Modeling and Characterization of Urban Vehicular Mobility using Web Cameras

Gautam S. Thakur[‡§], Pan Hui[‡] and Ahmed Helmy[§]
[§]CISE, University of Florida, Gainesville, Florida
[‡]Deutsche Telekom Laboratories, Berlin, Germany.
Email: gsthakur@cise.ufl.edu, pan.hui@telekom.de, helmy@cise.ufl.edu

***Abstract*—Realistic design and evaluation of vehicular mobility has been particularly challenging due to a lack of large-scale real-world measurements in the research community. Current mobility models and simulators rely on artificial scenarios, random connectivity, and use small and biased samples. In this paper, we perform a combined study to learn the structure and connectivity of urban streets and modeling and characterization of vehicular traffic densities on them. Our dataset is a collection of 154 thousand routes and 12 million vehicular mobility images from 730 online web cameras located in four different cities. First, our study shows that driving routes and visiting locations of cities demonstrate power law distribution, indicating a planned or recently designed road infrastructure. Second, we represent cities by network graphs in which nodes are camera locations and edges are urban streets that connect the nodes. Such representation exhibits small world properties with short path lengths and large clustering coefficient. Third, traffic densities show 80% temporal correlation during several hours of a day. Finally, modeling these densities against known theoretical distributions show less than 5% deviation for Log-logistic and Gamma distribution. We believe this work will provide a much-needed contribution to the research community for *realistic* and *data-driven* design and evaluation of vehicular networks.**

I. Introduction

Research in the area of vehicular networks has increased dramatically in the recent years. With the proliferation of mobile networking technologies and their integration with the automobile industry, various forms of vehicular networks are being realized. These networks include vehicle-to-vehicle [1], vehicle-to-roadside [2], and vehicle-to-roadside-to-vehicle architectures. Realistic modeling, simulation, and informed design of such networks face several challenges, mainly due to the lack of two main factors: (i) Underlying topology, and (ii) Large-scale community-wide libraries of vehicular data measurement.

Topological understanding is important in accurately modeling vehicular mobility. It involves intersections, roads and their connectivity. It comes as no surprise those topological constraints like speed limits, direction, etc. impact traffic congestion, density, scenario generation, and mobility, which in turn affect the performance of any routing protocol [3]. Thus, for accurate evaluation of a vehicular network, one should have a better knowledge of its topology.

Earlier studies in mobility modeling have clearly established a direct link between *vehicular density* distribution and the performance of vehicular networks' primitives and mechanisms, including broadcast and geocast protocols[4]. Initial efforts to capture realistic vehicular density distributions were limited by a lack of availability of sensed vehicular data[5]. Hence, there is a need to collect and conduct vehicular density modeling using larger scale and more comprehensive datasets. Furthermore, commonly used assumptions, such as Exponential distribution [4], [6], have been used to derive many theories and conduct several analyses, the validity of which bears further investigation.

In this paper, we first study the structure and connectivity of the urban streets of four major cities and second, perform a large scale systematic analysis and modeling of vehicular traffic density distributions. Earlier studies have been done to study the centrality of major cities from a different perspective [7]. Recently, the departments of transportation of several cities have started to deploy traffic web-cameras to critical intersections and highways to study traffic patterns. We collected the geo-graphical coordinates of these locations and created a graph $G(V, E)$ as mentioned before. To avoid the limitations of sensed vehicular data, we also utilize the existing global infrastructure of tens of thousands of video cameras providing a continuous stream of street images from dozens of cities around the world. Millions of images captured from publicly available traffic web cameras are processed using a novel density estimation algorithm to build an extensive measurement dataset of spatio-temporal vehicular traffic densities. We perform a comprehensive analysis of this data to study the structure and connectivity of urban streets and characterize the underlying statistical patterns of traffic density at individual intersections and highways of major cities. In short, we find (i) Visits to locations follow power-law distribution; (ii) Temporal correlations of vehicular traffic density for individual camera locations are nearly 80% between consecutive hours, but go down to 30% for a 3-4 hours lag difference. We also investigate traffic modeling by comparing the frequencies observed in the empirical density distribution to the expected frequencies of the theoretical distribution. The result of this activity shows that the empirical values closely follow (less than 3% deviation on KS-test) 'Log-logistic' and 'Gamma' distributions. The contributions of this work are:

- We provide, to the best of our knowledge, by far the largest and most extensive dataset for future vehicular network analysis. This potentially addresses a severe shortage of such datasets in the community.

