Is it Possible to Find Social Properties in Vehicular Networks?
Felipe Domingos da Cunha, Aline Carneiro Viana, Raquel A. F. Mini,
Antonio A. F. Loureiro

To cite this version:

HAL Id: hal-00981127
https://hal.inria.fr/hal-00981127
Submitted on 21 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Is it Possible to Find Social Properties in Vehicular Networks?

Felipe D. Cunha*,†, Aline Carneiro Vianna‡, Raquel A. F. Mini†, Antonio A.F. Loureiro*

*Federal University of Minas Gerais, Brazil - {fdcunha, loureiro}@dcc.ufmg.br
†Pontifical Catholic University of Minas Gerais, Brazil - raquelmini@pucminas.br
‡INRIA, France - aline.viana@inria.fr

Abstract—Everyday, vehicles transit in a city and along their trajectories, they encounter other vehicles. The frequency of these encounters is influenced by many factors, such as: vehicle speed, destinations, traffic conditions, and the period of the day. However, these factors are justified by the public roads limits and the driver's behavior. The people present daily routines and similar behaviors that have a great impact in the daily traffic evolution. In this work, we present a numerical analysis of real and realistic data sets that describe the mobility of a set of vehicles. Social metrics are computed, and the results obtained are compared to random graphs in the direction to verify if vehicular network presents a social behavior. Finally, we discuss new social perspectives in vehicular networks.

I. INTRODUCTION

Vehicular Ad Hoc Networks (VANETs) are a special type of Ad Hoc networks, formed by vehicles with processing and wireless communication ability, traveling on streets or highways. Commonly, the vehicles can communicate directly or by the use of a roadside unit (RSU). Through this infrastructure, vehicles can access network services and obtain data from other networks, such as the Internet. Due to this nature, VANETs can be established in different environments such as in urban centers and highways [1], [2].

VANETs are proposed to assist the drivers in the traffic, to avoid collisions with sending alerts, to provide entertainment to passengers, and to contribute to create an intelligent traffic system. An important application of this network is to alert the drivers about road conditions and imminent crashes. Recent studies show that 60% of car crashes could be avoided if the drivers were warned just few seconds before the collision [2]. In this context, the interaction among vehicles can strongly contribute to broadcast an alert message reducing the crash ratio [3]. Another example of application of VANETs is the assistance to drivers to find the best route. In this scenario, it is important that VANETs monitor and predict the traffic conditions on streets, to avoid traffic jams decreasing the trip time [4].

The communication in VANETs is highly influenced by the vehicles density and mobility patterns. Vehicles move and stop according to the speed limits, and directions imposed by the public roads. Furthermore, the period of the day can influence the vehicles mobility. In rush hours, the traffic can be slower and traffic jams can force the driver to stop the car, or to change the route. This is a scenario where the vehicles density can be higher. On the other hand, weekends and early mornings are considered quiet traffic scenarios, with low vehicles density. These density variations in different hours reinforce the characteristic of dynamic topology of VANETs, making the communication a challenging task in vehicular networks.

Also, the mobility is influenced by the behavior of the driver and its routine. During the weekends, destinations such as camping, malls, churches are chosen to leisure and entertainment, and are frequently visited. However, on weekdays people tend to repeat their paths at the same time and to the same destination, such as: school, work, university, restaurant, etc. During their trajectories, vehicles encounter others vehicles, pass in the same streets, and suffers the same traffic conditions. These presented features suggest the study of the vehicular mobility under a social perspective, in order to understand better its evolution, and to apply the social concepts improving the services and the connectivity in VANETs.

Social analysis can be used to monitoring the traffic evolution during the day aiming to the understanding of the human routines, the similar trajectories, and the rush times. In literature, many works present the mobility analysis of vehicular networks in order to finding social features and to characterize this mobility [5], [6], [7]. However, these works are based only in analytical mobility models, and the evaluation was performed in short periods. A initial work [8] presents a social analysis of one realistic trace with duration of one day. In this work, we focus on the investigation of the following question: What are the characteristics that a vehicular network needs to present in direction of to prove the existence a social behavior? Following we summarize the contributions of this paper, which intents to answer this question:

- A social analysis of two traces that describe features of different group of vehicles.
- A comparison of the results with random graphs, with the objective to increase the validity of the analysis.
- One definition about the social metrics that can be used to evaluate the mobility behavior in VANETs.
- A discussion about the different perspectives to study the social mobility in VANETs, and what the factors can influence this analysis.

