Complex Networks: Study and Performance Evaluation with Hybrid Model for Wireless Sensor Networks

Vincenzo Curia, Mauro Tropea, Peppino Fazio, Salvatore Marano
Department of Computer Science, Engineering, Modellistic, Electronics and Systems (DIMES)
University of Calabria, Cosenza (Italy)

curiavincenzo86@gmail.com, {mtropea, pfazio, marano}@dimes.unical.it

***Abstract*—Recent studies have demonstrated that telecommunication networks can be well modeled as complex ones, instead of the classic approach based on graph theory. The study of complex networks is a young and active area of scientific research inspired largely by the empirical study of real-world networks such as computers and social networks. In this paper, a new approach for modeling Wireless Sensor Networks (WSNs) is proposed, with the purpose of ensuring an improvement in terms of clustering coefficient and average network diameter. With the proposed hybrid idea, the major part of nodes will not be necessarily adjacents but, for each pair of them, there will exists a relatively short path for a stable connection. The effectiveness of the proposed idea has been validated through a deep campaign of simulations.**

Keywords-Small world, Complex Networks, Data Mules, WSN, Wireless Sensor Networks, Clustering.

I. INTRODUCTION

Internet is today part of the life of everybody thanks to the great progress done in the technologies that allow the use of networks through different type of devices. Despite their different nature, technologies are born with the purpose of ensuring connectivity everywhere: to send information, for example, we can use electromagnetic radiations in infrared frequency band, electric transmission lines or wireless devices, which take advantage of the ether. Since the ether is present everywhere in the world, it is reasonable to exploit the potentiality of this medium in order to use it as a mean of communication on which information can travel. This is a very good choice, although there are some limitations, due to its intrinsic nature and physical barriers to be overcame. The majority of limitations are given by the limited coverage radius that the current transmission standards fail to ensure and by the supply management of the devices. Wireless Sensor Networks (WSNs) [1], [2], [3] are the subject of some kind of recent studies: they fall into the world of Complex Networks (CNs) [4], [5], since each network present in nature, whether artificial (as the national water supply network), based on communication, natural (such as brain synapses network), and molecular interactions can be seen as a CN if appropriately modeled. It is suitable to find a modeling technique which allows us to represent a real network in a Small World Network (SWN, a particular branch of CNs) [5], [6], [7], while maintaining a low network diameter and a high clustering coefficient. With the effort and research activity carried on in this paper, a new idea applied to a urban WSN scenario has been studied and evaluated, increasing the range of action of nodes and the overall communication efficiency. The proposed scheme is based on the utilization of public transport means (such as buses, trams and subways) considered as special nodes and defined Data Mules (DMs), in order to carry data within the network and to interconnect various network clusters, relatively distant one from each other. The considered DMs are equipped by "special" Wi-Fi routers, using two different transmission standards, guaranteeing a "double" coverage range, which permit to expand the scope of the entire network. The introduction of special nodes within the network contributes to the improvement of network scalability, allowing the addition of new sensors nodes, without substantial changes to the structure of the network. This paper is structured as follows: section II gives a deep overview of the existing related works, section III describes the considered scenario, section IV illustrates the proposed idea and it is evaluated in section V; section VI concludes the paper.