TABLE I
GLOBAL WEBCAM DATASETS

City	# Cameras	Duration	# Images	# Routes
Connecticut	274	21/Nov/10- 20/Jan/11	7.2 million	74,801
London	181	11/Oct/10 - 22/Nov/10	1 million	32,580
Sydney	67	11/Oct/10 - 05/Dec/10	2.0 million	4,422
Toronto	208	21/Nov/10 - 20/Jan/11	1.8 million	43,055
Total	**730**	**-**	**12 million**	**154,858**

- We introduce a new and more practical way to look into urban street networks based on driving routes. A network graph of routes and locations depict small world properties.
- We establish Log-logistic and Gamma distributions as the most suitable fits for modeling vehicular traffic density.

We believe our work helps 'fill a gap' between the expected and realized necessity for the 'design and evaluation of realistic and data-driven models' for future generations of vehicular networks.

In section II, we discuss measurements and pre-processing, in section III we discuss topological analysis of urban street maps for four different cities. In section IV, we statistically model vehicular traffic and characterize it. In section V, we show the impact and challenges on the vehicular networks. Finally, we conclude in section VI with future work.

II. Measurements and pre-processing

In this section, we give details of the collected geo-location information of cameras used for the analysis of topological properties and recorded vehicular images captured from these cameras to model and characterize the vehicular traffic.

A. Topology Data of Camera Locations

Traffic web cameras are deployed on key intersections and highways within every city. Thus, we can assume these locations are representative of urban streets of that city. We start by recording cameras' geo-coordinates and location information to study the topological properties of urban streets. This includes latitude and longitude, zipcode, state, directional view, and camera location. Later on, in section III-A we use this data to create a network graph of urban streets.

B. Vehicular imagery data collection

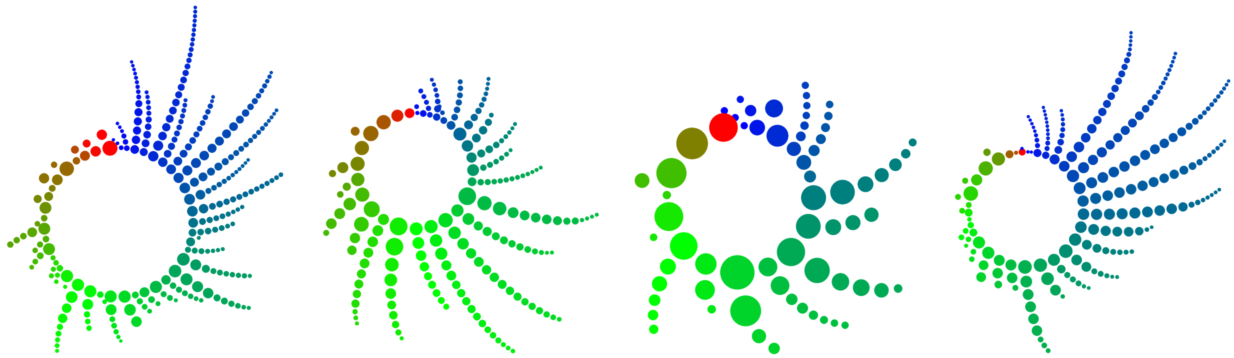
There are thousands, if not millions, of outdoor cameras currently connected to the Internet, which are placed by governments, companies, conservation societies, national parks, universities, and private citizens. We view the connected global network of webcams as a highly versatile platform, enabling an untapped potential to monitor global trends or changes in the flow of the city, and providing large-scale data to realistically model vehicular, or even human mobility. Majority of these webcams are deployed by a city's Department of Transportations (DoT). Although, it's not possible to deploy them at every intersection or highway, nonetheless they are strategically placed to capture the traffic trends at critical locations. At regular intervals of time, they capture still pictures of on-going road traffic and send them in the form of feeds to the DoTs media server. We have developed crawlers that collect vehicular mobility traces from these servers. For the purpose of this study, we have also made agreements with DoTs with large city coverage to collect these vehicular imagery data for several months. We cover cities in North America, Europe, Asia, and Australia. Overall (here only four out of ten cities are presented with details in Table-I), we download 15 Gigabytes of imagery data per day from over 2700 traffic web cameras, with an overall dataset of 7.5 Terabytes containing around 125 million images. Since these cameras provide better imagery during the daytime, we limit our study to only those hours. Table-I gives a high level statistics of the dataset used in this study. Each city has a different number of deployed cameras and a different interval time that captures images. We believe our study is comprehensive and reflects major trends in traffic movement. Next, we discuss the algorithm to extract traffic information from images.

