Energy-efficient and Fault-tolerant Evolution Models for Large-scale Wireless Sensor Networks: A Complex Networks-based Approach

Haixia Peng*†, Shuaizong Si*, Mohamad Khattar Awad‡, Nan Cheng†,
Haibo Zhou†, Xuemin (Sherman) Shen†, Hai Zhao*
* School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning, China
†Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada
‡ Department of Computer Engineering, Kuwait University, Kuwait City, Kuwait
Email: h24peng@uwaterloo.ca, sishuaizong@neuera.com, mohamad@ieee.com, n5cheng@uwaterloo.ca,
h53zhou@uwaterloo.ca, sshen@uwaterloo.ca, zhaoh@mail.neu.edu.cn

***Abstract*—In this paper, we present three network evolution models for generating fault-tolerant and energy-efficient large-scale peer-to-peer wireless sensor networks (WSNs) based on complex networks theory. Being scale-free is one of the intrinsic features of complex networks-based evolution models that generates fault-tolerant topologies. In this work, we argue that fault-tolerant topologies are not necessarily energy efficient. The three proposed energy-aware evolution models are energy-aware common neighbors (ECN), energy-aware large degree promoted (ELDP) and energy-aware large degree demoted (ELDD). ECN considers neighborhood overlap, whereas ELDP and ELDD consider topological overlap for node attachment. The ELDP model promotes the establishment of links to nodes with a large degree, whereas the ELDD model demotes this strategy. Performance evaluations demonstrate that the proposed models outperform a candidate clustering-based model, thereby providing greater energy savings and fault-tolerance. Among the proposed models, ECN is the winner in-terms of energy efficiency, ELDD performs best in-terms of fault-tolerance, and ELDP conveniently provides balance between the two.**

***Index Terms*—Large-scale wireless sensor networks; complex network; energy-efficient networks, fault-tolerant networks.**

I. Introduction

Large-scale wireless sensor networks (WSNs) consist of hundreds to thousands of small-size, low-cost, battery-powered, and data-processing-capable sensors and cluster heads deployed over a large geographic area. Such networks are self-organized and communicate in a peer-to-peer ad-hoc fashion which make them well suited for long-term data acquisition. Unlike conventional sensor networks, reliable sensing in this type of networks requires dense deployment. Owning to limited battery power, sensors die over time, and the network loses density which degrades the reliability of the aggregation results. This situation calls for the design of energy-efficient and fault-tolerant network topologies capable of generating reliable aggregation results over a long lifespan.

Recent advances in the field of complex networks science have motivated a research direction that adopts complex networks-based evolution models for generation scale-free WSN topologies [1], [2] . The indispensable feature of scale-free complex networks is fault tolerance. Therefore, researchers have focused on incorporating energy awareness in such models to prolong the network lifespan and generate not only fault-tolerant but also energy-efficient topologies. To this end, an extension of the Barabási-Albert (BA)[1] model was proposed in [3]. The authors modulated the node degree with a function of the sensor's residual energy for computing the selection probability of other nodes. Similarly, the BA model was extended in [4] by considering both connectivity and residual energy in the selection mechanism. Furthermore, a link compensation method was also introduced in [4] to compensate for the loss of exhausted sensors. The preferential selection mechanism of the BA model was extended in [5] by a weighted combination of interconnectivity and energy consumption; adjusting the weights gives preference to one over the other in making selection decisions. An additional phase that balanced the energy consumption of different sensors was added to the BA model in [6]. In the added phase, the traffic flow of a lost link is distributed over multiple links to neighboring sensors.

Despite the importance of and the performance improvement in these efforts toward the formation of energy-aware and scale-free topologies, these approaches are limited to applications in conventional wireless sensor networks for three reasons. First, large-scale WSNs consist of both cluster heads and small-size sensors, but conventional WSNs consist only of larger more energy-empowered sensors. Therefore, it is unclear how topologies can be formed in the presence of cluster heads. Second, preferential selection in these models focuses only on the established connectivity of nodes and ignores the potential connectivity of the nodes and nodal positions in the network. Third, in conventional WSNs, the effect of transmission distance on nodal energy dissipation can be ignored; however, this issue is significant in large-scale WSNs.

