Lifetime Analysis of Random Event-Driven Clustered Wireless Sensor Networks

Moslem Noori, *Student Member*, *IEEE*, and Masoud Ardakani, *Senior Member*, *IEEE*

Abstract—Considering event-driven clustered wireless sensor networks, a probabilistic approach for analyzing the network lifetime is presented when events occur randomly over the network field. To this end, we first model the packet transmission rate of the sensors, using the theory of coverage processes and Voronoi tessellation. Then, the probability of achieving a given lifetime by individual sensors is found. This probability is then used to study the cluster lifetime. In fact, we find an accurate approximation for the probability of achieving a desired lifetime by a cluster. Our proposed analysis includes the effect of packet generation model, random deployment of sensors, dynamic cluster head assignment, data compression, and energy consumption model at the sensors. The analysis is presented for event-driven networks, but it comprises time-driven networks as a special case. Computer simulations are used to verify the results of our analysis.

Index Terms—Wireless sensor networks, cluster, lifetime, event-driven, random deployment.

1 Introduction

A wireless sensor network (WSN) consists of a set of sensing devices which collaborate to collect data from an area. The limited energy available at the sensors makes the network lifetime one of the most critical issues in the design of WSNs. An analytical lifetime analysis can significantly help the network design step. For example, if the analysis takes into account the effects of different network parameters (e.g., initial energy of sensors, network density, sensors transmission range, etc.), these parameters can in turn be adjusted to assure a desired lifetime by the network.

The main focus of this work is on the lifetime analysis of randomly deployed clustered networks. Clustered WSNs are suggested for extending the network scalability and ease of data processing, [1], [2]. In clustered networks, usually a representing node, called cluster head (CH), is assigned to each cluster. CH is in charge of collecting data from other nodes of the cluster, performing data aggregation and compression, and relaying it to the sink. Depending on the network application, various clustering protocols are proposed and studied in the literature, e.g., [3], [4].

In this paper, the lifetime of a randomly deployed clustered network is modeled using a probabilistic approach. The probabilistic approach is motivated by the stochastic nature of the network lifetime which is mainly due to the randomness of the sensors deployment. For event-driven networks, the random traffic pattern of the sensors can further intensify the stochastic behavior of the network lifetime. Consequently, the expected lifetime cannot completely express the behavior of the network.

We consider a rather general network setup and discuss how the analysis can be applied to other variations of the considered setup. As a result, our analysis can effectively handle a wide range of network configurations. For example, we study adjustable transmission power [5] for sensors, but we show that our analysis applies as well to multilevel or fixed transmission power.

Similar to many clustering protocols, we assume that data transmission within each cluster is single hop (mainly because of the short transmission distance). For data transmission from CH to the sink both single-hop and multihop transmission schemes are considered in the literature (see [3] and [4] as two examples). We study both cases.

To summarize, the following points distinguish our work from existing work on WSN lifetime analysis.

1. Unlike most existing results, lifetime of the sensors is not assumed to be known. Instead, we start with a lifetime study at the sensor level. In other words, the lifetime of an individual sensor is viewed as a random variable whose distribution is studied using central limit theorem (CLT).
2. Our approach is probabilistic. More specifically, we find a quite accurate approximation for the complementary cumulative distribution function (ccdf) of cluster lifetime. Using the clusters lifetime, network lifetime can be derived thereafter.
3. We study event-driven networks. Time-driven networks—that have been the focus of most existing results—can be treated as a special case in our analysis.
4. For event-driven networks, a novel packet transmission rate analysis is provided which relates the network coverage to its lifetime.
5. Finally, we show that clustering is not always advantageous in terms of energy consumption.

Specifically, we find a condition on the data compression ratio of CH to make clustering advantageous.

- *The authors are with the Department of Electrical and Computer Engineering, University of Alberta, ECERF Building, 9107 116 Street, Edmonton, AB, Canada T6G 2V4.*
E-mail: {moslem, ardakani}@ece.ualberta.ca.

Manuscript received 20 Oct. 2009; revised 13 Oct. 2010; accepted 19 Oct. 2010; published online 17 Dec. 2010.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2009-10-0442.
Digital Object Identifier no. 10.1109/TMC.2010.254.

