Toward Energy-Efficient and Robust Large-Scale WSNs: A Scale-Free Network Approach

Haixia Peng, *Student Member, IEEE*, Shuaizong Si, Mohamad Khattar Awad, *Member, IEEE*, Ning Zhang, *Member, IEEE*, Hai Zhao, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*— Due to the limited battery power of sensor nodes and harsh deployment environment, it is of fundamental importance and a great challenge to achieve high energy efficiency and strong robustness in large-scale wireless sensor networks (LS-WSNs). To this end, we propose two self-organizing schemes for LS-WSNs. The first scheme is the energy-aware common neighbor scheme, which considers the neighborhood overlap in link establishment. The second scheme is energy-aware low potential-degree common neighbor (ELDCN) scheme, which considers both neighborhood overlap in topology formation and the potential degrees of common neighbors. Both schemes generate clustering-based and scale-free-inspired LS-WSNs, which are energy-efficient and robust. However, the ELDCN scheme shows higher energy efficiency and stronger robustness to node failures, because it avoids establishing links to hub-nodes with high potential connectivity. Analytical and simulation results demonstrate that our proposed schemes outperform the existing scale-free evolution models in terms of energy efficiency and robustness.**

***Index Terms*— Large-scale wireless sensor networks, self-organizing scheme, scale-free network, energy efficiency, robustness.**

I. Introduction

WIRELESS Sensor Networks (WSNs) have been widely employed in military [2], agriculture [3], environment monitoring [4], and health monitoring [5]. WSNs consist of wireless sensor nodes with the capabilities of sensing, processing, and communications [6], [7]. The sensor nodes are usually battery powered and therefore energy limited. In most scenarios, they are deployed in unattended and hostile environments, where it is usually difficult to recharge or replace their batteries; thus, a node fails permanently when it exhausts the battery. Furthermore, the more important the role a node plays in a network (denote this node as *hub-node*), the more traffic it relays or aggregates, and the faster it exhausts its battery. Hub-node failures, leading to the failures of their connected links, deteriorate the network connectivity. Therefore, the design of energy efficient and robust WSNs becomes crucial, especially for Large-Scale WSNs (LS-WSNs) where the large network scale and monitored region result in a heavier packet load and a higher rate of energy consumption.

Multi-hop communication is considered one of the effective approaches to mitigate the high energy consumption of wireless transmission over a long distance [8]. However, the non-uniform dissipation of energy in multi-hop networks leads to an energy hole problem [9]. Hence, it is of great importance for LS-WSNs to balance the energy consumption at each node to prevent energy holes and prolong the lifespan of LS-WSNs. To this end, various protocols and schemes have been proposed for WSNs, such as MAC protocols [10] and routing protocols [11]. Many existing works have shown that clustering-based routing protocols are promising for improving energy efficiency in WSNs [12], especially in LS-WSNs [13]. Thus, in the past decades, many clustering-based protocols have been proposed, such as the stable and energy-efficient clustering (SEEC) protocol [14]. Furthermore, some related issues in clustering-based WSNs have been analyzed, such as the trade-offs between the network lifetime and transport delay [15]. However, most clustering-based protocols are suitable for WSNs of small scale, and do not consider the network physical topology, which significantly affects the network performance [16], [17].

In terms of robustness, in addition to the limited battery-powered energy and unbalanced energy consumption, hub-node failures may also be caused by harsh surroundings, natural disasters, and hostile forces [18]. With the increase of network size and monitored area, the negative impact of hub-node failures becomes more significant, leading to more communication link failures and more disjoint partitions [19]; hence, lower robustness. There have been some approaches to improve the robustness of LS-WSNs [19], [20]. In [19], the relay nodes are carefully placed to maximize network robustness. In [20], a new environmental monitoring framework is proposed. Most works only focus on network robustness. However, energy efficiency of a WSN is also very important.

Manuscript received February 23, 2016; revised May 15, 2016 and August 16, 2016; accepted October 16, 2016. Date of publication October 27, 2016; date of current version December 29, 2016. This work was supported in part by the Kuwait Foundation for the Advancement of Sciences under Project P314-35EO-01 and in part by the NSERC of Canada, National Nature Science Foundation of China, under Grant 60973022. The work of H. Peng was supported by the China Scholarship Council. This work was presented at the 2015 IEEE Globecom.

H. Peng is with the School of Information Science and Engineering, Northeastern University, Shenyang, China, and also with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: h24peng@uwaterloo.ca).

S. Si and H. Zhao are with the School of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: sishuaizong@neuera.com; zhaoh@mail.neu.edu.cn).

M. K. Awad is with the Department of Computer Engineering, Kuwait University, Kuwait City, Kuwait (e-mail: mohamad@ieee.com).

N. Zhang and X. Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: n35zhang@uwaterloo.ca; sshen@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2016.2621618

0733-8716 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Recent advances in complex networks have motivated investigation in scale-free WSNs [21], [22]. A scale-free network is the one 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. The indispensable features of scale-free networks are high energy efficiency [22] and strong robustness [23], which provide a great motivation for generating scale-free WSNs [17], [21], [22], [24]. The Barabási-Albert (BA) model[1] has been applied to generate WSNs in [26]. However, practical limitations of WSNs, i.e., limited energy capacity and transmission range, have not been considered. To improve the energy efficiency, an extension of the BA evolution model has been proposed in [22], wherein the length of connections and algebraic connectivity have been taken into account. The scale-free property of this model has been verified experimentally.

In this paper, two novel self-organizing schemes are proposed to simultaneously improve the energy efficiency and robustness of LS-WSNs by employing the scale-free network approach. The first scheme is called an energy-aware common neighbor (ECN) scheme. Taking the evolution processes of BA model as a basis, the ECN constructs a clustering-based, degree-constrained, and scale-free-inspired LS-WSN. To improve the energy efficiency of LS-WSNs, several intrinsic characteristics of LS-WSNs (including residual energy and communication range) and a typical topology information regarding the common neighbor are exploited when establishing links between any two nodes. Meanwhile, the clustering-based structure and scale-free property are maintained in the generated LS-WSN. With a large number of neighbors, i.e., *potential degree*, nodes tend to exhaust the energy rapidly, which significantly degrades the network performance. Thus, we further propose another scheme, called energy-aware low potential-degree common neighbor (ELDCN) scheme. This scheme avoids establishing links to hub-nodes with large potential degrees, which further improves energy efficiency and balances the connectivity of nodes, i.e., *node degree*, thereby providing strong robustness. The main contributions of this paper are as follows:

1) Two self-organizing schemes, ECN and ELDCN, are proposed, considering the intrinsic characteristics of LS-WSNs. To the best of our knowledge, this is the first work that topology information is exploited in generating LS-WSNs.
2) The proposed schemes construct clustering-based, degree-constrained, and scale-free-inspired LS-WSNs to achieve high energy efficiency and strong robustness.
3) An analysis model based on the heterogeneous mean-field approximation theory is provided to analyze the scale-free property[2] of our generated LS-WSNs.

The rest of this paper is organized as follows. First, the network model and energy consumption model are presented in Section II. In Section III, we present the proposed self-organizing schemes in detail, including the self-organizing steps and the corresponding preferential selection probability models. Then, we theoretically analyze the scale-free property and energy consumption of our generated LS-WSNs in Section IV. Section V is devoted to performance validation of our generated LS-WSNs, where the scale-free property, energy efficiency, and robustness are demonstrated through extensive computer simulation. Finally, in Section VI, we conclude this work.

[1]A well-known model in complex network theory that generates scale-free networks [25].

[2]A network with scale-free property has been shown some advantages, such as energy efficiency and strong robustness [22], [23].

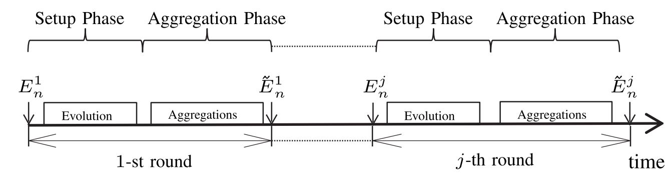


Fig. 1. Proposed schemes operation phases over time.

II. System Model

In this section, we introduce the heterogeneous LS-WSN model first, and then describe the energy consumption model.

A. Network Model

We consider a heterogeneous LS-WSN consisting of a sink node b and a set of N nodes denoted by $\mathcal{N}$. The number of nodes in LS-WSN, N, is on the order of thousands and is larger by at least one order of magnitude compared with the number of nodes in the traditional WSNs. These N nodes can be classified into two groups, cluster-head-nodes (CHs) $\mathcal{C} = \{1, \cdots, c, \cdots, C\}$ with $C = |\mathcal{C}|$, and member-sensor-nodes (SNs) $\mathcal{S} = \{1, \cdots, s, \cdots, S\}$ with $S = |\mathcal{S}|$. The CHs usually have higher hardware requirements, such as initial energy and transmission power, since they need to aggregate and relay the sensed data from either other CHs or SNs to the sink node. The number of CHs, C, is smaller than that of SNs, S; thus, the ratio $\rho = \frac{C}{N} < 1$. The sink node b is located at the center of a $U \times U$ two-dimensional monitored region, and CHs and SNs are randomly distributed over this region. Then the average node density can be described as $D = \frac{N}{U^2}$. Nodes within either the subset $\mathcal{C}$ or $\mathcal{S}$ have the same specifications, such as sensing range, transmission range, and computational power, and the specifications for CHs are higher than these for SNs as usual.