This work extends the use of the BA model to large-scale networks, thus producing scale-free and energy-aware

[1] A well-known model in the field of complex networks theory that generates scale-free networks using a preferential selection mechanism based on nodes degrees. New nodes added to the topology select other nodes for link establishment with a probability computed based on nodes interconnectivity.

topologies of both cluster-heads and sensors. Three novel evolution models are proposed to improve the fault-tolerance, survivability and energy efficiency of the network by considering the nodes neighborhood [7] and topological overlap [8] in the preferential selection process. Performance evaluations support the theoretical claims and demonstrate that the proposed models outperform the candidate clustering algorithm from [9].

The remainder of the paper is organized as follows. Section II introduces the large-scale WSNs network model. The proposed evolution models are presented in Section III together with a brief overview of the BA model. The performance of the proposed models is evaluated in Section IV and is followed by conclusions in Section V.

II. Network Model

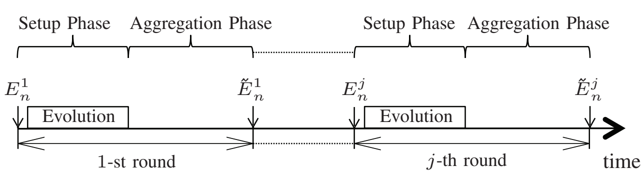


Figure 1. Proposed model operations over time.

We consider a large-scale WSN with a hierarchical network structure of nodes that form the set $\mathcal{N}$. This set consists of three types of nodes arranged in three subsets: a single element subset $\mathcal{K} = \{k\}$ of the sink node, a subset $\mathcal{C} = \{1, \cdots, c, \cdots, |\mathcal{C}|\}$ of cluster heads (CHs), and a subset $\mathcal{S} = \{1, \cdots, s, \cdots, |\mathcal{S}|\}$ of sensor nodes (SNs), i.e., $\mathcal{N} = \mathcal{K} \cup \mathcal{C} \cup \mathcal{S}$. The cardinality $|\mathcal{C}|$ is of the order of tens to hundreds, whereas $|\mathcal{S}|$ is of the order of thousands. Nodes are uniformly distributed in a two-dimensional monitored region. The sink node that initiates the data aggregation process, e.g., sum aggregation or average aggregation, is placed at the center of the monitored region. The CHs collect sensed data from SNs and relay them to the sink node. Nodes within the same subset are assumed to have the same sensing range, transmission range and computational power. However, these parameters are different for the different subsets.

In this work, we propose network evolution models to form a geometric graph $G(\mathcal{V}, \mathcal{E})$ of the available $|\mathcal{N}|$ nodes in which $\mathcal{V}$ is the set of nodes (sink, SNs, and CHs) added to the graph over the evolution process, and $\mathcal{E}$ denotes the set of communication links established among them. The evolution process begins with the sink node only, i.e., $\mathcal{V} = \{k\}$, and iteratively adds nodes of $\mathcal{N}$ to $\mathcal{V}$ until it terminates with $\mathcal{V} = \mathcal{N}$. The communication link between two nodes exists if one of them relays the sensed data to the other during the data aggregation phase, as shown in Fig. 1.

The communication links can be classified into four types, the bidirectional link between CHs, the directional link from a CH to the sink node, the directional link from a SN to a CH, and the bidirectional link between SNs. A node n could have many neighboring nodes within its coverage area denoted by $\Gamma(n)$ but connects only to nodes that has already joined the graph. The set of candidate nodes for establishing a link within the range of node n that has joined the graph already is denoted by the optional set $\mathcal{O}_n$, i.e., $\mathcal{O}_n = \Gamma(n) \cap \mathcal{V}$. For a given node $n \in \mathcal{N}$, d_n denotes the node degree, i.e., node links established to the given node, and E_n^0 denotes its initial energy. The node's initial energy E_n^0 is assumed to follow a uniform distribution in the range $[\underset{\smile}{E_n}, \widehat{E_n}]$. The CHs are assumed to have higher initial energies than the SNs; thus, $\underset{\smile}{E_s} \leq \widehat{E}_s \leq \underset{\smile}{E_c} \leq \widehat{E}_c$, $\forall s \in \mathcal{S}$ and $\forall c \in \mathcal{C}$.

III. Evolution Models

The large-scale WSN consists of thousands of nodes that are battery powered. The large number of nodes and the large area over which they are distributed, make battery replacement impractical. Furthermore, accurate data aggregation requires participation from all SNs in the aggregation phase. Therefore, an energy-efficient evolution model that prolongs the life time of all network nodes is critical for this type of network.

