Characterizing the topology of an urban wireless sensor network for road traffic management

Caractérisation de la topologie d'un réseau urbain de capteurs sans fil pour la gestion de la circulation routière

Sébastien Faye
Claude Chaudet

2014D002

mai 2014

Département Informatique et Réseaux
Groupe RMS : Réseaux, Mobilité et Services

Caractérisation de la topologie d'un réseau urbain de capteurs sans fil pour la gestion de la circulation routière

Sébastien Faye, Claude Chaudet
Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI UMR 5141, Paris, France
{*prenom.nom@telecom-paristech.fr*}

Mai 2014

Abstract

Dans un futur proche, les réseaux sans fil seront l'une des technologies clés pour la gestion de la circulation routière dans les villes intelligentes. Les véhicules et systèmes routiers devraient être reliés, par exemple à travers l'extension IEEE 802.11p. En parallèle, nous pouvons nous attendre à ce que les feux de circulation et les signaux routiers aient leur place dans cette architecture, formant un réseau distribué à grande échelle composé essentiellement de petits appareils peu coûteux. En ce sens, ce réseau partage beaucoup de points communs avec les réseaux de capteurs sans fil classiques, y compris sont organisation autour d'un centre de contrôle unique. Cependant, la topologie de ce réseau est fortement influencée par les caractéristiques propres à chaque ville.

Dans cet article, nous classons et caractérisons les topologies probables de ces réseaux. Le but de ce travail est de fournir des modèles de réseaux qui peuvent être utilisés pour évaluer des protocoles et algorithmes sur un scénario réaliste, plutôt que sur des graphes aléatoires génériques. Nous appliquons des méthodes de déploiement du réseau sur plus de 52 cartes de villes extraites de OpenStreetMap et caractérisons les graphes en résultant, en terme d'échelle (nombre de noeuds, diamètre), de densité (distribution de degré), de qualité de liaison (distance entre les noeuds) et de connectivité (nombre de composantes connexes). Les résultats montrent que les villes peuvent raisonnablement être classées en trois catégories qui constituent une base pour les outils aléatoires de génération de scénarios. Les outils, notre jeu de données complet et les modèles OMNeT++ en résultant sont disponibles en ligne.

Characterizing the Topology of a Urban Wireless Sensor Network for Road Traffic Management

Abstract

In a near future, wireless networks will be one of the key technologies for road traffic management in smart cities. Vehicles and dedicated roadside units should be interconnected for example through the IEEE 802.11p extension. In parallel, we can expect that traffic light and road signs will also take their place in this architecture, forming a distributed large-scale network composed essentially of small inexpensive devices. In this sense, this network shares many similarities with classical wireless sensor networks, including its organization around a single control center. However, the topology of this network shall be strongly influenced by each city characteristics.

In this article, we classify and characterize the probable topologies of these networks. The aim of this work is to provide network models that can be used to evaluate protocols and algorithms over a realistic scenario, rather than on generic random graphs. We apply network deployment methods over 52 city maps extracted from OpenStreetMaps and characterize the resulting graphs in terms of scale (number of nodes, diameter), density (degree distribution), link quality (inter-nodes distance) and connectivity (number of connected components). The results show that cities can reasonably be classified into three categories that form a basis for random scenarios generation tools. The tools, the complete datasets and the resulting OMNeT++ models are available online.

1 Introduction

As more and more people move to cities, the drastic increase in urban population density has begun to pose several challenges regarding vehicle traffic management. Town planners and researchers are designing digital monitoring and control systems to help reduce congestion, prevent accidents, limit the environmental cost of transportation, as well as to reduce nuisances. This evolution towards smart cities is an active research domain [30, 31, 32, 7] with ongoing large experimental projects. For example, the city of Pittsburgh, Pennsylvania, acts as a pilot deployment for the Traffic21 research program, which includes the design and deployment of an adaptive traffic light control system. This program has successfully proven that such systems are efficient and has received full support from the city mayor and from the U.S. Secretary of Transportation.

If such systems, in the past, were limited by the cost of vehicle detection systems, today, the success of embedded systems allows deploying a dense network of detectors and actuators that communicate using wireless interfaces. Tiny devices capable of counting vehicles with a magnetometer or a camera, but also to measure CO_2, micro-particle and noise level, can be installed at traffic lights and on urban lighting systems without complex roadwork.

Today's cheap devices are powerful enough to auto-organize, report measurements to a central SCADA software such as SCOOT ([26]) or SCATS ([27]) and receive global policies in return. This pledges for a dense deployment of nodes over traffic lights and even road signs. Deploying such a large-scale network would, in addition, provide a fixed infrastructure to foster the development of vehicular applications that need a minimal amount of users to form an infrastructure.

Indeed, when the vehicles traffic load is high, a small event can easily and rapidly escalade into a severe congestion [13] and communicating with a central decision point may not be the most efficient solution. Taking advantage of the distributed computing results, such devices can easily communicate together and rapidly adapt the traffic light plans to solve a situation. Finally, the density of devices increases fault tolerance, as measurements and communication paths impaired by the failure of a device can be replaced by neighbor devices.

However, the huge amount of research in ad-hoc, mesh and sensor networks has shown that the network topology has a strong effect on the network performance and identified which network protocols suite was most adapted to various situations. The network density has an effect on local congestion and on nodes energy consumption. It influences the medium access control protocol and the possibility of deploying a few autonomous nodes. Path diversity influences fault tolerance and the network global capacity. The network diameter has an effect on the end-to-end delay. The network partitioning defines whether the distributed network can work in autonomy or needs to be interconnected to a cellular or wired backbone.

No real city-wide network is deployed yet, and such deployments will happen at a very slow pace until the technology prove itself efficient. In this paper, we characterize plausible network topologies and we derive graph models that are more realistic than the generic random graph models to serve as a basis for protocols and applications performance evaluation. Based on a few deployment strategies that we explain in Sec. 3, we create the communication graphs that result from the sensors deployment over 52 city maps extracted from *OpenStreetMap*, as explained in Sec. 4. We then analyze the resulting graphs structural properties in Sec. 5 and discuss the networking aspects in Sec. 6.

2 Related Works

Using embedded devices to help managing smart cities is not a novel idea. Press-covered research results show the interest of traffic monitoring (through a sensor or a vehicular network) for suppressing instability in traffic flows [12] and for reducing the congestion level. The ramp metering systems that have been widely deployed show that an active management of traffic lights could reduce drastically the traffic jams, even though the Minneapolis ramp meters were highly contested by users who had the *feeling* that waiting times increased. This practical case shows that simulations and experiments are necessary prior to operational deployments.

In parallel, several projects and initiatives have reached an experimental phase and a few medium-scale deployment have begun. CitySense [20] is an urban wireless network testbed deployed all over the city of Cambridge (MA, USA), forming a mesh network. It is composed of 100 linux-based computers that can be programmed directly by end users. Even though the primary focus was to foster mesh networks applications development, nodes have been augmented with environmental and pollution sensors.

In the Cambridge experiments, nodes were deployed to provide a good wireless coverage of the area. However, for application-specific networks, the questions of the best deployment strategy has attracted the attention of the scientific community. Corredor *et al.* [3] look at the deployment of magnetometers for monitoring road traffic over smart highways. They propose to deploy such sensors on every lane to maximize vehicles detection probability and couple the sensors with roadside units to solve connectivity problems. Hu *et al.* [14] proposes to deploy sensors across the 2nd ring road of Beijing (China) for road traffic monitoring. They influence the deployment so that the resulting topology conforms to a small world graph in order to take advantage of this type of structures. The article proposes to optimize transmission radiuses of the nodes and to refine the location of high coverage nodes using an evolutionary algorithm. CitySee [17] is a project to deploy a sensor network in the city of Wuxi (China) to measure the CO_2 level in realtime. The paper models the deployment issue as a relay node placement problem and evaluates the number of additional nodes deployed for connectivity purposes.

All these papers propose different deployment strategies, and the resulting connectivity graphs should be slightly different. In the literature, it is commonly assumed that city maps are scale-free networks. Besides, the complex networks analysis methods that are widely used in social networks analysis are also applied in urban networks [24, 4, 23]. However, the topology of the *network* deployed over a city infrastructure depends on the deployment method and this topology has a strong effect on the network protocols performance. Ishizuka and Aida [15] examine the effect of the sensor topology on fault tolerance and on the event detection probability. [28] study the performance of various congestion control algorithms for wireless sensor networks over simple topologies. [25] evaluate the impact of the topology on the data collection process in a sensor network. Ducrocq *et al.* [5] evaluate the impact of the network topology on geographic routing. All these studies concern different aspects of the communication process, but they unanimously conclude that the structural properties of the network has a strong impact on the algorithms performance.

