# The Connectivity of Cologne's Large-scale Vehicular Network

Diala Naboulsi
INSA Lyon, INRIA

Marco Fiore
INSA Lyon, INRIA

## I. Introduction

More than a decade after the allocation of a dedicated frequency bandwidth in the USA, vehicular communication networks have finally abandoned their long-standing status of fundamental research exercise and have started developing into real-world systems. On-going standardization efforts, including IEEE 802.11p, IEEE 1609, OSI CALM-M5 and ETSI ITS, jointly with the growing interest of the automobile industry, have played a major role in speeding up the process over the last few years. Thus, it will be soon time to field test the vast amount of vehicular networking solutions proposed by the scientific community during the last decade, whose scope ranges from medium access to transmission power control, from data rate adaptation to mobile address management, from multi-hop routing to reliable data transfers.

However, the real-world performance of many of such protocols risks to fail the expectations, for the simple reason that they will be confronted by a network they are not designed for. As a matter of fact, dedicated solutions have been proposed for vehicular networks whose major topological features remain largely unknown. In urban environments in particular, even basic questions stay unanswered, such as: *is the vehicular ad hoc network well connected or highly partitioned? Which size can components of multi-hop connected vehicles attain? How do the network connectivity features vary in time? How do they depend on the geographical location?*

The responses to these questions directly determine the capabilities of a spontaneous vehicular network, and shall thus be among the main drives to the design of dedicated protocols. Moreover, they are the key to quantifying the availability of the network, one of the main concerns of car manufacturers when it comes to vehicle-to-vehicle multi-hop communication.

The aim of this paper is to shed some light on topological features of a large-scale urban vehicular network, providing answers to the questions above and highlighting their impact on vehicular networking solutions. To that end, we leverage a specific mobility dataset, presented in Sec. II, that is representative of a typical middle-sized European city and yields unprecedented realistic microscopic and macroscopic mobility features. We model the dataset as a set of instantaneous connectivity graphs, as described in Sec. III, and analyze their features from a communication network perspective, in Sec. IV. Finally, conclusions are drawn in Sec. V.

## II. Vehicular mobility dataset

Our analysis is based on a realistic vehicular trace from the TAPASCologne project. This project aims at reproducing car traffic in the city of Cologne in Germany on a typical workday with a high level of realism. Using these traces, the dataset we study has been generated by combining different tools as detailed in [1].

**Road topology**. The road topology is obtained from the OpenStreetMap (OSM) database. The OSM project provides highly detailed free maps of the world. It is based on the collection of users' data and GPS traces. This combination makes of OSM the finest source of publicly available maps.

**Microscopic mobility**. The microscopic mobility of vehicles is simulated with the SUMO software. SUMO is an open-source traffic simulator, it provides a high level of accuracy in modeling the behavior of vehicles as it takes into account several factors such as the car-to-car distance, traveling speed, and acceleration/deceleration profiles. Due to its high level of accuracy and scaling capacity, SUMO is the most advanced open-source microscopic vehicular mobility generator.

**Traffic demand**. To obtain the traffic demand, the Travel and Activity Patterns Simulation (TAPAS) methodology is applied. This technique provides the source and destination points of each vehicular trip, by combining information about the population and the land use.

**Traffic assignment**. Gawron's algorithm is used for the traffic assignment. This algorithm is capable of providing a dynamic traffic equilibrium. It computes for each vehicle the fastest route to reach its destination, and assigns to each road a cost based on the traffic intensity. By making several iterations in which vehicles take less congested routes, the algorithm offers a realistic road capacity utilization.

Finally, the obtained dataset spans over 24 hours of a typical workday and encompasses 4500 km of roads in an area of 400 km$^2$, with per-second information on the position and speed of vehicles involved in more than 700000 trips.

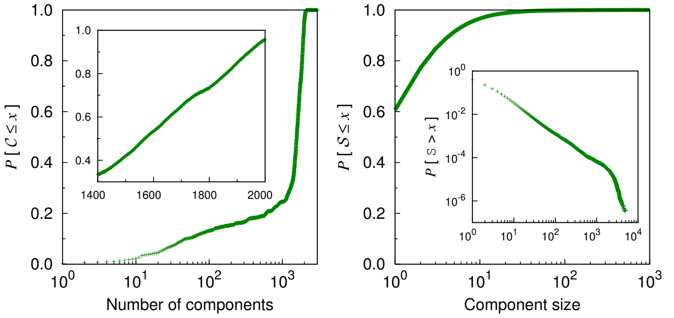
## III. Network model and metrics

We model the vehicular network topology as a graph where the nodes represent vehicles and the edges constitute communication links between them. The set of communication links depends on the RF signal propagation model adopted, which we assume to be a simple unit disc model. The unit disc model allows to capture the average behavior of the network and drastically reduce the computational complexity of the analysis. We employ a communication range equal to 100 m, as this is the value referenced by field tests as a typical distance for reliable DSRC vehicle-to-vehicle communication [2].

In this study, we focus on the instantaneous connectivity of the network: we analyze the instantaneous graphs derived from the vehicular mobility dataset by considering a sampling process with a 10-seconds period. We use $\mathcal{N}$ to denote the instantaneous number of nodes in the network. We define a component as a group of nodes that can communicate together via direct or multi-hop links. The size of a component $\mathcal{S}$ represents the number of nodes belonging to it. A component is considered a giant component if $\mathcal{S} \geq 0.1 \cdot \mathcal{N}$. We use $\mathcal{C}$ to refer to the number of components that appear in the network at a specific time instant.