II. RELATED WORKS

CNs represent a new paradigm to which the world (seen in its various disciplines of humanities, physical and scientific studies) is shifting, because their ability of modeling complex relations among objects. In the modern literature, there are various known models, which belong to this type of networks, as the one introduced by Erdös-Rényi and Watts-Strogatz [8], [9], from which the concept of Small World Networks has born [9]. Authors found that many already studied networks and systems (such as biological oscillators, excitable media, neural networks, spatial games, genetic control networks and many other self-organizing systems) can be highly clustered, like regular lattices, with a small characteristic path length, as for random graphs. These kind of networks are called SWNs (by analogy with the small-world phenomenon, popularly known as six degrees of separation). For example, the neural network of the Caenorhabditis Elegans worm, the power grid of the western United States and the collaboration graph of film actors are shown to be small-world networks. Albert and Barabasi [10] proposed their model, in which an analogy with the structures of biological and social systems have been made, with the associations between ideas and concepts and with some artificial networks, including Internet and air routes; the model is famous for its ability to explain the evolution of these systems in terms of adding, removing and editing nodes [11], [12]. A SWN can contain billions of nodes, as an Internet connection: in order to move from one node to another one, few intermediate nodes should be crossed and the entity of this property can be measured by evaluating some objective parameters of the network. SWNs are called Random SWNs (R-SWNs) when the probability density function (pdf) related to nodes distribution is approximately Poissonian and can be modeled by an exponential function ($P(k) \sim exp(k)$), or Scale-Free (SF), if the distribution can be modeled by a power function ($P(k) \sim k^c$),where c is a constant value). According to Barabasi model, networks such as Internet, should be modeled by a power-law model with $2<c<3$. This model shows that a network can be classified as CN, more precisely SWN, if it has a high CC (average value of local CCs, where local CCs are evaluated as ratio between the number of connections among neighbor nodes and the number of neighbor nodes). The described condition is necessary but not sufficient; the treatment can be completed if a second condition is added: the network must have a small diameter (the longest of all the calculated shortest paths), with a consequent average path length equal, approximately, to six (quantified in hops number). Given these parameters, it is possible to model an existing network such as a CN, ensuring an improvement in communication.

By the empirical experiments on CNs of the pioneers Euler and Milgram [13], it has been shown that, in a network of any kind and form, the communication between two random nodes occurs (with information forwarding) through an average of six nodes before reaching its destination (small-world phenomenon). This is a useful concept which allows us to analyze and evaluate the performance of a telecommunication network, in terms of efficiency and speed of data exchange.

In [5], authors demonstrated how a WSN can be well approached by a SWN with two different approaches; as known, in a WSN there is a special node, called sink node, that is either the origin or the destination of a message, while the other types of data communications(relay activities for example) happen between arbitrary communicating entities. Authors found that sink node exhibits the most interesting tradeoff between energy and latency, allowing the design of strict applications that demand a small latency and energy consumption.

In this work, a novel approach for modeling WSNs as SWNs is introduced; it is based on the concept of DMs (called also ferries), that act as special nodes for the enhancement of the small-world properties of the considered network. In particular, we carried-out some analysis in urban environments, considering some ways for increasing the overall connectivity and enhancing network diameter.

III. SCENARIO

Before showing the idea of the proposed model it is necessary to study different parameters and learn about the distribution degree of nodes within the network and understand the functionality of a generic network. The parameters that have to be taken into account for complex networks are basically the following two: the average network diameter (evaluated as hop-number) and the clustering coefficient (that is the probability of a node A to be connected to the node B, having as common neighbor a node C). For SWNs, the most suitable clustering coefficient needs to be about 1. The other parameters (such as the Degree Centrality) represent secondary elements when defining a CN. The Degree Centrality (DC) is a fairly simple metric which describes a network, providing a general estimation of the structure of its graph, based on the number of incident edges to a node. Given an undirected graph $G=<V, E>$,where V is the set of vertex and E is the set of edges, the degree of a node $v \in V$, indicated as$\delta(v)$, is defined as the number of incident edges in$v$:

$$\delta(v) = \sum_{i \in V} e_{v,i},\ with\ e_{v,i} = \begin{cases} 1\ if\{v,i\} \in E \\ 0\ else \end{cases} \quad (1)$$

So, the DC parameter provides an estimation of the importance of node v, related to the analyzed system domain and based on the number of nodes directly related with it. This concept introduces the distribution degree of nodes within a network, considering the DC for all single nodes. Generally, distributions like the Normal one are used, but many real networks cannot follow this trend, so the Power Law (PL) is used: it is radically different from the previous, because it has a non-negligible fraction of nodes with very high degree (called hubs) and it has the property of being scale-free (its typical trend is illustrated in fig. 1). In a graph G, the probability that a generic node $v \in V$ has degree $\delta(v)=k$ is proportional to $\frac{1}{k^\gamma}$, where $\gamma > 1$ is called scaling exponent (typically $2 < \gamma < 3$):

$$P(k) \cong k^{-\gamma} \quad (2)$$

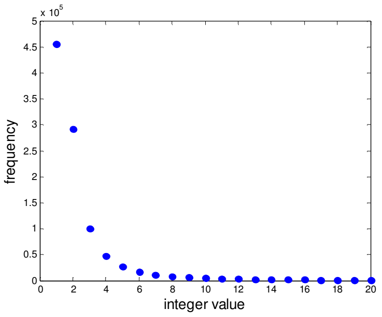

Figure 1. A sample distribution for a PL with γ=3.