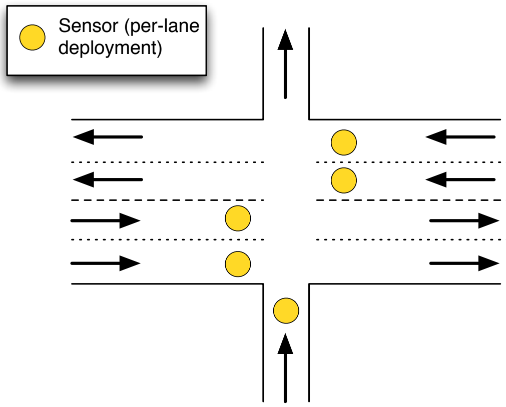
Yet, very few contributions really tried to propose realistic models of large scale urban sensor networks. [21] examines the topology of a vehicular network, hence a mobile network, in the city of Cologne (Germany). The authors show the weaknesses that vehicular protocols may encountered: mobility as an additional constraint has the effect of creating a very volatile and fragmented network. However, no contribution to our knowledge, has characterized the topology of a fixed distributed network of sensors and actuators that would be deployed and managed by the city itself, even though the applications of such networks for traffic lights and adaptive speed limits management is obvious.

3 Deploying Sensor Nodes in Cities

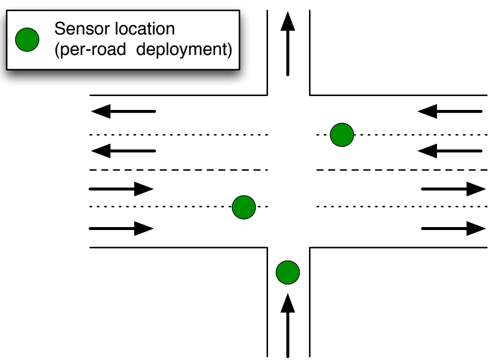
3.1 Basic Strategy

Let us begin by detailing the sensors deployment methods we assume. We begin by acquiring a city map that we suppose accurate enough to identify the intersections with and without traffic light, the traffic directions, the lanes and the distance between two intersections. Such data can be obtained from pubic geographic information systems such as *OpenStreetMap*.

The primary goal of the network we are building is to count vehicles to feed an intelligent transportation system. We did not assume any intelligent data correlation algorithm and we hence start by placing one sensor node at the end of each lane. In other words, we deploy, at each intersection, a number of sensors equal to the number of *incoming* lanes, as illustrated by the yellow dots on Fig. 1(a). We focus on lanes rather than roads because we have in mind magnetometer-like sensors which can accurately count vehicles passing over them.



(a) 1 sensor per lane (magnetometers)



(b) 1 sensor per road (cameras)

Figure 1: Sensors deployment strategies illustrated on an intersection between a major road (4 lanes) and a minor street (1 lane, 1 way).

We also considered an alternate strategy that consists in deploying one sensor per *road*, as illustrated on Fig 1(b). This strategy corresponds to the case where sensors are overhead cameras, backed up by a video analysis software, that are able to capture all lanes simultaneously. Both strategies are possible, and they differ on accuracy. The choice is beyond the scope of this article.

3.2 Reducing the number of sensors

Both strategies induce deploying a large number of sensors and an infrastructure cost that would be considered too high by city planners. To reduce the number of sensors without impairing the monitoring capability, we decided to avoid deploying sensors between two intersections that are too close. We tested different threshold values (from 10 m to 100 m) and found that the global number of sensors decreases linearly, as the threshold increases. We therefore choose a value that encompasses modern urbanism recommendations in application in France and in Quebec city and only positioned sensors on intersections that are more than 50 m away.

Figure 2 represents the number of nodes in six different representative cities that belong to our dataset. The *each lane* case, displayed in red, corresponds to the placement of one sensor per-lane at all intersections. The second (orange) bars represent the number of sensors that remain when removing one out of two close sensors. The third bars, in yellow, represent the scenarios in which one sensor is deployed per road rather than one per lane and the fourth (green) bars combine both optimizations. The fifth and final bars (in blue) represent the scenario in which only one sensor is deployed per intersection (e.g. with a fisheye camera).

3.3 Creating the connectivity graph

Each sensor positioning method produces a set of nodes, N, with geographic coordinates. We then create an undirected weighted graph $G = (N, E)$, whose edge set (E) is created by confronting inter-sensor distances to nodes transmission range. $E = \{(i, j, \delta)\}$ is a set of unordered weighted pairs of nodes, i and j, whose weight, $\delta \in]0 : 1]$ represents the strength of the connection (e.g. the wireless link quality).

The edges weight is calculated based on the *Sensys Networks VSN240* sensors[1] model, which are used on roads all around the world and can be deployed densely [11]. These nodes use a nominal output power of $0\ dBm$ and have a receiver sensitivity of $-95\ dBm$ in the 2.4 GHz band. We confront these values to a simplified propagation model that corresponds to a 2.4 GHz IEEE 802.15.4 network interface ([1, 18]). This model defines the path loss (in dB) across a distance of d meters as follows:

$$PL(d) = \begin{cases} 40.2 + 20\log_{10}(d), & 0.5m \leq d \leq 8m \\ 58.5 + 33\log_{10}(d/8), & d > 8m \end{cases} . \quad (1)$$

Note that this model, which simply defines a transmission range at this level of analysis, should fit most technologies that operate in the S-band (2 GHz to 4 GHz), which covers possible technologies and it can easily be adapted to other narrow frequency bands such as the 5.9 GHz band utilized by IEEE 802.11p (WAVE).

[1] http://www.sensysnetworks.com/products/sensor/

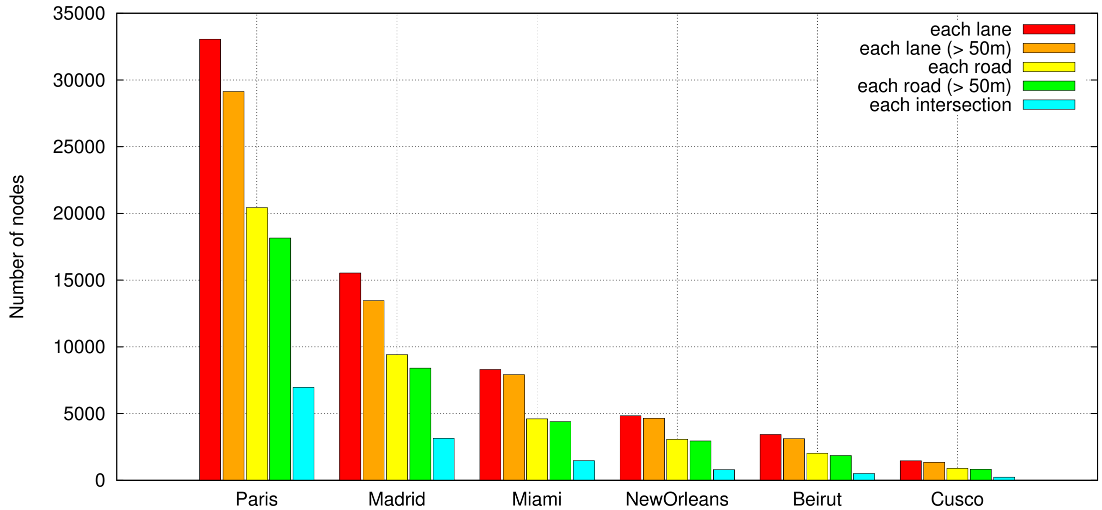


Figure 2: Comparison of the number of sensors deployed in different strategies on various scenarios.

3.4 Dealing with network partitioning

A graph can be partitioned into one or multiple *connected components*. A connected component is a sub-graph in which all couples of nodes are connected by a path, and such that no path exists between nodes that belong to the component and nodes outside the component. In our deployment, a connected component models a group of nodes that are connected together but disconnected from the rest of the network, as represented on Figure 3. On this figure, four intersections are monitored by one sensor per lane and the resulting graph is partitioned due to a too large distance between the left and right groups of sensors.

