# Network Topological Analysis Utilizing Significant Profiles for Sensor Network Applications

Lun Zhang, Lam Ling Shum, Lionel Sacks

University College London

**Abstract:** In the paper Average Paths Length and Clustering Coefficient are used to analyse different generated network topologies. We further introduce Significance Profiles (SP) to analyse the networks in search for the network motifs, which are the dominant subgraphs of different topologies, as a preparation of our future work. Our results can be utilised in designing an effective network protocol to maintain integrity and robustness, and to ensure connectivity in a wireless sensor network..

## 1. Introduction.

In a wireless sensor network, the underlying network structure can varies from network to network and from time to time. Recently, much research has been done on efficient routing algorithms for sensor networks [1][2][3]. They lack a common metric to evaluate the structure of the network, and adopt the algorithms accordingly to maximise their performance.

Complex networks are structures consisting of vertices inter-connected by links or edges. The ubiquity of complex networks in science and technology has led to a set of common research problems to find out how network structure can facilitate or constrain network connectivity and integrity[4][5][6]. These studies are particularly important in self-organising networks including wireless sensor networks. Common graph theory tools for analysing network structures are *Average Degree, Average Paths Length* and *Clustering Coefficient*. Recently a lot of works have been done on classifying diverse networks into distinctive superfamilies using a quantity called *Significance Profile* (SP) [7][8]. In this paper we investigate SPs of several common network structures, as a first step to explore the possibility to use SP in a distributed manner for evaluating the integrity of the sensor networks.

Sensor Networks have certain characteristics that are very different than other types of networks. Normally sensor nodes are deployed based on the requirement of monitoring and therefore, there is no fixed network structure and topology. Most sensor applications do not require the nodes to be mobile. Sensor nodes communicate with each other across wireless channels and the connectivity may vary over time dependent on the environment. In some applications sensor nodes are deployed in a hazard environment that retrieval and manual intervention is particularly difficult. The open issues of the networking of sensor networks are coverage, fault tolerance, power consumption constraints, optimal placement, *etc.* We emphasize on the following issues on the development of algorithms.

- Network Integrity. The network should be reliable and well-connected. The network should be robust against failure of sensor nodes and communication links, and be adaptive to additions of nodes.
- Scalability. The network design should be scalable across the network.
- Distributed. Algorithms must be distributed and require no central processing and management nodes.
- Topological flexibility. Different networks topologies can be formed and overlayed on the same nodes for different purposes.

Our ultimate goal is to develop a routing or information dissemination algorithms, which utilise a common metric calculated locally in a sensor network and represent the local connectivity, and adjust its parameters to adapt to the part of the network. Network integrity is achieved as a whole by local interaction.

## 2. Methodology.

In this paper, we have generated and evaluated five different types of common network topologies in search for their representative qualities which may be useful in evaluating and maintaining network integrity in sensor networks. The networks are:

- **Random network**: nodes are positioned randomly. The degree of connectivity is fixed but nodes can be connected to any other nodes in their communication distance with a preset probability *p*. See Figure 1A.
- **Ring topology**: all nodes are connected to its nearest two neighbours bi-directionally within the radio range. .See Figure 1B.

- **Tree topology**: A root node is randomly chosen among the network. The root node expands a defined number of branches to the nearest nodes within a distance threshold. .See Figure 1C.
- **Small world graphs**: the properties and generation of a small-world graph is defined in [9]. In a small-world network, nodes are highly clustered similar to regular graphs, *e.g.*, lattices; yet have typically short distances between arbitrary pairs of vertices similar to random graphs. See Figure 1D.
- **Scale free network**: Many large real networks are scale free, that the degree distributions of their nodes follow a power law for large *k*. This means that the networks contain relatively few highly interconnected *super nodes* or *hubs*: the vast majority of nodes are weakly connected, and the *connectedness* ratio of the nodes remains the same whatever size the network has attained. If a network is scale-free, it is also a small world. .See Figure 1E.

In our analysis a number of sensor nodes are placed at random locations within the space for each network topologies. Each sensor has an effective transmission radius defining the nearest-neighbour relationships, or in graph theory terminology edges between two nodes. In the case where the transmission radio of sensor nodes are not equal or links are not bi-directional; the connections are represented by directed links. However, for simplicity we only consider un-directed links in this paper.

### 3. Basic Topological Analysis.

In this section we evaluate our generated networks using common graph theory terminologies [See Figure 1]. It helps us obtain an understanding of the basic parameters and the characteristics of the topologies and assists us in the explanation of the significance profiles in the next section.