The rest of this paper is organized as follows: Section 2 reviews the previous work on the lifetime of clustered networks. The system model is introduced in Section 3. Section 4 presents our method for determining sensor packet generation model. The proposed lifetime analysis for the single-hop case is discussed in Section 5. Section 6 discusses the lifetime analysis for the multihop networks. Some remarks are provided in Section 7. Simulation results are reported in Section 8. Finally, Section 9 concludes the paper, and all proofs are presented in the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254.

2 Related Work

Owing to the significance of the lifetime in WSNs, there has been considerable research on the lifetime analysis of WSNs under various setups. For example, see [6], [7], [8], [9] and the references therein.

In [6], the lifetime of a clustered network, where the sensors are spread over a hypercube, is studied. The lifetime is defined based on coverage and connectivity where it is assumed that the lifetime of a single sensor, T_s, is known. Authors provide an asymptotic analysis for the network lifetime when the number of sensors tends to infinity. The analysis results in a lower bound on the lifetime as a function of T_s.

Zhang and Hou [7] focus on the lifetime analysis for WSNs where the lifetime is defined based on maintaining k-coverage in the network. First, the necessary and sufficient condition on the network density is derived guaranteeing the complete k-coverage in the network when the network size approaches infinity. In addition, it is proved that by having this density over the network, the lifetime of the network is almost surely upper bounded by kT_s when the network area approaches infinity. Here, T_s is the lifetime of a single sensor.

Estimating the lifetime of a heterogeneous WSN is studied in [8]. It is assumed that a group of more powerful sensors are overlayed in the area which act as CHs. All communications initiated from sensors to CHs or from CHs to the data sink are assumed to be done in a single-hop fashion. Assuming that the network is clock-driven (time-driven), the average consumed energy by each type of nodes is found and as a consequence the expected network lifetime is obtained. The authors also discuss efficient energy distribution over each type of nodes and find the optimal number of clusters via simulation.

Ferrari and Martal [9] study the lifetime of a clustered network where the lifetime of each sensor is assumed to obey a known distribution (i.e., exponential, uniform, lognormal, and Rayleigh). The network lifetime is defined based on the quality of service and the decision error at the network access point. Fixing the mean of the abovementioned distributions, the authors evaluate the network lifetime for different distributions. Also, an adaptive reclustering algorithm is proposed to extend the network lifetime.

A set of distributed scheduling algorithms is proposed in [10] in order to prolong the total network lifetime. The scheduling algorithm forms minimal cover sets to cover the local targets in the network. Since different cover sets result in different network performance, authors introduce lifetime dependency graph to determine the priority of each cover set. Using the proposed set of algorithms in [10], the network lifetime can be improved.

Work reported in [11] focuses on the energy consumption and end-to-end delay of QoS-based clustered networks where the computationally- and energy-intensive tasks, e.g., cluster formations and routing, are left for high-power base stations. First, the authors propose a dynamic QoS-based clustering protocol and then provide its performance analysis in terms of the average energy consumption and expected network lifetime. For the sake of the energy efficiency, it is assumed that the CH role rotates among different nodes within a cluster. Random networks are not considered in this study.

The mean and the variance of the consumed energy and network lifetime of a clustered network is studied in [12]. For this purpose, the authors use the kernel density estimation method to approximate the consumed energy by each sensor when no knowledge of the sensor distribution is given beforehand. However, to apply kernel density estimation, the actual location of a small number of deployed sensor is required in order to obtain the approximated distribution function of the deployment.

Zhuang et al. [13] investigate the energy consumption of a grid-based clustered sensor network where all clusters have a square shape and nodes are distributed uniformly within them. Then, by considering the random distances between the nodes and CHs and also between CHs, they find the average consumed energy by a sensor. They also consider the case where clusters may all be square while having different sizes.

Notice that in all of the above mentioned works, the network is time-driven and the average consumed energy or expected lifetime of the network is considered. In this work, we study the probability distribution of the lifetime. Moreover, instead of assuming a time-driven network, we consider event-driven networks, for which we first develop a packet generation model.

3 System Model

3.1 Clustering Model

Clustering is proposed for WSNs to decrease the energy consumption and ease the network management, [1], [2]. Usually, the nodes within a cluster send their data to the CH and then CH performs necessary data processing and aggregation before relaying it toward the data sink.