The average length of the shortest path, can be calculated on all pairs of nodes of the network as :

$$L = \frac{N}{N(N-1)} \sum i,j \text{, with } i,j \in N, i \neq j \quad (3)$$

where N=||V||is the total number of nodes; obviously a low value of L (≈ 6) is related to a network with a small diameter and, then, a low latency coefficient. The first problem comes precisely from the respect of that condition in a network with a high clustering coefficient, given that this condition tends to necessarily increase the diameter of the network. The clustering coefficient that, by definition, measures the tendency of two nodes adjacent to a common node to be connected each other, is another important parameter when observing network properties. In practice, it indicates (locally) the connection ability of network nodes and it increases significantly the network's diameter. The clustering coefficient of the *i-th* node is:

$$c_i = \frac{e_i}{k_i(k_i-1)/2} \quad (4)$$

where e_i is the number of edges which interconnect *i's* neighbors and k_i is the number of *i's* neighbors. The clustering coefficient of the entire network is constructed with the average of all c_i calculated as :

$$C = \frac{1}{N} \sum_i c_i \quad (5)$$

Since 0≤C≤1, a value tending to 1 indicates a high connection between neighboring nodes (obviously in a fully connected network this value is exactly equal to 1). The second problem is that a low clustering coefficient excludes a fundamental characteristic of CNs, that is navigability and high interconnection between neighboring nodes. The evaluation and measurement of some basic properties of a CN, such as the average path *L*, the clustering coefficient *C* and the degree distribution *P(k)*, are the first steps to understand its structure. The next step, therefore, is to develop a mathematical model with a "similar statistical properties" topology, obtaining a platform on which it is possible to perform some mathematical analysis. Except the average network diameter and the clustering coefficient, all the other parameters are considered, in general, when evaluating the obtained results, with the main aim of giving more emphasis to them. With the proposed hybrid model, we start from a real regular network (already existing) and model it in order to consider the obtained model as a CN (SWN in particular): it will have a value of L near to 6 (as Milgram's theory) and an average clustering coefficient near to 1 (as desired). As stated before, natural systems (regular or random) presents complex properties related to internal connections of nodes: traditional network models are not able to express the "complexity" of their connections, so CNs are chosen to do that, by modifying the layout of the starting network. The most common techniques are: Kleinberg's theory, link adjustments, DM insertion, etc. The Random Graphs Networks (RGNs) are systems in which the connections, with N>>1, are created in a random way and with probability *p*. The cluster coefficient for this type of network is, therefore:

$$C = p = \langle k \rangle / N << 1 \quad (6)$$

If the connectivity follows a Poisson distribution, the network topology is in agreement with the SWN definition. The known variant is represented by Regular Networks (RNs), in which each node is interconnected with its neighbors on a regular basis, forming a lattice. Obviously, in these networks, the cluster coefficient →1, but the small-world effect cannot be observed. It is easy to note that the two models mentioned above fail to reproduce many features found in real networks. It is therefore necessary to introduce the SWNs concept: it can be obtained by translating a real network in a random one and vice-versa. In fact, using known techniques such as the rewiring of links or adding shortcuts, a random or regular network can be modeled in a SWN.

