The Dynamics of Vehicular Networks in Large-Scale Urban Environments

(Invited Paper)

Nicholas Loulloudes, George Pallis, Marios D. Dikaiakos
Department of Computer Science, University of Cyprus
Nicosia, CY1678, Cyprus
{loulloudes.n, gpallis, mdd}@cs.ucy.ac.cy

***Abstract*—Vehicular Ad hoc NETworks (VANETs) have emerged as a platform to support Intelligent Transportation Applications. A key to the development of protocols and services for IVC lies in the knowledge of the topological characteristics of the VANET communication graph. This article aims provide a "higher order" knowledge in the time-evolving dynamics of vehicular networks in large-scale urban environments through complex network analysis. In addition, we examine VANET dynamics from a spatial viewpoint in order to identify where/why in the underlying road network, particular network phenomena occur. We make reference and correlate our findings with the main communication paradigms employed in VANETs, and discuss whether the emerging network phenomena can indeed support (and to what extend) these paradigms.**

I. Introduction

Inter-vehicle communication (IVC) has emerged as a promising field of research and development, where advances in wireless and mobile ad-hoc networks, global positioning systems and sensor technologies can be collectively applied to automobiles and result to great market potential. The concept of employing wireless communications in vehicles dates back to the '80s, however with the recent proliferation of wireless hardware that support communication standards designed exclusively for the vehicular environment (IEEE 802.11p), wide deployment of Vehicular Ad Hoc NETworks (VANETs) is one step closer to realization.

The development of IVC and its utilization in the deployment of numerous Intelligent Transportation System (ITS) applications has been the target of research investigations in the recent literature, as well as of industrial projects run by large government-sponsored consortia (Car-2-Car Communication - USA and CVIS - EU [1, 2]), along with field trials such as the Connected Vehicle Safety Pilot Program [3]. Due to the nature of vehicular mobility, VANETs are characterized by highly dynamic topologies, which are frequently prone to disconnection and severe fragmentation. Inevitably, these inherent characteristics put at stake the availability and quality of services of those intelligent applications for which VANETs were envisioned. Therefore, the establishment of robust VANETs that could effectively support communications and applications on a large geographical scale, remains an open challenge.

A. Motivation

The motivation for this work stems from the fact that the study of the structural properties of large, real-world, dynamic systems (i.e. Internet topology, biological and online social networks) has led to crucial observations having significant influence in information and social sciences. For instance, the discovery of power-laws in the Internet topology by Faloutsos et al.[4], revolutionized the computer science field by enabling the design of efficient information routing protocols that capture the Internet topology characteristics. Consequently, the study of VANET communications from the *complex network science* aspect can provide crucial insights to their extremely dynamic nature. A vehicular network is a very challenging domain given that it combines opportunistic ad hoc communications among vehicles and, oftentimes, fixed road-side units. Despite the fact that it presents similarities with traditional MANETs, mobile nodes (i.e. vehicles) have ample energy and are highly mobile in spite of spatial constrains imposed by the underlying road network topology.

Grasping the dynamics of the underlying VANET and its inherent capacity to support message dissemination using different communication paradigms, is fundamental during the design and deployment of any innovative ITS application. For instance, the Virtual Traffic Lights (VTL) application [5], which recently spun out into a startup by a team of researchers, aims to provide optimized traffic light signaling by leveraging IVC capabilities that will soon be available in modern vehicles. Synchronization phases will be decided in a distributed and self-organized manner through the exchange of periodic *1-hop broadcasts* of traffic signal message.

To this end, VTL and other ITS application stakeholders would like answers to questions such as: "Does the topology support message spread with minimal re-broadcasts reducing latency and network overhead?". In case that *multi-hop communication* is preferred, "How far in the topology can data packets travel?", and "Which are the 'best' nodes that can be en-charged with the data forwarding process, and where are they located?". On the other hand, if relying on *geocasting* for pushing alert messages to remote geographic areas, "What is the spatial coverage of the VANET and its components", and "What are the characteristics of bridge nodes that can be assigned with the task of message ferrying?". Finally, in *information hovering*, "does the VANET incorporate structures that facilitate temporal information maintenance?"

All these questions and many more require knowledge of the characteristics of the *VANET communication graph*, wherein vehicles correspond to vertices and wireless ad hoc communication links correspond to edges. In [6], we set out to gain and provide "first-order" knowledge about the topological features of small urban-based VANET graphs. Later in [7, 8] we extended our initial study and aimed to provide a "higher-order" knowledge by studying the time-evolving topological characteristics of the VANET, along with how such knowledge could eventually be utilized towards improving the performance of VANET routing protocols. More recently, [9] reported new and interesting insights by examining the structure of large-scale, urban VANETs. This work focused on the topological characteristics of the network and made key observations regarding the overall level and quality of inter-vehicle connectivity.