As we will see in section 6.1, the basic strategy described above produces a partitioned network with a high number of connected components. Network partitioning is not an issue *per se*, as the components can be interconnected together by a cellular network or by a metropolitan wired network. However, the number of independent network components should remain reasonable to limit the backbone complexity. That's why we also evaluate a strategy to interconnect close connected components that we detail in section 7, after analyzing the raw graphs.

4 Dataset creation

4.1 Method and tools

We applied the graph creation method described in section 3.1 on a set of 52 city maps extracted from *BBBike.org*[2], a service that offers to retrieve *OpenStreetMap*[3] maps data from more than 200 cities and regions worldwide. OpenStreetMap is an international project started in 2004 that intends on creating an open access map of the world. The maps have gone through several modifications, thanks to crowdsourcing, and are now accurate enough for navigation software [10].

In order to filter the information contained in these complete maps by removing elements that are not relevant to our study (e.g. bike lanes, pedestrian areas), we use *NETCONVERT*, a tool provided by the SUMO microscopic traffic flow simulator[4] [16] (version 0.19). The resulting maps are easier to utilize than their OpenStreetMaps counterparts. Besides, the SUMO simulator can easily be coupled to a network simulator like *OMNeT++*[5] [8]. In order to avoid overloading the network, we kept only the main and the secondary streets, as defined on the OpenStreetMap wiki[6].

4.2 Selection of 6 representative scenarios

We created the graphs for 52 cities that we extracted from the public database. Among these 52 scenarios, we selected 6 representative cities to illustrate our points. All the datasets, the results and the OMNeT++ models are however available online at `http://g.sfaye.com/`. The scripts to generate the graph, invoking the different tools in sequence with different configurable parameter (path-loss model, deployment method, etc.), is also available for use through a web interface and for download at

[2]`http://download.bbbike.org/osm/`

[3]`http://www.openstreetmap.org/`

[4]`http://sumo-sim.org/`

[5]`http://www.omnetpp.org/`

[6]`http://wiki.openstreetmap.org/wiki/Key:highway`

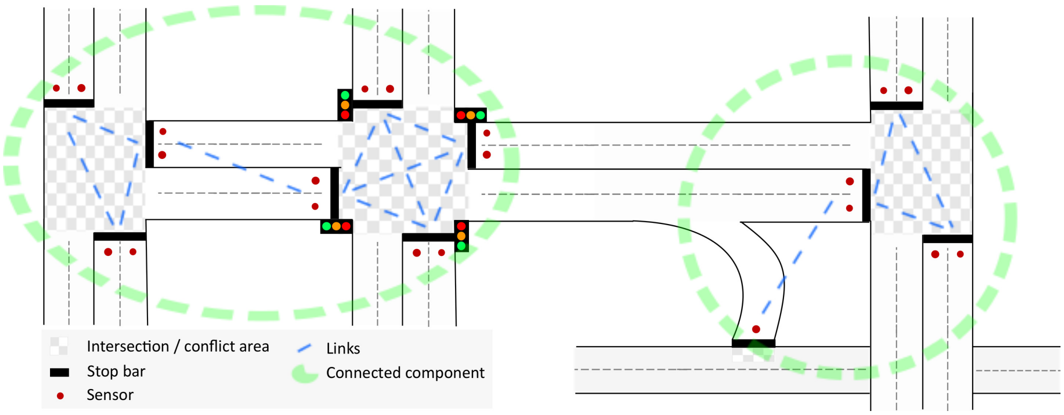


Figure 3: A 4-intersections network monitored by 1 sensor per lane

the same address.

The first property that influenced our choice is the area covered by the city. It is independent of the deployment method. We wanted to include large cities as well as small cities to account of the diversity of urbanism rules. The surfaces covered by all 52 cities are represented on Fig. 4. We selected the largest and the smallest cities: *New Orleans* and *Beirut* respectively.

The second selection criterion is the nodes density produced by the basic method. Density has a direct effect on network performance, as it influences collision probability at the medium access level, and congestion probability. Figure 5 represents the average number of nodes deployed per square kilometer in all 52 cities. We included the densest and the sparsest networks in our dataset: *Miami* and *Cusco* respectively.

Finally, we added to the dataset two more networks of average size and density, with different profiles of distribution of the inter-distances between intersections. Figure 6 represents the diversity of distributions we found in our datasets on a few scenarios. Analyzing all the graphs, we came across three main distribution profiles:

1. Most cities (37 in our dataset) exhibit a unimodal and asymmetric distances distribution skewed to the left. Miami (Fig.6(a)), Beirut (Fig. 6(b)) and Paris (Fig. 6(c)) belong to this category. This type of distribution indicates that these cities have a relatively uniform intersections repartition and density. This is typically the case for geometric cities (e.g. Miami), sparse cities (Beirut) or uniformly dense cities (Paris). The width of the peak gives an indication on how regular the city structure is. Its shift towards smaller values is more pronounced in denser road networks.

2. Some cities (11 in our dataset) show a bimodal distances distribution, like Madrid (Fig. 6(d)) or New Orleans (Fig. 6(e)). This usually means that the urban rules are different for the city center and for the peripheral area, and that the frontier between the two zones is abrupt.

3. Finally, a few distributions (4 in our dataset) are quasi-uniform. For example Bagdad (Fig. 6(g)) follows this type of profile and to some extent Cusco (Fig. 6(f)) does too, even though a few peaks appear at smaller distances. Note that if this type of distribution is independent of the city size or density, it is more frequently found for lower density networks.

Based on the inter-distance criterion, we added Madrid and Paris to our set of representative maps. Figure 7 represents the number of nodes deployed in all scenarios. In terms of network implementation, the network size has a direct influence on the addressing scheme and on the memory required for routing tables, as well as on the deployment cost. The 6 scenarios that we selected reflect quite well the diversity of the dataset, as we have large networks (Paris), small networks (Cusco) and average ones.

5 Connectivity graphs analysis

5.1 Nodes degrees

Figure 8 shows the average node degree for each network, i.e. the average number of other nodes that are within its transmission range. In terms of networking, node degree represents the number of contenders each node has to compete with for accessing the wireless channel. As a node has to share the channel bandwidth with all its neighbors, network planning should aim for a relatively low degree. Yet, a too small value is not desirable, as a fair degree offers path diversity and redundancy.

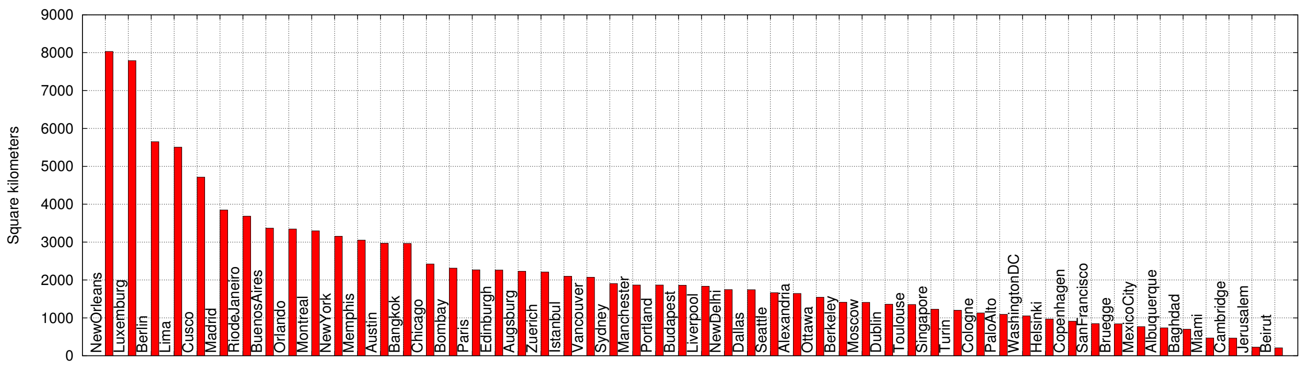


Figure 4: City size (km^2)