Here, we assume that N sensors are deployed randomly over the cluster. First, we focus on the case where all transmissions are in single-hop mode. Later on, the multihop mode for intercluster transmissions (CH to sink) will be discussed. For multihop intercluster transmissions, it is assumed that a CH forwards its data toward the sink through other CHs, i.e., other nodes are not involved in intercluster data communication.

It is also assumed that the network has a static clustering, meaning that the shape of the clusters are fixed during the network operation [14], [15]. While dynamic clustering allows for a more flexible network design, its overhead to form the clusters is considered a serious drawback [3], [16]. For many practical situations, therefore, static clustering is an attractive solution [13].

Another important issue in the network clustering is to rotate the CH role among the sensor nodes. Since the intercluster transmissions consumes a vast amount of energy, rotating the CH role results in a more balanced energy consumption among sensors. As a consequence, the early death of nodes is avoided. Here, we assume that the CH role rotates among the nodes within a cluster based on a predetermined periodic schedule which reduces/removes the need for the overhead packets and can result in energy efficiency in the network [16].

3.2 Energy Consumption Model

Sensors consume energy for sensing, receiving, transmitting, data processing, and also during the idle mode when no data sensing, processing or exchange happens. Data transmission can be accomplished with a fixed or adjustable power. Utilizing the power adjustment mechanism can benefit the energy conservation in the network [5] at the expense of a more complex hardware. The proposed analysis in this paper is mainly based on adjustable transmission power. Extending the analysis to multilevel transmission power (e.g., for Mica Mote sensors [17]) or fixed transmission power is straight forward and will be briefly discussed in Section 7.

Assuming adjustable transmission power proposed in [5], the transmission energy for one packet can be modeled as

$$e_t(d) = kd^{\alpha} + c, \tag{1}$$

where α stands for the path loss exponent and d refers to the distance between the sensor and destination. Also, k and c represent the loss coefficient and the overhead energy for one packet, respectively. Usually, α is considered to be two for small distances and four for large distances [3], [18]. Assuming that the intracluster distances are small and intercluster distances are large, similar to [3], we use $\alpha = 2$ for intracluster and $\alpha = 4$ for intercluster transmissions.

Consumed energy for receiving, e_r, is almost independent of the distance between the transmitter and receiver and depends on the electronic parts of the receiver. In addition, we assume that the idle mode consumed power, used to keep the radio part on to listen to the channel, is almost fixed.

3.3 Event Occurrence Model

There exist three main models for packet generation in WSNs, namely, event-driven, time-driven, and query-driven [19]. In the time-driven case, sensors send their data periodically to the sink. Event-driven networks are used when it is desired to inform the data sink whenever a random events occurs. In query-driven networks, sink sends a request of data gathering when needed.

Based on the network application, data characteristic, and the type of data inquiry, usually one of these models is utilized to characterize the sensors traffic generation. Here, the proposed analysis mostly concentrates on the event-driven networks. As discussed in Section 7, for the purpose of our analysis, event-driven networks comprise a rather general case. In fact, time-driven networks can be studied as a special case of the proposed analysis.

Depending on the nature of the events, events may or may not have temporal/spatial correlation. For instance, in a target tracking network, the events, i.e., the target location, show a high degree of correlation. On the other hand, impact detection networks, e.g., [20] and [21], typically deal with uncorrelated events.

Here, we assume an event-driven network where the events occur randomly and independently over the network field. More specifically, it is assumed that the distribution of the number of events over the field follows a spatial-temporal Poisson distribution. Poisson distribution has been widely used in the literature to model events whose time/place of occurrence are random and independent from each other, [22], [23], [24]. Also, in radiation detection applications, Poisson distribution is suitable to model the event occurrence [25].

Here, we employ a stationary spatial-temporal Poisson model where the number of events, $\mathcal{P}$, occurring during T time units in an area with size S has a probability distribution function (pdf) as follows:

$$P(\mathcal{P} = m) = \frac{e^{-\lambda ST}(\lambda ST)^m}{m!} \quad m = 0, 1, 2, \ldots. \tag{2}$$

In (2), λ shows the average number of events per time unit over a unit area.

