Connectivity Analysis of Wireless Sensor Networks Deployments in Smart Cities

Sébastien Faye*, Claude Chaudet†

***Abstract*—Recent advances in the field of intelligent transportation systems have focused on the use of wireless networks to link vehicles and road infrastructure. Applications that might result from such networks range from the adaptive management of traffic lights to the detection of traffic jams and accidents. Whatever the case may be, it seems important to explore the possibilities and limitations of such networks, which the literature often portrays in a somewhat idealistic way (e.g. no packet loss, fully connected sensors, etc.). In this paper, we study the deployment of wireless sensor networks at intersections in some of the world's major cities and characterize their topologies. Using a propagation model that corresponds to a 2.4GHz IEEE 802.15.4 network interface, we focus our study on the global connectivity of graphs resulting from different networks. By deploying this type of network over 52 city and region maps extracted from OpenStreetMap, we show that cities can reasonably be classified into three network structure categories of low connectivity (i.e. a high number of connected components) and that it should be feasible to improve the networks by adding sensors. All the tools and the complete dataset are freely available online.**

***Index Terms*—Smart Cities, WSN, Network Topology, Graphs**

I. Introduction

Smart cities constitute an active research domain, driving experimental projects. The success of distributed systems allows us to consider the deployment of sensor networks over large areas such as cities. The use of wireless communication links between devices allows them to be quickly installed on the roads and thereby open the way to the gradual introduction of intelligent transportation systems. These devices are powerful enough to organize themselves and decide a policy locally [1]. Rather than wasting time communicating to a central entity that can be compromised, saturated, or simply not reactive enough to process requests within a workable time frame [2], this latter approach allows for the creation of fully autonomous areas with the ability to respond rapidly to sudden, unexpected events such as accidents. Finally, in addition to providing a fixed infrastructure, this type of deployment allows the development of vehicular networks which at present are difficult to integrate [3].

Selecting a real deployment scenario – as well as algorithms and protocols – requires to study the characteristics of the network topologies. In [4], we studied the deployment of a wireless sensor network over 52 cities whose maps had been extracted from OpenStreetMap. Based on graph theory and the IEEE 802.15.4 communication standard, we concluded that traditional graph models are not able to represent these networks, whose degree distribution follows a gamma distribution accurately enough to be able to generate random graphs approaching it. In addition, the graphs appear to be highly partitioned and comprise a large number of isolated nodes. Here, we want to go further by focusing on: (1) identifying network categories in our data set; (2) studying global network partitioning; (3) studying the maximum connected components in order to deduce properties on the most covered areas; and (4) improving the connectivity of these networks.

After describing a state-of-the-art in Section II, we briefly recall our deployment strategy and present the tools we used in Section III. The following sections offer a subsequent analysis of the structure of the networks and of their partitioning (Sec. IV and V). Finally, in Section VI, we discuss a strategy for improving connectivity.

II. Related Works

Sensor networks experimental platforms are legion today, but most of them are limited to one or a few buildings (e.g. FlockLab [5]). In contrast, CitySense [6] is an urban wireless network testbed deployed all over the city of Cambridge (MA, USA), forming a mesh network of 100 Linux-based computers. Even though the primary focus was to foster mesh networks applications development, nodes have been augmented with environmental and pollution sensors. Corredor *et al.* [7] 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.* [8] proposes to deploy sensors across the second ring road of Beijing (China) for road traffic monitoring. They influence the deployment so that the resulting topology forms a small world graph to take advantage of this type of structures, by optimizing transmission radiuses of the nodes and refining the location of high coverage nodes using an evolutionary algorithm. CitySee [9] is a project to deploy a sensor network in the city of Wuxi (China) to measure the carbon dioxide level in real-time. The paper models the deployment issue as a relay node placement problem and evaluates the number of additional nodes deployed for connectivity purposes. Some authors in the literature define the deployment of traffic light control algorithms that act locally on each intersection of a road infrastructure [10], [11], [1]. Their algorithms are based

* Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 4 rue Alphonse Weicker, L-2721 Luxembourg (e-mail: sebastien.faye@uni.lu).

† Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI UMR 5141, Paris, France (e-mail: claude.chaudet@telecom-paristech.fr).

978-1-4673-9907-4/15/$31.00

on sensors deployed at an intersection for the purpose of calculating a timed sequence of green lights corresponding to the level of traffic. By defining the roles and hierarchy of the sensors, [1] uses communications between the adjacent intersections to create green waves (paths of successive green lights). This process has the potential to be used in large urban areas and studying the resulting graph would enable us to verify at which level.