IV. SWN Hybrid-Modeling (SWN-HM) Proposal

The modeling technique suggested in this article has the purpose of shaping a real WSN in a CN with the SW property, with all the benefits that this kind of modeling can carry to the network, in terms of safety, speed and connection stability. This model is called Hybrid because, although based on DMs insertion, it considers two types of nodes: one type composed by traditional DMs (urban buses for example) and the second type composed by nodes with double coverage range (for example, based on two interfaces with different power or two different technologies); we demonstrate that, in this way, it is possible to increase *C*, with a decrease in the average length of the route L. If we are considering regular networks in which, for example, urban buses follow standard roads, we can decide what route these nodes should follow in order to obtain a complex structure (without deleting routes casually); if we are considering a random network in which, for example, accidents or fires happen in random points or in which sensors are deployed casually, we should investigate on the number (and/or the position) of mules to insert, for reducing battery consumptions [1], [14], [15] and coverage radius; DMs movements of can be described by prefixed paths or by markovian processes. By the application of Watts-Strogatz's algorithm, it is possible to reduce the average path-length from L=N to L=log(N), where N is the number of nodes. During the analysis of real cases it is possible to achieve this goal by exploiting the urban buses and the mobile routers (as DMs), equipped with a double radio coverage system. As shown later, for different *p* values, it is possible to reach high values of *C* and lower values of *L*. Intuitively, the number of mobile nodes to insert within the network has an upper limit, which varies according to the physical size of the entire network. So, DMs are means that allow to rewire links, adding new shortcuts, without removing any existing link. Relations $L(p)/L(0)$ and $C(p)/C(0)$ (where $L(0)$ and $C(0)$ are the values obtained before the insertion of ferries) represent a decrease of network's average length and a decrease of clustering coefficient, with the increase of probability of rewiring the links, accordingly to the addition of ferries. The objective of the HM proposal is to model any existing network as a SWN, with cluster coefficient C → 1 and route's average length≈6, as learned from Milgram's experiences [13]. DMs, in fact, are not only mobile nodes, but are transmission devices with an increased range: this additional condition makes possible to create some long-range links between different DMs added in the network, drastically reducing the average path length L, bringing it at the right condition (near to 6). Special Nodes can be implemented by Wi-Fi router with MIMO standard, which use different IEEE standard transmissions as IEEE 802.11n for short-range links (coverage radius r) and UMTS for long-range links (coverage radius r_u). However, this approach has some practical restrictions, as the energy consumption and interferences. In such a network, two nodes i and j are directly connected in the graph if they meet one of the following conditions: the distance $d_{ij}<r$ or, if i and j are special nodes, $d_{ij}<r_u$. However, connectivity also depends on other factors as interference, noise and energy constraints, therefore cannot be always guaranteed.

V. PERFORMANCE ANALYSIS

In order to carry simulation analysis, a Java simulator has been implemented. It is designed to be able to simulate the behavior of a network and to verify the described studies, regarding the modeling of a Ad-hoc network through the use of CNs and, in particular, by SWNs, with special mobile devices. The first part of the implementation involved the development of a MultiThread system, creating the class objects: a) MobileNode, which creates objects able to change each time their position in an area in according to predetermined paths; b) FixedNode, which generates objects able to simulate the behavior of real sensors placed in the system in order to create a WSN. Both classes extend the class "Parent Node". In order to generate the adjacency matrix corresponding to the graph, Manhattan Model has been used. Nodes movements are bounded or restricted by the implementation of the simulator in a real urban environment (Corigliano Calabro Scalo urban area has been considered). Subsequently the Graph class is created, recording every movement of each node and saving the information in a matrix adjacency structure. Then, it was developed a graphical tool using java Graphics class in order to display on video the movements of special nodes, with the creation of appropriate links. By simulator we have conducted many simulations with different settings, before and after the insertion of special nodes. Simulation campaigns have been performed for: a) the calculation of the density, b) the calculation of average path length with only Sensors Nodes, c) the calculation of the average path length to vary the number of special nodes. It has been considered a cover range of each node from 70 meters to 100 meters; this characteristic was established after careful research, carried out in the literature on the nature of the wireless devices, without any distinction between classic wireless devices and wireless sensors.