We consider two situations for data reporting by sensors. In the first situation, we consider that all sensors, which detect an event, forward a packet containing the event information to CH. Evidently, in this case, CH may receive multiple packets regarding a single event. We refer to this situation as multiple event reporting. In the second scenario, we assume only the closest sensor to the event reports the event to CH. Notice that this sensor usually has the most reliable data about the event. This scenario is referred to as single event reporting.

We will use both models to determine the sensors packet generation model in Section 4.

3.4 Lifetime Definition

The network lifetime is the time interval over which the network can operate effectively. Clearly, a more specific definition of the lifetime is possible only based on the network application. Hence, different lifetime definitions are used in the literature. For instance, in [26], and in [6] the network connectivity is used to define the lifetime.

Another commonly used definition is based on the percentage of the dead nodes [8], [27], i.e., the network lifetime is the moment when the number of live nodes drops below $(1 - \beta)N$ where $0 \leq \beta < 1$ and N stands for the total number of nodes in the network.

Here, we adopt the last definition for the cluster lifetime. The lifetime definition based on the percentage of the dead nodes also includes the lifetime definition based on the first node death [28], [29]. Moreover, since the number of dead nodes can reflect the quality of the network coverage and/or connectivity [30], this lifetime definition can also be considered as an approximation of the lifetime based on the network coverage and/or connectivity.

3.5 Media Access Control (MAC) Protocol

It is common to use time division multiple access (TDMA) technique for MAC in clustered WSNs [16], [31]. As an example, low-energy adaptive clustering hierarchy (LEACH) uses TDMA to enable multiple access within a cluster [3]. In this case, CH usually coordinates the time scheduling among the nodes. To avoid intercluster interference, code division multiple access (CDMA) might also be used [3]. Here, we assume that TDMA is used in the clusters. Thus, collisions are avoided and transmissions can be assumed ideal [31].

4 Packet Generation Rate of the Sensors

In this section, we first characterize the packet transmission rate of sensors for both multiple event reporting and single event reporting. Our approach for modeling the packet generation rate of the sensors is based on an asymptotic analysis. That is, we first assume that the clusters are large. Later, in Section 8, we observe that the analysis provides accurate results for studying the network lifetime in nonasymptotic cases.

4.1 Multiple Event Reporting

In multiple event reporting, we assume a circular sensing region with radius r_s for each sensor. When an event happens in the area, all the sensors whose sensing region covers the event location report the event to CH. Then CH aggregates all received packets and forwards a single packet to the central data sink.

Since each sensor reports all events happening in its sensing region, the packet transmission rate of each sensor obeys the following model in time domain when we look at a T time slot

$$P(\mathcal{P}=m)=\frac{e^{-\lambda\pi r_s^2 T}(\lambda\pi r_s^2 T)^m}{m!}\quad m=0,1,2,\ldots. \tag{3}$$

Later, we work with mean and variance of $\mathcal{P}$ in our analysis which are

$$\mu_{\mathcal{P}}=\sigma^2_{\mathcal{P}}=\lambda\pi r_s^2 T. \tag{4}$$

Notice, however, that the number of transmitted packets by the sensors are not independent and there can exist a possible high correlation between neighboring sensors. Total number of packets received by CH, $\mathbf{P}'$, is the summation of all packets transmitted by all sensors within the cluster. However, CH does not necessarily forward all these packets to sink and just forwards packets associated with different events. For the number of packets forwarded by CH to sink, $\mathbf{P}$, in multiple event reporting case, we state the following theorem.

Theorem 1. *In an arbitrary cluster $\mathcal{C}$, with size $\|\mathcal{C}\|$, the number of transmitted packets by CH to sink during T time units, $\mathbf{P}$, has the following distribution*

$$P(\mathbf{P}=m)=\int\frac{e^{-\lambda xT}(\lambda xT)^m}{m!}f_c(x)\,dx\quad m=1,2,3,\ldots, \tag{5}$$

where

$$f_c(x)=\frac{1}{\sqrt{2\pi\sigma_c^2}}e^{-\frac{(x-\mu_c)^2}{\sigma_c^2}}. \tag{6}$$