All these papers propose different deployment strategies, and the resulting connectivity graphs are expected to 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 (e.g. [12]). 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 at all levels of the communication process [13]. Indeed, the network density has an effect on local congestion and on nodes energy consumption. A dense network is very challenging for the medium access control layer. It generally utilizes poorly the channel capacity, but it provides diversity that contributes to fault tolerance. The median end-to-end delay increases with the network diameter, which depends on the number of deployed nodes, but also on the effort made to reduce network partitioning. A partitioned network, on the other hand, requires cellular or wired gateway to let autonomous clusters exchange information, which influences the traffic patterns. Finally, 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.

III. Deployment Strategy

In the scenario we imagine, a city operator wishes to deploy sensors to monitor all the intersections of an urban road network to count vehicles and feed an intelligent transportation system. The results presented in this paper assume that the sensors are deployed individually on each *incoming* lane, as illustrated by the yellow dots on Fig. 1. We focus on this strategy because we have in mind magnetometer-like sensors which can accurately count vehicles passing over [14], and which is one of the most popular deployment.

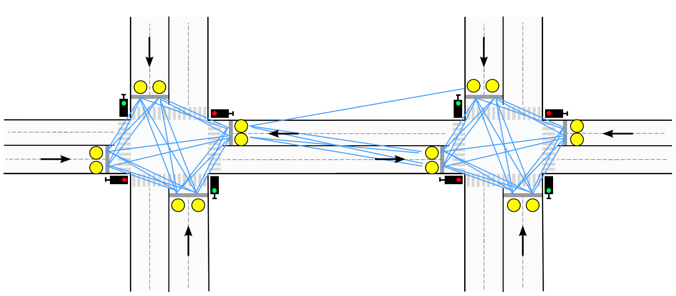


Fig. 1. One sensors deployment strategy on two intersections

In order to study the topology of the network formed by this type of deployment, we chose to extract 52 maps of different cities from *OpenStreetMap* maps data, using *BBBike.org*. These maps give the GPS coordinates of each intersection, as well of the characteristics of the roads that connect these intersections. Note that certain selected maps are not strictly confined to the boundaries of the cities: sometimes, it may be a region around the city, including the outskirts (e.g. Paris and its suburbs). In order to filter the information contained in these 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 (0.19) microscopic traffic flow simulator [15]. In order to avoid overloading the network, we kept only the main and the secondary streets[1]. We thus eliminate roads that are inaccessible to motorized vehicles, as well as a number of minor roads (e.g. residential areas).

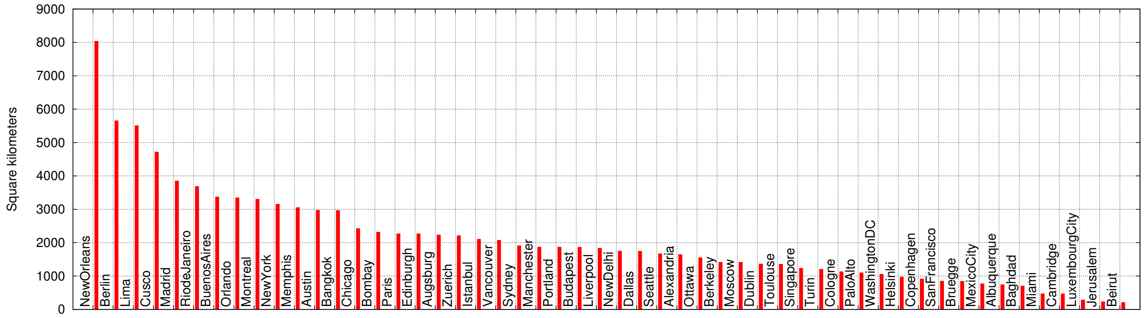
The network formed on each map can be described as a set of nodes that possess geographic coordinates and a set of associated undirected edges, created by confronting the Euclidean distances between each couple of sensors to a distance modeling the nodes transmission range of a 2.4 GHz IEEE 802.15.4 network interface [16]. This scenario, as well as others with different levels of accuracy, is described in [4]. Moreover, the full dataset comprising the 52 city maps, the results for each strategy and the scripts to generate the graphs are available online at `http://g.sfaye.com/`. These scripts invoke the different tools in sequence with configurable parameters (path-loss model, deployment method, etc.) and also generate OMNeT++ simulation models [17], [1].