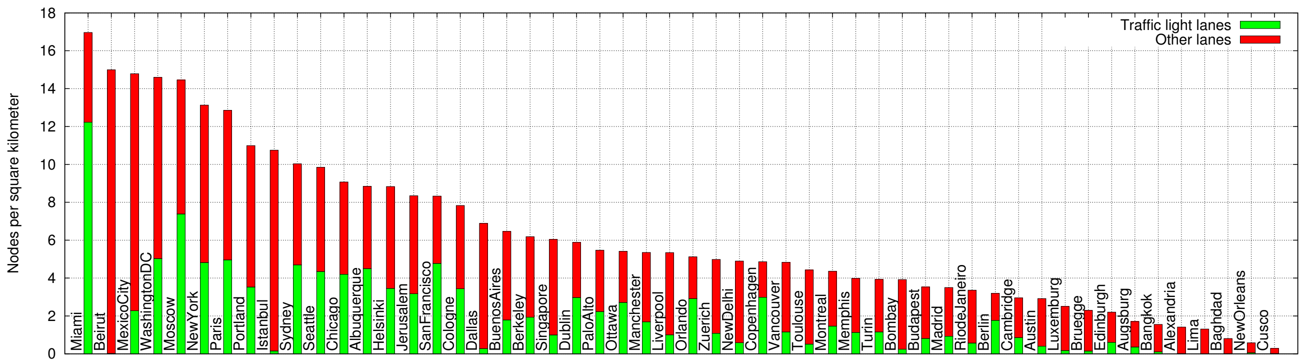
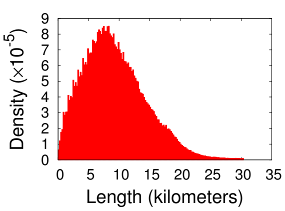
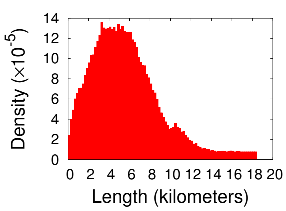


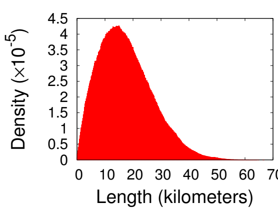
Figure 5: Networks density (nodes per square kilometer)



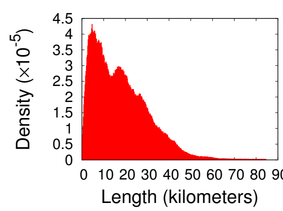
(a) Miami



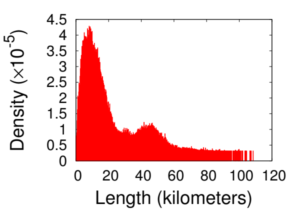
(b) Beirut



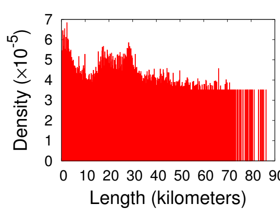
(c) Paris



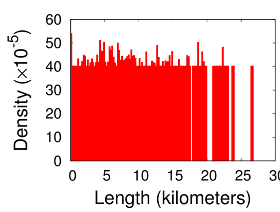
(d) Madrid



(e) New Orleans



(f) Cusco



(g) Bagdad

Figure 6: Distribution of the distances between intersections in representative cities.

Figure 8 shows that all graphs have a similar average degree that lies between 5 and 7 neighbors. Given the considerable amount of performance evaluations realized on various wireless technologies and considering the technological choices that standards (Bluetooth, Zigbee, etc.) usually make, this fits quite well the classical use case of today's wireless standards. We can notice that the average degrees of all networks are quite close from each other. Cities like Beirut, whose road network is relatively uniform, have a higher average degree than other

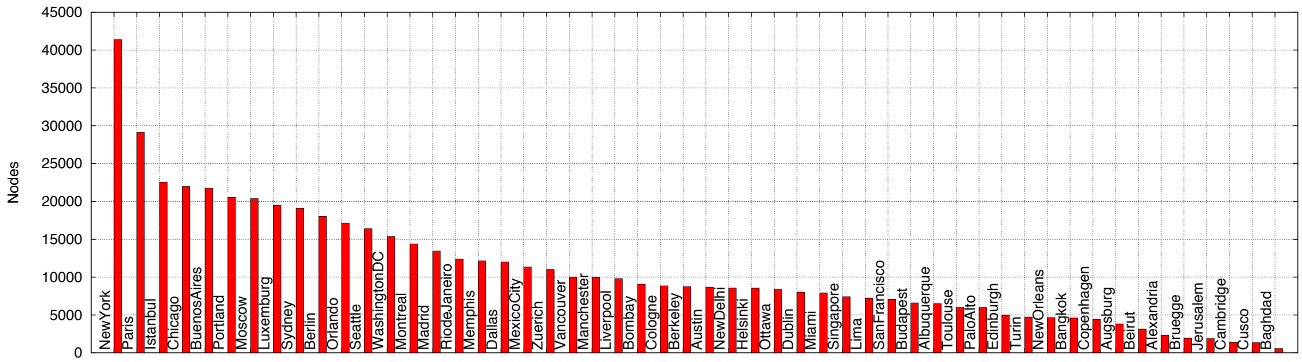


Figure 7: Number of nodes

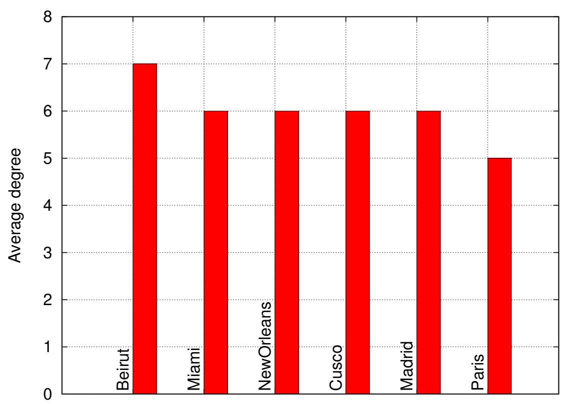


Figure 8: Average node degree

City	Shape (k)	Scale (θ)
Beirut	2.397899	3.312492
Cusco	0.9380117	6.666784
Madrid	1.347767	4.933085
Miami	2.183021	3.192106
New Orleans	2.380301	2.52304
Paris	2.26065	2.558382

Table 1: Gamma distributions parameters

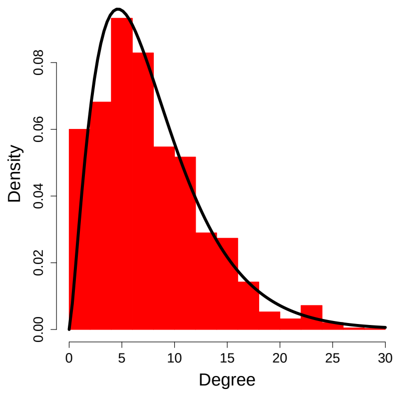
cities like Paris, for example, who have a wide suburban area.

The degree distribution is a classical measure to characterize large graphs. The bar graphs on Figure 9 represents the empirical degree distributions measured on the 6 chosen scenarios. We can first notice on these graphs that the maximum degree in the graph is relatively high (close 35 for Cusco and Madrid), but not too high. The empirical average and standard deviation are very different in all degree distributions, therefore these distributions cannot be fitted by Poisson distributions. They do not correspond to a power law distribution either, as the log-log representation of the degree distribution is far from linear, as shown on Figure 10 for Paris. We finally approximated the empirical degree distribution by a gamma distribution whose *scale* parameter (θ) is calculated, for each graph, as the ratio between the empirical variance (σ^2) and the empirical average (μ), and the *shape* parameter (k) is calculated as the ratio between the empirical average and the scale parameter: $\theta = \sigma^2/\mu$ and $k = \mu/\theta$.

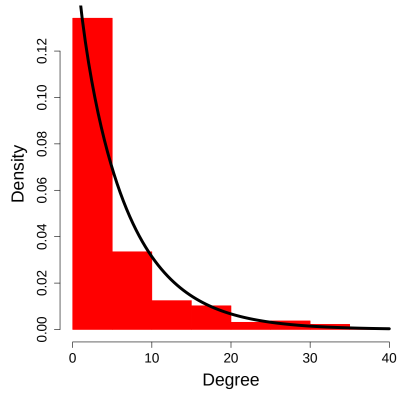
The fitted distributions are represented by the black curves on Fig. 9 and the values of the parameters are reported in Table 1. Except for a few high values (around degree 23 in the Beirut scenario e.g.), the fitting is quite accurate. Figure 11 represents the quantile vs. quantile plot that compares the empirical and fitted distributions. The closer the dots are of the diagonal line, the better the matching is. This graph tells us that the fitting using a gamma distribution is relatively accurate and deviates slightly for higher degrees. If the lowest values and the tail of the different distributions differ, their modes are all located around a value of 5 nodes, which gives a good idea on the expected congestion level in most parts of the network.