The remainder of this paper is organized as follows. In Section II, we define the main statements used to build the
temporal and random graphs, and the evaluated metrics. In Section III, we present the traces data set, and the quantitative results according to each calculated metric. In Section IV, we discuss about the factors that may affect the social mobility analysis in vehicular networks. Finally, in Section V, we present the final conclusions and future perspectives.

II. MODELING

This section details the modeling used to create the temporal graph, the characteristics of the random graphs used as reference in the results and the definition of the evaluated metrics used in the social analysis.

A. Temporal Graph

The model considers the encounter among vehicles, i.e., when two vehicles are within communication range of each other. In our evaluation, we consider communication range of 100 m among vehicles, according to the protocol 802.11p and when two vehicles are within communication range of the whole trace into discrete time slots of duration \( t = 1h \). Since vehicle traces usually describe fast changes in the topology, short time slots are recommended to better capture traffic changes. Thus, with the data in each slot we build a temporal graph. The temporal graph at the time \( t \) is an undirected, and can be formally defined as a graph \( G(t) = (V, E) \), where \( V \) represents the set with all vehicles \( v_i \) and \( E \) represents the set of edges \( e_{ij} \). In \( G(t) \), an edge \( e_{ij}(t) \) exists between the vehicles \( v_i \) and \( v_j \) during time \( t \), with \( i \neq j \). Each metric is evaluated hourly, considering the temporal graph \( G(t) \) that takes into account all the encounters in the period \( t \). We aggregate the encounters that happen during the slot window, and build the graph using the Growing Time Window technique [9]. For each trace, 24 graphs \( G(t) \) are generated describing the vehicles encounters during the day.

B. Random Graphs

A random graph \( G^R \) is a graph obtained from a set of vertex and its respective degree distribution. The criteria to add an edge between two vertex in the graph \( G^R \) is made in a random way, based in probabilities models. In this work, the model used to add edges uses urn algorithm that create random structures [10]. In this algorithm, the first step is to insert \( d_i \) balls in an urn, marked with the identifier \( i \) of the vertex. The amount of balls of each vertex is equal to the degree of the vertex \( i \). The total amount of balls in the insertion process in the urn, matches with the degree distribution of the original graph \( G \). Since this step, two balls \( b_i \) and \( b_j \) are randomly selected and their identifiers are compared. If \( i \neq j \) and there are no edge between these vertex, an edge \( e_{ij} \) is added. The selection step is made until there is no ball inside the urn, or there is no possibility to add more edges in the graph, according to the insertion criteria. In this work, for each temporal graph \( G(t) \) we generated its random graph \( G^R \), where the metrics are computed.

C. Metrics Description

We use social metrics to characterize the vehicles mobility in metropolitan areas. We divide the social metrics in macroscopic and microscopic group, according to the classification proposed in [11]. The macroscopic metrics represent measures of the network global state. We use the Distance, the Diameter, and the Edge Persistence as macroscopic metrics. The microscopic metrics define individual values for the vehicle representing the behavior of a unique vehicle. We select the Node Degree, the Cluster Coefficient and the Closeness Centrality as microscopic metrics.

III. RESULTS

In this Section we present the trace evaluation results. In the Section III-A, we describe details about the two traces evaluated. The Section III-B presents the results to the macroscopic metrics, and the Section III-C the results to the microscopic metrics.

A. Traces Evaluated

In this work, we analyze two public data set traces that describe the vehicle mobility along a city. The first trace picked is the Zurich’s Trace [12], a data set realistic based in mathematical models, and generated in a mobility simulator that considers the census information about the Zurich’s city. This trace represents the traffic evolution over the day including rush times and daybreak. The Zurich’s trace describes the trajectory of 260,000 vehicles over the day, and the Figure 1-(a) represents the number of vehicles in each day time. It is possible to see two periods with the higher number of vehicles that we consider as the rush times: 6 – 9am and 2 – 5pm, and the two hours with the major number of vehicles: 173,470 vehicles at 7am and 132,633 vehicles at 3pm. The analysis of this trace takes into account the traffic in all the day, with focus on the rush times.