However, given that a VANET graph is realized from the movement and opportunistic interactions of vehicles while being driven on the underlying road network, it can be "seen" from the viewpoint of real-world networks that are well-embedded in low-dimension geometric structures(e.g power grid on surface of the earth). Consequently, besides its temporal evolution, the VANET has a spatial dimension and should thus be studied accordingly. With a concurrent understanding of the spatio-temporal characteristics of V2V communication, researchers will ultimately gain the necessary knowledge that will enable the design and implementation of data dissemination protocols that exploit effectively and efficiently the different VANET communication paradigms.

B. Paper's Contributions

The objective of this work is twofold: a) to examine and supplement the knowledge in time-evolving topological characteristics of the VANET communication graph. In contrast to the previous studies reported in [6, 10–12] we improve our analysis testbed by considering large-scale realistic vehicular mobility traces based in real-world urban environments and dedicated communication technologies models for generating the VANET graphs; and b) to extend the findings of [9] by examining the VANET dynamics from a more specific spatial viewpoint in order to identify where and/or why particular network phenomena occur on the underlying road network. The present work continues and improves upon our preliminary efforts in [7, 8] towards deriving insightful implications for the vehicular networking community. Furthermore, we correlate our findings with the main communication paradigms employed in VANETs, and discuss whether the emerging network phenomena can indeed support (and in what extend) these paradigms.

The rest of the article is organized as follows: Section II surveys the Related works; Section III presents the System Model and the measures used to characterize the spatio-temporal evolution of the VANET communication graph; Section IV provides information concerning the source of the data studied and the analysis testbed. Section V records the findings of the study. Finally, Section VI concludes the article.

II. Related Works

There is a rich body of work currently on the literature that deals with scientific frameworks for studying the temporal evolution of a number of real-world systems [13]. These graphs span a broad range of domains (autonomous systems, e-mail networks, citations, etc.) and their study has lead to significant implications since most of real-world dynamic networks (online social networks, Internet, etc.) have been proved to follow some topological statistical features (i.e. scale-free, small-world, power-laws).

The value of the connectivity analysis of ad hoc networks is so fundamental that a competition-experiment was initiated [11] to study MANET network connectivity, degree distribution, clustering, topology change frequency, route length distribution, route change frequency, etc. The obtained results show a high degree of topology and route changes, even when mobility is low, and a prevalence of asymmetric routes, both of which contradict assumptions commonly made in simulations. Also, well-known concepts from social network analysis, such as the betweenness centrality index have been used as primitives to design advanced protocols for routing and caching in ad hoc sensor networks.

In the context of vehicular networking, Fiore et. al [10] study the node degree distribution, link duration, clustering[1] coefficient and number of clusters for VANET graphs under various mobility models. The objective was to study the topological properties of different mobility models and explaining why each leads to dissimilar network protocol performance.

Viriyasitavat et. al [12], presented a comprehensive analytical and simulation framework for network connectivity of urban VANETs, utilizing system parameters such as link duration, re-connection frequency, and re-healing time. The authors use data from a cellular automata-based traffic mobility model based on a grid-like road topology. Monteiro et. al [14] also make use of the above traffic model to analyze simple properties such as degree distribution, average shortest path length and clustering co-efficient. Similar to our previous efforts [7], the authors extend a known VANET-dedicated protocol, with preliminary results showing benefits (network overhead reduction) when using topological information during information broadcast.

Despite the thorough evaluation, the above research works base their observations in the study and analysis of a small-scale VANET deployed on a rectangular road topology. Given that the majority of real-world urban environments do not pertain to the simplistic layout and characteristics of such topologies, the findings on network connectivity could be biased [15]. To the best of our knowledge, our preliminary work in [6] was the first effort to study the structural evolution of a large-scale VANET topology over a real-world road network. Using realistic mobility traces available at the time, we

[1]The terms "cluster" and "component" are used interchangeably in this manuscript.

analyzed the communication graph emerging from thousands of vehicles driven in Zurich. Despite the valuable findings of that study, the employed synthetic traces were characterized by highly variable vehicle density and a rather non-uniform traffic distribution, which hindered the precise capture of the topological features. Recently, [16] employed the same dataset in a quest to find social properties in VANETs. Macroscopic (distance, diameter, density) and microscopic (degree, cluster coefficient and closeness centrality) measures were examined to identify social behavior emergence from opportunistic vehicle interactions. Similarly to our findings in [6], they characterize VANETs as scale free, since they identify (during peak-traffic hours) the presence of a degree distribution that follows a Power Law. Grzybek et. al [17] use synthetic mobility traces for the city of Luxembourg to evaluate the presence and characteristics of large-scale communities in urban environments.

More recently [9] provided a more complete analysis of the instantaneous VANET communication graph by examining the opportunistic encounters of vehicles driven in Cologne, Germany. The authors employ the TAPAS-Cologne [18] dataset that describes the mobility of several thousand vehicles moving in the urban and sub-urban regions of Cologne. Their analysis unveiled that the VANET is composed of a large number of disconnected components that hinder delay-sensitive, multi-hop communication among vehicular peers. Furthermore, it revealed that large components are unreliable in terms of their internal dynamics (frequent re-wiring of multi-hop paths), ascribed to their significant temporal variability. The VANET also lacks navigability since it does not exhibit the properties of scale-free networks. Although [9] provided new and interesting findings regarding the VANET topology, these were summarized over a 24-hour time-frame analysis, which included both peak and off-peak traffic periods. However, all the interesting phenomena appear during morning and afternoon peak-hours, at which time, the prevailing traffic conditions shape the vehicular network in a way that inter-vehicular communication can effectively take place.