IV. Network Structures

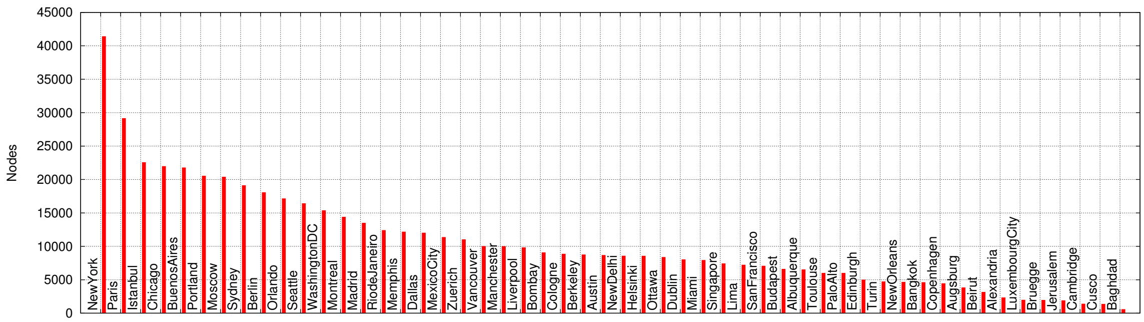
Analyzing all the graphs, we came across various distribution profiles. Fig. 3 represents three categories of inter-sensors distances distributions we found in our datasets, a criterion which reflects the morphology of road structures. Most cities seem to exhibit a *unimodal and asymmetric* distances distribution skewed to the left, as it is the case for Paris and Madrid (Fig. 3(a)-3(c)). This type of distribution indicates that these cities have a relatively uniform intersections repartition and density. 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. Some other cities show a *bimodal* distances distribution, like New Orleans (Fig. 3(b)). Finally, some distributions with a low density of sensors are almost *uniform* and present profiles whose representation deviates from that of the main two, for example Bagdad 3(d).

Among the 52 cities that constitute our full dataset, we selected 6 representative ones to illustrate our analysis. The first property that influenced our choice is the area covered by the city. We wanted to include large cities as well as (relatively) small cities to account of the diversity of urbanism rules. As illustrated in Fig. 2(a), we chose to include the largest and the smallest cities within the set of representative cities: *New Orleans* and *Beirut* respectively. The second selection criterion is the nodes density, i.e. the number of nodes –

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

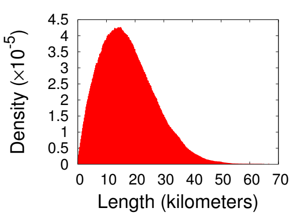


(a) City size (km^2)

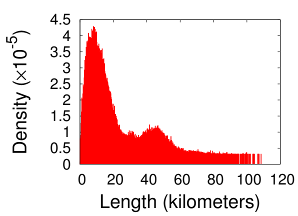


(b) Number of nodes

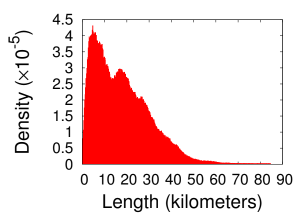
Fig. 2. Scale metrics for the 52 cities and regions that compose our dataset



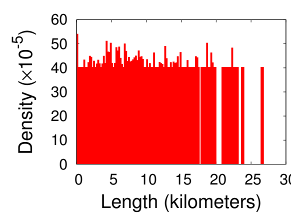
(a) Paris



(b) New Orleans



(c) Madrid



(d) Bagdad

Fig. 3. Profiles generated from the inter-sensors distances distributions

Fig. 2(b) – per city size). Density has a direct effect on network performance, as it influences collision and congestion probability at the medium access level. 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: *Madrid* and *Paris* because of the characteristics of their inter-sensors distances distributions.

V. Connectivity graphs analysis

A. Connected components and network partitioning

To evaluate the global connectivity of the networks, we analyze its partitioning by looking the number of *connected components* in the resulting graph. A connected component models a group of nodes that are connected together, but disconnected from the rest of the network. Red bars on Fig. 4 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. Paris has for example more than 5,500 components for 29,000 nodes. This means that the network, without additional relays, is composed of many areas and hence has limited interaction possibilities. Green bars on Fig. 4 shows the number of *biconnected components* in each network. A biconnected component is a connected component in which there are at least two node-disjoint paths between each couple of nodes. It reflects the proportion of sub-networks that can tolerate any single node failure. Note that the biconnected components, as reported here, are created from connected components formed by at least three nodes. Under three nodes, there is only one path between the nodes. Note also that a connected component may include several biconnected components. Fig. 4 indicates that relatively few additional nodes need to be deployed to comply with the classical N-1 reliability criterion (i.e. the loss of any single element does not break a connected component in two).