C. Traffic Information Extraction

We aim to estimate traffic density(d) on roads considering the number of vehicles or pedestrians crossing the road. We have a sequence of images captured by webcams. Considering our problem, we have to be able to separate information we need, e.g., number of vehicles and pedestrians from the background image, which is normally road and buildings in that image. We apply background subtraction techniques [8] and dynamic filters [9] to extract relevant traffic information. One could then use regular object detection techniques to identify and count number of vehicles in the high pass filtered image. However, this is computationally expensive and unnecessary. As an alternative, we count the number of pixels and sum their values (with a value higher than a certain threshold). The cumulative sum is represented as *traffic density*. This is much faster than detecting and counting objects in an image [10]. At the same time, it is more effective, because we are looking at the percentage of the street (road), covered by vehicles (as an indicator of how crowded the street is), rather than number of vehicles. The number of vehicles is not a good indicator of crowdedness, as a long vehicle may introduce more traffic than a small one. Second, our method overcomes the issues that object detection face in case of severe congestion. Thus, our use of vehicular traffic densities instead of vehicle counts is inspired from fluid dynamics, for analyzing trends in traffic distribution instead of tracking individual vehicles for studying micro-mobility. In order to ascertain the accuracy of our algorithm, we compare our results with that of ground truths recorded of total counted cars. It shows 87.2% correlation between them [11]. Due to limited space, we ask interested readers to also refer our extended report [10].

III. Analysis of Topological Properties

In this section, we examine degree distribution and small world properties of four different cities and states in order to study the structure and connectivity of their urban street network. We represent this network by a graph $G = (V, E)$, where V is the set of camera locations as nodes and E is a set of driving segments as edges, inter-connecting the nodes of set V of the network graph G. The degree k_i of a node



(a) Connecticut (b) London (c) Sydney (d) Toronto

Fig. 1. Radial axis layout of urban street location show a two dimensional representation of degree distribution. Each radiating axis (spar) is grouped by similar degree distribution $\langle k \rangle$. The clockwise varying of color dots from blue to red mean increase in the value of $\langle k \rangle$ for nodes. The varying sizes of dots are respective average weighted degree $\langle \hat{k} \rangle$. A larger size dots mean more weight. For Connecticut and Sydney the distribution of $\langle \hat{k} \rangle$ show power law distribution with $2 < \alpha < 3$ and for Toronto show the same for its $\langle k \rangle$ distribution. London is an old city, with lots of small streets and intersections (hence more than one ways to reach destination), show no power law distribution ($\alpha = 3.5$).

TABLE II

PARAMETER AND DETAILS

G, $\hat{G}$	Unweighted, Weighted graph	k_m	Largest degree	d	traffic density	P	Exponential Distribution
V	Total number of nodes (camera locations)	$\langle \hat{k} \rangle$	Average weighted degree per node	L	Characteristic path length	M	Gamma Distribution
E	Total number of edges (streets)	$\hat{k_m}$	Largest weighted degree	C	Clustering coefficient	LL	Log logistic Distribution
k	Degree of a node	α	Power law exponent	L_r	Random graph characteristic path length	N	Normal Distribution
$\langle k \rangle$	Average degree per node	ρ	Correlation coefficient	C_r	Random graph clustering coefficient	W	Weibull Distribution

TABLE III

REPORT FOR DEGREE DISTRIBUTION, POWER LAW EXPONENT, PATH LENGTHS, CLUSTERING COEFFICIENTS AND MODEL FITTING OF EMPIRICAL DATA