In this work, we focus on the VANET communication graph realized during the morning peak-period and we commit to a *fine-grained* analysis of the respective topology features and dynamics by employing additional measurements from complex network science. Furthermore, the VANET can be "seen" from the viewpoint of real-world networks that are embedded in low-dimension geometric structures. Consequently, besides its obvious temporal evolution, it also exhibits a spatial dimension and should thus be studied accordingly. As a consequence, in contrast to the above research works, we aim to identify where in the underlying road network particular network phenomena occur and which are the driving factors.

III. System Model and Network Measures

We consider a VANET in a large-scale urban environment, comprised of vehicles equipped with an IEEE 801.11p capable transceiver. All vehicles share identical nominal line-of-sight and non-line-of-sight transmission range (R_{LOS} and, R_{NLOS}). Each vehicle has its own mobility profile defined as an ordered-sequence of spatio-temporal points $MP = \langle P_1 \ldots P_n \rangle$, where $P_c = (x, y, t)$ and x, y are defined as geospatial coordinates (transformed to the Euclidean space) and t is a distinct time instance. $S = t(P_1)$ and $D = t(P_n)$ are regarded as the start and destination location/times, respectively. Vehicles indicate their intention to participate in the VANET by frequently broadcasting beacon messages B_{msg}. We assume, vehicles are aware of the road travelling on through on-board maps and the Global Positioning System (GPS).

Each vehicle can be driven on a road obeying to the following taxonomy:

- *Urban Roads*: Low capacity roadways, providing access to residential areas with houses or amenities on either side. One-way or two-way with 1 lane per travel direction. The maximum speed $\upsilon_{(max)}$ is 30 Km/h.
- *Collector Roads*: Low-to-moderate capacity roadways, collecting traffic from urban roads to arterial roads or vice-versa. Traffic flow is managed via signalized intersections and stop-signs. Two lanes per direction with $\upsilon_{(max)}$: 30–50 Km/h, depending on the area.
- *Arterial Roads*: High-capacity, bi-directional urban roadways, delivering traffic between collector roads and freeways and also carrying long-distance traffic flows between activity areas. Considered as the back-bone of the road network and traditionally are arranged in concentric circles. Two lanes per direction and usually separated into two classes: (i) Low-speed, with $\upsilon_{(max)}$: 50 Km/h and (ii) high-speed, where $\upsilon_{(max)}$: 80 Km/h.
- *Freeways*: Multi-lane, bi-directional, high volume roadways dedicated to high speed vehicular traffic. Traffic flow is not obstructed by stop signs and traffic signals. Separation and connection with other roadways of lower grade (i.e arterial and collector roads) carried out via overpasses across the highway span. Entrance or exit to freeway are provided at overpasses by ramps that enable speed transitions. Freeways have 2 lanes per drive direction with $\upsilon_{(max)}$: 120 Km/h.

We model the VANET as an undirected graph $G(t)$, where vehicles correspond to the set of vertices $V(t) = \{u_i\}$ and communication links to the set of edges $E(t) = \{e_{ij}\}$. An edge $e_{ij}(t)$ exists, if u_i can communicate directly with u_j at time t (exchange B_{msg}) , with $i \neq j$. In turn, the following measures are utilized to characterize the spatio-temporal evolution of the VANET communication graph.

A. Link-level metrics

- **Number of connected periods**. The number of established links between a pair of vehicles within a given time period. A connected period is the continuous time interval during which a link is established between two vehicles.
- **Link duration**. The time duration of a connected period. Formally, the duration $l_{ij}(t)$ of the link from u_i to u_j at

time t is defined as $l_{ij}(t) = t_c - t_o$, if $\exists e_{ij}(t)$, where $t \in [t_o, t_c]$ and $\nexists e_{ij}(t')$, where $t' < t_o$ or $t' > t_c$.
- **Re-healing time**. The time span between two successive connected periods of a pair of vehicles.

B. Network-oriented metrics

- **Node degree**. The number of vehicles within the transmission range of a node. Formally, the degree of u_i at time t is defined as: $D_i(t) = \|\{u_j \mid \exists e_{ij}(t)\}\|$.
- **Effective Diameter**. The minimum distance in which the 90th percentile of all connected pairs of vehicles can communicate with each other. It is a smoothed form of network diameter which we use for our studies.
- **Network Density**. The ratio between the number of edges in the $G(t)$ and the maximum number of edges possible for $G(t)$.
- **Geographic Diameter**. The length of the effective graph diameter considering the real underlying geographic topology. It conveys the geographic span of the shortest path between the most distant nodes in the graph and is measured in meters instead of hops.
- **Network Connectivity [12]**. The maximum fraction of vehicles that are directly or indirectly connected in $G(t)$. Formally, network connectivity is defined as: $NC \stackrel{d}{=} \max_{i}\{\frac{1}{|V(t)|}\sum_{j} A(u_i, u_j)\}$, where $A(u_i, u_j)$ is a connectivity indicator which takes the value of 1 if a path is available from vehicle u_i to vehicle u_j at time t, and 0 otherwise.

