicle-based DTN network as the non-real-time IoT data transmission platform to enhance the IoT data collection capability and reducing the IoT data collection costs in smart cities.

According to the essential features of urban vehicles from the analysis of the trace data of the private cars, we propose a location-based urban vehicle network (LUV) for IoT data transmission in the urban environment. In LUV, the parked vehicles, the daily activity patterns of vehicles and the existing urban vehicular transportation network are fully utilized to reduce the network complexity and improve the network scalability for practical application.

Due to prohibitive costs and policy constraints, it is challenging to deploy sufficient sensors in the urban environment for acquiring public sensing data. Compared to sensor deployment, collecting sensing data generated by these sensors distributed in urban is even harder. Though message exchange only happens between parked vehicles within places, private vehicles can perform various other tasks like sensing urban environment from multi-sensor mounted on vehicle or gathering sensing data from the roadside sensors when moving. Relying on LUV, the urban sensing servers deployed in particular places can collect these sensing data on vehicles.

The works of this paper are based on our previous work over the last four years [24]–[26]. There are still many aspects for us to study in the future. For the parking places make a profound impact on vehicles, the studies on characteristics of parking place, such as the type of place, the spatial scale of places, and the vehicle capacity of parking place, is yet to come. Based on the place studies, we can utilize the features of places to design the details of LUV further. Furthermore, in this paper, there are two activity models of vehicles that are divided merely by workday and rest day which vehicle can detect by themselves. In the further work, the activity model of vehicle can be refined to improve the accuracy of predicting the current behavior of the vehicle.

REFERENCES

[1] H. Schaffers, N. Komninos, M. Pallot, B. Trousse, M. Nilsson, and A. Oliveira, ''Smart cities and the future Internet: Towards cooperation frameworks for open innovation,'' in *The Future Internet Assembly* (Lecture Notes in Computer Science), vol. 6656. Berlin, Germany: Springer-Verlag, 2011, pp. 431–446.

[2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, ''Internet of Things for smart cities,'' *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[3] R. Zhu, M. Ma, L. Liu, and S. Mao, ''A special section in IEEE Access: Cooperative and intelligent sensing,'' *IEEE Access*, vol. 5, no. 1, pp. 27824–27826, 2017.

[4] D. Puthal, S. Nepal, R. Ranjan, and J. Chen, ''Threats to networking cloud and edge datacenters in the Internet of Things,'' *IEEE Cloud Comput.*, vol. 3, no. 3, pp. 64–71, May/Jun. 2016.

[5] M. Agiwal, A. Roy, and N. Saxena, ''Next generation 5G wireless networks: A comprehensive survey,'' *E Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.

[6] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. Ten Brink, I. Gaspar, N. Michailow, A. Festag, and L. Mendes, ''5GNOW: Non-orthogonal, asynchronous wave-forms for future mobile applications,'' *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, 2014.

[7] M. Strohbach, H. Ziekow, V. Gazis, and N. Akiva, ''Towards a big data analytics framework for IoT and smart city applications,'' in *Modeling and Processing for Next-Generation Big-Data Technologies*. Berlin, Germany: Springer, 2015, pp. 257–282.

[8] A. Galati, *Delay Tolerant Network*. Saarbrucken, Germany: Lambert Academic, 2010, pp. 2–15.

[9] S. Mumtaz, L.-L. Yang, C. Wang, F. Adachi, and N. Ali, ''Smart-device-to-smart-device communications: Part 2,'' *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 18–19, Jun. 2014.

[10] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin, and S. Chen, ''Vehicular fog computing: A viewpoint of vehicles as the infrastructures,'' *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3860–3873, Jun. 2016.

[11] M. Gerla and L. Kleinrock, ''Vehicular networks and the future of the mobile Internet,'' *Comput. Netw. J.*, vol. 55, no. 2, pp. 457–469, 2011.

[12] P. Hui, J. Crowcroft, and E. Yoneki, ''BUBBLE rap: Social-based forwarding in delay-tolerant networks,'' *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1576–1589, Nov. 2011.

[13] Y. Mao, ''PASA: Passive broadcast for smartphone ad-hoc networks,'' in *Proc. IEEE 23rd Int. Conf. Comput. Commun. Netw.*, Shanghai, China, Aug. 2014, pp. 1–8.

[14] C. Liu, C. Chigan, and C. Gao, ''Compressive sensing based data collection in VANETs,'' in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 1756–1761.

[15] G. Grassi, D. Pesavento, G. Pau, R. Vuyyuru, R. Wakikawa, and L. Zhang, ''VANET via named data networking,'' in *Proc. INFOCOM WKSHPS*, Toronto, ON, Canada, Apr./May 2014, pp. 410–415.

[16] N. B. Truong, G. M. Lee, and Y. Ghamri-Doudane, ''Software defined networking-based vehicular adhoc network with fog computing,'' in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, May 2015, pp. 1202–1207.

[17] F. Qu, Z. Wu, F.-Y. Wang, and W. Cho, ''A security and privacy review of VANETs,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 2985–2996, Dec. 2015.