As shown in Fig. 1, the sensing operation is dynamically performed in two phases: setup phase and aggregation phase. During the setup phase, the LS-WSN is generated by our proposed schemes and represented by a geometric graph, $G(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes that have been added to the graph, and $\mathcal{E}$ is the set of communication links established to the $|\mathcal{V}|$ nodes. The setup phase starts with the sink node, i.e., $\mathcal{V} = \{b\}$, and other nodes in $\mathcal{N}$ are iteratively added to $\mathcal{V}$ until all nodes are added, i.e., $\mathcal{V} = \mathcal{N} \cup \{b\}$. Then, the aggregation phase starts and it consists of τ time instants of data aggregation. At each of these instants, SNs sense the monitored area and transmit the sensed data to the sink.

If one node relays the sensed data for another node during the aggregation phase, the communication link between these

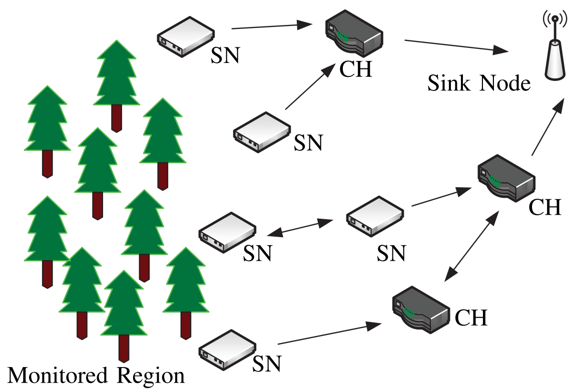


Fig. 2. Illustration of the network model with communication links

two nodes exists. As shown in Fig. 2, there are four types of communication links: bidirectional links among CHs and among SNs, respectively, directional links from CHs to the sink node and from SNs to CHs, respectively.

Denote $\Gamma(n)$ as the set of neighboring nodes within the communication range of node n. The number of neighboring nodes of node n is called as its *potential degree*, which is denoted by $|\Gamma(n)|$. The set of candidate nodes for establishing links to node n are those that have not been added to $\mathcal{V}$ and are within $\Gamma(n)$, and can be written as

$$\mathcal{O}_n = \Gamma(n) \cap (\mathcal{N} \setminus \mathcal{V}). \tag{1}$$

For a given node $n \in \mathcal{N}$, i.e., either a CH or a SN, E_n^0 denotes its initial energy, which is assumed to follow a uniform distribution in the range $[\underset{\smile}{E}_n, \widehat{E}_n]$. However, $E_n^0 \in [\underset{\smile}{E}_s, \widehat{E}_s]$ if n is a SN, i.e., $n \in \mathcal{S}$; $E_n^0 \in [\underset{\smile}{E}_c, \widehat{E}_c]$ if n is a CH, i.e., $n \in \mathcal{C}$. Heterogeneous network structure is considered in this work, where CHs are assumed to have higher initial energy and larger communication ranges than SNs. Therefore, $\underset{\smile}{E}_s \leq \widehat{E}_s \leq \underset{\smile}{E}_c \leq \widehat{E}_c$, $\forall s \in \mathcal{S}$ and $\forall c \in \mathcal{C}$, and $R_c \geq R_s$, where R_c and R_s are the communication ranges of CHs and SNs, respectively.

B. Energy Consumption Model

During the aggregation phase, both SNs and CHs drain their energy on data processing, reception and transmission. According to the radio model in [8], the energy consumption of transmitting and receiving l bits of data over a distance d can be expressed by

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + ld^2\varepsilon_{fs}, & d \leq d_0, \\ lE_{elec} + ld^4\varepsilon_{mp}, & \text{otherwise}, \end{cases} \tag{2}$$

and

$$E_{Rx}(l,d) = lE_{elec}, \tag{3}$$

respectively, where E_{elec} is the transmitter or the receiver circuit energy consumption to transmit one bit of data, d_0 is the distance reference point, which can be given by $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$, and ε_{fs} and ε_{mp} denote the energy for power amplifications in free space channel model and multi-path fading channel model, respectively. The free space channel model is adopted when $d \leq d_0$; and the multi-path fading channel model is used otherwise.

In addition to collecting sensed data from other nodes, CHs also aggregate it before forwarding it to the sink. Let E_{Ag} be the energy consumption in each CH for aggregating one bit of data. The SNs that depend on a given CH to relay their data to b are called as the child nodes of this CH. Let $\Lambda(c)$ be the number of child nodes of CH c. The energy consumption of the CH for receiving, aggregating, and transmitting data can be rewritten as

$$E_c(l,d) = \Lambda(c)lE_{Rx}(l,d) + \Lambda(c)lE_{Ag} + E_{Tx}(l,d). \tag{4}$$

In addition, adding a node to $\mathcal{V}$ requires establishing a link to another node, which consumes a constant amount of energy, β, on these two connected nodes as in [26] and [27]. In this work, we assume that the sensed data of each SN is a packet with l bits and each CH aggregates its received sensed data into a packet with l bits.

III. Self-Organizing Schemes

The BA model has been proposed by Wang *et al.* [25] to generate scale-free networks. In a scale-free network, the node degree exhibits a power-law distribution. The degree distribution, $P(k)$, is defined as the probability that a node is with degree k over the whole network. This probability is independent of the network size. As in [25], $P(k)$ of the scale-free network can be written as $P(k) \approx \alpha k^{-\gamma}$, where α is a normalization constant and γ is the characteristic exponent.

Taking the evolution processes of the BA model and the clustering-based structure as a basis, we present two novel self-organizing schemes, ECN and ELDCN, to generate energy efficient and robust LS-WSNs. Due to high connectivity of the hub-nodes in the scale-free LS-WSNs, they may die faster than other nodes [29]. This is particularly true when the network size increases and hard cutoff limit is not set on the node degree. Therefore, it causes unbalanced energy consumption and makes the network vulnerable to the energy hole problem, let alone the degradation in robustness. Unlike most existing works [22], [24], we limit the number of links established to nodes based on their residual energy. Furthermore, the network topology is dynamically updated as shown in Fig.1, which facilitates recovery from structural changes such as node failures.

Next, we introduce self-organizing steps of the proposed schemes. Then, we present the proposed preferential selection probability models for ECN and ELDCN. These probability models are used to establish links among nodes in the proposed schemes.

A. Self-Organizing Steps

Each setup phase in Fig. 1 starts from the sink node only, i.e., $\mathcal{V} = \{b\}$. Then, in each iteration, a new-incoming node, either SN or CH, is added to $\mathcal{V}$ based on either ECN or ELDCN selection probability model until all nodes are added to $\mathcal{V}$. The proposed self-organizing schemes consists of the following two major steps and a preferential selection probability:

- *Initialization*: N nodes are randomly deployed over the monitored region and sink node b is placed in the center of this region. Beginning with the sink node, i.e., $\mathcal{V} = \{b\}$, we compute its neighborhood overlap with the CHs within its communication range. Here, the neighborhood overlap is measured by either $\lambda_{(n,b)}$ or $\mu_{(n,b)}$ between the sink node and its neighboring CHs. Whereas $\lambda_{(n,b)}$ is a simple measure of neighborhood overlap, $\mu_{(n,b)}$ measures the neighborhood overlap but depresses the contribution of common neighbors with high potential degrees. The M CHs with highest neighborhood overlap are nominated to establish links to the sink node and added to the set $\mathcal{M} = \{1, \cdots, m, \cdots M\}$, where $M = |\mathcal{M}|$. The nominated CHs establish links to the sink node when they are randomly selected in the *Growth* step.
- *Growth*: In each iteration, a new-incoming node n, CH or SN, is randomly selected from $\mathcal{N}$ and is added to $\mathcal{V}$. A new-incoming node is added by establishing a direct communication link to b if it belongs to $\mathcal{M}$. Alternatively, it is added by establishing a link to one node within its set of candidate nodes O_n given in equation (1).
- *Preferential selection*: One node in O_n is preferred over others when establishing links in *Growth* step based on the selection probabilities in equations (8) and (10) of ECN and ELDCN, respectively.

B. Preferential Selection Probability Models

Nodes drain their batteries during sensing, receiving and transmitting data over time; thus, their residual energy change over the network lifespan. In the proposed schemes, the network structure is regularly updated in every setup phase based on the current residual energy of each node. At the beginning of the setup phase in the j-th round, let E_n^j denote the residual energy of node n, $\forall n \in \mathcal{N}$. Each round consists of one setup phase and the following aggregation phase. The elapsed time between any two consecutive setup phases is called as a *round*.

Nodes consume energy at a rate proportional to their degrees and die at a faster rate when overloaded with a large number of links. In order to avoid such scenarios and the energy hole problem, we limit the number of communication links established to a given node according by its *degree capacity*. The *degree capacity* is defined as the maximum number of communication links that can be established to a given node [1]. For a node n in the j-th round, the *degree capacity*, $\tilde{k}_n^j$, evaluates to

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

where $\widehat{k}$ is the maximum number of communication links that can possibly be established to the node with maximum initial energy $\widehat{E} = \max\{E_n^0\}\ \forall n \in \mathcal{N}$. In each round, the maximum energy consumption for link establishment at node n is limited by its *degree capacity*, and can be described as $\beta\tilde{k}_n^j$. The residual of $\beta\tilde{k}_n^j$ in the j-th round is denoted by $\tilde{E}_n^j$ and can be given by

$$\tilde{E}_n^j = \beta\left(\tilde{k}_n^j - k_n^j\right) \qquad \forall n \in \mathcal{N}, \tag{6}$$

where k_n^j is the current degree of node n in setup phase of the j-th round.