A. Background

Data transmission is the most energy hungry function of wireless sensor nodes [3]. To preserve the residual energy of the modes, a reduction in the number of connections established to a node is desirable, which implies a larger number of hops from the source node to the sink. Interest in a topology with a smaller number of highly connected nodes has motivated us to consider an evolution model that generates a scale-free topology. A scale-free network is a network with a small portion of nodes that have a large number of connections and a large portion of nodes that have a small number of connections. Scale-free networks enjoy strong robustness to node failures, which is also important for WSNs [6]. One of the most well-known complex networks evolution models with this property is the BA model [10]. The BA model consists of the following three phases:

- *Initialization phase*: The network consists of a few vertices v_0.
- *Growth phase*: In each iteration, a new node is added to the graph with e connections, where $e \leq v_0$.
- *Preferential selection phase*: The new node selects the i-th node in its range to connect to with probability $\Pi(d_i) = \frac{d_i}{\sum_n d_n}$. The selection probability is proportional to the node degree.

The node degree distribution follows a power-law distribution. The BA model places high emphasis on interconnectivity, and nodes with high degrees are preferred. Although networks with high interconnectivity are fault tolerant, their energy-constrained nodes with higher degrees deplete energy more rapidly. At a certain point in the WSN life-time, nodes with higher degrees die before other nodes, which significantly degrades the accuracy of aggregation. Considering the characteristics of large-scale WSNs, in the following, we extend the BA model and propose three energy-aware evolution models.

B. Proposed WSNs Evolution Models

We introduce the "residual energy" parameter into our models to balance the energy consumption in the WSNs. By limiting the number of communication links, we control the rate of energy consumption for every node. In addition, in forming the communication links between any two nodes in the large-scale WSN, the local topology structures of these two nodes are taken into account.

Nodes drain their batteries during sensing, processing and transmitting data over time. Network structures should be updated regularly based on the current residual energies of the nodes. The proposed model operates over many rounds, as shown in Fig. 1. Each round begins with a setup phase over which the network structure evolves and is followed by an aggregation phase during which the environment is sensed and data are transmitted. The node energy is measured at the beginning of each setup phase and is denoted by $E_n^j \; \forall n \in \mathcal{N}$. As the network evolves, nodes are added to the graph G and links are established among them. Therefore, the node residual energy in the j-th round is given by

$$\tilde{E}_n^j = E_n^j - \beta d_n^j \qquad \forall n \in \mathcal{N} \backslash \mathcal{K}, \tag{1}$$

where β is the constant energy consumption per link. Note that the superscript $(\cdot)^j$ is introduced in all notations to indicate the round number, j.

Residual energy is critical for WSNs and should be considered in the preferential selection phase. Adding a node to the graph and connecting it to an existing vertex with higher residual energy prolongs the connection life. Furthermore, connections over long distances consumes more energy and drain the nodes' batteries at a faster rate. Thus, connecting the node to neighboring vertices conserves its energy. In the following, three preferential selection probabilities are proposed:

1) *Energy-aware Common Neighbors (ECN) selection probability*: The larger the number of common neighbors, the higher the proximity of the nodes. The number of common neighbors between a node n and vertex v can be written as $|\Gamma(n)^j \cap \Gamma(v)^j| \; \forall n \neq v$ and $v \in \mathcal{V}, n \in \mathcal{N} \backslash \mathcal{V}$. Scaling the number of common neighbors, by the residual energy of vertex v gives higher preference to vertices with higher residual energy. In round j, the ECN selection probability of node v is

$$\Pi_{v-ECN}^j = \frac{|\Gamma(n)^j \cap \Gamma(v)^j| \tilde{E}_v^j}{\sum_{l \in \mathcal{O}_n^j \backslash \{v\}} |\Gamma(n)^j \cap \Gamma(l)^j| \tilde{E}_l^j}, \tag{2}$$

2) *Energy-aware Large Degree Promoted (ELDP) selection probability*: In addition to the neighborhood overlap, the topological overlap between nodes can be considered in the preferential selection phase. The topological overlap is defined by $\frac{|\Gamma(n)^j \cap \Gamma(v)^j|}{\min |\Gamma(n)|, |\Gamma(v)|} \; \forall n \neq v$ and $v \in \mathcal{V}, n \in \mathcal{N} \backslash \mathcal{V}$. A topological overlap of 1 implies that n and v can possibly connect to the same set of nodes, whereas 0 implies that the possibility that they share links to the same set of nodes is quite limited. This process promotes linking of n nodes with large number of links to its neighbors because the denominator is determined by the lower number of neighbors [11]. The ELDP selection probability of node v becomes

$$\Pi_{v-ELDP}^j = \frac{\frac{|\Gamma(n)^j \cap \Gamma(v)^j|}{\min\{|\Gamma(n)|, |\Gamma(v)|\}} \tilde{E}_v^j}{\sum_{l \in \mathcal{O}_n^j \backslash \{v\}} \frac{|\Gamma(n)^j \cap \Gamma(l)^j|}{\min\{|\Gamma(n)|, |\Gamma(l)|\}} \tilde{E}_l^j}. \tag{3}$$