[18] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, ''A comprehensive survey on vehicular ad hoc network,'' *J. Netw. Comput. Appl.*, vol. 37, pp. 380–392, Jan. 2014.

[19] S. Yousefi, M. S. Mousavi, and M. Fathy, ''Vehicular ad hoc networks (VANETs): Challenges and perspectives,'' in *Proc. Int. Conf. ITS Telecommun.*, Chengdu, China, Jun. 2006, pp. 761–766.

[20] Q. Su, K. Jaffres-Runser, G. Jakllari, and C. Poulliat, ''An efficient content delivery infrastructure leveraging the public transportation network,'' in *Proc. MSWiM*, Valletta, Malta, Nov. 2016, pp. 338–347.

[21] D. Shoup, ''The high cost of free parking, American planning associatio,'' *J. Planning Edu. Res.*, vol. 17, no. 1, pp. 3–20, 2005.

[22] M. Haklay and P. Weber, ''OpenStreetMap: User-generated street maps,'' *IEEE Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct. 2008.

[23] *ArcGIS Desktop, Release 10.2.1*, Environ. Syst. Res. Inst., Redlands, CA, USA, 2014.

[24] H. Li, Y. Liu, S. Zhou, and H. Xu, ''Toward location based vehicular networks in urban environments,'' in *Proc. 11th Int. Conf. Mobile Syst. Pervasive Comput.*, Niagara Falls, ON, Canada, Aug. 2014, pp. 442–449.

[25] H. Li, Y. Liu, Y. Sun, and R. Yu, ''Place identification in location based urban VANETs,'' in *Proc. Proc. Int. Conf. Mobile Ad Hoc Sensor Syst.*, Dallas, TX, USA, Oct. 2015, pp. 621–626.

[26] H. Li, Y. Liu, and S. Zhou, ''Power in silence: Revealing the stationarity of private-owned vehicle in urban vehicular ad hoc networks,'' in *Proc. ACM Symp. Develop. Anal. Intell. Veh. Netw. Appl.*, Valletta, Malta, 2016, pp. 131–138.

[27] H. Teng, Y. Liu, A. Liu, N. N. Xiong, Z. Cai, T. Wang, X. Liu, ''A novel code data dissemination scheme for Internet of Things through mobile vehicle of smart cities,'' *Future Gener. Comput. Syst.*, vol. 94, pp. 351–367, May 2019.

[28] T. Li, K. Ota, T. Wang, X. Li, Z. Cai, and A. Liu, ''Optimizing the coverage via the UAVs with lower costs for information-centric Internet of Things,'' *IEEE Access*, vol. 7, pp. 15292–15309, 2019.

[29] H. Teng, W. Liu, T. Wang, A. Liu, X. Liu, and S. Zhang, ''A cost-efficient greedy code dissemination scheme through vehicle to sensing devices (V2SD) communication in smart city,'' *IEEE Access*, vol. 7, pp. 16675–16694, 2019.

[30] B. Huang, W. Liu, T. Wang, X. Li, H. Song, and A. Liu, ''Deployment optimization for data centers in vehicular networks,'' *IEEE Access*, vol. 7, no. 1, pp. 20644–20663, 2019.

HENG LI was born in Changsha, Hunan, China, in 1977. He received the M.S. degree in computer science from Hunan University, Changsha, in 2001, where he is currently pursuing the Ph.D. degree in computer science. From 2017 to 2018, he was a Visiting Scholar with the University of Texas at Arlington, USA. His main research interests include vehicle networks, wireless sensor networks, urban sensing, and smart city.

YONGHE LIU received the B.S. and M.S. degrees from Tsinghua University, in 1998 and 1999, respectively, and the Ph.D. degree from Rice University, in 2004. He is currently an Associate Professor with the Department of Computer Science and Engineering, University of Texas at Arlington, USA. His research interests include wireless networking, sensor networks, security, and system integration.

ZHENG QIN received the Ph.D. degree from Chongqing University, in 2001. He is currently a Professor and the Associate Dean of the College of Computer Science and Electronic Engineering, Hunan University, China. He is the Director of the Key Laboratory of Big Data Research and Application, Hunan. His research interests include big data, information security, and smart city.

HUIGUI RONG received the Ph.D. degree in information management from Wuhan University, China, in 2010. He is currently a Full Associate Professor with Computer Science Department, Hunan University. He is also the System Analyst and the Senior Project Manager certified by The Ministry of Industry and Information. He has published over 20 research papers in international conferences and journals, such as ACM CIKM, IEEE-TBD, and the IEEE-HPCC. His major research interests include data mining, sensing networks, and cloud computing. He is a member of ACM.

QIN LIU was born in Changsha, Hunan, China, in 1982. She received the Ph.D. degree from Central South University, China, in 2012. She is currently an Associate Professor with the College of Computer Science and Electronic Engineering, Hunan University. Her major research interests include cloud computing security, big data security, and privacy protection in social networks.