In our proposed self-organizing schemes, $\tilde{E}_n^j$ is considered in preferential selection of the *Growth* step. Linking a new-incoming node to the node with higher $\tilde{E}_n^j$ can balance the energy consumption in the network. However, a communication link over a long distance consumes more energy, thus drains the battery at a faster rate. Moreover, the chance of interference is higher over long distances, which degrades the communication performance. Therefore, it is necessary to limit the communication range, and link the new-incoming node to its neighboring nodes to conserve residual energy.

It has been proved in [15] and [16] that network physical topology significantly affects the network performance. Connecting a node to its nearest node improves the energy efficiency of WSNs [22]. At the same time, from [30], the larger the number of common neighbors that two nodes have, the shorter the distance between them. To simultaneously take the distance and network topology information into account, a common neighbor which is a typical local topology information in complex network [31], is introduced into our preferential selection probability models as follows.

- *Selection probability of ECN:* The more common neighbors that two nodes have, the higher proximity to one another. In the j-th round, the number of common neighbors between new-incoming node n and node o within $O_n^j = \Gamma(n)^j \cap (\mathcal{N} \setminus \mathcal{V})$ can be described by
$$\lambda_{(n,o)}^j = |\Gamma(n)^j \cap \Gamma(o)^j| \qquad \forall n \neq o, n \in \mathcal{N} \setminus \mathcal{V}, o \in O_n^j. \tag{7}$$
Scaling $\lambda_{(n,o)}^j$ by the residual energy for link establishment, $\tilde{E}_o^j$, gives higher preference to nodes with higher residual energy. In the j-th round, the ECN selection probability of node o, i.e., the probability that n establishes one communication link to o is given by
$$\Pi_{o-ECN}^j = \frac{\lambda_{(n,o)}^j \tilde{E}_o^j}{\sum_{w \in O_n^j} \lambda_{(n,w)}^j \tilde{E}_w^j}. \tag{8}$$

The potential degree of the n^{th} node, $|\Gamma(n)|$, is determined once N nodes are deployed in the monitored area. A node with higher potential degree are expected to have a larger degree and faster depletion of its battery. Consider a pair of connected nodes n and n', and a node z being a common neighbor to both of them, i.e., $z \in \Gamma(n) \cap \Gamma(n')$. A larger $|\Gamma(z)|$ results in a higher chance that one of the nodes n or n' is isolated when the other node dies. In addition, the effect of residual energy in each node may lead node n or n' to connect node z, $z \in \Gamma(n) \cap \Gamma(n')$, thereby resulting in higher connectivity and unbalanced energy consumption in node z, $z \in \Gamma(n) \cap \Gamma(n')$. To avoid this scenario, the following selection probability is proposed.

- *Selection probability of ELDCN:* We extend the simple measure of the neighborhood overlap in $\lambda_{(n,o)}^j$ to depress the contribution of the common neighbors with high potential degree as follows

$$\mu^{j}_{(n,o)} = \sum_{z\in\Gamma(n)^{j}\cap\Gamma(o)^{j}} \frac{1}{|\Gamma(z)^{j}|}, \quad \forall n \neq o, n \in \mathcal{N} \setminus \mathcal{V}, o \in \mathcal{O}^{j}_{n}. \tag{9}$$

We scale $\mu^{j}_{(n,o)}$ by $\tilde{E}^{j}_{o}$ to give a higher preferential selection probability to the node with higher residual energy for link establishment. Then, in the j-th round, the ELDCN selection probability of node o can be given by,

$$\Pi^{j}_{o-ELDCN} = \frac{\mu^{j}_{(n,o)}\tilde{E}^{j}_{o}}{\sum_{w\in\mathcal{O}^{j}_{n}} \mu^{j}_{(n,w)}\tilde{E}^{j}_{w}}. \tag{10}$$

Algorithm 1 provides a concise description of the ECN self-organizing schemes. For ELDCN, lines marked with $\star$ are replaced with "compute $\mu^{j}_{(b,c)}\tilde{E}^{j}_{c}$" and lines marked with $\star\star$ are replaced with probability given in (10).

Algorithm 1: Proposed ECN Self-Organizing Scheme

Data: $\mathcal{N}$, $\widehat{k}$, E^{0}_{n} $\forall n \in \mathcal{N}$
Result: $G = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \mathcal{N} \cup \{b\}$.
repeat for every round j **until** *all nodes die*;
`/* Initialization */`
$\mathcal{V} = \{b\}$ **forall the** *elements of* $\mathcal{N}$ **do**
  estimate $\tilde{k}^{j}_{n}$, E^{j}_{n}, and $\tilde{E}^{j}_{n}$
foreach c *in the set* $\mathcal{C}$ *and sink node* b **do**
$\star$ compute $\lambda^{j}_{(c,b)}\tilde{E}^{j}_{c}$;
$\star$ sort computed $\lambda^{j}_{(c,b)}\tilde{E}^{j}_{c}$;
$\star$ add M nodes with largest $\lambda^{j}_{(c,b)}\tilde{E}^{j}_{c}$ to the set of nominated nodes $\mathcal{M}$;
`/* Growth and Preferential Selection */`
repeat
  foreach i *in the set* $\mathcal{N} \setminus \mathcal{V}$ **do**
    if $i \in (\mathcal{C} \setminus \mathcal{M})$ **then**
      find the set of CHs that have not been added to $\mathcal{V}$ and in the range of CH i,
      $\mathcal{O}^{j}_{i} = \Gamma(i)^{j} \cap \mathcal{C} \cap (\mathcal{N} \setminus \mathcal{V})$;
      foreach o *in the set* $\mathcal{O}^{j}_{i}$ **do**
$\star\star$ **if** $k^{j}_{o} < \tilde{k}^{j}_{o}$ **then**
          compute Π^{j}_{o-ECN}
$\star\star$ **else**
          set Π^{j}_{o-ECN} to 0
$\star\star$ link i to o with Π^{j}_{o-ECN}, add i to $\mathcal{V}$ and update $\mathcal{E}$
    else if $i \in \mathcal{M}$ **then**
      link i to b, add it to $\mathcal{V}$ and update $\mathcal{E}$
    else if $i \in \mathcal{S}$ **then**
      find the set of CHs or SNs that have not been added to $\mathcal{V}$ and in the range of SN i,
      $\mathcal{O}^{j}_{i} = \Gamma(i)^{j} \cap (\mathcal{N} \setminus \mathcal{V})$;
      foreach o *in the set* $\mathcal{O}^{j}_{i}$ **do**
$\star\star$ **if** $k^{j}_{o} < \tilde{k}^{j}_{o}$ **then**
          compute Π^{j}_{o-ECN}
$\star\star$ **else**
          set Π^{j}_{o-ECN} to 0
$\star\star$ link i to o with Π^{j}_{o-ECN}, add i to $\mathcal{V}$ and update $\mathcal{E}$
until *all nodes in* $\mathcal{N}$ *are added to* $\mathcal{V}$;

IV. ANALYSIS

In Section III, we have proposed two self-organizing schemes for constructing degree-constrained and scale-free-inspired LS-WSNs. In this section, we apply heterogeneous mean-field approximation to analyze the scale-free property of the generated LS-WSNs [32]. In addition, the energy consumption and robustness of LS-WSNs are investigated in this section.

A. Degree Distribution

For some complex networks, the concepts of statistical physics, such as the clustering coefficient and degree distribution, are very important and useful to analyze the network features. Here, we prove that the generated LS-WSNs display a scale-free property. One traditional and direct way to prove the scale-free property is checking whether the degree distribution asymptotically follows a power law [21].

There exist several methods to evaluate the degree distribution of complex networks, such as rate equation, master equation, and mean field theory [32]. Similar to [17], the degree distribution of the generated LS-WSN $P(k)$ is analyzed via applying the heterogeneous mean-field approximation in this section. The following theorem, shown in Appendix A, demonstrates the power law property of the degree distribution.

Theorem 1: In a LS-WSN constructed based on ECN, the node degree distribution, $P(k)$, follows a power law. The power-law exponent, γ, is a function of the network parameters and given by

$$\gamma = \frac{\pi R_s^2 R_c^2(\bar{E} + \beta)}{U^2\beta(R_c^2 - \rho R_c^2 + \rho^2 R_s^2)} + 1. \tag{11}$$

where $\bar{E}$ is the mean value of $\beta\tilde{k}_n$, $n \in \mathcal{N}$.

From Theorem 1, the degree distribution follows a power law. Therefore, LS-WSNs generated by ECN display the scale-free property. Similar to the analysis of degree distribution for the ECN, we can obtain the degree distribution in a LS-WSN generated based on ELDCN, which is the same as (11).

B. Energy Consumption and Network Robustness

According to the characteristics of the hierarchical structure network and the energy consumption model given in Subsection II-B, the total energy consumption for aggregating sensed data can be expressed as,

$$E_{Ta} = (N - M)lE_{Ag}. \tag{12}$$

Furthermore, the energy consumption on data reception and transmission depends on the number of hops between two nodes. Let $\bar{H}$ denote the average path length in terms of hops between a node, either SN or CH, and the sink node. $\bar{H}$ can be written as

$$\bar{H} = \frac{1}{N} \sum_{n \in \mathcal{N}} h_{(n,b)}, \tag{13}$$

where $h_{(n,b)}$ is the path length from node n to the sink node b. Then the following theorem, proved in Appendix B, demonstrates the range of $\bar{H}$ in topologies generated by the proposed schemes.