5.2 Clustering coefficient

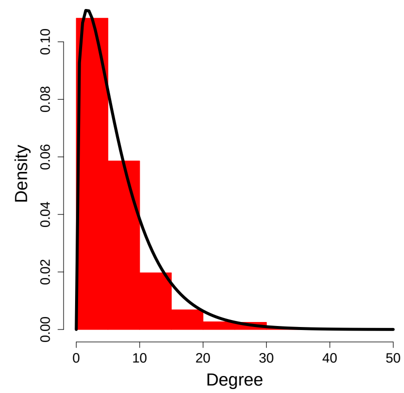
The *clustering coefficient* of a node in a graph is the probability that two neighbors of the node are themselves mutual neighbors. It accounts for the presence of communities in the graph and it is a classical measure to characterize graphs. Small world networks, that model several interaction graphs usually exhibit a high clustering coefficient. Figure 12 represents the CDF of the clustering coefficients in the different scenarios. Comparing these values with the network density, we have the confirmation that the graphs we generate do not possess the small world property.



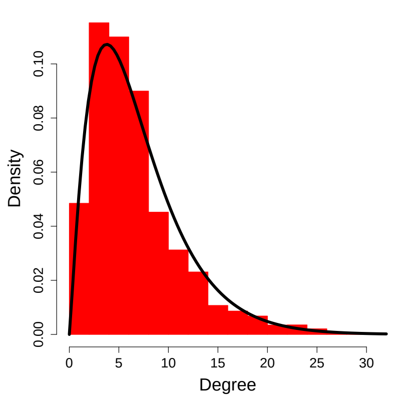
(a) Beirut



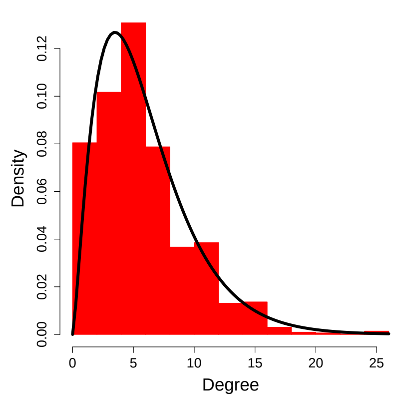
(b) Cusco



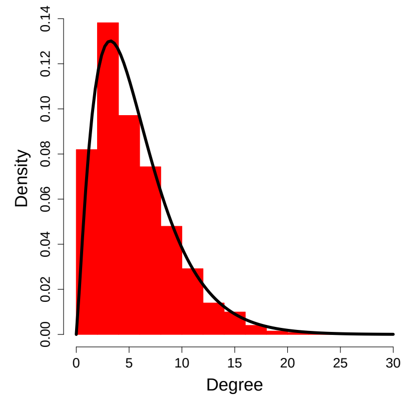
(c) Madrid



(d) Miami



(e) New Orleans



(f) Paris

Figure 9: Empirical and fitted degree distributions

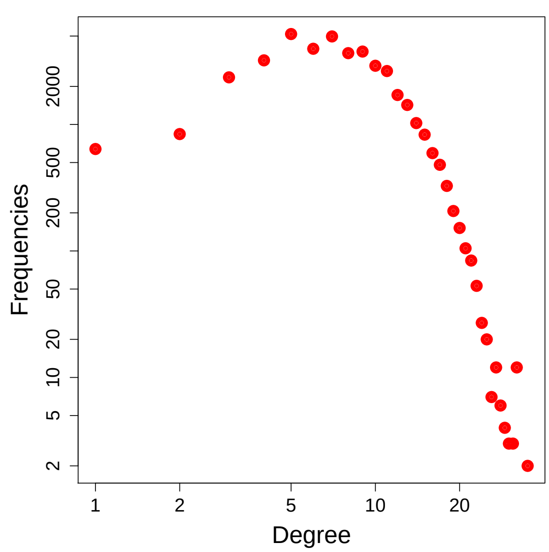


Figure 10: Degree distribution in the Paris scenario (log-log scale)

5.3 Creating random graphs

As the degrees tend to follow a gamma distribution, none of the state of the art random graphs model really fits this type of networks. The models from Gilbert [9] produces a degree distribution that corresponds to a binomial distribution. Erdös and Rényi [6] model generates graphs whose degree distribution follows a Poisson distribution, as well as the random geometric graph model [22], which is classically used to generate random wireless networks. The preferential attachment method proposed by Barabasi and Albert [2], as well as the Watts and Strogatz model [29] both produce scale-free networks whose degree distribution follows a power law.

Generating graphs that correspond to such deployments therefore requires to use models such as the Molloy and Reed [19] method that allows to use an arbitrary degree distribution. The experimenter should first decide of the type of city he wishes to generate and decide of the shape and scale parameters of the gamma distribution. Smaller shape values shift the distribution towards low degrees and hence model cities in which intersections are

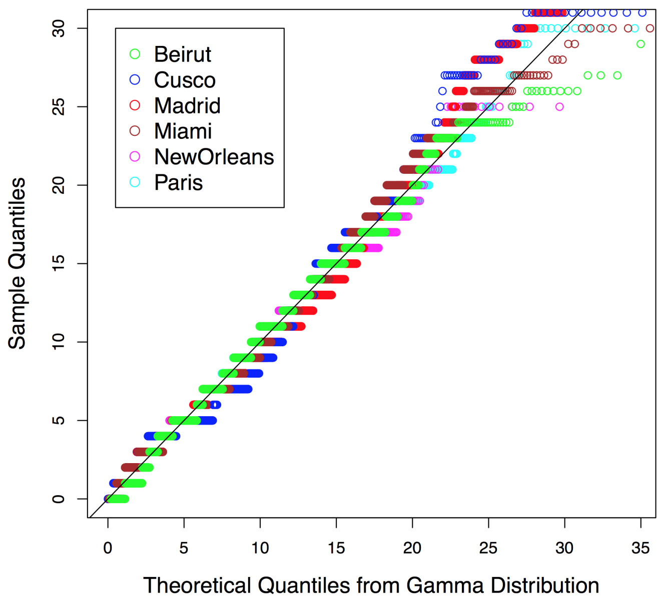


Figure 11: Quantile-Quantile Plot of Degree distributions vs. Gamma distributions

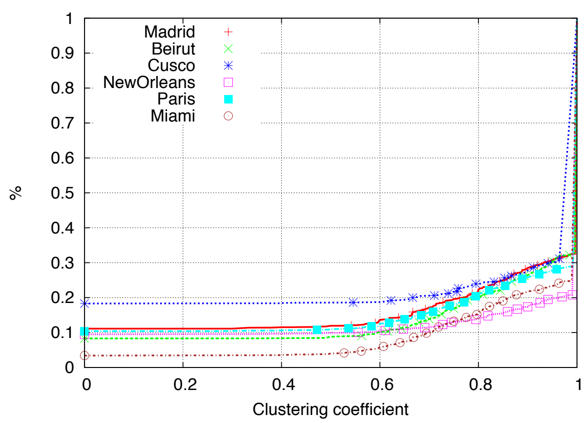


Figure 12: CDF of the clustering coefficient

far away from each other. The scale parameter defines the height of the peak and hence models how uniform the degrees will be. It accounts to some extent for the regularity of the distances between the intersections.

6 Network analysis

6.1 Connected components

The degree distribution accounts for local connectivity. To evaluate the global connectivity of the networks, we analyze the number of connected components that form the network. This either means that each of these areas will be fully autonomous and disconnected from the control center, or that it is necessary to deploy nodes or links solely for connectivity purposes. The network manager may choose to place, in each sub-network a gateway through a cellular network or through a metropolitan wired network, and/or to deploy additional sensors. We will study the effect of such an extension in section 7.

Figure 13 shows the number of connected components in the different networks. This number depends directly on the dimension of the different networks as well as on the number of nodes. We can see that Paris has more than 5 500 components, which reflects a highly fragmented network and show a real difficulty in implementing protocols that require the entire network without further interconnection.