Also, $\mu_c=\|\mathcal{C}\|(1-e^{-\pi\delta})$ and

$$\sigma_c^2=\pi r_s^2\|\mathcal{C}\|e^{-2\pi\delta}\left(8\int_0^1\left[e^{\delta B(x)}-1\right]x\,dx-\pi\delta\right), \tag{7}$$

where

$$\delta=\frac{Nr_s^2}{\|\mathcal{C}\|}, \tag{8}$$

and

$$B(x)=\begin{cases}\pi-2x\sqrt{1-x^2}-2\sin^{-1}(x), & 0\le x\le 1,\\ 0, & x>1.\end{cases} \tag{9}$$

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

As we mentioned, there are multiple packets associated with one event among the received packets by CH. Thus there is a redundancy and CH aggregates these multiple packets to one single packet in order to avoid transmitting redundant packets. To evaluate the efficiency of the data aggregation in multiple event reporting case, we define the redundancy as the ratio of the average number of received packets to the average number of transmitted packets by CH.

Corollary 1. *The redundancy rate of the packet sensing by sensors in a multiple event reporting scheme is*

$$r=\frac{N\pi r_s^2}{\|\mathcal{C}\|(1-e^{-\pi\delta})}. \tag{10}$$

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

4.2 Single Event Reporting

In single event reporting scheme, only the sensor that is closest to an event location reports the event to CH. That is, the cluster is tessellated by Voronoi cells generated by sensors located in it and each sensor reports only events within its associated Voronoi cell. It is noteworthy that identifying the Voronoi cells within a WSN can be accomplished in a distributed manner, e.g., [32].

Since the sensors are placed randomly over the area, a Voronoi cell, $\mathcal{V}$, has a random size. To determine the packet transmission rate of sensors, we need to consider the random size of the Voronoi cells which results in the following theorem.

Theorem 2. *In single event reporting scheme, the number of packets that a sensor transmits toward CH during T time units obeys a distribution as follows:*

$$P(\mathcal{P}=m)=\frac{343}{15}\left(\frac{N}{\|\mathcal{C}\|}\right)^{\frac{7}{2}}\frac{(\lambda T)^m}{m}\times\left(\frac{7N}{2\|\mathcal{C}\|}+\lambda T\right)^{-\frac{7}{2}-m}\Gamma\left(\frac{7}{2}+m\right)\quad m=1,2,\ldots, \tag{11}$$

where $\Gamma(\cdot)$ represents gamma function.

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

Since we mainly deal with the mean and variance of $\mathcal{P}$, we state the following corollary.

Corollary 2. *In single event reporting scheme, we have*

$$\mu_{\mathcal{P}} = \lambda T \mu_{\mathcal{V}} = \lambda T \frac{\|\mathcal{C}\|}{N}, \tag{12}$$

and

$$\sigma^2_{\mathcal{P}} = \lambda T \frac{\|\mathcal{C}\|}{N} + \frac{2}{7}\lambda^2 T^2 \left(\frac{\|\mathcal{C}\|}{N}\right)^2. \tag{13}$$

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

In the case of single event reporting scheme, since all events occurred within the cluster are reported by the sensors, the number of packets received by a CH from other nodes obeys a Poisson distribution with mean $\lambda T\|\mathcal{C}\|$ during T time units. This is equal to the number of packets transmitted by CH to sink, because in this case there is no redundancy in the collected packets by sensors and each event is sensed exactly once. Thus, in single reporting scheme we have

$$P(\mathbf{P} = m) = \frac{e^{-\lambda T\|\mathcal{C}\|}(\lambda T\|\mathcal{C}\|)^m}{m} \quad m = 0, 1, 2, \ldots. \tag{14}$$

5 Lifetime Analysis for Single-Hop Clustered Networks

In this section, we propose a probabilistic analysis to model the lifetime of a cluster in a clustered WSN that employs single-hop transmissions for both intracluster and intercluster communications.