Theorem 2: The range of $\bar{H}$ in our considered LS-WSNs can be written as

$$\bar{H}_{min} \leq \bar{H} \leq \frac{1+N}{2}, \tag{14}$$

where

$$\bar{H}_{min} = \left(1 - \frac{M(1-\bar{K}^{\omega})}{N(1-\bar{K})}\right)(\omega + 1) + \frac{M(1+\omega\bar{K}^{\omega+1} - (1+\omega)\bar{K}^{\omega})}{N(1-\bar{K})^2}. \tag{15}$$

Here, $\omega = \lceil \log_{\bar{K}}[1 - \frac{N}{M}(1-\bar{K})] \rceil - 1$ is the maximum number of hops between a node and the sink node, with $\bar{K} = (\bar{E}/\widehat{E})\widehat{k}$ being the average degree capacity of all nodes in the LS-WSN.

The following theorem can be directly derived from Theorem 2.

Theorem 3: Denote E_{To} as the total energy consumption for data aggregation, reception, and transition over one time instant in the aggregation phase, then we have

$$\bar{H}_{min} N l(E_{Tx} + E_{Rx}) + E_{Ta} \leq E_{To} \leq \frac{1+N}{2} N l(E_{Tx} + E_{Rx}) + E_{Ta}. \tag{16}$$

From the degree distribution in (11), the probability that a given node has a high degree is small, and therefore only a small proportion of nodes have high degrees in the whole LS-WSN. Therefore, only a small number of nodes drain their batteries at a fast rate. In addition, because nodes are uniformly distributed, they have equal probability of being overloaded. Since the energy consumption rate in each node is updated dynamically during each setup phase, the energy consumption at each node is balanced in our generated LS-WSNs.

Due to the balanced energy consumption in each node, hub-node failures decreased over the network lifespan. Each node has a same small probability to be a hub-node; thus, there are only a few hub-nodes in the network and the network robustness to hub-node failures is improved.

V. Simulations and Performance Evaluations

The proposed self-organizing schemes are evaluated under different LS-WSN parameters and are compared with four candidate models that generate scale-free networks. The first one is the BA model [25] and is the classical model for generating scale-free networks. Based on the BA model, the new-incoming node n establishes a link to one of the nodes that are connected to the sink node and with the highest degrees. The second one is an extended model of the BA model, where we consider the limited communication range of both CHs and SHs in order to make it comparable to the proposed schemes.

TABLE I
Simulation Parameters

Parameter	Value	Parameter	Value
N	6000	β	1×10^{-15}
$U \times U(m^2)$	$550m \times 550m$	ρ	0.3
$\underset{\smile}{E}_s$	0.6	$\overset{\frown}{E}_s$	0.8
$\underset{\smile}{E}_c$	0.8	$\overset{\frown}{E}_c$	1
$\overset{\frown}{E}$	1	$\overset{\frown}{k}$	22
R_c	120 m	R_s	80 m
τ	20	M	10
l	4000 $bits$	E_{Ag}	5×10^{-9}
ε_{fs}	10×10^{-12}	ε_{mp}	0.0013×10^{-12}
E_{elec}	50×10^{-9}	d_0	87.7 m

We refer to this model as the modified-BA (M-BA) model. In the M-BA model, the new-incoming node n establishes a link to the node with the highest degree in O_n. The third model is the distributed solution (DS) in [33]. In DS, the new-incoming node n establishes a link to the node which is connected to the sink node, has higher degree, and therefore incurs less communication energy consumption. A trade off factor has been introduced in DS to balance the degree and communication energy consumption. The fourth model is the semi-randomized growth algorithm (SRA) proposed in [34]. SRA is an evolution model for generating scale-free peer-to-peer networks, where a hard cutoff has been introduced to limit the node degree. Based on SRA, the new-incoming node n establishes a link to the node with a degree matching a randomly generated number, and at the same time connected to the sink node.

Our comparison is performed under the same network setup and energy consumption model. The dynamics of our proposed schemes are simulated in Myeclipse.[3] The LS-WSNs topologies generated in Myeclipse are imported to Gephi[4] where the scale free property is evaluated. Then, by analyzing the topological data in MATLAB, the average transmission length, the energy efficiency and network robustness are evaluated. All major parameters in our simulation are lifted in Table I unless stated otherwise. For DS, the trade off factor is 0.4. For SRA, the maximum degree of each node is set to $\widehat{k}$, and degree distribution follows the power law with an exponent of 3.

A. Scale-Free Property

Degree distribution is considered one of the most important metrics for analyzing complex networks. It has a great impact on the transmission dynamics and network reliability [24], [34]. The theoretical analysis in Subsection IV-A shows that the degree distribution in our generated LS-WSNs follows the power law. In this subsection, we present the linear-regression analysis results on the degree distribution in topologies generated by the proposed schemes. The results show that the generated topologies are approximately scale-free.

Fig. 3 and Fig. 4 show the degree distribution of ECN and ELDCN, respectively, for $\rho = 0.3$ and $\rho = 0.5$. The results

[3]A powerful, full-featured Java IDE for network simulation.

[4]Gephi is a software for analyzing complex networks

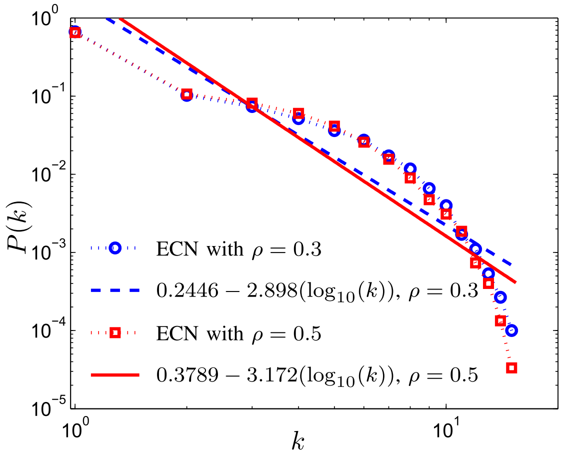


Fig. 3. Degree distributions and regressing analysis results for ECN with $\rho = 0.3$ and $\rho = 0.5$

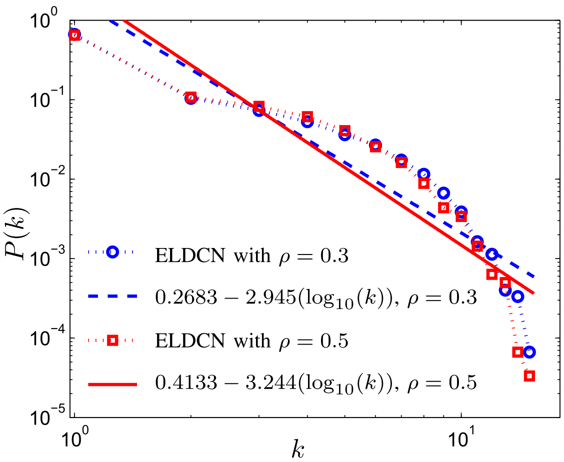


Fig. 4. Degree distributions and regressing analysis results for ELDCN with $\rho = 0.3$ and $\rho = 0.5$

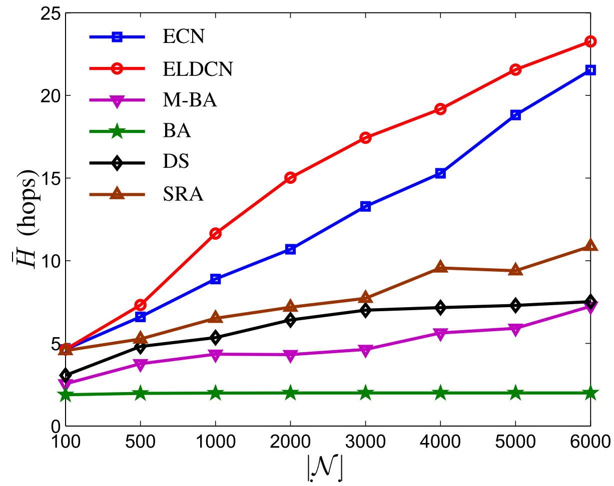


Fig. 5. Average transmission length versus network size.

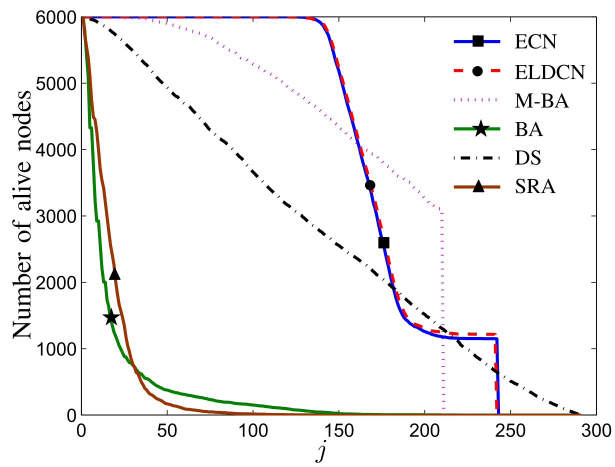


Fig. 6. Number of alive nodes vs. number of rounds in the network lifespan.

shown are averaged over 10 trials. The slopes of regression lines in Fig. 3 and Fig. 4 show that the simulated γs for ECN and ELDCN are 2.898 and 2.945, respectively, with $\rho = 0.3$. γ increases to 3.172 and 3.244, respectively, for ECN and ELDCN with $\rho = 0.5$. Given the simulation parameters in Table I and analytical results, γ for ECN and ELDCN is 2.5916 for $\rho = 0.3$ and 3.0230 for $\rho = 0.5$. Due to the heterogeneous mean-field approximation applied in our theoretical derivation in Subsection IV-A, there is a small gap between the theoretical and simulated γ. Both of the theoretical and simulated results show that γ increases with the increase of ρ. The adjusted R^2 values[5] corresponding to the four fitting lines of ECN with $\rho = 0.3$, ECN with $\rho = 0.5$, ELDCN with $\rho = 0.3$ and ELDCN with $\rho = 0.5$ are 0.8566, 0.8264, 0.8443, and 0.8168, respectively. The four R^2 show that the LS-WSNs topologies generated by ECN and ELDCN display approximate scale-free properties [34].