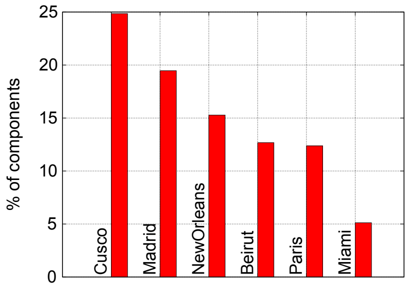


Fig. 5. Percentage of isolated nodes

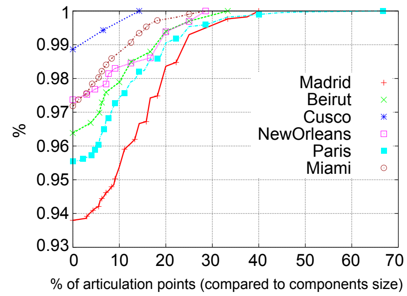


Fig. 6. Articulation points (CDF)

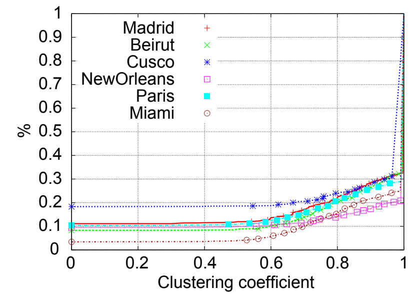
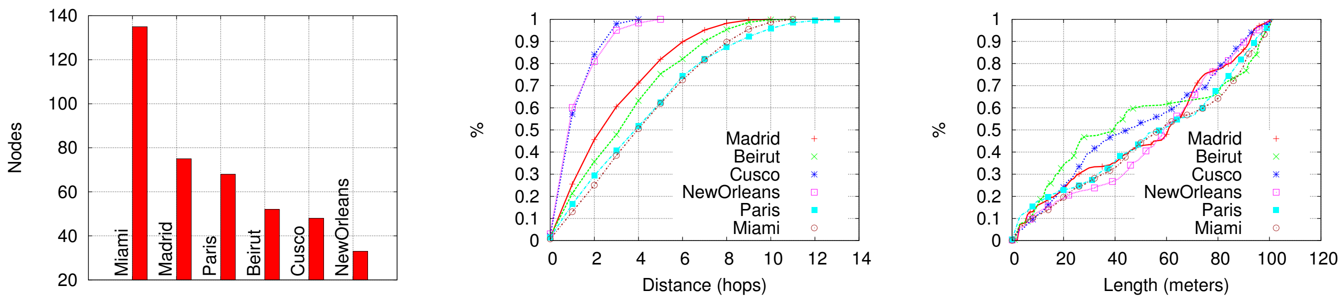


Fig. 7. Clustering coefficient (CDF)



(a) Number of nodes

(b) Hop distance between two nodes (CDF)

(c) Edge length (CDF)

Fig. 8. Analysis of the maximal component

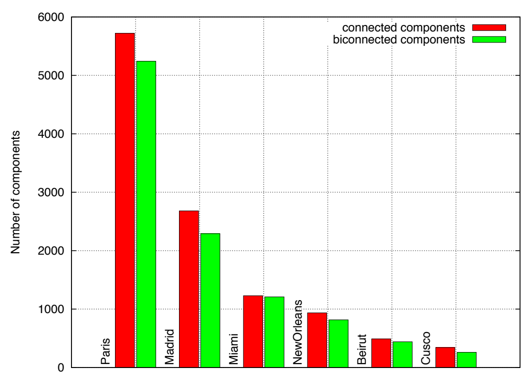


Fig. 4. Number of connected and bi-connected components

At a more detailed level, Fig. 5 shows the percentage of connected components that are composed of only 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 wireless LAN technology. This proportion naturally 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%. Fig. 6 shows the cumulative distribution function (CDF) of the number of components formed by at least 3 nodes 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. The graphs are fairly redundant: in the worst case (Madrid), almost 94 % of the components have no articulation point. 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 center. In this case, the increase rate of articulation points is not as sharp.