City	V	E	$\langle k \rangle$	k_m	$\langle \hat{k} \rangle$	$\hat{k_m}$	$\alpha(G)$	$\alpha(\hat{G})$	L	L_r	C	C_r	Dominant distribution as Best Fits (By Ranking)			Dominant distributions as Best Fits (By % Deviation KS-Test)	
													1^{st} Best Fit	2^{nd} Best Fit	3^{rd} Best Fit	≤3%	≤5%
Connecticut	274	2128	15	51	8994	32119	3.5	2.41	3.6	2.33	0.52	0.05	LL[87%]	M[11%]	P[0.5%]	LL[62%], M[15%], W[3%]	LL[94%], M[44%], W[19%]
London	181	1252	13	32	2180	7089	3.5	3.5	3.04	2.23	0.5	0.07	LL[42%]	M[39%]	W[16%]	M[34%], LL[34%], W[10%], N[0.5%]	LL[82%], M[70%], W[47%], N[7%]
Sydney	67	319	9.5	26	322	985	3.5	2.98	2.73	2.05	0.56	0.137	LL [62%]	M[32%]	N[2%]	LL[88%], M[61%], W[4%], N[2%]	LL[98%], M[88%], W[44%], N[18%]
Toronto	208	1128	10	44	7435	21323	2.8	3.5	5.02	2.5	0.6	0.05	M[46%]	W[31%]	LL[21%]	M[75%], W[58%], LL[34%]	M[94%], W[88%], LL[87%], P[4%], N[1%]

i in G is the number of edges incident with the node. In an undirected and unweighted G (weight = 1), the degree can be written in terms of the adjacency matrix A as

$$k_i = \sum_{j=1}^{n} A_{ij}$$

The weighted degree of each node i in undirected graph, $\hat{G}$ is $\hat{k}$, and can be written in terms of the adjacency matrix W as

$$\hat{k}_i = \sum_{j=1}^{n} W_{ij}$$

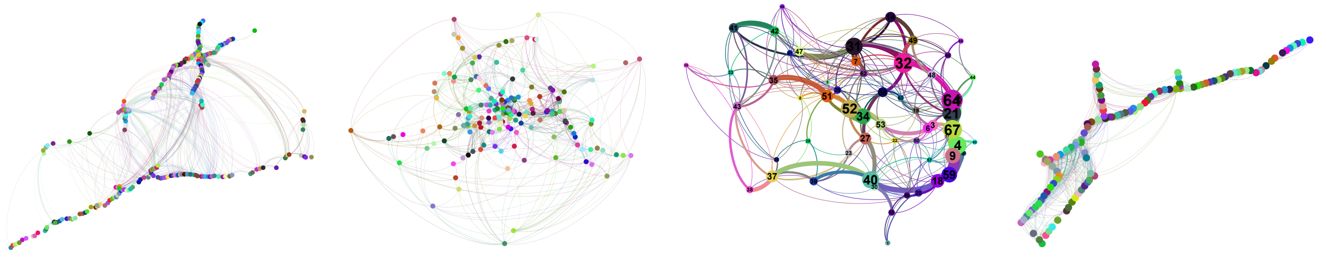
Next, we explain the graph generation process of urban street network of cities using Google Maps, and then analyze their degree distribution for unweighted and weighted cases, and finally examine their small world properties.

A. Network of Urban Streets

Segment: A path (edge in the network graph) that directly connects two locations.

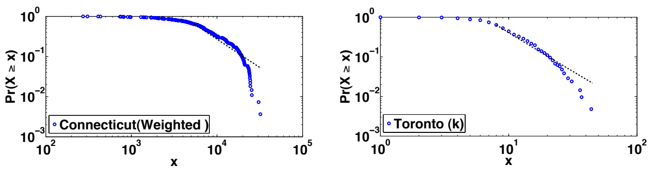
Route: A route is a set of segments appearing in the order of the increasing distance from source to destination.