A B C D E

**Figure 1: All network graphs are generated with 32 random deployed nodes. (A) connection Probability =0.5 for random graph; (B) ring topology; (C) there are 3 branches per nodes in the tree topology; (D) Rewiring probability =0.3 for small world graph. (E) Scale-free network**

Graph theory studies the properties of the probability space associated with different combination of nodes and their links(edges), which can be determined using probabilistic arguments. It assists in understanding of the underlying network structure. Spectacular concepts—the average degree, the average path length and clustering coefficient play a key role in the recent study and development of complex networks theory[4]. Table 1 illustrates some results obtained from altering the parameters of the topologies. The quantities investigated are average degree (AD), which is the average number of links per node, average path length (APL), which is the average number of hops between any 2 nodes, and clustering coefficient (CC), which is defined as the number of links exist in the neighbourhood divided by the possible number of links. It is a quantity that measures how clustered the neighbourhood is.

**Table 1: Feature analysis of networks with different topologies**

| Topologies | Parameters | AD* | APL** | CC*** |
|---|---|---|---|---|
| **Random** N=32 | P=0.1 | 3.1344 | 2.9490 | 0.1793 |
| | P=0.5 | 15.2547 | 1.5081 | 0.7413 |
| | P=1.0 | 31 | 1 | 1 |
| **Tree** Maximum 3 branches, N=32 | P=0.1 | 1.9375 | 4.5585 | 0 |
| | P=0.5 | 1.9375 | 4.1452 | 0 |
| | P=1.0 | 1.9375 | 4.1169 | 0 |
| **Tree** Maximum 4 ranches,N=32 | P=0.1 | 1.8125 | 4.2345 | 0 |
| | P=0.5 | 1.9375 | 3.7319 | 0 |
| | P=1.0 | 1.9375 | 3.6996 | 0 |
| **Ring** | P=0.1 | 1.5000 | 3.8547 | 0 |

** AD - Average Degree*

$AD = \langle k \rangle / N$ , *Where,k denotes degrees of different nodes.*

*** APL - Average Paths Length, means average hops between any pair of nodes in the network.*

**** CC - Clustering Coefficient*

| | | | | |
|---|---|---|---|---|
| N=32 | P=0.5 | 1.9375 | 11 | 0 |
| | P=1.0 | 1.9375 | 11 | 0 |
| **Small World** k=6 N=32 | P=0.1 | 6.0625 | 2.7742 | 0.3844 |
| | P=0.5 | 7.2500 | 1.9677 | 0.2289 |
| | P=1.0 | 10.3750 | 1.6452 | 0.3202 |
| **Scale Free** N=32 | m=5,m0=3 | 2.1818 | 2.5181 | 0.2072 |
| | m=10,m0=6 | 3.9394 | 2 | 0.3246 |
| | m=20,m0=12 | 12.1818 | 1.4215 | 0.6667 |

$$CC = 2E_i / k_i(k_i - 1)$$

*where $E_i$ is the real existed numbers of edges among neighbors, $k_i$ is the number of neighbors.*

We see that for a random network, *p* denotes the probability for a link exists between any two nodes within communication range and therefore the average degree increases with *p*. While APL decreases with increasing *p*, CC increases slightly. We can visualise that as number of links increases in the network the average separation of any 2 nodes will be smaller and the nodes would be slightly more clustered.

Both tree topologies with maximum 3 and 4 branches exhibit similar characteristics that their average degrees are slightly under 2 regardless of *p*, *p* here denotes the possibility of any branches exist within the communication range. We can see that APLs of tree networks are significantly higher than that of random graphs because there is no short-cut in the tree topology. CC of any tree networks is very low meaning that tree topology is not clustered.

In Ring topology, *p* denotes the probability of the links of any two nodes in the network , ADs of ring topology is close to 2 and it has high APL that average hops between any 2 nodes is high and that the network is not clustered.

Small world and scale free networks exhibit similar characteristics. *p* in small world topology denotes the rewrite probability depending on the evolution process of a Small World network, and *m* and *m0* in scale free topology denotes the initial existed nodes and the edges of a new added node could be linked to the network separately. While AD increases with the parameters, APLs of both topologies are low and CCs are high in comparison of other topologies aforementioned. Small world and scale free networks are characterised by small characteristic path length and high clustering coefficients.