3) *Energy-aware Large Degree Demote (ELDD) selection probability*: This selection probability is similar to ELDP but has opposite effects on nodes with large degrees because the denominator is determined by the maximum number of neighbors. Therefore, the selection probability becomes

$$\Pi_{v-ELDD}^j = \frac{\frac{|\Gamma(n)^j \cap \Gamma(v)^j|}{\max\{|\Gamma(n)|, |\Gamma(v)|\}} \tilde{E}_v^j}{\sum_{l \in \mathcal{O}_n^j \backslash \{v\}} \frac{|\Gamma(n)^j \cap \Gamma(l)^j|}{\max\{|\Gamma(n)|, |\Gamma(l)|\}} \tilde{E}_l^j}. \tag{4}$$

The CHs surrounding the sink node deplete their energy at a rapid rate proportional to their degrees. The WSNs with highly connected CHs suffer from the energy hole problem in which CHs around the sink die before other nodes [12], and the sink node is isolated from other nodes. To avoid such scenarios, we limit the number of links established to any node according to its *degree capacity*. The *degree capacity* is defined by the maximum number of links that can be established to a given node n in the j-th round, is denoted by $\tilde{d}_n^j$, and can be calculated by [3],

$$\tilde{d}_n^j = \widehat{d}_n^0 \frac{E_n^j}{\widehat{E}_n} \quad \forall n \in \mathcal{N} \tag{5}$$

where $\widehat{d}_n^0$, i.e., the maximum number of links that can be supported at initialization. Nodes that reach their *degree capacity* are removed from optional sets of other nodes.

To summarize, the phases of the proposed large-scale WSNs evolution model are described as follows:

- *Initialization phase*: A total of $|\mathcal{N}|$ nodes are distributed over the region of interest and the sink node is placed in the center of the region. Beginning with the sink node, we compute either its neighborhood overlap or topological overlap with the CHs in its range. A number of CHs $\bar{C}$ with the highest overlap are selected to establish connections to the sink node. Therefore, these CHs are added to the graph and its set of vertices $\mathcal{V}$.
- *Growth phase*: A new node (CH or SN) is added to the graph and establishes links to existing nodes that have not reached their *degree capacity*.
- *Preferential selection phase*: The new node links to one of the existing nodes in its optional set based on of the probabilities ECN, ELDP or ELDD given in equations (2), (3) or (4), respectively.

A concise description of the proposed ECN model is shown in Model 1. Lines marked with $\star$ are replaced with "compute $\frac{|\Gamma(k)^j \cap \Gamma(c)^j|}{\min |\Gamma(k)|, |\Gamma(c)|} \tilde{E}_c^j$" for both ELDP and ELDD. However, lines marked with $\star\star$ are replaced with probabilities given in (3) and (4) for ELDP and ELDD, respectively.

IV. Performance Evaluations

We evaluate the performance of the proposed evolution models compared with a candidate clustering-based algorithm known as Low-Energy Adaptive Clustering Hierarchy