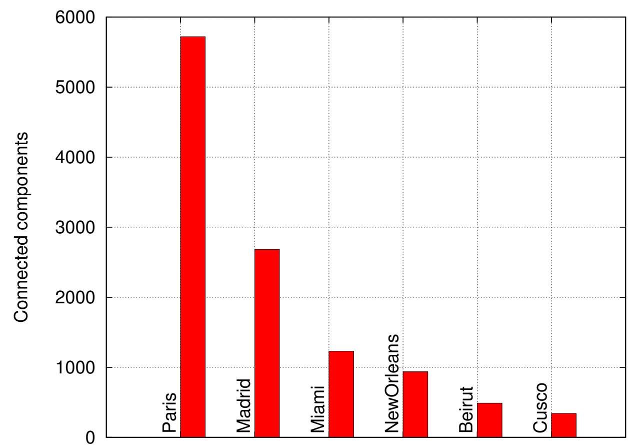


Figure 13: Number of connected components

Figure 14 shows the percentage of connected components that are only composed of one single node, i.e. the number of sensors who are too far away to be connected directly to the rest of the network through the same wireless technology. This proportion increases as the nodes density decreases. Cusco, for example has around 25% of connected components composed of a single node, while Paris has around 12.5%, which represents 750 nodes.

Figure 15 shows the CDF of the distance between a component and its neighbor component. To identify the closest components, we first compute the coordinates of the centroid of each connected component. This produces a set of points in the plan and we build the Voronoi diagram of this set of points. A Voronoi diagram separates the plan in zones centered on each node. A zone is composed of all the points that are closer to the central node than any other node. We then consider that two components are neighbors if their Voronoi cells have a common frontier. Figure 15 shows that a few components are very far from the rest of the network, but that most components are relatively close to each other, which indicates that reducing the number of components

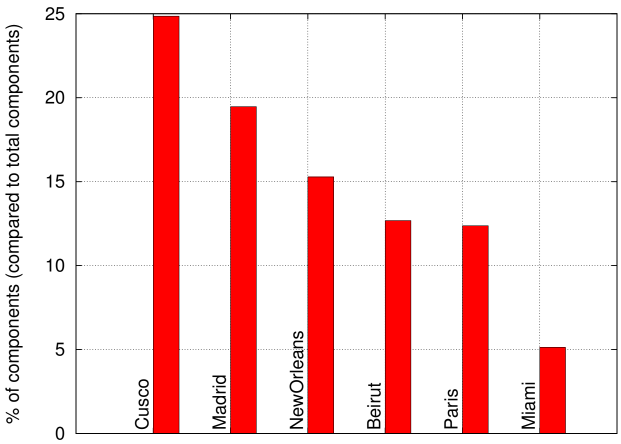
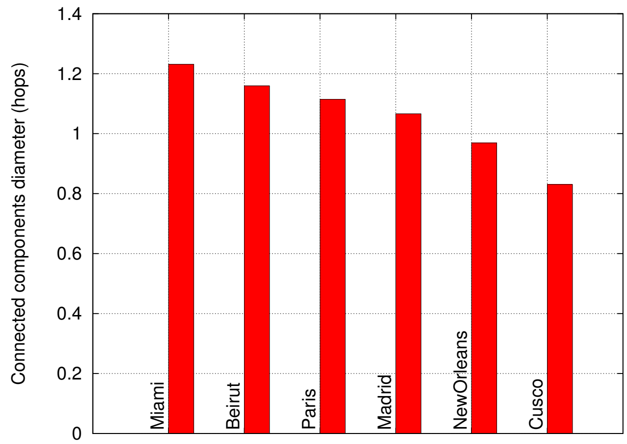


Figure 14: Percentage of isolated nodes



(a) Average

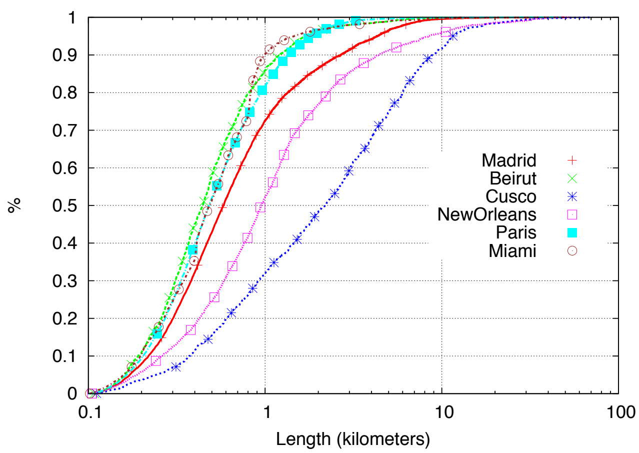
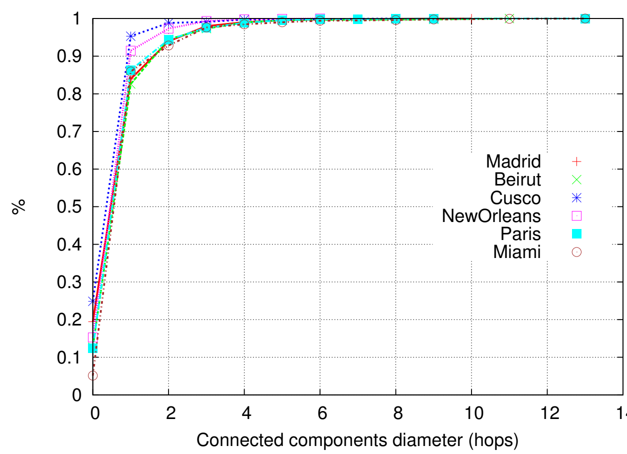


Figure 15: CDF of the distance between close connected components



(b) CDF

Figure 16: Connected components diameter

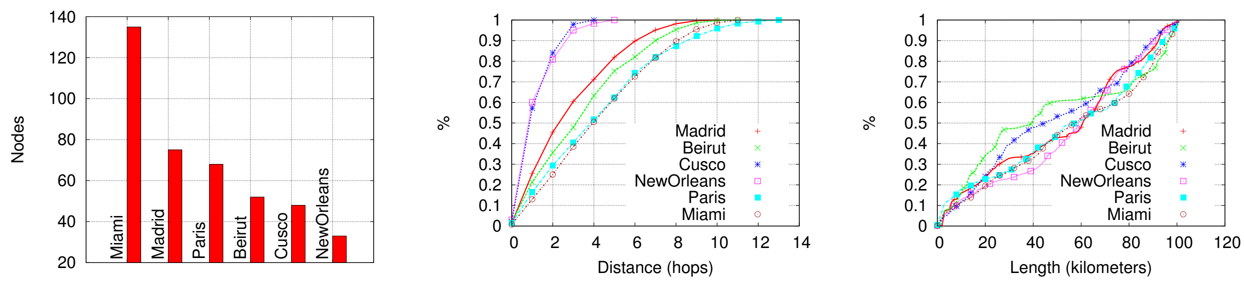
by inserting intermediate relays should be efficient.

6.2 Inside connected components

Figure 16(a) represents the average diameter of the connected components of each network. The diameter is the length of the longest of the shortest paths between couples of nodes that belong to the same component, expressed in number of hops. This distance accounts for the transmission delay between pairs of nodes within the connected component. It depends on the component size and, to a lower extent on the nodes density. We can see that this diameter remains very low, essentially due to the presence of several small sized components. Figure 16(b), which represents the CDF of the diameters, confirms this result. Networks are mainly composed of small-sized connected components and a few large ones.

Let us now focus on the *maximum connected component* (*max-component*), which is the connected component that contains the highest number of nodes. Figure 17(a) represents the number of nodes that belong to this max-component and ranges from 33 nodes (New Orleans) to more than 130 nodes (Miami). Nodes that belong to the same connected component can be seen as belonging to the same broadcast domain, hence this figure gives an indication on the cost of broadcasts and on how many nodes can be reached by control packets (ARP, routing protocols, etc.).

Figure 17(b) shows the CDF of the distances (in number of hops) that separates couples of nodes within this max-component. It gives an indication on the delays. We can see here that the distributions range from low diameter components (about 4 hops) to larger components (10 hops) and that these distributions do not always follow the trend defined by the size of the component, or from the average density. Madrid, for example, is sparser than



(a) Number of nodes

(b) CDF of the hop distance between two nodes

(c) CDF of the edge length

Figure 17: Analysis of the max-component

Paris (Fig. 5) but its maximum connected components has shorter path for a comparable number of nodes. The answer lies in the inter-distances distribution (Fig. 6): the mode of the Paris distribution is located at a higher distance than the first mode of the Madrid distribution, which indicates that the intersections density is higher in downtown Madrid than in Paris.