C. Component-oriented metrics

- **Number of Components**. The number of co-existing, non-connected clusters of nodes at a given instant. A cluster is a sub-graph of the network such that there is a path between any pair of nodes. In addition we examine the spatial distribution of components and their coverage (Km^2) in terms of geography of the underlying urban area.
- **Number of Communities**. The number of existing communities at a given instant. A community is as a dense sub-graph where the number of intra-community edges is larger than the number of inter-community edges. To identify communities, we transform $G(t)$ to directed graph so as $D_i^{in}(t) = D_i^{out}(t) = D_i(t)$, where $D_i^{in}(t), D_i^{out}(t)$ is the in-degree and out-degree of node i at time t. Formally, a sub-graph $U(t)$ of a VANET graph $G(t)$ at time t constitutes a community, if it satisfies: $\sum_{i \in U}(D_i^{in}(t))(U(t)) > \sum_{i \in U}(D_i^{out}(t))(U(t))$ i.e., the sum of all connections within the community $U(t)$ is larger than the sum of all connection toward the rest of graph the $G(t)$.

D. Centrality metrics

- **Betweenness Centrality**. The fraction of the shortest paths between any pair of nodes that pass through a node. The betweenness centrality of a vehicle u_i at time t is: $BC_i(t) = \sum_{j \neq k} \frac{sp_{jk}(u_i, t)}{sp_{jk}(t)}$, where sp_{jk} is the number of shortest paths linking vertices j and k at time t and $sp_{jk}(u_i, t)$ is the number of shortest paths linking vertices j and k that pass through u_i at time t. BC is a measure of the extent to which a vehicle has control over information flowing between others.

IV. Analysis Testbed

We consider a corrected and enhanced version of the initial TAPAS-Cologne dataset [18], following the results of [19]. The dataset covers approximately an area of $33 \times 35Km$ with 42148 intersections and 134645 roads. For the purpose of our study, we focus on the morning traffic rush-period between 6:00am and 8:00am, taking snapshots of the network every 1 second. During this 7200 second time window, approximately 75600 unique vehicles enter and travel within the aforementioned area.

We assume the model of line-of-sight(LOS) and non-line of sight (NLOS) wireless communication, wherein vehicles establish a wireless link if they are within 250 and 140 meters of each other, respectively. For more information on when two vehicles might have LOS or NLOS communication, the reader is referred to [12].

We provide the necessary input files (road map, route definitions, etc) obtained from [19] to the SUMO simulation framework and instruct it to run 10 times, each one with a different random seed generator. We allow a 1000 second warm-up period at the beginning before obtaining any measurements in order to achieve some level of stability in the network. Consequently, we study snapshots of the VANET communication graph taken every 1 second. Overall, approximately 436000 snapshots of the VANET communication graph were recorded.

V. Observations

This section presents the findings of our study related to the laws governing the dynamics of vehicular connectivity by considering the realistic testbed presented previously. We note that, although important, our work does not examine the dynamics of the VANET communication graph under the presence of medium access problems such as contention and interference.

A. Link-level Analysis

Our observations begin with the link-level analysis of the VANET communication graph, which contributes to a close estimation of the network-link lifetime. The number and duration of connected periods between two vehicles, as well as the duration between successive connected periods is influenced by microscopic mobility characteristics as well as the underlying road network. Table I presents an overview of the link-level statistics of the IVC members in the city of Cologne.

Links are categorized by recording their respective start-end road types. Links that start and finish in freeways or arterials, are tagged as outskirt links. Respectively, links that start and

	Min	Max	Mean	Median
Connected Periods	1 (1)	3 (8)	2.2 (6.4)	2 (6)
Link Duration (sec.)	1 (1)	560 (224)	322.7 (47)	301 (42)
Re-Healing Period (sec.)	1 (1)	980 (450)	121.16 (48)	117 (46)

TABLE I: Link statistics for vehicles in the city outskirts (freeways/arterials). In parenthesis for vehicles driven in the inner city (urban/collector roads).

finish in urban or collector roads, are tagged as inner city links. Higher traffic densities result to smaller inter-vehicle distances and thereby increase the time period in which two vehicles are in range of each other. This allows established links to have a longer duration and smaller number of connecting periods. Increased durations are therefore encountered particularly in arterial and freeways, where the characteristics of these roadways do not cause links to break often. We observe the number of connected periods for vehicles driven on urban and collector roadways to be almost 3x as their outskirt counterparts, however with smaller re-healing durations.

B. Network-level Analysis

We refer to the size of the VANET as the number of edges that are present in the communication graph at any given time instance. The VANET size is not sensitive to changes in the arrival and departure rates of vehicles. Specifically, although in several time instances we monitor an increase in the number of departing vehicles can cause a decrease of the VANET size, this effect is quickly mitigated by the set of vehicles still in the network, whom their collective mobility profile (MP) can cause the sustainability of current links and even the establishment of new ones. Interestingly, the VANET core[2] geography, is very well embedded (overlaps) to the underlying geography of the city center. Averaging across time, the core includes 18% of the total edges and 7% of the nodes available in the network, respectively.