Model 1: Proposed ECN Evolution Model

Data: $\mathcal{N}$, $E_n^0\ \forall n \in \mathcal{N}\backslash\mathcal{K}$, $\widehat{d}_n^0\ \forall n \in \mathcal{S}$, and $\widehat{d}_n^0\ \forall n \in \mathcal{C}$
Result: $G = (\mathcal{N}, \mathcal{E})$ with $\mathcal{V} = \mathcal{N}$.
repeat for every round j **until** *all nodes die*
  /* Initialization */
$\mathcal{V} = \{k\}$;
forall the *elements of* $\mathcal{N}\backslash\mathcal{K}$ **do**
  estimate $\widehat{d}_n^j$, E_n^j, and $\tilde{E}_n^j$;
foreach c *in the set* $\mathcal{C}$ *and sink node* k **do**
⋆ compute $|\Gamma(k)^j \cap \Gamma(c)^j| \tilde{E}_c^j$;
  sort computed values;
  select largest $\bar{C}$ values and add them to $\mathcal{V}$;
/* Growth and Preferential Selection */
repeat
  foreach i *in the set* $\mathcal{N}\backslash\mathcal{V}$ **do**
    if $i \in \mathcal{C}$ **then**
      find the set of CHs that joined the graph already and in the range of CH i;
      $\mathcal{O}_i^j = \Gamma(i)^j \cap \mathcal{C} \cap \mathcal{V}$;
      foreach h *in the set* $\mathcal{O}_i^j$ **do**
⋆⋆   **if** $d_h^j \leq \widehat{d}_h^j$ **then**
          compute Π_{h-ECN}^j
⋆⋆   **else**
          set Π_{h-ECN}^j to 0
⋆⋆   link i to h with Π_{h-ECN}^j add i to $\mathcal{V}$
    else if $i \in \mathcal{S}$ **then**
      find the set of CHs or SNs that joined the graph already and in the range of SN i;
      $\mathcal{O}_i^j = \Gamma(i)^j \cap \mathcal{V}$;
      foreach z *in the set* $\mathcal{O}_i^j$ **do**
⋆⋆   **if** $d_z^j \leq \widehat{d}_z^j$ **then**
          compute Π_{z-ECN}^j
⋆⋆   **else**
          set Π_{z-ECN}^j to 0
⋆⋆   link i to z with Π_{z-ECN}^j add i to $\mathcal{V}$
until *all nodes in* $\mathcal{N}$ *are added to* $\mathcal{V}$;

(LEACH) [9]. The LEACH method is an adaptive clustering algorithm that uses randomization to balance the energy consumption of nodes. Nodes elect to serve as CHs with a given probability, and nodes choose to connect to CHs that require minimum communication energy. Our performance evaluations focus on several performance measures. First, the degree of the nodes distribution is used to verify whether the generated topologies are scale-free, which implies fault-tolerance. Second, the average length of the shortest path among nodes is correlated with energy-efficiency and aggregation reliability. Third, the network's life-time is used as a measure of energy efficiency. The fourth measure is the balance of energy consumption among the network's nodes.

To provide a clear visual representation of sample topologies generated by the proposed models and LEACH, we generate topologies for a small set of nodes. Fig. 2 shows the topologies of 100 nodes. It is clear that different selection probabilities generate different topologies and LEACH generates a topology with two hops at most, a large distance between linked nodes, and highly connected CHs.

A. Simulation Setup

We simulate 3000 nodes with node IDs ranging form 1 to 3000. Nodes are randomly and uniformly distributed over a $300m$ by $300m$ square region; 10% of the nodes are CHs while the remainder are SNs. The SNs initial energy $E_n^0\ \forall n \in \mathcal{S}$ follows a uniform distribution in the range $0.3J$ to $0.7J$ with a maximum communication range of $30m$. However, the CHs have higher energies and $E_n^0\ \forall n \in \mathcal{C}$ follow a uniform distribution in the range $0.7J$ to $1J$; their maximum communication range is $60m$. In LEACH, CHs are not chosen a priori, and any SN can elect to serve as a CH. For fair comparison, the SNs in the LEACH network are assigned the average energy of all nodes in the proposed network, i.e., $0.535J$. Other simulation parameters are $\bar{C} = 3$, $\widehat{d}_n^0 = 7\ \forall n \in \mathcal{S}$, and $\widehat{d}_n^0 = 20\ \forall n \in \mathcal{C}$. We run the simulation experiment over a number of rounds until all nodes die. The aggregation phase consists of four measurements.