B. Are these networks small-world?

Given the characteristics of the graphs we detailed above, we have the intuition that the generated graphs indeed possess the small world property, as many interaction graphs do. However, the small world property is generally not verified in networks with strong geographical constraints, such as urban street networks, where the graphs are created by elements of the road infrastructure (intersections, roads) [18]. To evaluate this hypothesis, let us examine how the clustering coefficient is distributed in these networks. The *clustering coefficient* of a given node in a graph is defined as the probability that two neighbors of this node are themselves mutual neighbors. This classical metric accounts for the presence of communities in the graph. In terms of networks, it indicates the presence of dense areas, ultimately cliques, that yield to the creation of shorter communication paths. Fig. 7 represents the CDF of the clustering coefficients in the selected cities. As mentioned in section V-A, the graph is partitioned, but this result shows that most of the connected components have a high clustering coefficient. Indeed, the graphs are influenced by geographic constraints, but are biased by the high density of sensors deployed at each intersection. We can conclude that unlike road-intersection graphs, as the clustering coefficient is independent of the total number of nodes, our graphs possess the small world property. Besides, the small world property translates into a short average distance between each couple of nodes in the graph. This is confirmed by looking at the diameter of the connected components (Fig. 9), which is, on average, lower than the logarithm of the total number of nodes.

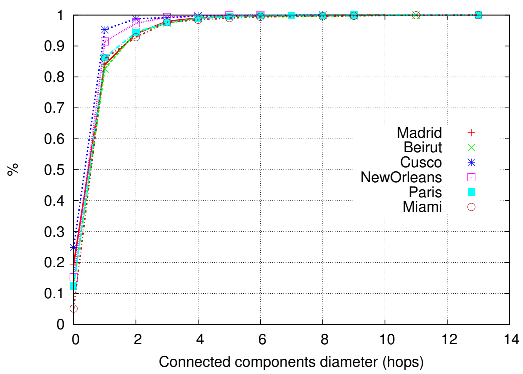


Fig. 9. Components diameter (CDF)

C. Anatomy of the connected components

Fig. 9 represents the CDF of the diameters 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. We can see that this diameter remains very low, essentially due to the presence of several small sized components. It should be underlined that some components have a diameter that is less than 1: this case occurs when a component has only one node (sec. V-A). Networks are mainly composed of small-sized connected components and a few large ones.

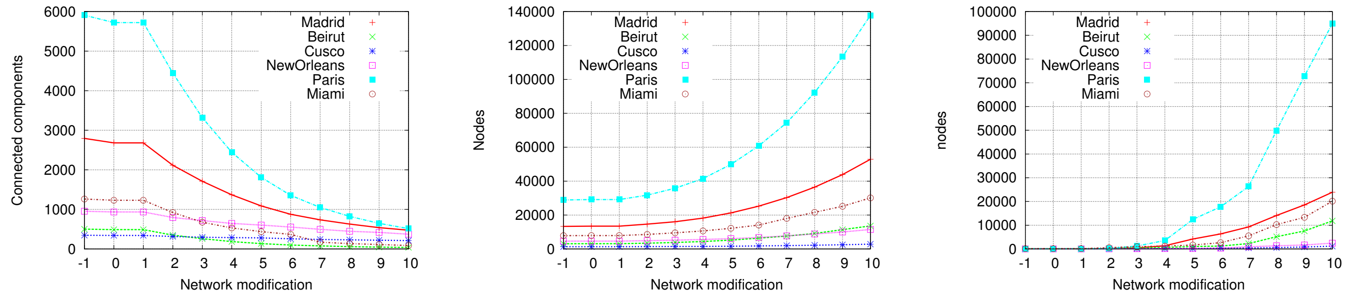
Let us now focus on the *maximum connected component* (*maximal component*), which is the connected component that contains the largest number of nodes. Fig. 8(a) represents the number of nodes that belong to this maximal component. This number ranges from 33 nodes (New Orleans) to more than 130 nodes (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.). Fig. 8(b) shows the CDF of the hop distances that separates couples of nodes within this maximal 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 Paris (Fig. 2) but its maximum connected components has shorter path for a comparable number of nodes. This means that the intersections density is probably higher in downtown Madrid than in Paris. Fig. 8(c) shows the CDF of the edge lengths within the maximal component, in meters. This parameter influences 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.

VI. 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: 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. Knowing the transmission range of a node, we place a chain of nodes between two neighboring connected components. Let us suppose that the operator imposes a limit on the maximum number of intermediate nodes that could be deployed for interconnection purposes between two components and let us study the effect of setting this limit from 1 to 10 relays. For example, if a chain of nodes is sufficient to connect two connected components, we add it. We determine whether this chain is sufficient based on the propagation model used to simulate the deployment, and separate each node in the chain with a distance equal to half their maximum range, to prevent the transmitted signals from being completely attenuated by the distance. 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.