[5]Adjusted R^2 is a fitting optimization index which is used to measure how well terms fit a curve or line, which falls in [0, 1]. A larger value of adjusted R^2 implies greater resemblance to scale-free property and vice versa.

B. Average Path Length

Fig. 5 shows $\bar{H}$ in the LS-WSNs generated by ECN, ELDCN, M-BA, BA, DS, and SRA. Results are averaged over 20 trials. From the Fig. 5, the $\bar{H}$ for the BA model is the smallest in comparison to the other considered five models, because a node operating under the BA model prefers to establish a communication link to the node with the higher degree without considering any other factors. Compared with the BA model, $\bar{H}$ of the M-BA is larger due to the limited communication range of each node. However, the higher degree nodes still have higher preference in the M-BA model; its $\bar{H}$s are less than those of ECN, ELDCN, DS, and SRA.

The number of communication links established to a node is limited by its degree capacity in the proposed schemes and SRA, which results in a high average path length. This explains the larger $\bar{H}$ for ECN, DCN, and SRA in comparison to the other three models. According to equation (2), the energy consumption for transmitting the sensed data increases at a faster rate with d for $d > d_0$. Thus, for ECN, DCN, M-BA, and DS, which have considered the constrained communication range, a smaller $\bar{H}$ implies more links with longer distances

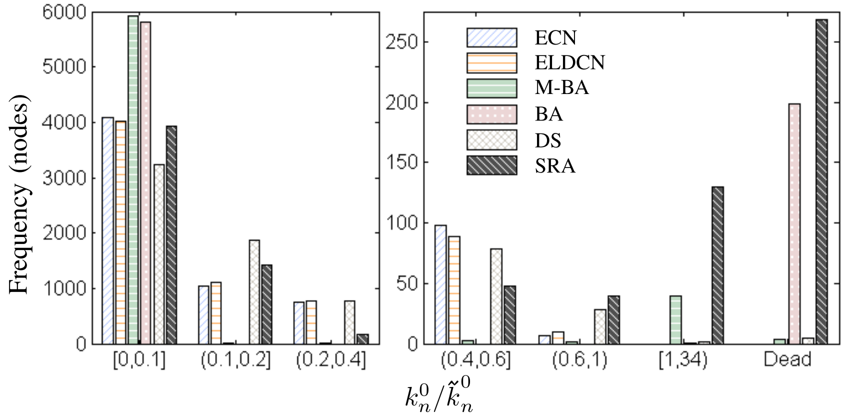


Fig. 7. Histogram of the ratio of node degree to their *degree capacities*, $k_n^0/\tilde{k}_n^0$, $\forall n \in \mathcal{N}$.

and higher energy consumption for transmitting data. Consequently, some nodes are overloaded with a large number of connections, thereby exhausting their batteries.

C. Energy Efficiency

Higher energy efficiency implies longer network lifespan and vise versa. The network lifespan is evaluated based on the following metrics, the round in which the first node dies (FND), the round in which half of nodes die (HND), and the round in which the last node dies (LND) [8].

Fig. 6 shows the number of alive nodes throughout the lifespan of LS-WSNs constructed based on ECN, ELDCN, M-BA, BA, DS, and SRA. It is clear that ECN, ELDCN and M-BA establish links in such a way that energy consumption is balanced and network lifespan is extended in comparison to the other models. The FND is 127th, 129th and 14th round for the ECN, ELDCN and M-BA, respectively; whereas, the first node dies in the first round for BA, DS, and SRA. The HND for the ELDCN and ECN are very close and are about 17 times of BA, 2 times of DS, and 10 times of SRA. Although the HNA for the M-BA is 210 and is larger than our proposed schemes, its LND is only 210, which is 30 rounds less than our proposed schemes. The great difference in results for the BA and M-BA suggests that the energy efficiency in LS-WSN can be significantly enhanced by limiting the communication range of nodes. This has contributed to the improvement of energy efficiency.

Fig. 7 shows the ratio of the node degree to their degree capacities in the first round, $k_n^0/\tilde{k}_n^0$. The histogram for the M-BA model shows that there are 40 nodes with degrees outweighed their degree capacities and the degrees of almost all of the rest nodes are one, which indicates that the energy dissipation is significantly unbalanced among all nodes in the M-BA model. This is due to the unlimited number of links that can be established to a given node. Since nodes prefer to establish communication links to the nodes with higher degree for the BA model, some nodes are overloaded with a large number of links to relay data and exhaust their batteries even in the first round. Due to the trade-off between energy and degree metrics in DS, the value of $k_n^0/\tilde{k}_n^0$ is balanced among the N nodes while there are still 5 nodes that die in the first round. The histogram for SRA shows

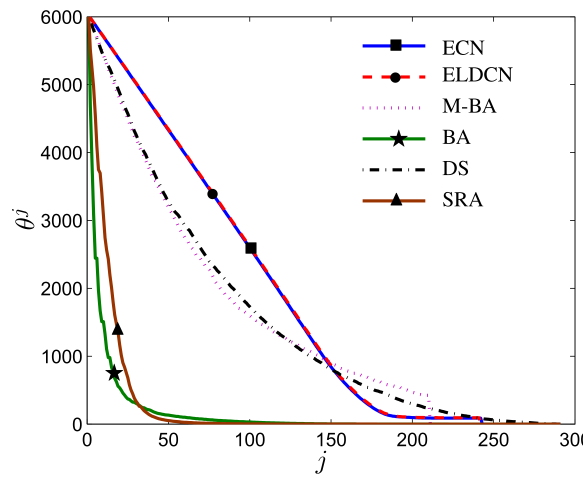


Fig. 8. Proportionality of energy dissipation in LS-WSNs measured by θ^j over the network lifespan.

that there are 130 nodes with degrees outweighed their degree capacities, and 268 nodes die in the first round. In contrast, energy dissipation in our proposed schemes is more balanced, none of the nodes is overloaded; thus, the lifespans of our generated LS-WSNs are longer. In addition, ELDCN avoids hub-nodes overload resulting from high potential connectivity, which further balances the node degree with the average transmission length, as shown in Fig. 5 and Fig. 7. Thus, ELDCN outperforms ECN.

We measure the proportionality of energy dissipation in LS-WSNs using the following index

$$\theta^j = \sum_{n=1}^{N^j} \frac{E_n^j}{E_n^0}, \tag{17}$$

where N^j is the number of alive nodes in the j-th round. For round j, the larger the ratio of a nodes' residual energy to its initial energy, the longer it lives. For the whole network, more alive nodes in each round implies more balanced energy dissipation. Thus, a larger θ^j indicates a higher energy efficiency and a longer network lifespan [35].

Fig. 8 shows θ^j over the network lifespan. The largest θ^j are shown for ELDCN, and ECN, followed by DS and M-BA. Results show far less values of θ^j for SRA and BA, which

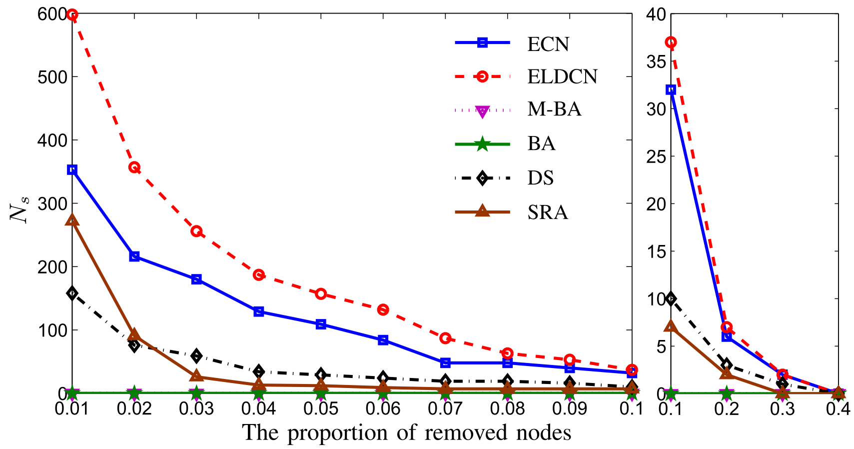


Fig. 9. Number of nodes connected the sink node vs. the proportion of removed nodes.

supports our earlier conclusion that the proposed ELDCN and ECN outperform all other candidate models with respect to energy efficiency.

D. Network Robustness

Nodes in LS-WSNs fail due to unbalanced energy consumption or external hostile forces. To evaluate the robustness of the generated LS-WSNs to hub-node failures, we simulate the hub-node failures in our generated LS-WSNs. The hub-node failures are simulated as follows. First, the nodes are sorted by their degrees. Then, hub-nodes with largest degrees are removed from the network one-by-one. Next, links to the removed nodes are removed from the network at the same time. This simulation approach maximizes the hub-node failures impact on the network connectivity. We use the number of nodes connected to the sink node, N_s, as a measure of robustness to hub-node failures.