B. Scale-free

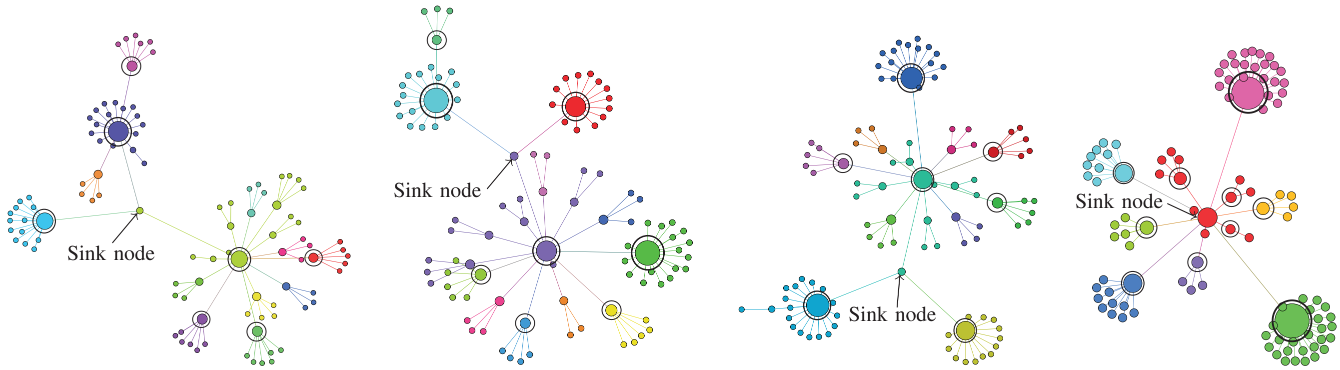
A scale-free network is a network whose degree distribution denoted by $P(d_n^j)\ \forall n \in \mathcal{N}$ follows a power law, i.e., for large d_n^j, $P(d_n^j) \sim {d_n^j}^{-\lambda}$, where λ is the power law exponent. On a log-log plot, we perform linear regression analysis to verify whether the generated topologies exhibit scale-free properties. Our regression analysis shows that the λ values for ECN, ELDP, ELDD and LEACH are 1.457, 1.506, 1.635 and 1.311, respectively. The larger the value of λ, the more obvious the power law characteristics. The corresponding adjusted R^2 values are 0.7681, 0.7896, 0.8983 and 0.6502 for ECN, ELDP, ELDD and LEACH, respectively. Adjusted R^2 is a parameter in the period $[0, 1]$, which is used to measure the goodness of fit to the logarithmic distribution curve. A larger value of adjusted R^2 implies greater resemblance to scale-free properties, where a small portion of nodes have a large degree and vice versa [13]. All four topologies display the scale-free characteristics, within which ELDD is the most obvious and LEACH is the most unapparent. The constraint imposed on the node degrees in (5) slightly reduces the network interconnectivity and its resemblance to scale-free properties; however, it improves the network energy efficiency.

C. Average Length of Shortest Paths

The average length of the shortest paths measures the average number hops of shortest paths among any pair of nodes and can be calculated by [14]

$$\bar{L} = \frac{1}{\frac{1}{2}|\mathcal{N}|(|\mathcal{N}| - 1)} \sum_{i \neq j} h_{ij}, \quad \forall i, j \in \mathcal{N}, \tag{6}$$

where h_{ij} is the length of the shortest path between any two nodes i and j. Fig. 3 shows the average length of the shortest path for the different topologies.



(a) ECN-generated topology (b) ELDP-generated topology (c) ELDD-generated topology (d) LEACH-generated topology

Figure 2. Topologies generated by the proposed models and LEACH in which the node sizes represent their current degree and cluster heads are circled. The topologies are not drawn to scale.

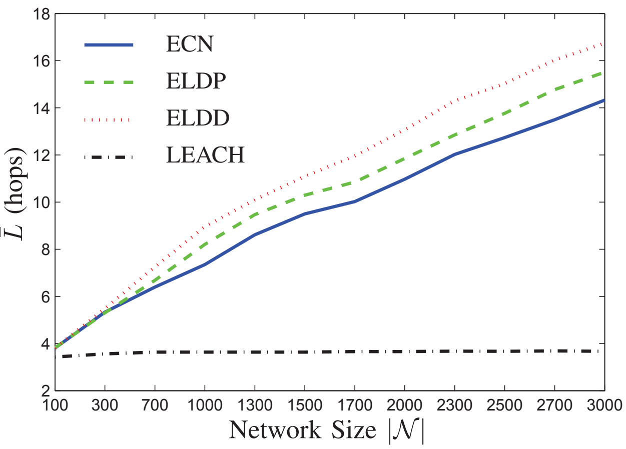


Figure 3. Average length of shortest paths among nodes versus network size of topologies generated by the proposed models, compared with the average length of the shortest path of the LEACH-generated topology.