Initial insights on the global connectivity conditions that prevail in the VANET, report that this is indeed a very sparse network. Even in the case of our study, where we consider 100% V2V technology penetration ratio and high vehicle density (due to rush-period), the corresponding network density never exceeds 40%. Fig. 1(b) plots the global network connectivity for the 6200s we analyze. NC is upper bounded roughly at 35% of the total number of nodes that coexist in the area. In other words, even under optimal global connectivity settings, only 35% of all vehicles in the topology can exchange information either using a direct or a multi-hop communication paradigm. The nature of NC provides a birds-eye view of the VANET capacity to facilitate direct or multi-path inter-connectivity for its member vehicles, in the presence of various constrains induced by a large-scale urban environment (morphology of road topology, inter-vehicle distances, mobility patterns etc). *These findings serve as a precursor of the VANET fragmentation level and indicate the necessity of utilizing supporting mechanisms for city-wide, multi-hop and geocasting communication paradigms. For instance, the installation of fixed roadside infrastructure can improve link stability and also serve as gateway to long range V2V information exchange. In addition, caching solutions can be an option given that the application requirements permit it.*

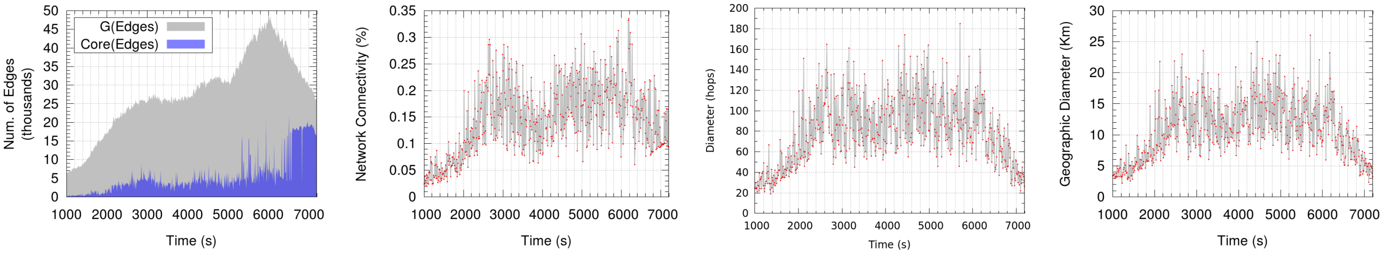
Next, to comprehend how far in the topology data packets can potentially travel, we examine the effective diameter of the network. Since a network hop in the VANET graph is dependant on the geographic distance between the two connected vehicles, inevitably the largest geodesic pertains also a geographic dimension. Fig 1 plots both diameter measures, network (c) and geographic (d), against time. Evidently, the network diameter of the communication graph, as well as its geographic counterpart, exhibit significant variability in the arrow of time. Interestingly, we have identified several time instances where the diameter can oscillate in excess of 50 hops ($\sim 15Km$) within just a few seconds, denoting a severe or complete rewiring in the structure of the largest shortest path in the VANET.

Our spatial analysis has shown that in contrast to real world temporal networks - where graph diameter is known to be a function of network size [13] - the changes in the VANET diameter are attributed down to one microscopic vehicle feature, that of geographic position. Specifically, by studying closely the MP of vehicles comprising the geodesic, we conclude that the observed fluctuations in both diameter measures are the effect of individual vehicles' mobility. Several nodes unawarely function as bridges among weakly-connected components and even slightest changes in their geographic position is enough to fragment those components and trigger a strong "butterfly effect" over global network connectivity. It is important to note that this variability of the shortest path is not an isolated phenomenon, rather it can be observed also in the top-5 shortest paths. Although, there seem to exist long paths that could potentially transfer information among distant city parts, the emerging volatility in their structure, can hinder even short-range multi-hop data dissemination.

Fig. 2 exhibits a snapshot of vehicle positions while being driven on the road network of Cologne. Each vehicle is colored according to its individual node degree $D_i(t)$. Circles denote underpasses/overpasses or exit ramps that enable vehicles transition to higher/lower grade roadways. Likewise to [9], our analysis demonstrates that the spatial distribution of node degree is not uniform throughout the topology. The VANET is a very heterogeneous network, where isolated nodes, and nodes with over 100 neighbors that resemble to traditional "hubs"[3], can coexist in unity. Particularly, the analysis indicates that D_i is dependent on the proximity of vehicles to different features of the underlying road network. The majority of vehicles with lower degree (green to light blue

[2]Core is defined as the largest biconnected component existing in the network.

[3]Hubs are defined as those nodes with degree equal or larger to an order of magnitude higher than the average node degree. However, in the Cologne such degree magnitudes were not identified.