Fig. 10(a) represents the evolution of the number of connected components in function of the maximum number of relays. The x-axis value of -1 represents the inverse situation in which all the articulation points in the graph 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 effect in very scattered graphs. All scenarios seem to converge to comparable values close to 200 components. Fig. 10(b) represents the evolution of the number of deployed nodes in function of the threshold. We can notice that the value tends to increase 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. Finally, Fig. 10(c) shows the evolution of the number of nodes that belong to the maximum 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.



(a) Number of connected components (b) Number of nodes (c) Percentage of nodes in the maximal component

Fig. 10. Improving the connectivity of the networks

Obviously, this strategy is only one option and other connection methods could be imagined. Adding a chain of nodes to link two connected components makes the structure fragile and congestion is more likely to appear on paths involving nodes with a high centrality. One could think of an algorithm that adds enough relay nodes to merge connected components with a strict constraint on the resulting betweenness centrality. $K-1$ long backhaul links could also be created to interconnect the K components. Using a vehicular network to ferry messages between these components could also be an interesting alternative. However, the goal here is simply to demonstrate that with a naive strategy, it is possible to improve network connectivity without too much difficulty.

VII. Conclusions and Future Works

In this paper, we examined and characterized the connectivity of a wireless sensor network deployed at the intersections of various cities. Examining the partitioning in connected components, we show that the resulting graph is highly disconnected and comprises up to 25 % of isolated nodes. Nevertheless, the network presents a good redundancy level within connected components. The average diameter of connected component is low, but can rise to fair values. Finally, we show that a real deployment should be feasible and that a moderate proportion of relay nodes is required to let the maximum connected component encompass most of the network and cover flagship urban areas (e.g. downtown) with a single sub-network.

In future work, it would be interesting to study other deployment strategies. Our complete results show, for example, that if we deploy a node at the center of each intersection while adopting the assumptions made in this article, a maximum of 14 adjacent intersections could be covered in the city of Paris without an additional relay node. Before analyzing networking aspects, it would also be interesting to use a more complex propagation model and consider the presence of buildings as a means of creating more realistic graphs.

References

[1] 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.

[2] S. Lämmer and D. Helbing, "Self-control of traffic lights and vehicle flows in urban road networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 04, p. P04019, 2008.

[3] 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.

[4] S. Faye and C. Chaudet, "Characterizing the topology of an urban wireless sensor network for road traffic management," *IEEE Transactions on Vehicular Technology*, 2015.

[5] R. Lim, F. Ferrari *et al.*, "Flocklab: A testbed for distributed, synchronized tracing and profiling of wireless embedded systems," in *Proceedings of the 12th international conference on Information processing in sensor networks*. ACM, 2013, pp. 153–166.

[6] R. N. Murty, G. Mainland *et al.*, "Citysense: An urban-scale wireless sensor network and testbed," in *Technologies for Homeland Security, 2008 IEEE Conference on*, 2008.

[7] I. Corredor, A. García *et al.*, "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, Jun. 2008.

[8] J. Hu, J. Song *et al.*, "Topology optimization for urban traffic sensor network," *Tsinghua Science & Technology*, vol. 13, no. 2, 2008.

[9] X. Mao, X. Miao *et al.*, "CitySee: Urban CO2 Monitoring with Sensors," in *Infocom*, 2012.

[10] 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*, vol. 26, no. 3, May 2010.

[11] 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.

[12] P. Crucitti, V. Latora, and S. Porta, "Centrality measures in spatial networks of urban streets," *Physical Review E*, vol. 73, no. 3, 2006.

[13] T. Ducrocq, M. Hauspie *et al.*, "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.

[14] M. Tubaishat, Q. Qi *et al.*, "Wireless sensor-based traffic light control," in *5th IEEE Conference on Consumer Communications and Networking (CCNC 2008)*, Las Vegas, USA, Feb. 2008.

[15] D. Krajzewicz, J. Erdmann *et al.*, "Recent development and applications of SUMO - Simulation of Urban MObility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3&4, December 2012.

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

[17] A. Varga and R. Hornig, "An overview of the omnet++ simulation environment," in *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, p. 60.

[18] G. Csányi and B. Szendrői, "Fractal–small-world dichotomy in real-world networks," *Physical Review E*, vol. 70, no. 1, p. 016122, 2004.