As mentioned before, we generate a network graph with nodes acting as camera locations and set of edges as driving routes connecting them. These routes are basically the driving segments between the locations, as returned by Google Maps API. In order to generate this graph, we start by taking the geo-coordinates of a pair of camera locations, calculating the driving distance between them. Next, we check for a possible subset of other camera locations that might lie en-route from source to destination. All such locations are inserted in order of their occurrences and connected through intermediate segments (as edges). For example, driving from New York to San Francisco, we drive through Iowa City, Omaha, Salt Lake City, and Sacramento in that order. If no such locations exist, the source and destination are directly connected by an edge. We iterate this process for all pairs of camera location, total $V*(V-1)$. While doing so, we also maintain a *between* count for the traversed edges, connecting two nearest locations. Every time they are traversed, we increase their respective between count by one. Finally, we end up with a weighted graph showing the locations and segments taken with laters' weight equal to the times they have appeared on any route. For example, the most frequently taken edge segment has the largest between count. For simplicity, we assume each street allow bi-directional traffic. In general, locations are connected by a maximum of 3-5 roadways, in our case we ease this assumption for investigating the connectivity patterns. In Fig.2(c), we show the example of a weighted graph of Sydney generated with 67 camera locations. The underlying process of generating this network graph is computationally expensive [12], nonetheless it has many benefits: (i) We use Google Maps API to calculate all possible routes and intersections, today, anyone planning to travel, accesses maps via Google or like services. (ii) We are assured that resultant graph filters-out



(a) Connecticut (b) London (c) Sydney (d) Toronto

Fig. 2. A geo-laid network of urban streets of four cities is shown. The nodes are camera locations and edges are routes connecting these locations. Out of the four, we have shown Sydney with its weighted degree $\langle\hat{k}\rangle$ network.



(a) Connecticut ($\langle\hat{k}\rangle$, $\alpha = 2.41$) (b) Toronto ($\langle k\rangle$, $\alpha = 2.8$)

Fig. 3. A CDF $P(x)$ and its maximum likelihood power law fits for two locations.

non-frequent routes, which help to better explain the cause of traffic congestion on frequently taken routes and locations. (iii) The recommendations can be made to generate dynamic routes from diverting the traffic on already congested segments. There are several variations of *between* count.

- *Unweighted Graph*: We baseline the *between* score of a street (edge) to one if it has ever appeared.
- *Weighted graph by distance:* The weights on the edges can be replaced by actual driving distance. Thus, recommendation can be made in case shortest path is available between a pair of source and destination.
- *Weighted graph by distance and between score*: The weights on the edges are a combination of distance and *between* score. It helps to discover overhead and congested segments in the network.

In this paper, we focus our study on the structure and connectivity of the urban streets using unweighted and weighted (as discussed above, by counting the occurrences) cases.

B. Analysis of Degree Distribution: Unweighted Case

We study the number of connections that camera locations have with one another. It helps to analyze the connectivity and probability of taking alternate routes to the destinations. A network analysis of Toronto's cameras shows their degree distribution is powerlawed. As evident in Fig.1, the radial axis graph of Toronto clearly shows only 2-3 locations with large degree distribution. In Fig.3, an exponent value of $\alpha = 2.8$ shows that the maximum likelihood power law fit for the degrees of its few locations have a long right-side tail of values that are above the mean. A value x obeys a power law if it is drawn from a probability distribution p as:

$$p(x) \propto x^{-\alpha}$$

α a constant known as exponent parameter. The usual value of exponent lies in the range $2 < \alpha < 3$ with some exceptions.

Above results indicate that such locations have much higher connectivity with rest of the one-hop far locations. On the other hand, if they are removed from the network, average path length will increase, and location pairs will become disconnected and traveling between them will become impossible.

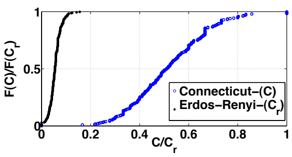
C. Analysis of Degree Distribution: Weighted Case

The weighted degree of a camera location is calculated based on the frequency of its connected edges that have appeared between any pair of source and destination. Using Google Maps, we have calculated shortest path between all pairs of locations, and the list of locations that are on en route. Therefore, it is possible that few locations have been traversed more often than other, making them the most visited locations. In our study, we find the locations belonging to Sydney and Connecticut demonstrate a power-law distribution, which means they create an hour glass model, making most of the traffic to pass through few locations. It also makes them susceptible to traffic congestion and closures. In Fig.1, we see the distribution of node sizes representing weighted degrees for Connecticut and Sydney, with power-law exponent $\alpha = 2.41$ and 2.98 respectively in Table-III. In Fig.3, a cumulative distribution function for maximum likelihood fit for Connecticut is shown.