The second trace picked is the San Francisco’s Trace. This is a real data set with GPS information of 551 taxis in San Francisco’s city, during four weeks [13]. This information is captured from GPS devices in each vehicle, at each minute, over its trajectory. In the Figure 1-(b), it is showed the traffic evolution over the day. It is possible to see that during the daybreak there is a decrease in the number of taxis in service. At 6am, this behavior changes and the value increases again, just in time which people start their workday. Thus, since this hour, 400 taxis in average keep in service during the day, and we note a constant behavior in the graphic.

B. Macroscopic Metrics

Distance: To analyze the distance between two vertices in the graph, it is important to remark that in this work we consider the distance as the number of edges, which belongs to the path connecting these vertices. In a vehicular network, this metric can represent the existence of common interest between two
vehicles, i.e., if two vehicles cross the same routes or visit near places, the distance between them tends to be short. Otherwise, vehicles that do not pass the same streets in the same period, or vehicles that transit in distinct regions tend to present long distances. The Figure 2 presents the hourly average for the distance in the two traces.

To the Zurich’s trace, Figure 2-(a), we attribute the distance 0 to the hours where there is no traffic. However, in hours with high traffic it is possible to see one variation in the distance value. The number of vehicles in transit increase and the encounters probability increase too. Despite this traffic variation, the distance does not suffer a huge increase. This happened because in rush hours, in downtown there are many traffic jams that increase the encounters probability contributing to keep a short average distance. Also, we note that the random graphs $G^R$ present a little difference. These are graphs generated in random way, where the edge between two vertices obeys a probability; differently of the Zurich’s graphs where the existence of an edge respects the direction limits of the public roads. Thus, in the random graphs the edge insertion can create shortcuts in the graph. In addiction, in the Figure 2-(b) we can see the average distance to San Francisco’s trace. As expected, the distance the trace presents a small distance during all the day, and the curve perform a constant behavior. In the same mode, the random graph $G^R$ presents the same behavior; all the vehicles can reach one each other with just one hop. It is happening because this trace particularly describes taxis mobility, due to the mobility pattern and the high vehicles density, during one hour, many taxis encounter one each other and the graph is very dense. As the random graph is generated from the San Francisco’s graph, the density is equal and the distance keeps close, just a little variation in the daybreak.

**Density:** The Density is one metric that quantifies the number of encounters happen during one hour. The baseline of this metric is the complete graph, and we will compare the number of edges of the graph $G(t)$ with it. In the Figure 4 we can see the daily Density evolution. The first observation when we compare the two traces confirms the mobility pattern of the two different data sets. Taxis transit all the day with varied targets, and the particular vehicles have defined routines, with schedule and destinations defined. In the Figure 4-(a), we can see the Density to Zurich’s trace, and it is possible to verify a low density. This trace has many vehicles crossing a large area with distinct destinations, and because these characteristics it presents a small numbers of encounters. In another hand, the Density of the San Francisco’s Trace showed in the Figure 4-(b), and we can observe a higher density. Due the small number of taxis and the continuous movements per hour, the density remains around 80% during the day, and 60% in daybreak in according to the traffic evolution; many number of taxis in transit, major is the encounter probability. Also, many of encounters of taxis can have happen in the moment to leave and to return to the parking, considering taxis that belong to the same enterprise.
**Edge Persistence:** In this work, we compute the edge persistence as the number of times one edge repeat during the time, i.e., whether an encounter between two vehicles repeat. Thus, we evaluate the edged persistence in accordance to the format of each trace. To the Zurich’s trace we calculate the persistence in the two rush times, where the concentration of vehicles is elevated. We compare the encounters in the start of the day (6 – 9am) with the end of the day (2 – 5pm). In the San Francisco’s trace we calculate the persistence in each day hour during one week, and we capture the encounters repetition during the week. Giving the traffic and encounter distribution in Zurich we analyze 340,000 edges in each period, and the persistence for this trace is 43%, i.e., nearly 150,000 edges repeated, which indicate similar routines among this vehicles. In the same context, we analyze the San Francisco’s trace and we found an edge persistence of 75%. In spite of taxis did different trajectory; it is necessary to return to initial point to wait for new passengers, which increase the encounter probability. When we look to both results, it is interest to observe that this is a good metric to explain the presence of similar routines and behavior in the graph.