(a) (b) (c) (d)

Fig. 1: (a) VANET Size and Core, (b) Network Connectivity, (d) Effective Diameter. (d) Geographic Diameter - vs Time.
(Best viewed in color)

color) are encountered on freeways and collector roads. Multi-lane, high capacity freeways, enable neighbouring vehicles with the same direction and relative velocities to travel in platoons for a prolonged period of time. This mobility pattern fosters the establishment of a number of connections (on average between 13–17) among vehicles driving in the same or adjacent lanes towards the direction of travel, resulting to the degree distribution visible in the outer parts of the city. On the other hand, the reduced 1-hop connectivity on collector roads comes as a result of the low-to-moderate capacity of the roadway itself and also the presence of intersections and stop-signs which interrupt the continuous traffic flow.

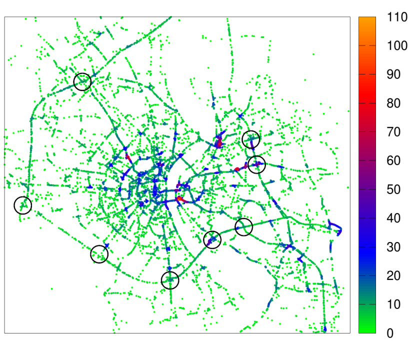
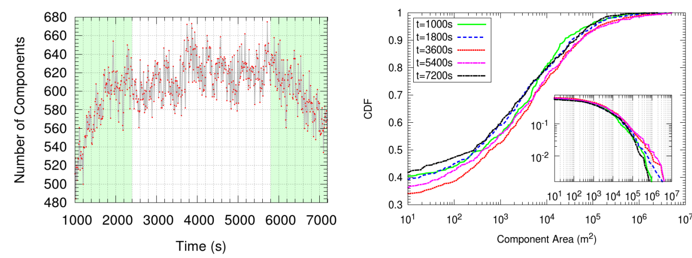


Fig. 2: Spatial representation of vehicle location colored according to node degree.
(Best viewed in color)

Nevertheless, the degree of vehicles travelling on high capacity roadways is more stable over time, due to the fact that such connections are long-lived; vehicles tend to use freeways and arterial roads to cover large distances. An exception to the above occurs whenever vehicles approach various overpasses or exit ramps. Such high density road features cause spikes in the temporal distribution of degree and in certain occasions some vehicles are able to establish a short connection with up to 122 neighbors. The above pattern is recorded repeatedly to several such ramps, nearby which we also identify the presence of vehicle communities with dense inter-connection among their members. *These geography bounded dense structures allows them to act as "data-islands", facilitating thus spatio-temporal information replication and maintenance during broadcast of geographic hovering communications.*



(a) (b)

Fig. 3: (a) Number of Components, (b) Component Geographic Coverage Distribution.
(Best viewed in color)

Moving towards to the city center and particularly on urban roads, the node degree tends to increase substantially (darker green - blue) primarily due to the increased traffic density and lower speed limits. These factors allow vehicles to connect with an increased number of other vehicles being driven in the same or opposite directions, or even on various intersecting roads. However, the dense road topology with rapid transition in-between LOS/NLOS communication, makes these connections short-lived. Table II presents a summary of vehicle degree statistics for each of the road features. Statistics are aggregated for the whole 6200s of our analysis.

Road Type	Min	Max	μ	σ
Freeway	0	101	13.7	34.0
Arterial	0	122	17.0	34.0
Collector	0	114	9.9	22
Urban	0	115	8.1	19

TABLE II: Aggregate Vehicle Degree statistics for the different road network features.

C. Component Analysis

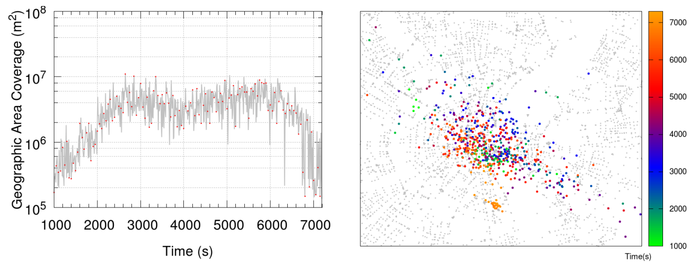
In contrast to [9], our work does not consider isolated vehicles (i.e singletons) in component analysis, since communication of a vehicle with itself is meaningless. As depicted by Fig. 3(a), the network exhibits noteworthy fragmentation, since no less than 480 components are present in the network at any given instance. The VANET consists of both small and large components, with 40% of those being components established among just two vehicles, while 90% of them are comprised by 15 vehicles or less. By studying the CCDF of the component size distribution, and specifically its long tail, we identify the simultaneous presence of certain large components that host between 1000 and a giant component that surprisingly can facilitate up to 2500 vehicles.

Overall, we observe 3 temporal phases in component population evolution. The 1st phase (left green shaded area, Fig. 3(a)) exhibits a growth trend, wherein the arrival of new vehicles (~ 7500 vehicles) in the topology sparks the creation of additional small components. During this period the majority of components are comprised of 3–4 vehicles only and cover, on average, a geography of less than $15m^2$.