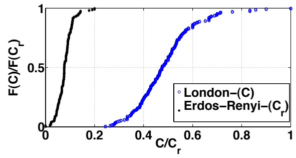
Thus, while Toronto is skewed on connectivity, Connecticut and Sydney are skewed on visiting same locations again and again. We can say that traffic congestion in Toronto appears because of geometry of locations, while for Connecticut and Sydney its because specific routes have been traversed. The city of London appears to have even distribution for both metrics, as evident in its radial layout in Fig.1 and Table-III. We can say that London network is more resilient than other cities, with lot of small and inter-connecting streets, exhibiting properties of an historic city's growth.

D. Small World Analysis

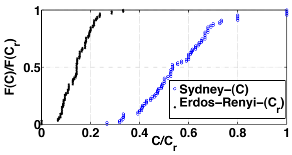
We investigate that network of urban streets of all four cities clearly exhibit small world properties. In general, a network with small world should have small average path length ($L < 6$) and large clustering coefficient ($0.4 < C < 1$). We make a basis for a fair comparison, by using Erdos-Renyi G(n,M) [13] model to generate a random graph for each city separately, with $n = V$ and $M = E$. To ascertain our structure, we examine C against C_r for each city - for C to be extreme in that distribution and greater than the ninety-fifth percentile.



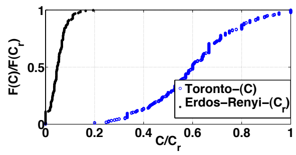
(a) Connecticut



(b) London

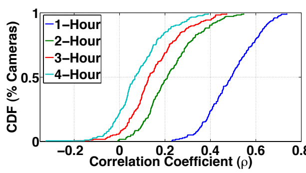


(c) Sydney

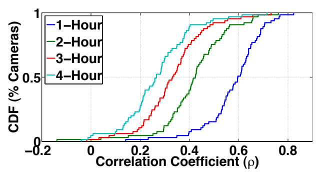


(d) Toronto

Fig. 4. A CDF of C and C_r show large values of clustering coefficient for cities network to random graphs, indicating priors' network structure exhibiting small world properties.



(a) London



(b) Sydney

Fig. 5. CDF showing correlation of traffic densities between hour differences of the day.

Next, we calculate the average path length (L) and clustering coefficient (C) of the four cities' networks and compared them against L_r and C_r of random graphs respectively, as shown in Table-III. We find that networks of all four cities have small average path length ($\forall L < 6$) and large clustering coefficient ($\forall C \to 1$, $C \gg C_r$), with Toronto having largest value of clustering, $C = 0.6$. The CDFs of clustering coefficient are shown in Fig. 4 - (i) a quick convergence for the random graph, indicating very small clustering coefficient values of C_r (ii) All values of $C > 0.3$, and large gaps in curves indicating network of cities exhibiting strong small world properties.

IV. Traffic Modeling and Characterization

We studied connectivity of urban streets, now we turn to model and characterize the traffic density on these streets. We will see, how the traffic is correlated with itself for several hours of the day. Later, we will use known theoretical distributions to model traffic densities.

A. Traffic Flow Auto-Correlation

We investigate correlation coefficients (ρ) to measure the degree to which traffic from a camera is linearly associated with itself for 42 days. In our case, we are using this to analyze the change in traffic densities. We analyze the correlations for 1-4 hour lags for each camera against itself during 12 hours of the day, from 7 AM to 6 PM. For example, we investigate what the correlation is between the traffic at 7 AM and 8 AM (1-hour lag), 1 PM and 3 PM (2-hour lag) etc. In Fig.5, we show CDF for various hours lag of the day. For the city of Sydney the hourly traffic change is highly correlated, almost 80% of cameras' next hour traffic is 70% correlated to its current hour. For next two hours from the current, the traffic

(a) $d = 2023, 0.28$

(b) $d = 5400, 0.55$

(c) $d = 9230, 0.93$

Fig. 6. Traffic with varying densities[(a)low/(b)medium/(c)high] is shown. The first value is the result of background subtraction and later is the normalized value.