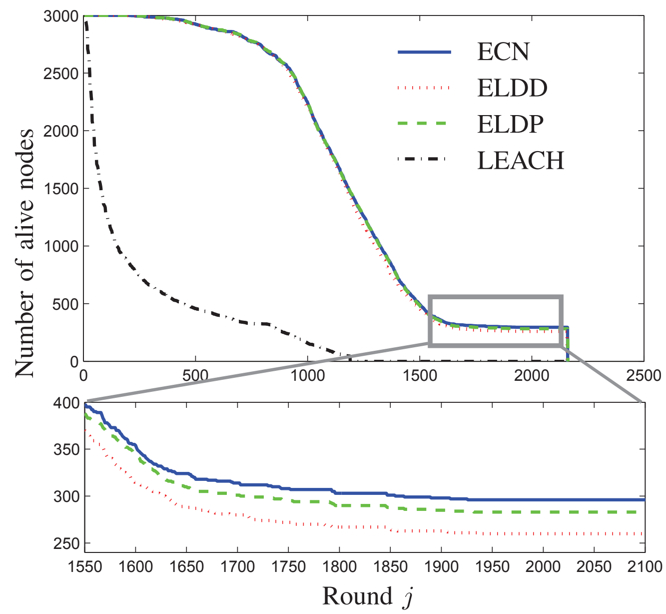


Figure 4. Number of live nodes vs. number of rounds in the network's lifespan.

The LEACH limits the maximum number of hops connecting a node and the sink to two, and thus, the longest path between two nodes is four. Among the plots shown in Fig. 3, the LEACH topology has the shortest average path length ranging from 3.526 to 3.773 hops. Among the proposed models, the smallest $\bar{L}$ is shown for ECN, followed by those ELDP and ELDD, respectively. Topologies with relatively small $\bar{L}$ are desired to reduce data processing and incurred forwarding energy consumption. However, rather small $\bar{L}$ implies long distance links that require higher transmission power and higher energy consumption. Furthermore, much smaller $\bar{L}$ overloads the CHs with large number of connections and exhausts their batteries, as can be observed clearly from the sample LEACH topology shown in Fig. 2(d) and Fig. 5.

D. Energy Efficiency

The central issue of this work is generating energy-efficient topologies that prolong the large-scale WSNs lifespan. In this section, we evaluate and compare the energy efficiency of the proposed models and LEACH. The number of live nodes over the lifespan of various topologies are shown in Fig. 4. The network begins with 3000 nodes and persists until all nodes batteries are exhausted, i.e., all nodes die.

The simulation results show that the proposed models outperform LEACH and persist for nearly twice the lifespan of the LEACH topology. This improvement is attributed to the balance between the average shortest path length and node degrees in the proposed evolution models. The constraint imposed by LEACH on the number of hops from the sensors to the sink forces the nodes to transmit over large distances and overloads them with large numbers of connections relative to their *degree capacities*, as demonstrated in Fig. 5. This extra energy consumption out weights the data-processing-related energy savings achieved by limiting the number of hops. Therefore, distant sensors are exhausted at a faster rate, as demonstrated by the steep slope of the LEACH curve in Fig. 4.

Fig. 5 shows the histogram of the ratio of nodes degrees to their *degree capacities* for the different topologies in the first round $j = 0$. The histogram for the LEACH topology is skewed right, and the energy dissipation is not balanced among the nodes. Approximately, 2700 nodes use only 20% of their *degree capacities* whereas 136 nodes are over loaded with links beyond their capacities, as shown by the pie chart in Fig. 5. In every round, these nodes die and are replaced with other nodes until the entire network is brought down. In contrast, energy dissipation in the proposed topologies is more balanced, none of the nodes is overloaded, and thus, they live longer.

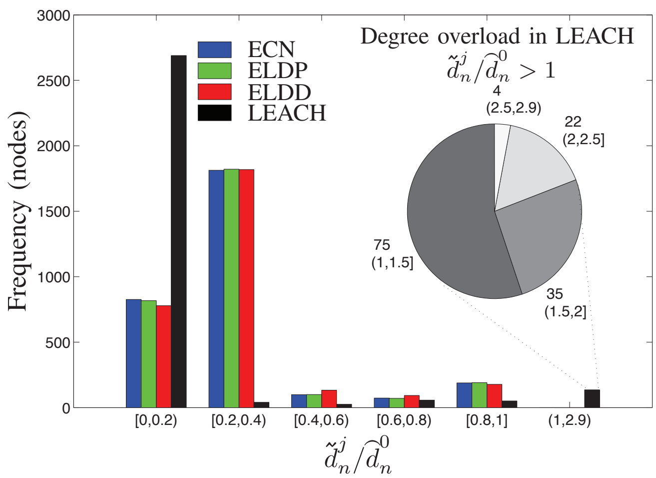


Figure 5. The histogram of the ratio of nodes degrees to their *degree capacities*, $\tilde{d}_n^j/\widehat{d}_n^0$ $\forall n \in \mathcal{N}\backslash\mathcal{K}$ and $j = 0$. The pie chart shows the number of nodes with $\tilde{d}_n^j/\widehat{d}_n^0 > 1$ for LEACH, i.e. overloaded nodes.