In our analysis, we first consider energy consumption and lifetime at sensor level and then extend the lifetime analysis to the cluster level. Our analysis is according to the following main steps:

- Step 1: We first divide the energy consumption at the sensor level to two portions. The first portion represents the energy consumed for intracluster transmissions by a sensor, and the second portion is the energy consumed for intercluster transmission, i.e., when a sensor functions as a CH.
- Step 2: We find the mean and variance of energy consumed for intercluster and intracluster transmissions by a given sensor whose location is known. Notice that the derived mean and variance are conditioned on the sensor location.
- Step 3: Using CLT and also the mean and variance associated with consumed energy for intercluster and intracluster transmissions, we determine the conditional pdf of the total consumed energy by the given sensor. This also gives us the conditional probability of achieving the desired lifetime by that sensor.
- Step 4: To complete the lifetime analysis at the sensor level, by averaging over the location of the sensor on the cluster, we remove the condition on the sensor location. This gives the probability of achieving the desired lifetime by an arbitrary sensor.
- Step 5: Using the probability of achieving the desired lifetime by an arbitrary sensor, we find the probability of achieving the same lifetime by the cluster.

In the following, analytical details of our proposed method are discussed.

Here, it is assumed that N nodes, each with initial energy E_0, are spread independently and randomly inside the cluster. The CH role rotates among the sensor nodes each T time units, i.e., each sensor serves as a CH for T time units each round. Also, CH can compress l data bits to γl bits, where $0 < \gamma \leq 1$. We call the set of data packets transmitted by a CH during T time units *effective packets* since no event is represented by multiple packets. These effective packets are derived after removing repetitive packets in multiple event reporting scheme. For single event reporting scheme, effective packets are the same as the total number of packets transmitted by all sensors within a cluster. Here, it is assumed that the data compression happens on the effective data packets.

We start our analysis by finding the probability of achieving a desired lifetime T_ℓ by a single node. For this purpose, we study the total consumed energy by a node after T_ℓ, called E_t. It is clear that a sensor is still alive after T_ℓ if its consumed energy does not exceed E_0.

Assume that the data sink is located at $\mathbf{s}_s = (x_s, y_s)$. If the desired lifetime is long,

$$k = \left\lfloor \frac{T_\ell}{NT} \right\rfloor \approx \frac{T_\ell}{NT}, \tag{15}$$

is the number of rounds that the CH role is assigned to a sensor during the desired lifetime.

The consumed energy by a sensor can be expressed as $E_t = E_n + E_c$, where E_n reflects the amount of energy depleted by the sensor for intracluster transmissions and E_c refers to the energy used by the sensor when it serves as CH. Since the sensors are deployed randomly and all of them have the same traffic generation model, without loss of generality, E_t can be found for sensor 1 located at $\mathbf{s}_1 = (x_1, y_1)$.

The energy that sensor 1 consumes to send data to sensor i during T_ℓ is

$$e_i = \mathcal{P}_i(m_1\|\mathbf{s}_1 - \mathbf{s}_i\|^2 + c_1) \quad i = 2, 3, \ldots, N, \tag{16}$$

where $\mathcal{P}_i$ stands for the total number of packets sent to sensors i by sensor 1 when sensor i functions as CH. Also, $\|\cdot\|$ denotes the euclidean distance and $\mathbf{s}_i = (x_i, y_i)$ represents the position of sensor i. In addition, m_1 and c_1 are the parameters of transmission model (1) when $\alpha = 2$. Notice that $\mathcal{P}_i$ is a random variable which is identically distributed for all i, thus we use $\mathcal{P}_k$ to refer to all $\mathcal{P}_i$s where $i = 2, 3, \ldots, N$. It can be shown that $E[\mathcal{P}_k] = kE[\mathcal{P}]$ and $\mathrm{VAR}[\mathcal{P}_k] = k\mathrm{VAR}[\mathcal{P}]$. Considering all rounds that sensor 1 is not a CH, one can state $E_n = \sum_{j=2}^{N} e_j$.