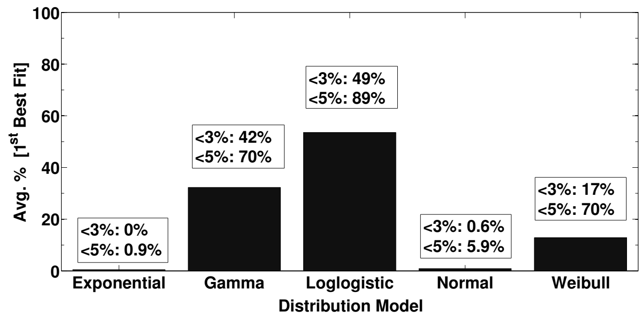
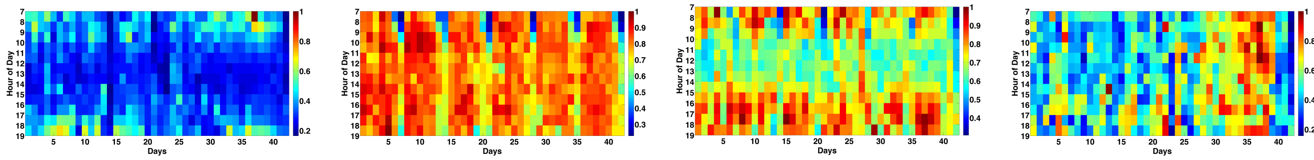


Fig. 8. Best fits for four cities. The values in the box show deviation.

for 80% of the cameras are only 50% or less correlated. And around 60% cameras have only 30% correlation for a time lag of 3-4 hours. While in case of the city of London, the next hour traffic density for 80% cameras is close to 60% correlated to the current hour. It goes further down to 30% for next two hours and around 15-20% for a 3-4 hour difference. *Thus, vehicular traffic has temporal richness, which in-turn affects the mobility of vehicles and therefore, have an impact on the performance of routing protocols [3].* Similar trends are observed in other cities, but omitted here for brevity.

B. Traffic Modeling and Characterization

Here, we focus on modeling the arrival process of traffic (traffic density value) in equal intervals of times against known theoretical distributions. In Fig.6, we show three traffic scenarios of varying intensities from low to fully congested location, captured by the density parameter(d).The objective of this study is to help understand the underlying statistical patterns. We already filtered these for the purpose of showing the maturity in our studies to select and identify the statistical patterns without much deviations. To ensure the validity, we also performed several goodness of fit test using Maximum likelihood estimation (MLE) and Kolmogorov-Smirnov test to measure average deviation and compare the values in the density vector to known distribution. We systematically model individual locations' empirical traffic density distribution against well known theoretical ones. In Fig.7, we show four different locations with changing traffic densities during 12 hours for 42 days. *This result invalidates a general notion of 'rush hours' that traffic is relatively higher only during morning and early evening hours.* In order to match, we use five theoretical distributions: Exponential, Gamma, Log-Logistic, Normal and Weibull. We find that traffic at individual cameras can vary a lot, but in general log-logistic, Gamma and Weibull distribution can capture some of the key features. We rank these distributions (based on KS-tests) in Table-



(a) Low Traffic (b) High Traffic (c) Periodic Traffic (d) Random Traffic

Fig. 7. Several variations in traffic densities across 42 days traffic monitoring are shown. Fig-(a) show relatively mild traffic during various hours of the day, while (b) show high traffic recording for the full trace periods. In Fig-(c) we find a regularity patterns during the morning and evening hours when the traffic is relatively higher to afternoon hours. A random traffic pattern is recorded in the last.

II, with three out of four cities' individual locations have log-logistic as the 1^{st} best fit, while Toronto has Gamma distribution. In Table-II, we show dominant distributions at 3% and 5% deviation using the KS-test. In Fig.8, results show the dominance of distributions for all the locations from all four cities. Overall, the empirical data closely matches log-logistic and Gamma distributions. We find that even on cities' aggregate traffic levels, the log-logistic distribution provides a good estimate of empirical data. These results are realistic scenarios, and can be used as input for simulators to evaluate the performance of vehicular routing protocols.