ECN is the most energy-efficient of the three models, followed by ELDP and ELDD. A comparison between Fig. 3 and Fig. 4 indicates the existence of a negative correlation between the average length of the shortest path and the energy efficiency of the proposed models. However, this negative correlation not upheld if the node degrees exceed their *degree capacity*, which is the case in LEACH. The ECN model attaches nodes based on their neighborhood overlap and produces topologies with smaller $\bar{L}$ than those of ELDP and ELDD, which instead consider the topological overlap for node attachment. In the ECN topology with smaller $\bar{L}$, the data-processing energy consumed is reduced, and the node degrees are bounded by their *degree capacities*, which makes it the most energy efficient among all of the considered models. In contrast, ELDD demotes the existence of nodes with large degrees, which comes at the cost of increased $\bar{L}$ required to connect all nodes, thus, producing greater scale-free resemblance but less energy efficiency. The ELDP performance falls between that of ECN and ELDD.

The above performance evaluations also provide insight into the effect of networks characteristics (i.e., scale-free and interconnectivity) on the energy efficiency of large scale WSNs.

V. Conclusions

This paper proposed three evolution models to generate fault-tolerant and energy-efficient large-scale WSNs. The three models inherit the scale-free characteristics of complex networks-based models to different degrees. The performance evaluations demonstrate that the three proposed models outperform a clustering-based model in both fault-tolerance and energy-efficiency. A negative correlation was observed between the average shortest path length $\bar{L}$ among the nodes and energy efficiency. A comparison of the proposed models performances revealed that the ECN model is the most energy efficient and that the ELDD is the most fault-tolerant, whereas the ELDP model shows a balanced performance. Furthermore, all proposed models extended the lifespan of the WSNs and generated topologies with twice the lifespan of the LEACH topology. These results are of great importance because they provide insights into the performance of models developed based on complex network theory.

Our plan for future work includes analytical evaluation of the performance of the proposed models; in addition, we plan to evaluate the performance of these models in a Network Simulator - version 2 (NS-2) environment, in which real network conditions can be simulated.

Acknowledgments

The project was partially funded by the Kuwait Foundation for the Advancement of Sciences under project code: P314-35EO-01 and NSERC, Canada, and Haixia Peng is also financially supported by the China Scholarship Council.

References

[1] L. Liu, Y. Liu, and N. Zhang, "A complex network approach to topology control problem in underwater acoustic sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3046–3055, 2014.

[2] C. Asensio-Marco and B. Beferull-Lozano, "Energy efficient consensus over complex networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 292–303, 2015.

[3] H. Zhu, H. Luo, H. Peng, L. Li, and Q. Luo, "Complex networks-based energy-efficient evolution model for wireless sensor networks," *Chaos, Solitons & Fractals*, vol. 41, no. 4, pp. 1828–1835, 2009.

[4] Y. Wang, E. Liu, X. Zheng, Z. Zhang, Y. Jian, X. Yin, and F. Liu, "Energy-aware complex network model with compensation," in *IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2013, pp. 472–476.

[5] Y. Jian, E. Liu, Y. Wang, Z. Zhang, and C. Lin, "Scale-free model for wireless sensor networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 2329–2332.

[6] D. Wang, E. Liu, Z. Zhang, R. Wang, S. Zhao, X. Huang, and F. Liu, "A flow-weighted scale-free topology for wireless sensor networks," *IEEE Communications Letters*, vol. 19, no. 2, pp. 235–238, Feb 2015.

[7] M. E. Newman, "Clustering and preferential attachment in growing networks," *Physical Review E*, vol. 64, no. 2, p. 025102, 2001.

[8] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, "Hierarchical organization of modularity in metabolic networks," *science*, vol. 297, no. 5586, pp. 1551–1555, 2002.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *IEEE Proceedings of the 33rd annual Hawaii international conference on System sciences*, 2000, pp. 3005 – 3014.

[10] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[11] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.

[12] Y. Xue, X. Chang, S. Zhong, and Y. Zhuang, "An efficient energy hole alleviating algorithm for wireless sensor networks," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 3, pp. 347–355, 2014.

[13] O. Gaudoin, B. Yang, and M. Xie, "A simple goodness-of-fit test for the power-law process, based on the duane plot," *IEEE Transactions on Reliability*, vol. 52, no. 1, pp. 69–74, 2003.

[14] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, 2002.