Fig. 9 shows the average number of nodes connected to the sink node over 20 trials for all considered schemes. In ECN, ELDCN and DS, the numbers of nodes connected to the sink node fall to 0 when 40% of the nodes are removed from the network. For SRA, no node remains connected to the sink node after removing only 30% of the nodes. As the proportion of removed nodes increases from 1% to 50%, ELDCN outperforms all other models and achieves the highest robustness. ELDCN is followed by ECN, DS, and SRA. Limiting the number of links to each node in ELDCN and ECN, not only balances the energy consumption in each node, but also improves robustness to hub-node failures. It should be noted that due to the unlimited node degree in the M-BA and the BA models, some nodes are overloaded, as shown in Fig. 7. Therefore, LS-WSNs constructed based on the M-BA and the BA models are prone to node failures and become unavailable when as low as 1% of nodes are removed from the network.

VI. Conclusions

In this paper, we have proposed two self-organizing schemes, ECN and ELDCN, for generating clustering-based, degree-constrained and scale-free-inspired LS-WSNs. Our schemes exploit more comprehensive intrinsic characteristics of LS-WSNs, and the topology information regarding common neighbors, and the potential degree of common neighbors, which significantly balance the node degree and the energy consumption over LS-WSNs. Both of our theoretical analytical and simulation results have demonstrated that, our proposed schemes have achieved high energy efficiency and strong robustness for LS-WSNs compared to M-BA, BA, DS, and SRA.

Appendix A
Proof of Theorem 1

Proof: Denote the degree of the o^{th} node in iteration t by $k_{o,t}$. To conveniently analyze the problems, without losing rationality, label the new-incoming node at iteration $t = n$ with n. We focus on the node o when analyzing $P(k)$, i.e., the expression of $P(k)$ is derived based on $P(k_{o,t})$. Note that, the superscript j is dropped here because we analyze the degree distribution of LS-WSN in the first round.

Algorithm 1 shows that a new-incoming CH establishes links only to the CHs in $\mathcal{N}\backslash\mathcal{V}$, while a new-incoming SN can connect to any node in $\mathcal{N}\backslash\mathcal{V}$, which means the probability that a CH is connected by a new-incoming node n is larger than that of a SN within $\mathcal{N}\backslash\mathcal{V}$. Thus, for iteration t in the growth step of the ECN scheme, the growth rates of $k_{o,t}$ for SN and CH are different, in the following, we treat them separately.

A. The Case of CH o in $\mathcal{N}\setminus\mathcal{V}$

Considering the communication range of CH o and the ECN selection probability Π_{o-ECN}, equation (8), the growth rate of $k_{o,t}$ can be evaluated by

$$\frac{dk_{o,t}}{dt} = \frac{|\Gamma(n) \cap \Gamma(o)|\tilde{\tilde{E}}_o}{\sum_{w \in O_n} |\Gamma(n) \cap \Gamma(o)|\tilde{\tilde{E}}_w}. \tag{18}$$

According to the definition of O_n, given in Subsection II-A, we get

$$O_n = \begin{cases} \pi D R_s^2 \frac{N-t}{DU^2} = \frac{\pi R_s^2 (N-t)}{U^2}, & n \in \mathcal{S}, \\ \pi D R_c^2 \frac{N-t}{DU^2} = \frac{\pi R_c^2 (N-t)}{U^2}, & n \in \mathcal{C}. \end{cases} \tag{19}$$

The previous analysis shows that the neighborhood overlap between the new-incoming node n and the CH o is different in terms of the type of the new-incoming node, and the value of $|\Gamma(n) \cap \Gamma(o)|$ can be described as

$$|\Gamma(n) \cap \Gamma(o)| = \begin{cases} \Omega, & n \in \mathcal{S}, R_s \le R_c < 2R_s, \\ \pi D R_s^2, & n \in \mathcal{S}, 2R_s \le R_c, \\ 2.1627 D R_c^2, & n \in \mathcal{C}, \end{cases} \tag{20}$$

where $\Omega = R_s^2 \arccos(1.25 - \frac{R_c^2}{R_s^2}) + R_c^2 \arccos(\frac{R_c}{R_s} - \frac{3R_s}{4R_c}) - 0.5R_s^2 \sin\left(\arccos(1.25 - \frac{R_c^2}{R_s^2})\right)$. Additionally, with out loss of generality, we replace $\tilde{E}_w$ with the average value of all $\tilde{E}_w$ $\forall w \in \mathcal{N} \backslash \mathcal{V}$ denoted by $\bar{E_w}$. Taking $R_s \le R_c < 2R_s$ as an example, since the probability that the new-incoming n is a SN is ρ, substituting equations (19) and (20) back into equation (18) gives

$$\begin{aligned} \frac{dk_{o,t}}{dt} &= \frac{\Omega \tilde{E}_o}{\frac{\pi R_s^2 (N-t)}{U^2} \Omega \bar{E_w}}(1-\rho) + \frac{2.1627 D R_c^2 \tilde{E}_o}{\frac{\pi R_c^2 (N-t)}{U^2} 2.1627 D R_c^2 \bar{E_w}} \rho \\ &= \frac{[R_c^2(1-p) + R_s^2 \rho] \tilde{E}_o U^2}{\pi R_s^2 R_c^2 (N-t) \bar{E_w}}. \end{aligned} \tag{21}$$

The value of $\bar{E_w}$ is given by

$$\bar{E_w} = \frac{\bar{E}(N-t) - \beta t}{(N-t)}. \tag{22}$$

Thus, combining with equation (6), the growth rate of $k_{o,t}$ for a CH can be rewritten as

$$\frac{dk_{o,t}}{dt} = \frac{[R_c^2(1-\rho) + R_s^2 \rho](\tilde{k}_o - k_{o,t})\beta U^2}{\pi R_s^2 R_c^2 [\bar{E}(N-t) - \beta t]}. \tag{23}$$

According to equation (5) and the energy distribution of these N nodes given in Subsection II-A, $\bar{E}$ evaluates to

$$\bar{E} = \beta \widehat{k} \frac{(1-\rho)(\underset{\smile}{E}_s + \widehat{E}_s) + \rho(\underset{\smile}{E}_c + \widehat{E}_c)}{2\widehat{E}}. \tag{24}$$

B. The Case of SN o in $\mathcal{N} \backslash \mathcal{V}$

The growth rate of $k_{o,t}$ for the o^{th} SN in $\mathcal{N} \setminus \mathcal{V}$ is same as equation (18). Unlike CHs, the $k_{o,t}$ of SN o is affected only when the new-incoming node is a SN. Thus, the O_n for SN o in $\mathcal{N} \backslash \mathcal{V}$ can be written as

$$O_n = \pi D R_s^2 \frac{N-t}{DU^2} = \frac{\pi R_s^2 (N-t)}{U^2}. \tag{25}$$

Since the new-incoming CH does not establish links to SNs within $\mathcal{N} \setminus \mathcal{V}$, the neighborhood overlap between the new-incoming node n and the SN o, $|\Gamma(n) \cap \Gamma(o)|$, can be written as

$$|\Gamma(n) \cap \Gamma(o)| = \begin{cases} 2.1627 D R_s^2, & n \in \mathcal{S}, \\ 0, & n \in \mathcal{C}. \end{cases} \tag{26}$$

Substituting equations (25) and (26) back into equation (18), and replacing $\tilde{E}_w$ with $\bar{E_w}$ given in equation (22), we get

$$\begin{aligned} \frac{dk_{o,t}}{dt} &= \frac{2.1627 D R_s^2 \tilde{E}_o}{\frac{\pi R_s^2 (N-t)}{U^2} 2.1627 D R_s^2 \bar{E_w}}(1-\rho) \\ &= \frac{U^2 \tilde{E}_o}{\pi R_s^2 (N-t) \bar{E_w}}(1-\rho) \\ &= \frac{U^2(\tilde{k}_o - k_{o,t})\beta}{\pi R_s^2 [\bar{E}(N-t) - \beta t]}(1-\rho). \end{aligned} \tag{27}$$

For a given node o, $o \in \mathcal{N}$ in LS-WSN, the probability that $o \in \mathcal{C}$ is ρ and $o \in \mathcal{S}$ is $(1-\rho)$. Thus, based on the above analysis, equations (23), and (27), the growth rate of $k_{o,t}, o \in \mathcal{N}$ can be written as

$$\begin{aligned} \frac{dk_{o,t}}{dt} &= \frac{[R_c^2(1-\rho) + R_s^2 \rho] U^2 \beta (\tilde{k}_o - k_{o,t})}{\pi R_s^2 R_c^2 [\bar{E}(N-t) - \beta t]} \rho \\ &\quad + \frac{U^2 \beta (\tilde{k}_o - k_{o,t})}{\pi R_s^2 [\bar{E}(N-t) - \beta t]}(1-\rho)^2 \\ &= \frac{U^2 \beta (R_c^2 - \rho R_c^2 + p^2 R_s^2)}{\pi R_s^2 R_c^2 [\bar{E}(N-t) - \beta t]}(\tilde{k}_o - k_{o,t}). \end{aligned} \tag{28}$$

For simplicity, we define function $P(t)$ as follows

$$\begin{aligned} P(t) &= \frac{U^2 \beta (R_c^2 - \rho R_c^2 + \rho^2 R_s^2)}{\pi R_s^2 R_c^2 [\bar{E}(N-t) - \beta t]} \\ &= \frac{U^2 \beta (R_c^2 - \rho R_c^2 + \rho^2 R_s^2)}{\pi R_s^2 R_c^2 \bar{E} N - \pi R_s^2 R_c^2 (\bar{E} + \beta) t} \\ &= \frac{a_1}{a_2 - a_3 t}, \end{aligned} \tag{29}$$

where a_1, a_2 and a_3 are given by

$$\begin{aligned} a_1 &= U^2 \beta (R_c^2 - \rho R_c^2 + \rho^2 R_s^2), \\ a_2 &= \pi R_s^2 R_c^2 \bar{E} N, \\ a_3 &= \pi R_s^2 R_c^2 (\bar{E} + \beta). \end{aligned} \tag{30}$$