**C. Microscopic Metrics**

**Node Degree:** When we evaluate the node degree in a VANET, it is important to remark that the traffic density suffers variation during the day, and the number of encounters that one vehicle can perform is influenced by the quantify of vehicle in transit, e.g., as more vehicles are circulating, higher is the probability of encounters. In the Figure 5 we can see the daily Node Degree evolution. It is possible to observe in the Figure 5-(a) the average, maximum and median value to the degree. During the rush hours, it is the moment that the degree reaches the high value (1.355 at 7am.), which can indicate that is the time where vehicles make long trips or cross high density regions. The same behavior can be observed at 3pm with high degree. Also, we can see the average and median value that presents the value to 200 encounters to the rush times. These are acceptable values, which with more traffic is expected that more encounters happen. Many of the social networks are considered scale free networks, because presents a power law distribution to the node degree. In this context, we evaluate the distribution for the two hours with major concentration of vehicles that we can see in the Figures 5-(b) and 5-(c). In these hours, 7am and 3pm respectively, there are in transit approximately 100,000 vehicles and the graphics showed that the degree distribution follows a power law distribution. Thus, we can consider the Zurich’s graphs as scale free: many vehicles with low degree and low vehicles with high degree.

The evolution for the San Francisco’s Trace is presented in the Figure 6. The values to the average, median and maximum values can be seen in Figure 6-(a). It is noted that during 5 – 7am the degree presents a reduction in according to reduction of the traffic density. Furthermore, according to the density, the average of degree is high, with values nearby 450 encounters. With this behavior, we observe that considering the time-window for one hour, practically all the vehicles encounter the others vehicles. To verify the existence of the power law distribution, we select three diverse day moments: 6am, 12am and 6pm (Figures 6-(b), 6-(c) e 6-(d)). We can note that, with the day progresses, the encounter probability increases, mainly in the 6pm where the degree of all nodes is between $10^2$ and $10^3$. Beside this, it is no possible to found the power law in the pattern mobility. Then, we should not consider this trace as a scale free network.

**Cluster Coefficient:** The cluster coefficient is a so important metric to analyze in the context of social network, because it indicates the presence of communities and users with common interests. In the context of vehicular network, a cluster can represent a group of vehicles that transit in the same direction, or that pass frequently in the same place, i.e., a group of vehicles that have common point in their trajectories. In the traditional paper [14], Watts and Strogatz present a discussion about the small worlds networks and comparison with random graphics. It is showed that the cluster coefficient of a social graph is one order of magnitude of a random graph $G^R$.

We can see the cluster coefficient to the Zurich’s trace and the respective comparison with random graphs in the Figures 7-(a) and 7-(b). When we look to the maximum value, it is possible to note that there are vehicles in graph with 100% of their neighbors connected. We can attribute this result to vehicles that are in traffic jams areas, in downtown, mainly in areas that receives traffic flows from many adjacent roads. In these regions, there is a strong tendency to form vehicles cluster that share the same routes and traffic conditions. In the same graphic, we can observe that the cluster coefficient average is 50%, which indicates that a half of the vehicles neighbors are connected in the rush hours. The comparison with $G^R$ can be seen in the Figure 7-(b), and we found a difference of one order of magnitude; the $G^R$ presents a low cluster coefficient compared with Zurich’s graphs. This result shows that social behaviors can be found in Zurich’s trace.

![Figure 7. Daily Cluster Coefficient evolution to the Zurich’s Trace.](image-url)
When we compare the cluster coefficient with the random graphs, Figure 8-(b), it is possible to see the same behavior with the two graphs, with very similar values. Thereat, with the high value of the cluster coefficient, it is possible to verify only a clique in the graph, i.e., a unique group with vehicles that have the same behavior. Taxis are vehicles that made random trips during all the day, i.e., they pick up clients in random locations in the city. In this sense, no routine is expected.

Contact happened in this slot window generates an edge in the analyzed temporal graph. This could explain the high density and node degree found, and consequently the complete graph. When we compare the cluster coefficient with the random graphs, Figure 8-(b), it is possible to see the same behavior with the two graphs, with very similar values. Thereat, with the high value of the cluster coefficient, it is possible to verify only a clique in the graph, i.e., a unique group with vehicles that have the same behavior. Taxis are vehicles that made random trips during all the day, i.e., they pick up clients in random locations in the city. In this sense, no routine is expected.