A. Density calculation

It was found that to obtain an optimum density, each graph must have a number of edges approximately equal to 6. The considered map covers an area of 1000 x 650 m^2 and the formula for the calculation of the density corresponds to the average of the degree of each node in the network:

$$\frac{\sum_{j=1}^{N} a_{i,j}}{N} \tag{7}$$

where N is the number of nodes. It is simply to understand, that when map dimension changes, the minimum number of nodes to obtain the optimum density gradually increases. The obtained results illustrated as the density, varying the number of nodes, always remains approximately equal to 6, obtaining an optimal value. These simulations on density are preliminary, in fact they have been carried out to determine the density of a graph, ensuring that in the subsequent simulations, scattered graphs are not taken into account, without compromising the optimal analysis of the network.

B. Average length of the shortest path without mobile nodes

In this section, the simulations carried out for the calculation of the average length of the shortest path in an ad-hoc network without the introduction of mobile nodes have been analyzed and described. We considered four different size of the map (areas): Area 1 (500x500), 2 (750x750), 3 (1000x1000) and 4 (1000x650) m^2. Node coverage range is 90 m. From fig. 2, it is possible to see, only as example, the typical values of path length for different area.

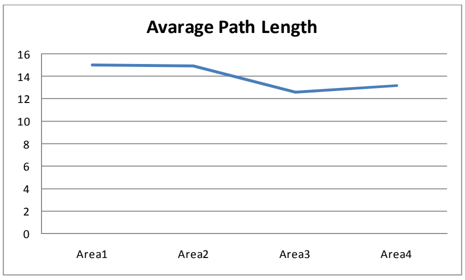

Figure 2. WSN Average Shortest path L with special hub.

C. Average length of the shortest path with additional DMs.

We can therefore observe that, indeed, we get the expected gains as we have a substantial decrease in the average length of the shortest path.

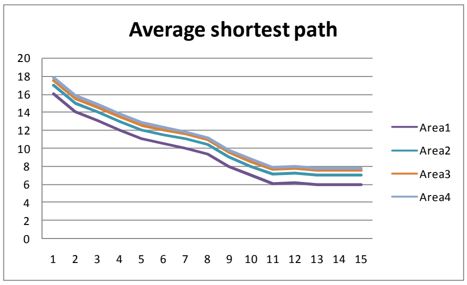

Figure 3. Average Shortest path L with special hub.

From Fig. 3, it can be seen how there exists an upper limit of the number of nodes to be included within a network of Ad-Hoc sensors. It is clear, furthermore, that the variation of the average paths L, is very similar in the different scenarios.

D. M*aximum length of minimum path.*

As shown in Fig. 4 a) and b), the variation of the maximum shortest path of several networks is taken into account. Considering areas with variable-size and a fixed number of nodes fixed a priori, obtained from the calculation of the optimum density.

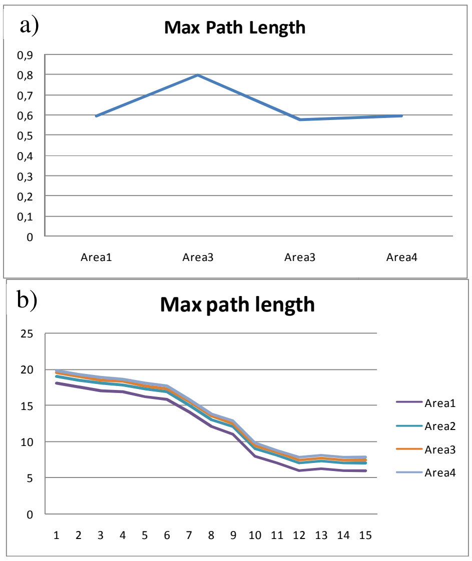

Figure 4. Max Path Length, without (a) and with (b) special hubs.

E. Clustering coefficient.

Another key feature of SWNs is the clustering coefficient of the network, which indicates the ability of a network to create a certain number of SWs.

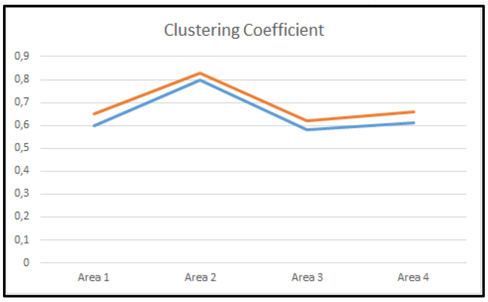

Figure 5. The trend of Clustering Coefficient.