In the 2nd phase we observe a relative stability in population growth, where newly arrived vehicles ($\sim 7500-8500$ vehicles) do not create additional components, but rather join existing ones. As day progresses and existing components become even larger, they are fused with other smaller components in the vicinity, extending their corresponding coverage. Finally, a shift point in vehicle population ($\tau = 5800s$) triggers the 3rd and decaying phase in the number of components. As the rush-period moves towards the end, vehicle density is not sufficient to accommodate strong connectivity across existing components, therefore further increasing fragmentation.

Nevertheless, these heterogeneous components scale differently in terms of the geographic area they cover. Fig. 3(b) shows the CDF of the coverage region, calculated using the *convex hull* of each component. 37% of the components present in the VANET cover an almost negligible geographic area less than $15m^2$. Approximately 70% of such small clusters are primarily encountered in urban roads near the outskirts of the city, where the specifics of the road-network topology encourage tight inter-vehicle distances. The remaining 30% of small clusters comprise vehicle platoons travelling on high speed freeways, with inter-platoon distances larger than 250 meters, which is the maximum transmission range in LOS. *Consequently, vehicles travelling in such areas, should not expect to establish and maintain communication sessions with a large number of other vehicles, either utilizing broadcast or multi-hop communication paradigms.*

Conversely, as the geographic coverage of components increases, we observe a different spatial distribution across the area of Cologne. While 90% of these larger components cover an area less than $56000m^2$, now they are found with higher probability on roads close to down-town. This transition period is in line with the commute habits of people during the morning rush hour, where the majority of vehicular traffic traverses initially arterial and collector roads prior to exiting and spreading to urban roads so as to reach various activity areas (workplaces, stores, etc) within the city core. The increased vehicle density encountered in these moderate-to-high capacity road structures, in conjunction with minimal travel flow interruptions, results to unhindered inter-vehicle



(a) (b)

Fig. 4: Temporal Evolution of Giant Component in terms of: (a) Geographic Area Coverage and (b) Coordinate Centroid. (Best viewed in color)

communication that in turn facilitates the establishment of larger moving components.

Focusing on giant component, we would like to know where such structures can be found in the underlying topology and how their location shifts with time. Having such knowledge at hand is beneficial to communication protocols, mainly in situations where complete or partial VANET topology information is unavailable. *Such details ultimately enable various protocols to approximate the geographic boundaries of the giant component, by employing only historical traffic statistics. Since the giant component is strongly connected internally, these approximations allow protocols to decide whether or not they are located within a region that favors multi-hop data dissemination.* Figure 4(b) provides a visualization of the underlying road topology, overlayed by different points that correspond to the geographic centroid of the giant component. Here, the centroid metric serves as an anchor point from which the giant component materializes. During the early morning hour, the giant component evolves adjacent to arterial and collector roads that transfer commuters from the outer parts of the city and the suburbs towards the center.

In these early times, the giant component centroids are spatially dispersed, and circulate both the eastern and western parts of the city where such road features are found. The alterations between the different parts of the city, come as a result of the border effects in terms of component size, where at one particular instance a component at the eastern part of Cologne is larger by a small margin than another component at the western part, or vice-versa. When vehicle density is increased and subsequently stabilized in the down-town area, the giant component centroids move in and remain concentrated in the center of the road network.

Figure 4(a) illustrates the geographic area coverage of the giant component across time. The coverage area is reported in m^2 and the y-axis is plotted in log-scale. Given that the whole road topology size is $33Km \times 35Km$, the giant component can cover between 1% and 4% of the city of Cologne. The spatial coverage grows in the early phase of the morning rush-hour when vehicle density is increased. As in the case of the component membership, after the mark of ~ 7500 vehicles ($\tau = 1900s$) in the network is achieved, the geographic area coverage starts to stabilize with very fine variations. In contrast to smaller components in its vicinity that continue after this instance to change both size and coverage, the giant component remains relatively constant and always encapsulates instances of the four road features.

D. Centrality Analysis

Considering the prominent presence and geographic coverage of this giant component, we then examine the centrality of its member vehicles in the communication graph. Figure 5 depicts (y-axis in log-scale) the centralization of the giant component in terms of the betweenness $BC_i(t)$ index of its members.

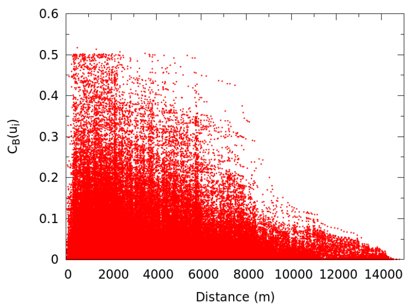


Fig. 5: Betweeness Centrality correlation with vehicle distance from city center.

Giant component members, are heterogeneous in terms of their BC_i. During this early phase of the morning rush hour, we identify a number of vehicles who control up to 52% of the geodesics present in the giant component. As it can be seen from the above plot, interestingly all of these high centrality vehicles are located on urban and collector roads in a distance of $\leq 6Km$ from the city center. Nevertheless, even during this period, mobility dynamics cause the giant component to alternate between states of centrality homogeneity and heterogeneity in a few seconds. As a result, vehicles that were previously identified as "central" to information communication can often get demoted to ordinary nodes within 10-20 seconds, or vice-versa.