**Closeness Centrality:** The evolution of the Closeness Centrality is significant to understand how close the vertices are in the graph. In VANETs, it is likely to evaluate this metric to appreciate the position of the vehicles in accordance to others and to know the mobility patterns of the vehicles. The Figures 9-(a) and 9-(b) show the results for the two traces analyzed. In the Figure 9-(a), we can verify that the Closeness modify according to the traffic evolution, with more traffic more is the number of vehicles to consider. However, it is possible to verify an increase in the value, indicating that the vehicles are more central. This happens because in the rush times there is a tendency to vehicles cross the downtown and encounter more vehicles. Also, more vehicles in transit contribute to generate more similar routes, which explain this small increase. Additionally, the random graphs present a similar behavior, with a little reduction with does not represent. Likewise, in the Figure 9-(b), we observe the values to closeness in the San Francisco’s trace. Again, it is verified that the random graphs present a similar behavior to the values for closeness, and then a high value. According to the average Distance for this trace, around to one hop, it is possible to certify, with this metric, that the vertices are so close.

**IV. Vehicular Social Networks Analysis**

In this work, we have analyzed two vehicular traces under a social perspective, aiming to characterize it according to metrics that bring us to comprehend the vehicular mobility and the driver’s routines. Therefore, to do this evaluation we chose some mobility parameters, and we do important assumptions to characterize the encounter among the vehicles (Section II-A). Through of these parameters, we create graphs and compute the metrics. However, when we look to the results we found social characteristics in the Zurich’s trace, with strong directions to characterize this network as a Social Network. Except, for the San Francisco’s trace we observed a high density and a huge number of encounters that collaborate to create a very dense graph, similar a complete graph. Thus,
we look to the literature in order to understand the parameters that can affect the encounters improving this analysis.

Many of parameters can influence the existence of one encounter, such as: vehicle speed, communication range, time slot window, and minimal time to consider as one encounter. These are important parameters to define in the analysis when the focus is evaluate the encounter probability in the Vehicular Network. In the work [9], the authors present a discussion about how create the social graph with the contacts in a Delay Tolerant Network. They argue how the selection of the parameters and the contact aggregation can affect to the social graph. Beyond that the authors show a performance evaluation of routing protocols considering different techniques of contact aggregation to create the social graph. Another example of social analysis is presented in [7]. In this work the authors presented a study of the social interactions of the vehicles. They create graphs, and aggregate the encounters by the growing time window technique with a threshold to define one relationship between two vehicles. The authors show the presence of the social properties in two real traces.

Although, the correct definition of the parameters to find social features in VANETs, it is interest to take account the scenario of the vehicles transit: urban, rural and highways. According to the scenario, the vehicles speed can suffer many fluctuations, which impacts the encounter ratio. Another important feature to consider is the nature of vehicles. The taxis cross the whole city and perform random trajectories, without fixed time and trip duration. The buses transit the same routes and they obey the fixed schedule to perform the routes. In other hand, common people use their vehicles to perform predetermined trajectories according to their routines. Therefore, in order to characterize and to execute a good analysis, it is important take into account the vehicle nature, its mobility pattern in order to define the evaluated parameters according to the characteristics of this vehicle.

V. CONCLUSION

In this work, we presented the social analysis of two vehicular traces describing the mobility of vehicles in the metropolitan area. We analyzed the Zurich’s trace and San Francisco’s trace that have different duration and routines. Considering the model that defines an encounter, the analyses of the Zurich showed the presence of the small world phenomenon. Also, in this trace the degree distribution follows the power law, which describes this network as scale free network. Moreover, it presents features that indicate the presence of communities and people with similar interest. In the San Francisco’s trace, we cannot find social properties. This is a graph with only one component and high edges density, which justifies the results (taxi mobility pattern and high encounters density). Besides all, according to results of the analysis we had shown that is possible to find social properties in VANETs, considering vehicles nature. As future works, we intend to extend the evaluation with the study of others social metrics and improve the analysis with variation of the parameters that define whether there is an encounter between two vehicles. Beyond that, we intend to perform a spatial and temporal metrics computation aiming to capture the difference of values in different points in the city. Also, we intend to design dissemination protocols aware of these social metrics evaluated.

REFERENCES