Special hybrid model, once applied, tends to slightly increase this value compared to the one of the original network. The change is about one percentage point. As stated before, a value of C near to 1 indicates a high connection between neighboring nodes (obviously, in a fully connected network, such value is equal to 1). Figure 5 shows the clustering coefficient before and after modeling, within the various areas.

VI. CONCLUSIONS

This work has shown good results about the model proposed for wireless networks. First of all, by considering any existing, real and appropriately modeled, network the hybrid model significantly reduces the average length of the shortest path, and increases (even if only slightly) the medium cluster coefficient, thus obtaining those characteristics that mark the complex networks as small-world. Then, modeling a network as a CN may be advantageous, in fact, reducing the shortest path, the communication within the network is faster, reducing latency and vital parameters such as RTT. Data Mules are also able to perform load balancing in the various paths of the network, avoiding overloads and preventing any performance degradation. Increasing the clustering coefficient also increases the resistance to network faults, in case of a fall of one or more links.

REFERENCES

[1] F. De Rango, S. Amelio, "Performance evaluation of scalable and energy efficient dynamic n-epidemic routing in delay tolerant networks," in 16th SPECTS 2013, Part of SummerSim 2013 Multiconference; Toronto, ON; Canada; 7-10 July 2013.

[2] F. De Rango, E. Bragina, "A weighted spatial correlation based strategy to improve the event estimation in a Wireless Sensor Networks," in IFIP Wireless Days, WD 2012; Dublin; Ireland; 21-23 Nov. 2012.

[3] E. Cassano, F. Florio, F. De Rango, S. Marano, "A performance comparison between ROC-RSSI and trilateration localization techniques for WPAN sensor networks in a real outdoor testbed," in Wireless Telecom. Symp., 2009; Prague; Czech Republic; 22-24 April 2009.

[4] C. Yan; Qiao Ji-Hong "Application analysis of complex adaptive systems for WSN", Computer Application and System Modeling (ICCASM), 2010 International Conference on (Volume:7).

[5] Guidoni, D.L., Mini, R.A.F., Loureiro, A.A.F., "Creating small-world models in wireless sensor networks", Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on.

[6] Huang, Jie, Xie, Qihui, Huang, Bei, "Creating Small-World Model for Homogeneous Wireless Sensor Networks, "Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on , Page(s): 1 - 4.

[7] Xiaojuan Luo, Huiqun Yu, "Constructing wireless sensor network model based on small world concept", Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on, Volume: 5, Page(s): 501-505.

[8] Erdős, Paul, A. Rényi, "On the evolution of random graphs". Pub. of the Mathematical Institute of the Hungarian Academy of Sciences: 17–61, 1960.

[9] Watts, Duncan J.;Strogatz, Steven H., "Collective dynamics of 'small-world' networks", Nature 393(6684): 440–442, 1998.

[10] R. Albert; A.-L. Barabási, "Statistical mechanics of complex networks", Reviews of Modern Physics74: 47–97, 2002.

[11] Boyd, D., Ellison, N., Social Network Sites: Definition, History, and Scholarship, Journ. of Computer-Mediated Comm., 13, article 11, '07.

[12] Petrou, M., Tabacchi, M.E.: Networks of concepts and ideas, the Computer Journal, 2009.

[13] S. Milgram, “The small world problem”; 1967.

[14] F. De Rango, F. Guerriero, P. Fazio, "Link-stability and energy aware routing protocol in distributed wireless networks," in IEEE Trans. on Parellel and Distributed Systems, Volume 23, Issue 4, 2012, pp.713-726.

[15] D. Amendola, F. De Rango, K. Massri, A. Vitaletti, "Neighbor discovery in delay tolerant networking using resource-constraint devices," in 6th IFIP/IEEE Wireless Days Conference, WD 2013; Valencia; Spain; 13-15 Nov. 2013.