Therefore, equation (28) can be simplified to

$$\frac{dk_{o,t}}{dt} = P(t)(\tilde{k}_o - k_{o,t}). \tag{31}$$

Through solving the first order linear non-homogeneous differential equation given in (31), we solve for $k_{o,t}$ as follows

$$k_{o,t} = \xi e^{-\int P(t)dt} + e^{-\int P(t)dt} \int (P(t)\tilde{k}_o) e^{\int P(t)dt} dt, \tag{32}$$

where ξ is a constant whose value can be determined by the initial condition of equation (31). Based on equation (29), we get

$$\begin{cases} e^{-\int P(t)dt} = e^{[a_1 \ln(a_3 t - a_2)]/a_3} = (a_3 t - a_2)^{\frac{a_1}{a_3}}, \\ \int (P(t)\tilde{k}_o) e^{\int P(t)dt} dt = \tilde{k}_o / (a_3 t - a_2)^{\frac{a_1}{a_3}}. \end{cases} \tag{33}$$

Substituting equation (33) into (32), and according to the initial condition $k_{o,o} = 1$, ξ evaluates to

$$\xi = (1 - \tilde{k}_o)(a_3 o - a_2)^{-\frac{a_1}{a_3}}. \tag{34}$$

Thus, we arrive at the solution of equation (31),

$$k_{o,t} = (1 - \tilde{k}_o)(a_3 o - a_2)^{-\frac{a_1}{a_3}} (a_3 t - a_2)^{\frac{a_1}{a_3}} + \tilde{k}_o. \tag{35}$$

In order to get the expression for $P(k)$, the cumulative probability $P(k_{o,t} < k)$ is calculated first. Suppose each new-incoming node is added to the graph in equal time intervals, then the probability density of o can be defined as $p(o) = \frac{1}{(N-t)}$, and we can write $P(k_{o,t} < k)$ as follows

$$\begin{aligned} &P(k_{o,t} < k) \\ &= P\left(o > \frac{1}{a_3}\left(\frac{\tilde{k}_o - k}{\tilde{k}_o - 1}\right)^{-\frac{a_3}{a_1}}(a_3 t - a_2) + \frac{a_2}{a_3}\right) \\ &= 1 - P\left(o \le \frac{1}{a_3}\left(\frac{\tilde{k}_o - k}{\tilde{k}_o - 1}\right)^{-\frac{a_3}{a_1}}(a_3 t - a_2) + \frac{a_2}{a_3}\right) \\ &= 1 - \frac{1}{(N-t)}\left[\frac{1}{a_3}\left(\frac{\tilde{k}_o - k}{\tilde{k}_o - 1}\right)^{-\frac{a_3}{a_1}}(a_3 t - a_2) + \frac{a_2}{a_3}\right]. \end{aligned} \tag{36}$$

Hence,

$$\begin{aligned} P(k) &= \frac{\partial P(k_{o,t} < k)}{\partial k} \\ &= \frac{a_2 - a_3 t}{a_1(\tilde{k}_o - 1)(N-t)}\left(\frac{\tilde{k}_o - k}{\tilde{k}_o - 1}\right)^{-\left(\frac{a_3}{a_1}+1\right)}. \end{aligned} \tag{37}$$

The power-law exponent γ of equation (37) is given by

$$\gamma = \frac{a_3}{a_1} + 1 = \frac{\pi R_s^2 R_c^2(\bar{E} + \beta)}{U^2\beta(R_c^2 - \rho R_c^2 + \rho^2 R_s^2)} + 1. \tag{38}$$

■

APPENDIX B
PROOF OF THEOREM 2

Proof: It is obvious that $\bar{H}$ peaks at the extreme case $k_o = 2, \forall o \in \mathcal{N}$ in the generated LS-WSN, in which all these N nodes connect to each other one-by-one, until a link is established to the sink node. In this case, there is

$$\bar{H} = \frac{\sum_{i=1}^{N} i}{N} = \frac{1+N}{2}. \tag{39}$$

At another extreme case that $k_o = \tilde{k}_o, \forall o \in \mathcal{N}$, i.e., each node reaches its degree capacity, $\bar{H}$ reaches the minimum. Then N can be a function of $\bar{K}$, ω, and M, shown below

$$M\sum_{i=0}^{\omega-2}\bar{K}^i \le N \le M\sum_{i=0}^{\omega-1}\bar{K}^i. \tag{40}$$

The proof for equation (40) is presented in Appendix C. Based on equation (40), we can write

$$\omega = \left\lceil \log_{\bar{K}}\left(1 - \frac{N}{M}(1-\bar{K})\right)\right\rceil. \tag{41}$$

Therefore, $\bar{H}$ can be written as

$$\begin{aligned} \bar{H} &= \frac{\sum_{i=0}^{\omega-1}[\bar{K}^i(i+1)] + [N - M\sum_{i=0}^{\omega-1}\bar{K}^i](\omega-1)}{N} \\ &= \left(1 - \frac{M(1-\bar{K}^\omega)}{N(1-\bar{K})}\right)(\omega+1) \\ &\quad + \frac{M(1+\omega\bar{K}^{\omega+1} - (1+\omega)\bar{K}^\omega)}{N(1-\bar{K})^2}. \end{aligned} \tag{42}$$

Thus, theorem 2 is proved. ■

APPENDIX C
PROOF OF EQUATION (40)

Proof: Denote $N_{\omega,\bar{K}}$ as the total number of nodes in the network with parameters ω and $\bar{K}$. The variables ω and $\bar{K}$ were introduced in Subsection IV-B. Given the initialization step of the proposed schemes, $\omega \ge 1$. Then proving equation (40) translates into proving the following equation

$$N_{\omega,\bar{K}} = M\sum_{i=0}^{\omega-1}\bar{K}^i. \tag{43}$$

Next, we use a recursive method to prove equation (43).

1) The case of $\omega = 1$ and $\bar{K} = 0$: only M CHs are in the network and there is

$$N_{1,0} = M. \tag{44}$$

Given $\omega = 1$ and $\bar{K} = 0$, then equation (43) reduces to equation (44). Considering the value of $\bar{K}$ is related to the energy distribution, the value of $\widehat{k}$ and ω can be uniquely determined under the fixed $N_{\omega,\bar{K}}$ and $\bar{K}$. Thus, in the following, we assume that ω is an integer which is greater than 1.

2) The case of $\omega > 1$ and $\bar{K} = 1$: The network consists of M clusters, and $(\omega - 1)$ SNs connect to each other one-by-one in each cluster; thus, there are ω nodes, $\omega - 1$ SNs and 1 CH, in each cluster. Therefore,

$$N_{\omega,1} = M\omega = M\sum_{i=0}^{\omega-1}1^i. \tag{45}$$

Hence, equation (45) is equivalent to equation (43) for $\omega > 1$ and $\bar{K} = 1$.

3) Suppose $\omega > 1$ and $\bar{K} = j$, the equation (43) can be re-written as

$$N_{\omega,j} = M\sum_{i=0}^{\omega-1}j^i. \tag{46}$$

When $\omega > 1$ and $\bar{K} = j+1$, comparing with the topology of the network with $\omega > 1$ and $\bar{K} = j$, and based on $N_{\omega,j}$, we get

$$\begin{aligned} N_{\omega,j+1} &= N_{\omega,j} + M\sum_{x=1}^{\omega-1}\sum_{y=1}^{x-1}j^{(x-1)-y}(j+1)^y \\ &= M\sum_{i=0}^{\omega-1}(j+1)^i. \end{aligned} \tag{47}$$

Therefore, (43) is proved for all values of $\omega \ge 1$ and $\bar{K}$. ■

REFERENCES

[1] H. Peng *et al.*, "Energy-efficient and fault-tolerant evolution models for large-scale wireless sensor networks: A complex networks-based approach," in *Proc. IEEE Globecom*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[2] X. Liu, "Atypical hierarchical routing protocols for wireless sensor networks: A review," *IEEE Sensors J.*, vol. 15, no. 10, pp. 5372–5383, Oct. 2015.

[3] Y. Shouyi, L. Leibo, Z. Renyan, S. Zhongfu, and W. Shaojun, "Design of wireless multi-media sensor network for precision agriculture," *China Commun.*, vol. 10, no. 2, pp. 71–88, Feb. 2013.

[4] J. Chen, Q. Yu, B. Chai, Y. Sun, Y. Fan, and X. Shen, "Dynamic channel assignment for wireless sensor networks: A regret matching based approach," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 1, pp. 95–106, Jan. 2015.

[5] F. Lin, C. Chen, N. Zhang, X. Guan, and X. Shen, "Autonomous channel switching: Towards efficient spectrum sharing for industrial wireless sensor networks," *IEEE Internet Things J.*, vol. 3, no. 2, pp. 231–243, Apr. 2015.

[6] J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. Shen, "Dynamic channel access to improve energy efficiency in cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3143–3156, May 2016.

[7] Z. Zheng, A. Liu, L. X. Cai, Z. Chen, and X. Shen, "Energy and memory efficient clone detection in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 5, pp. 1130–1143, May 2016.

[8] J. Ren, Y. Zhang, K. Zhang, A. Liu, J. Chen, and X. S. Shen, "Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 788–800, Apr. 2016.