V. Future Application to Vehicular Networks

The experience gained from the analysis and modeling of traffic densities potentially aids in future design and evaluation of vehicular networks. Today, most of the simulation tools input generic or random scenarios and disregard the challenges brought by mobility in vehicular networks [3], [14], [15]. In our case, the benefit of urban street analysis and large dataset of realistic traces, and its modeling results prove to be very helpful in developing rich scenarios for testing protocols, network dynamics, scalability of traffic, topology size estimation, and the analysis of traffic patterns. The data-driven realistic simulation tools and mobility models are necessary for accurate evaluation of vehicular routing protocols and services. However, our analysis shows that traffic characterization and communication network analysis tools (e.g., ns2) are separately developed and therefore lack a tight integration [14], [16]. Our gathering and analyzing real traffic data can aid in identifying metrics (e.g., spatio-temporal density) to develop data driven mobility models and simulators. The unique challenges (e.g., high speed, intermittent connectivity) in inter-vehicle [1] and car-to-roadside [2] communication require the development of robust and efficient routing protocols. We can use the cameras' geo-coordinates and their traffic density distribution to develop and test new performance metrics and protocols. In the future, we aim to focus on developing realistic and data-driven models. We have also plan to make this dataset available to the research community and extend our existing work to study centrality measure for all the cities.

VI. Conclusion

We know topological properties (like directions and lanes) impact the movement of vehicular traffic on roads. In this paper, first we have discussed an approach to create a network of urban streets from driving directions and second use of vehicular imagery snapshot images from freely available online cameras for traffic analysis. Our results have shown that for three cities, during several trips, visits to their locations and streets exhibit a power-law distribution. While the fourth city (London) has maze of roads indicating a historical growth and network resiliency. A temporal auto-correlation of 80% is evident for traffic densities in those three cities for consecutive hours (1-2 hours) of the day. In London, high and variable traffic pattern. We have observed a stable periodicity of traffic density for many days (42 days) corresponding to weekdays and weekends. This is an important result, and can aid in developing futuristic traffic prediction models. We have also found that empirical traffic densities closely follow (with less than 3% deviation) theoretical distributions like Log-logistic and Weibull. We believe our work will provide much needed contribution to the research community.

References

[1] J. Blum and et. al., "Challenges of intervehicle ad hoc networks," *ITS, IEEE Tran. on*, 2004.

[2] J. Jiru and D. Eilers, "Car to roadside communication using ieee 802.11p technology," *Industrial Ethernet Book Issue*, 2010.

[3] F. Bai, N. Sadagopan, and A. Helmy, "The important framework for analyzing the impact of mobility on performance of routing protocols for adhoc networks." *Ad Hoc Networks*, vol. 1, no. 4, pp. 383–403, 2003.

[4] F. Bai and B. Krishnamachari, "Spatio temporal variations of vehicle traffic in vanets: facts and implications," in *Vanet*, 2009.

[5] J. Yeo and et. al., "Crawdad: a community resource for archiving wireless data at dartmouth," *Sigcomm CCR*, 2006.

[6] N. Wisitpongphan and et. al., "Routing in sparse vehicular ad hoc wireless networks," *IEEE Comm.*, 2007.

[7] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: A dual approach," *Physica A: Statistical Mechanics and its Applications*, vol. 369, no. 2, pp. 853–866, Sep. 2006.

[8] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in *IEEE CVPR*, 1999.

[9] A. Elgammal and et. al., "Non-parametric model for background subtraction," in *Computer Vision*, ser. LNCS, 2000.

[10] G. S. Thakur, M. Ali, P. Hui, and A. Helmy, "Comparing background subtraction algorithms and method of car counting," *CoRR*, 2012.

[11] G. S. Thakur, P. Hui, H. Ketabdar, and A. Helmy, "Towards realistic vehicular network modeling using planet-scale public webcams," *CoRR*, vol. abs/1105.4151, 2011.

[12] G. S. Thakur. (2012) Generating Urban Street Network using Google Maps API. [Online]. Available: http://www.cise.ufl.edu/∼gsthakur/globalsense.shtml

[13] P. Erdos and A. Renyi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, pp. 17–61, 1960.

[14] R. Stanica, E. Chaput, and A. Beylot, "Simulation of vehicular ad-hoc networks: Challenges, review of tools and recommendations," *Computer Networks*, 2011.

[15] S. Yousefi and et. al., "Vehicular ad hoc networks (vanets): Challenges and perspectives," 2006.

[16] M. Piórkowski and et. al., "Trans: realistic joint traffic and network simulator for vanets," *Sigmobile CCR*, 2008.