Figure 17(c) shows the CDF of the edge lengths. This parameter is important for evaluating the attenuation on the wireless links and hence the links quality or the expected number of transmissions. The distribution is globally uniform, as the CDF is almost linear for all networks. Differences come from the architectural specificities of the cities.

6.3 Reliability analysis

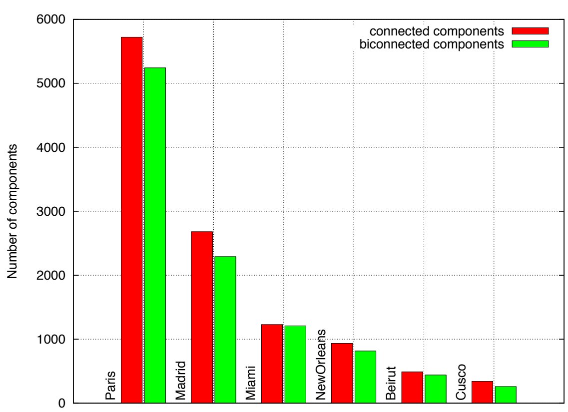


Figure 18: Number of connected and bi-connected components

Figure 18 compares the number of connected components (red bars) with the number of *biconnected components* (green bars) in each network. A biconnected component is a connected component in which there are at least two paths between each couple of nodes. It reflects the proportion of sub-networks that can tolerate any single node failure. The values show that there is only a small proportion of the sub-networks who exhibit such structural weakness and that few additional deployments will be required to comply to the classical N-1 reliability criterion.

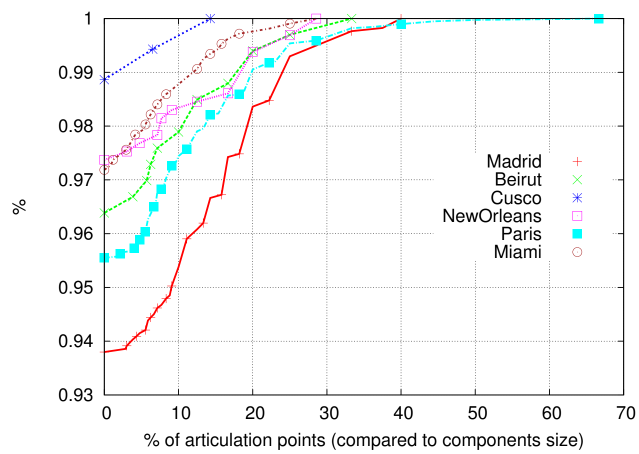


Figure 19: CDF of the number of articulation points (for each component composed of at least 3 nodes)

Figure 19 shows the cumulative distribution function (CDF) of the number of components that have a given percentage (represented on the X-axis) of their nodes that are *articulation points*. An articulation point is a node whose removal disconnects the component it belongs to, increasing the number of connected components. This definition implies that the graph only concerns components formed by at least 3 nodes.

First, the graphs are fairly redundant, as in the worst case (Madrid), almost 94 % of the components have no articulation point. In the worst case, a connected component had 50 % of its nodes who are critical for connectivity, which corresponds to a chain of nodes. A network like Paris, for example, tends to have a large number of articulation points, as the suburban area is large. Madrid has the same characteristics as the city of Paris, without the scattered suburbs, but with several areas of high density around the city centre. In this case, the rise in the rate of articulation points is not as sudden.

7 Improving connectivity

The analysis in the previous section was conducted on "raw" graphs, created by only positioning sensors that had a monitoring role. As no effort was made to improve connectivity, these graphs are composed of many connected components and an operator willing to acquire data or to disseminate policies across its whole network shall interconnect these components.

In this section, we examine the effect of such an interconnection strategy that relies on the insertion of relay nodes that we suppose identical to the sensor nodes. These relay nodes are positioned in order to merge two connected components. We define the distance that separates two arbitrary connected components as the minimum of the distance between couple of nodes that belong to each component. Depending on this distance, we would need one or more intermediate relays to merge both sub-graphs.

We supposed that the operator imposes a limit on the maximum number of intermediate nodes that could be deployed for interconnection purposes between two components and study the effect of setting this limit from 1 to 10 relays. Indeed, a value of 10 is most unlikely, as it would result in relying on chains of 10 nodes to interconnect components, knowing that the failure of any of these nodes would result in partitioning the component.

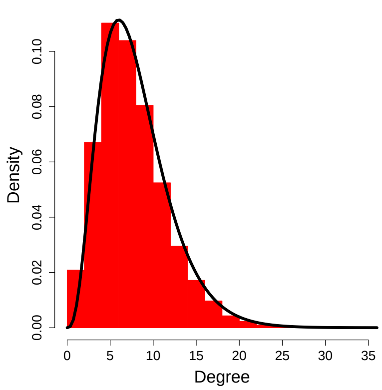


Figure 20: Paris. Empirical and fitted degree distributions when adding at most 4 relays between connected components

Fig. 20 shows the updated degree distribution for the Paris network. All scenarios go through similar evolution. We can see that the gamma distribution fitting is still valid. Table 2 shows the new gamma distribution parameter values for the 6 representative scenarios when 4 relays at most are added. Comparing these values with Table 1, we can notice globally a decrease in the scale parameter and an increase in the scale parameter. This indicates that the resulting networks are less uniform due to the presence of chains of nodes. The distribution also shifts towards the right value, which means that the average degree increases.

City	Shape (k)	Scale (θ)
Beirut	3.463977	3.030837
Cusco	1.222276	5.911105
Madrid	2.321423	3.584129
Miami	3.682748	2.584144
New Orleans	2.929175	2.5145
Paris	3.955912	2.024664

Table 2: Parameters of the fitted gamma distributions in the improved connectivity deployment

Figure 21(a) represents the evolution of the number of connected components. The X-axis value of -1 represents the inverse situation in which the articulation points in the graph (see Sec. 6.3) are removed. We can see that inserting a single relay has a limited impact, while increasing the threshold to 2 or 3 has a notable in very scattered graphs. All values seem to converge to comparable values around 200components.

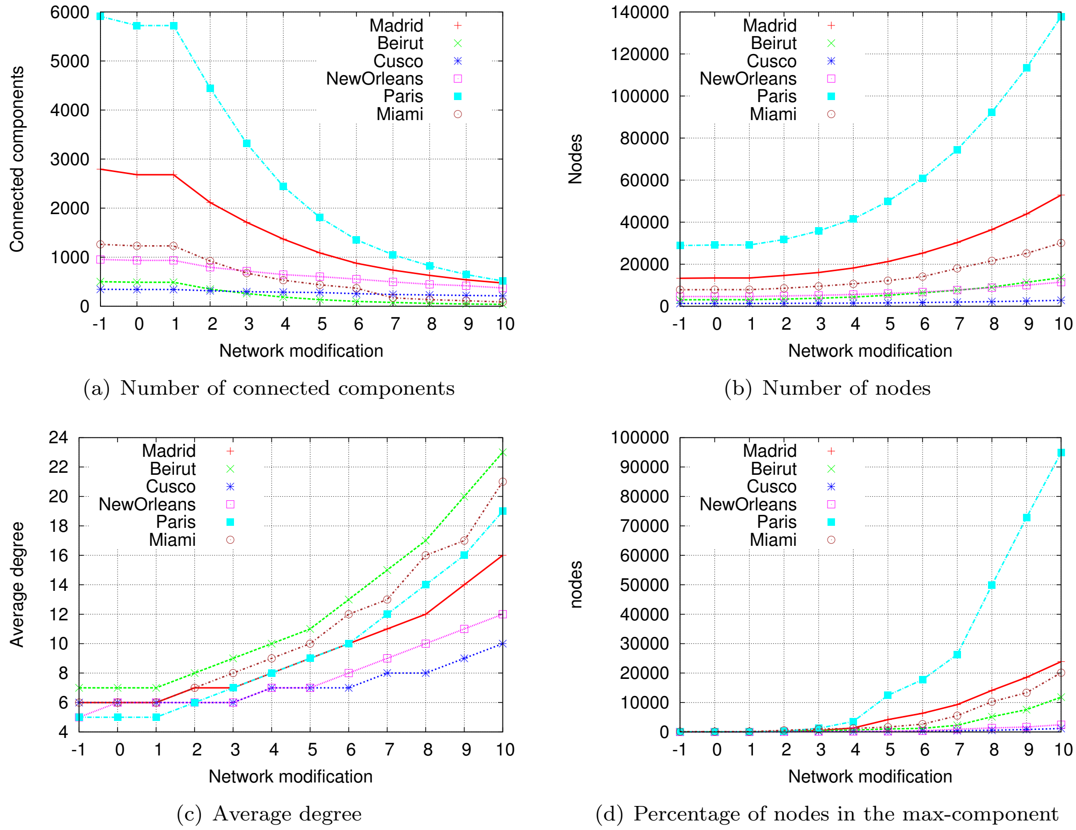
Figure 21(b) represents the evolution of the number of deployed nodes with the value of the threshold. We can notice that the value tends to increase faster and faster in scattered networks, as the reduction of the number of components slows down. In the case of Paris – the network with the most components – we need to add around 60 000 nodes to obtain less than 1 000 connected components. This indicates a strong diminishing returns effect.