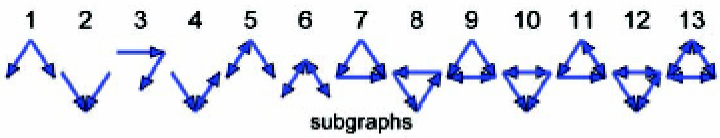
## 4. Significant Profiles

Significance profiles (SP) introduced in [7][8] are a way of identifying network motifs [8], which are subgraphs appear significantly more in a real network than that of a random network. SP is the vector of Z scores normalized to length 1:

$$Z_i = (N_{real_i} - \langle N_{rand} \rangle) / std(N_{rand})$$

$$SP_i = Z_i / (\sum Z_i^2)^{1/2} \quad \text{(Eq.1)}$$

We study the triad SP (connections between 3 nodes) in attempt to characterise the network topology with it. In Figure 2 we can see that subgraphs 4 to 6 (cascades) are the dominant features of tree and ring, while subgraphs 7 to 13 (loops) are almost non-existence for tree and ring networks. On the contrary, number of subgraphs 7-13 are significantly higher in scale free networks than that of random network. Subgraphs 1-6 of the small-world network is similar to the tree and ring topologies, however, while number of subgraphs 7-13 are higher than tree and ring, it remains lower than that of a random network.



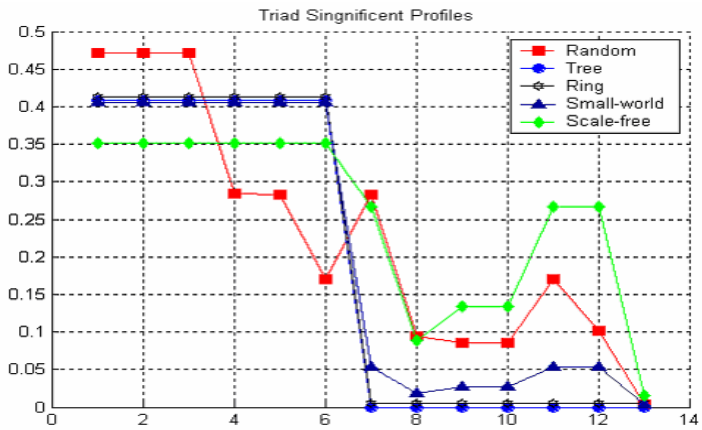



**Figure 2: Significance Profile of network topologies.**

It is observed that since ring and tree topologies, there is neither short-cut nor redundancy in the networks. The networks are formed of branches and rings. Loops do not exist in such networks. While in small-world network, a few short-cuts, which shorten the APL of the networks, has created a number of loops and therefore subgraphs 7-13 are slightly higher than zeros. Scale-free networks have significantly more short-cuts that loops become the dominant features of this type of networks.

## 5. Conclusion and Future Work

In this paper, we attempt to characterise the different generated network topologies including random, tree, ring, small-world and scale-free networks using graphs' average degree, average path lengths and clustering coefficients. We introduce a metric called significance profile to identify subgraphs that are dominant features of a particular topologies. We find out that subgraphs 4 to 6, which denotes cascades are the dominant features of tree and ring, while subgraphs 7 to 13, which denotes loops are the dominant features of scale-free networks. Small-world network is characterised by similar subgraphs 1-6 to that of tree and ring topologies with loops that are slightly less than that of a random network.

We hope to use the results presented in this paper to develop an efficient algorithm for routing or information dissemination in the network. SP is to determined in a neighbourhood such that local topological structure is understood. Parameters of the algorithms are tuned to optimise the performance of the network and ensure network integrity as a whole.

References


[1] I.F. Akyildiz, W.Su, Y. Sankarasubramaniam, E. Cayirci, *Wireless sensor networks: a survey*, Computer Networks, Vol 38, 2002, p.393-422.

[2] Wendi Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, *Energy-Efficient Communication Protocols for Wireless Microsensor Networks*, Proc. Hawaaian Int'l Conf. on Systems Science, January 2000.

[3] D.Brakinsky, D.Estrin, Rumor Routing Algorithm For Sensor Networks, WSNA 2002.

[4] Re´ka Albert* and Albert-La´ szlo´ Baraba´ si , Statistical mechanics of complex networks *Reviews of Modern Physics*, Vol.74, JANUARY 2002, p.47-97

[5] X.F. Wang and G. Chen, Complex Networks: Small-World, Scale-Free and beyond ,*IEEE Circuits and Systems Magazine* , First Quarter 2003 p.6-20

[6] A. Barrat1 and M. Weigt, On the properties of small-world network models, The European Physical Journal B 13, Springer-Verlag 2000, p. 547-560

[7] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, *Network Motifs: Simple Building Blocks of Complex Networks*, Science, Vol 298, Issue 5594, 824-827 , 25 October 2002

[8] Ron Milo, Shalev Itzkovitz,etl , *Superfamilies of Evolved and Designed Networks* , 5 MARCH 2004 VOL 303 SCIENCE pp 1538-1542.

[9] A. Hawick and H.A. James, *Small-World Effects in Wireless Agent Sensor Networks.*

[10] D. Watts and S. Strogatz. *Collective dynamics of small-world networks*. Nature, 363:202–204, 1998.