[9] S. Halder and S. DasBit, "Design of a probability density function targeting energy-efficient node deployment in wireless sensor networks," *IEEE Trans. Netw. Service Manage.*, vol. 11, no. 2, pp. 204–219, Jun. 2014.

[10] C. Zhu, V. C. M. Leung, L. T. Yang, and L. Shu, "Collaborative location-based sleep scheduling for wireless sensor networks integrated with mobile cloud computing," *IEEE Trans. Comput.*, vol. 64, no. 7, pp. 1844–1856, Jul. 2015.

[11] J. Luo, J. Hu, D. Wu, and R. Li, "Opportunistic routing algorithm for relay node selection in wireless sensor networks," *IEEE Trans Ind. Informat.*, vol. 11, no. 1, pp. 112–121, Feb. 2015.

[12] H. Lin, L. Wang, and R. Kong, "Energy efficient clustering protocol for large-scale sensor networks," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7150–7160, Dec. 2015.

[13] K. Yuan, Q. Ling, and Z. Tian, "Communication-efficient decentralized event monitoring in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 8, pp. 2198–2207, Aug. 2014.

[14] F. Farouk, R. Rizk, and F. W. Zaki, "Multi-level stable and energy-efficient clustering protocol in heterogeneous wireless sensor networks," *IET Wireless Sensor Syst.*, vol. 4, no. 4, pp. 159–169, 2014.

[15] M. Dong, K. Ota, A. Liu, and M. Guo, "Joint optimization of lifetime and transport delay under reliability constraint wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 1, pp. 225–236, Jan. 2016.

[16] W. M. Tam, F. C. M. Lau, and C. K. Tse, "Complex-network modeling of a call network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 2, pp. 416–429, Feb. 2009.

[17] J. Wu, C. Tse, F. Lau, and I. Ho, "Analysis of communication network performance from a complex network perspective," *IEEE Trans. Circuits Syst.*, vol. 60, no. 12, pp. 3303–3316, Dec. 2013.

[18] V. Karyotis and S. Papavassiliou, "Macroscopic malware propagation dynamics for complex networks with churn," *IEEE Commun. Lett.*, vol. 19, no. 4, pp. 577–580, Apr. 2015.

[19] F. Senel, M. F. Younis, and K. Akkaya, "Bio-inspired relay node placement heuristics for repairing damaged wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1835–1848, May 2011.

[20] C. Alippi, R. Camplani, C. Galperti, and M. Roveri, "A robust, adaptive, solar-powered WSN framework for aquatic environmental monitoring," *IEEE Sensors J.*, vol. 11, no. 1, pp. 45–55, Jan. 2011.

[21] L. Liu, Y. Liu, and N. Zhang, "A complex network approach to topology control problem in underwater acoustic sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3046–3055, Dec. 2014.

[22] C. Asensio-Marco and B. Beferull-Lozano, "Energy efficient consensus over complex networks," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 2, pp. 292–303, Mar. 2015.

[23] H. Zhang, E. Fata, and S. Sundaram, "A notion of robustness in complex networks," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 310–320, Sep. 2015.

[24] Y. Wang *et al.*, "NLL: A complex network model with compensation for enhanced connectivity," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1856–1859, Sep. 2013.

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

[26] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, p. 47, Jan. 2002.

[27] S. Li, L. Li, and Y. Yang, "A local-world heterogeneous model of wireless sensor networks with node and link diversity," *Phys. A, Statist. Mech. Appl.*, vol. 390, no. 6, pp. 1182–1191, 2011.

[28] Y. Wang *et al.*, "Energy-aware complex network model with compensation," in *Proc. IEEE 9th Int. Conf. WiMob*, Lyon, France, Oct. 2013, pp. 472–476.

[29] L. Wen, R. G. Dromey, and D. Kirk, "Software engineering and scale-free networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern*, vol. 39, no. 4, pp. 845–854, Aug. 2009.

[30] G. Wu, S. Wang, B. Wang, Y. Dong, and S. Yan, "A novel range-free localization based on regulated neighborhood distance for wireless ad hoc and sensor networks," *Comput. Netw.*, vol. 56, no. 16, pp. 3581–3593, 2012.

[31] C. Lin, D. Jiang, and A. Zhang, "Prediction of protein function using common-neighbors in protein-protein interaction networks," in *Proc. 6th IEEE Symp. BIBE*, Arlington, VA, USA, Oct. 2006, pp. 251–260.

[32] D. He, Z. Liu, and B. Wang, *Complex Systems and Complex Networks*. Beijing, China: Higher Education Press, 2009.

[33] C. Asensio-Marco, D. Alonso-Roman, and B. Beferull-Lozano, "Achieving energy-efficient distributed consensus in wireless scale free networks," in *Proc. IEEE ICC*, Sydney, NSW, Australia, Jun. 2014, pp. 2436–2442.

[34] E. Bulut and B. K. Szymanski, "Constructing limited scale-free topologies over peer-to-peer networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 4, pp. 919–928, Apr. 2014.

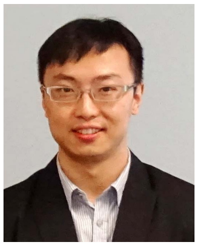
[35] R. Du, L. Gkatzikis, C. Fischione, and M. Xiao, "Energy efficient sensor activation for water distribution networks based on compressive sensing," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2997–3010, Dec. 2015.

Haixia Peng (S'15) received the B.S. and M.Sc. degrees from Northeastern University, Shenyang, China, in 2011 and 2013, respectively, where she is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering. She was a Visiting Scholar with the Broadband Communications Research Group, University of Waterloo, Waterloo, ON, Canada, for two years. Her current research focuses on wireless sensor networks and communication performances analysis of vehicular ad hoc networks in platoon. She served as a TPC member with the IEEE VTC-fall 2016, the IEEE Globecom 2016 conferences, and the IEEE ICC 2017 conference. She also served as a Volunteer in the IEEE VTC-fall 2016.

Shuaizong Si received the M.Sc. degree from Northeastern University, Shenyang, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Northeastern University. His current research interests include wireless communication and vehicular ad hoc networks.

Mohamad Khattar Awad (S'02–M'09) received the B.A.Sc. degree from the University of Windsor, Windsor, ON, Canada, in 2004, and the M.Sc. and Ph.D. degrees from the University of Waterloo, Waterloo, ON, Canada, in 2006 and 2009, respectively, all in electrical and computer engineering. From 2004 to 2009, he was a Research Assistant with the Broadband Communications Research Group, University of Waterloo. From 2009 to 2012, he was an Assistant Professor of Electrical and Computer Engineering with the American University of Kuwait. Since 2012, he has been with Kuwait University as an Assistant Professor of Computer Engineering.

His research interest includes wireless and wired communications, software-defined networks resource allocation, wireless networks resource allocation, and acoustic vector-sensor signal processing. He received the Ontario Research and Development Challenge Fund Bell Scholarship in 2008 and 2009, the University of Waterloo Graduate Scholarship in 2009, and a Fellowship Award from the Dartmouth College, Hanover, NH, USA, in 2011. In 2015, he received the Kuwait University Teaching Excellence Award.



Ning Zhang (S'12–M'16) received the B.Sc. degree from Beijing Jiaotong University in 2007, the M.Sc. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2010, and the Ph.D. degree from the University of Waterloo in 2015. From 2015 to 2016, he was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo. He is currently a Post-Doctoral Research Fellow with the University of Toronto. His current research interests include next-generation wireless networks, software-defined networking, green communication, and physical layer security.

Hai Zhao received the B.S. degree in electrical engineering from Dalian Maritime University, China, in 1982, and the M.S. and Ph.D. degrees in computer science from Northeastern University, China, in 1987 and 1995, respectively. He is currently a Professor with the Department of Computer Science and Technology, Northeastern University. He is also the Director of the Liaoning Provincial Key Laboratory of Embedded Technology. His current research interests include embedded Internet technology, wireless sensor network, vehicular ad hoc network, body area network, pervasive computing, operating system, data and information fusion and computer simulation, and virtual reality. He received allowance of the State Council due to his special contributions to the development of education.

Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree from Dalian Maritime University, China, in 1982, and the M.Sc. and Ph.D. degrees from Rutgers University, Piscataway, NJ, USA, in 1987 and 1990, respectively, all in electrical engineering. He is currently a Professor and a University Research Chair with the Department of Electrical and Computer Engineering, University of Waterloo, Canada.

His research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is a Registered Professional Engineer of Ontario, Canada. He is also a Fellow of the Engineering Institute of Canada, the Canadian Academy of Engineering, and the Royal Society of Canada. He is also a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He is an elected member of the IEEE ComSoc Board of Governors. He received the Excellent Graduate Supervision Award in 2006, the Outstanding Performance Award in 2004, 2007, 2010, and 2014 from the University of Waterloo, the Premiers Research Excellence Award in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. He was the Chair of Distinguished Lecturers Selection Committee. He served as the Technical Program Committee Chair/Co-Chair of the IEEE Globecom'16, the Infocom'14, the IEEE VTC'10 Fall, and the Globecom'07, the Symposia Chair for the IEEE ICC' 10, the Tutorial Chair of the IEEE VTC'11 Spring and the IEEE ICC' 08, the General Co-Chair of the ACM Mobihoc'15, the Chinacom'07, and the QShine06, the Chair of the IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief of the *IEEE Network, Peer-to-Peer Networking and Application*, and the *IET Communications*, a Founding Area Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Computer Networks*, and the *ACM/Wireless Networks* and the Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS, *IEEE Communications Magazine*, and *ACM Mobile Networks and Applications*. He is also the Associate Chair for Graduate Studies.