Figure 21(c) represents the evolution of the average degree in the graph. It shows that the effect of this nodes addition on the degree is very different depending on the network initial density (see Fig. 5). Networks with higher initial density (Beirut, Miami, Paris) see a steepest increase in the nodes degrees than others.

Finally, Figure 21(d) shows the evolution of the number of nodes that belong to the maximum connected component. This graph shows that even though the improvement is not the same for all cities, this component is able to gather up to 90 % of the nodes.

8 Conclusions and future works

In this paper, we examined a strategy to deploy a sensor network at the intersections of various cities. We



(a) Number of connected components

(b) Number of nodes

(c) Average degree

(d) Percentage of nodes in the max-component

Figure 21: Improving the connectivity of the networks

presented the graph generation method we used, that is based on operational constraints, and analyze the resulting graphs. We show that the classical random graphs models do not model these networks, whose degree distribution corresponds to a gamma distribution. We propose hints to generate random distributions and show that the resulting graph is highly partitioned and comprises up to 25 % isolated node. However, the resulting network presents a good redundancy level. The average diameter of connected component is low, but can rise to fair values. Looking at the distances between close connected components, we motivate a method to insert relays between connected components to increase the network connectivity.

The effect on various network protocols and algorithms remains to be evaluated for example through simulation. However, the few conclusions that we draw in this article should help selecting the most appropriate protocols for this class of scenarios. Besides, for a particular setup, the tools we developed, which are available online under the LGPL licence, allow to create models for various deployment strategies based on a real city map. OMNeT++ models can be directly generated and simulations using, for example, the MiXiM framework, can help city planners evaluate and compare protocols and algorithms.

In future works, we intend on studying formally correlations between geographic parameters and network graph parameters to improve the graph generation method we sketched. We also intend bringing the analysis to the networking level by comparing state of the art protocols and algorithms using simulation tools.

References

[1] Coexistence of wireless personal area networks with other wireless devices operating in unlicensed frequency bands. *IEEE Standard 802.15.2-2003*, 2003.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439), Oct. 1999.

[3] I. Corredor, A. García, J. Martínez, and P. López. Wireless sensor network-based system for measuring and monitoring road traffic. In *6th Collaborative Electronic Communications and eCommerce Technology and Research (CollECTeR 2008)*, Madrid, Spain, June 2008.

[4] P. Crucitti, V. Latora, and S. Porta. Centrality measures in spatial networks of urban streets. *Physical Review E*, 73(3), 2006.

[5] T. Ducrocq, M. Hauspie, N. Mitton, and S. Pizzi. On the impact of network topology on wireless sensor networks performances illustration with geographic routing. In *International Workshop on the Performance Analysis and Enhancement of Wireless Networks (PAEWN)*, Victoria, Canada, May 2014.

[6] P. Erdös and A. Rényi. On random graphs i. *Publicationes Mathematicae*, 6, 1959.

[7] S. Faye, C. Chaudet, and I. Demeure. A distributed algorithm for multiple intersections adaptive traffic lights control using a wireless sensor networks. In *UrbaNe Workshop (held in conjunction with ACM CoNEXT 2012)*, Nice, France, Dec. 2012.

[8] S. Faye, C. Chaudet, and I. Demeure. Influence of radio communications on multiple intersection control by a wireless sensor network. In *ITS Telecommunications (ITST), 2013 13th International Conference on*, 2013.

[9] E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30, 1959.

[10] M. Haklay. How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. *Environment and planning. B, Planning & design*, 37(4), 2010.

[11] A. Haoui, R. Kavaler, and P. Varaiya. Wireless magnetic sensors for traffic surveillance. *Transportation Research Part C: Emerging Technologies*, 16(3), 2008.

[12] B. Horn. "suppressing traffic flow instabilities". In *IEEE Intelligent Transportation Systems Conference (ITSC 2013)*, Den Haag, Netherlands, Oct. 2013.

[13] B. K. Horn. Suppressing traffic flow instabilities. In *IEEE Intelligent Transportation Systems Conference (ITSC 2013)*, Den Haag, Netherlands, Oct. 2013.

[14] J. Hu, J. Song, M. Zhang, and X. Kang. Topology optimization for urban traffic sensor network. *Tsinghua Science & Technology*, 13(2), 2008.

[15] M. Ishizuka and M. Aida. Performance study of node placement in sensor networks. In *24th International Conference on Distributed Computing Systems Workshops (ICDCS)*, Tokyo, Japan, 2004.

[16] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4), December 2012.

[17] X. Mao, X. Miao, Y. He, X.-Y. Li, and Y. Liu. CitySee: Urban CO2 Monitoring with Sensors. In *Infocom*, 2012.

[18] K. Marquess. Physical model sub-group discussion and questions. *IEEE 802.15/138R0*, 1999.

[19] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6(2-3), 1995.

[20] R. N. Murty, G. Mainland, I. Rose, A. R. Chowdhury, A. Gosain, J. Bers, and M. Welsh. Citysense: An urban-scale wireless sensor network and testbed. In *Technologies for Homeland Security, 2008 IEEE Conference on*, 2008.

[21] D. Naboulsi and M. Fiore. On the instantaneous topology of a large-scale urban vehicular network: the cologne case. In *Proceedings of the fourteenth ACM international symposium on Mobile ad hoc networking and computing*, 2013.

[22] M. Penrose. *Random Geometric Graphs*. Oxford Studies in Probability, 2003.

[23] S. Porta, P. Crucitti, and V. Latora. The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications*, 369(2), 2006.

[24] S. Porta, P. Crucitti, and V. Latora. The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design*, 33, 2006.

[25] D. Puccinelli, O. Gnawali, S. Yoon, S. Santini, U. Colesanti, S. Giordano, and L. Guibas. The impact of network topology on collection performance. In *8th European conference on Wireless sensor networks*, 2011.

[26] D. Robertson and R. Bretherton. Optimizing networks of traffic signals in real time-the scoot method. *IEEE Transactions on Vehicular Technology*, 40(1), Feb. 1991.

[27] A. Sims and K. Dobinson. The sydney coordinated adaptive traffic (scat) system philosophy and benefits. *IEEE Transactions on Vehicular Technology*, 29(2), May 1980.

[28] V. Vassiliou and C. Sergiou. Performance study of node placement for congestion control in wireless sensor networks. In *WSN Workshop, held in conjunction with NTMS 2009*, 2009.

[29] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 1998.

[30] K. M. Yousef, J. N. Al-Karaki, and A. M. Shatnawi. Intelligent traffic light flow control system using wireless sensors networks. *Journal of Information Science and Engineering*, 26(3), May 2010.

[31] B. Zhou, J. Cao, and H. Wu. Adaptive traffic light control of multiple intersections in wsn-based its. In *73rd IEEE Vehicular Technology Conference (VTC Spring)*, 2011.

[32] F. Zou, B. Yang, and Y. Cao. Traffic light control for a single intersection based on wireless sensor network. In *9th International Conference on Electronic Measurement & Instruments (ICEMI 2009)*, Beijing, China, Aug. 2009.

Dépôt légal : 2014 – 2e trimestre
Imprimé à Télécom ParisTech – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

Télécom ParisTech

Institut TELECOM - membre de ParisTech

46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.telecom-paristech.frfr

Département INFRES