VI. Conclusions

We perform an extensive analysis of the spatio-temporal characteristics of the VSN communication graph using Complex Network measures. We explore the dynamics of VSN in urban environments and investigate their impact in the use of dedicated communication paradigms. Specifically, the free flow nature of freeways, causes vehicles tend to form long-lived links with more neighbors in city-outskirts, with fewer disconnections. On the contrary, the inner-city road topology induces highly intermittent links among fewer vehicles. City-wide, multi-hop communication is hindered by the high variability of geodesics and network fragmentation, which in turn make the use of caching mechanisms and road-side infrastructure indispensable. We identify key road features (overpasses and exit ramps) in particular locations that induce high and dense IVC in their vicinity and can thus serve as data-islands in information hovering. A giant component, covering a large geography, is present most of the time, and its spatial evolution follows the commute habits of people during the morning rush hour. Central nodes exist within the giant component, however they are confined within an distance from the city-center.

References

[1] "Car-2-Car Communication Consortium," http://www.car-to-car.org/.
[2] "Cooperative Vehicle-Infrastructure Systems," http://www.cvisproject.org/.
[3] "U.S. Department of Transport (DoT) - Connected Vehicle Safety Pilot Program, Ann Arbor, Michigan," http://www.its.dot.gov/safety_pilot/.
[4] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *SIGCOMM '99*. Cambridge, Massachusetts, United States: ACM, 1999, pp. 251–262.
[5] "Virtual Traffic Lights (VTL)," http://www.virtualtrafficlights.com/.
[6] G. Pallis, D. Katsaros, M. Dikaiakos, N. Loulloudes, and L. Tassiulas, "On the Structure and Evolution of Vehicular Networks," in *Modeling, Analysis Simulation of Computer and Telecommunication Systems, 2009. MASCOTS '09. IEEE International Symposium on*, Sept. 2009, pp. 1–10.
[7] N. Loulloudes, G. Pallis, and M. Dikaiakos, "The Dynamics of Vehicular Networks in Urban Environments," *CoRR*, vol. abs/1007.4106, 2010.
[8] N. Loulloudes, G. Pallis and M. Dikaiakos, "Understanding V2X Communication Dynamics Through Complex Network Science," *ERCIM News*, vol. 2013, no. 94, 2013.
[9] D. Naboulsi and M. Fiore, "On the Instantaneous Topology of a Large-scale Urban Vehicular Network: the Cologne Case," in *The 14th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc '13, Bangalore, India*. New York, NY, USA: ACM, July 2013, pp. 167–176.
[10] M. Fiore and J. Harri, "The Networking Shape of Vehicular Mobility," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. New York, NY, USA: ACM, 2008, pp. 261–272.
[11] V. Srivastava, A. B. Hilal, M. S. Thompson, J. N. Chattha, A. B. MacKenzie, and L. A. DaSilva, "Characterizing mobile ad hoc networks: The MANIAC challenge experiment," in *Proceedings ACM WiNTECH*, 2008, pp. 65–72.
[12] W. Viriyasitavat, F. Bai, and O. Tonguz, "Dynamics of Network Connectivity in Urban Vehicular Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 3, pp. 515 –533, March 2011.
[13] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph Evolution: Densification and Shrinking Diameters," *ACM Trans. Knowl. Discov. Data*, vol. 1, no. 1, Mar. 2007.
[14] R. Monteiro, S. Sargento, W. Viriyasitavat, and O. K. Tonguz, "Improving VANET Protocols via Network Science," in *2012 IEEE Vehicular Networking Conference, VNC 2012, Seoul, Korea (South), November 14-16, 2012*. IEEE, 2012, pp. 17–24.
[15] N. Loulloudes, G. Pallis, and M. D. Dikaiakos, "On the Performance Evaluation of VANET Routing Protocols in Large-Scale Urban Environments (Poster)," in *Proceedings of the IEEE Vehicular Networking Conference, VNC 2012*. IEEE Computer Society, November 2012, pp. 129–136.
[16] F. Domingos Da Cunha, A. Carneiro Viana, R. A. F. Mini, and A. A. F. Loureiro, "Is it Possible to Find Social Properties in Vehicular Networks?" in *The 9th IEEE Symposium on Computers and Communications (ISCC)*, Ilha Madeira, Portugal, June 2014.
[17] A. Grzybek, G. Danoy, M. Seredynski, and P. Bouvry, "Evaluation of Dynamic Communities in Large-scale Vehicular Networks," in *Proceedings of the 3rd ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*. New York, NY, USA: ACM, 2013, pp. 93–100.
[18] "The TAPAS-Cologne Project," "http://sumo.dlr.de/wiki/Data /Scenarios/TAPASCologne".
[19] S. Uppoor and M. Fiore, "Large-Scale Urban Vehicular Mobility for Networking Research," in *Vehicular Networking Conference (VNC), 2011 IEEE*, November 2011, p. 62.