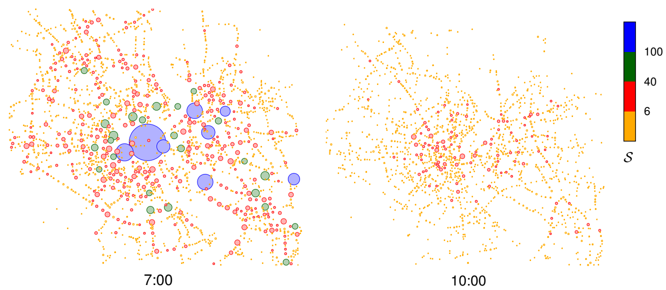
## IV. Network connectivity

The level of connectivity of the whole vehicular network is mainly characterized through two metrics: the number of




(a) (b)

Fig. 1. Distributions of (a) the number of components and (b) the component size, when aggregating all the samples of $\mathcal{C}$ and $\mathcal{S}$ over the whole day.




(a) (b)

Fig. 2. Geographical plots of the network components at time 7:00 (a) and time 10:00 (b).

components $\mathcal{C}$, that is an index of the level of network fragmentation, and the component size $\mathcal{S}$, which describes the heterogeneity of the fragmentation.

The Cumulative Distribution Function (CDF) of $\mathcal{C}$, aggregated over all the vehicular network graphs extracted from the whole 24 hours of road traffic, is portrayed in Fig. 1(a). In 80% of cases the vehicular network has more than 1000 disconnected components, and the inset plot shows that most of the probability is concentrated in a linear growth between 1000 and 2000 components. This suggests that the vehicular network is typically highly partitioned into thousands of separate node groups unable to communicate with each other.

The distributions of $\mathcal{S}$ help us clarify whether we face many components of similar size or a heterogeneous network of both large and small components. The CDF in Fig. 1(b) shows that the network is largely made of very small components, with 60% of vertices being *singletons*, i.e., isolated vehicles, and 95% of the components comprising 10 vehicles or less. However, by looking at the Complementary CDF (CCDF), in the inset plot, we can appreciate the heavy tail of the distribution, appearing as linear on a log-log scale. There is thus a non-negligible probability that the small components coexist with components including up to 2000 vehicles connected through direct or multi-hop links. After such a component size, the CCDF has an exponential decay, and components as large as 4500 nodes appear with significant lower probability.

Such a general view of a partitioned and heterogeneous network is however aggregated over time and space. In order to unveil the impact of the daytime and the geographical area, we study the temporal and spatial evolution of the network.

To that end, we capture the instantaneous vehicular network fragmentation through a series of geographical plots as shown

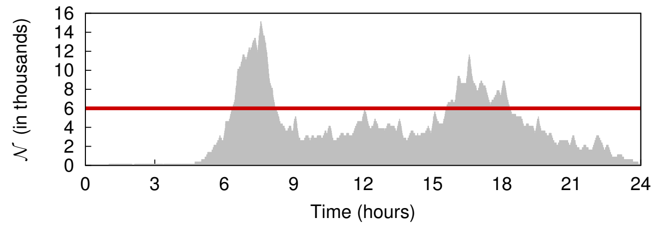



Fig. 3. Evolution of $\mathcal{N}$ over time

in Fig. 2 as well as the evolution of $\mathcal{N}$ over time in Fig. 3. In this paper, we only show two of the geographical plots at times 7:00 and 10:00 in Fig. 2(a) and Fig. 2(b), due to space limitations. In these plots, every circle corresponds to one component, its diameter and color mapping to the size of the component it represents: broader circles map to larger components, while the color code is reported on the bars at the right of the plots. Clearly, the morning traffic peak has very positive impact on the network topology as can be seen from Fig. 2(a), with the appearance of a very large component of thousands of vehicles and a diffuse presence of medium-sized components. The largest component forms a giant component and appears in the city center, where the traffic is denser. Later on as $\mathcal{N}$ decreases, only small components appear all over the city, as shown in Fig. 2(b). Clearly, giant components can be expected to appear in the vehicular network mainly during the morning and afternoon rush hours. Moreover, we observed that they appear in the network for time instants including more than 6000 vehicles as indicated in Fig. 3 by the threshold. This maps to an availability of the network around 4 hours/day.

We can infer that both time and space are paramount factors in the characterization of the vehicular network topology, which appears to be generally highly fragmented, with larger components only appearing at specific locations during the traffic peak hours. From a networking point of view, this indicates that multi-hop routing through the whole network is impossible and evidences the difficulty of routing data within a purely ad hoc vehicular network where only limited time windows are exploitable. Moreover, this also shows that the common practice of evaluating vehicular network protocols in small-scale scenarios characterized by arbitrary vehicular densities may be harmful.

## V. Conclusions

We present a study of the instantaneous topology of the vehicular network in the urban region of Cologne, Germany. Our large-scale analysis allows us to draw conclusions about the low connectivity and availability of the network. Overall, the vehicular network seems to best fit delay-tolerant transfers within localized areas. Moreover, to improve the overall network capabilities, our results let us advocate the adoption of carry-and-forward paradigms and/or the deployment of RSUs, and stress the role of MAC-layer protocols for channel access and rate/power control.

## References


[1] S. Uppoor, M. Fiore, "Large-scale Urban Vehicular Mobility for Networking Research", *IEEE VNC*, Amsterdam, Holland, Nov. 2011.

[2] D. Hadaller, S. Keshav, T. Brecht, S. Agarwal, "Vehicular Opportunistic Communication Under the Microscope", *ACM MobiSys*, San Juan, Puerto Rico, USA, Jun. 2007.