Since sensor 1 serves as CH for k rounds, the value of E_c can be found as follows:

$$E_c = \gamma \mathbf{P}_t(m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2) + e_r\mathbf{P}_r, \tag{17}$$

where $\mathbf{P}_t$ shows the total number of packets that sensor 1 transmits to the sink during the time it serves as CH. Also, m_2 and c_2 stand for the value of transmission parameters in (1) when $\alpha = 4$. In addition, in (17), $\mathbf{P}_r$ is the total number of packets received by sensor 1 from all sensors when sensor 1 functions as CH. For multiple event reporting scheme, $\mathbf{P}_r$ is the number of packets before removing the packets associated with a similar event. Again, remember that $E[\mathbf{P}_t] = kE[\mathbf{P}]$ and $\mathrm{VAR}[\mathbf{P}_t] = k\mathrm{VAR}[\mathbf{P}]$. Also, we similarly have $E[\mathbf{P}_r] = kE[\mathbf{P}']$ and $\mathrm{VAR}[\mathbf{P}_r] = k\mathrm{VAR}[\mathbf{P}']$. Recall that for single event reporting scheme, $\mathbf{P}$ and $\mathbf{P}'$ are equal.

The entire consumed energy by a sensor, therefore, is

$$E_t = E_n + E_c = \sum_{j=2}^{N} e_j + E_c. \tag{18}$$

In a randomly deployed WSN, positions of sensors are independent of each other. In addition, since it is assumed that the sensors packet generation occurs independently over the time, e_is are independent. E_c and e_is are also independent. Thus, we apply generalized CLT to approximate the pdf of E_t in (18) with a Gaussian distribution when $\mathbf{s}_1$ is given. The accuracy of the approximated Gaussian distribution depends on the number of terms in (18) (i.e., the number of sensors). There usually exist enough nodes in a cluster to have an acceptable approximation. Thus, pdf of E_c can be accurately approximated by a Gaussian distribution with mean

$$\mu_{\mathbf{s}_1} = \sum_{j=2}^{N} \mu_{e_j} + \mu_{E_c}, \tag{19}$$

where μ_{e_j} and μ_{E_c} represent conditional mean of e_j and E_c given $\mathbf{s}_1$. Owing to the independence of the packets generated by sensor 1 within different time units and having i.i.d. model for the nodes placement within the cluster, e_js are i.i.d. random variables. We use e, with mean μ_e, to refer to the distribution of e_is. Hence,

$$\mu_{\mathbf{s}_1} = (N-1)\mu_e + \mu_{E_c}. \tag{20}$$

In a similar way, one can arrive at

$$\sigma^2_{\mathbf{s}_1} = (N-1)\sigma^2_e + \sigma^2_{E_c}, \tag{21}$$

where σ^2_e and $\sigma^2_{E_c}$ are conditional variances of e and E_c, respectively.

To proceed with the analysis, we find $\mu_{\mathbf{s}_1}$ and $\sigma^2_{\mathbf{s}_1}$.

Lemma 1. *Mean and variance of the consumed energy by sensor 1 are*

$$\mu_{\mathbf{s}_1} = (N-1)E[\mathcal{P}_e]E[z] + k\gamma(m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2)E[\mathbf{P}] + ke_rE[\mathbf{P}'] \tag{22}$$

and

$$\begin{aligned}\sigma^2_{\mathbf{s}_1} &= (N-1)\big(\mathrm{VAR}[\mathcal{P}_e]E[z^2] + E^2[\mathcal{P}_e]\mathrm{VAR}[z]\big) \\ &\quad + k\gamma^2(m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2)^2\mathrm{VAR}[\mathbf{P}] + ke_r^2\mathrm{VAR}[\mathbf{P}'],\end{aligned} \tag{23}$$

where

$$z = m_1\|\mathbf{s}_1 - \mathbf{s}\|^2 + c_1, \tag{24}$$

and $\mathbf{s} = (x, y)$ represents an arbitrary point in the area referring to the position of a sensor within the cluster.

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

Notice that for each of two event reporting schemes, we found the distribution of $\mathcal{P}$, $\mathbf{P}$, and $\mathbf{P}'$ in Section 4. Thus, for each of these two schemes, $\mu_{\mathbf{s}_1}$ and $\sigma^2_{\mathbf{s}_1}$ can simply be evaluated.

Theorem 3. *If nodes have initial energy E_0, cluster $\mathcal{C}$ achieves lifetime T_ℓ with the following probability*

$$P_\ell = I_{P_a}(\lfloor(1-\beta)N\rfloor + 1, N - \lfloor(1-\beta)N\rfloor), \tag{25}$$

where $I_z(\cdot,\cdot)$ is the regularized incomplete beta function [33] and P_a, the probability of achieving T_ℓ by a single sensor, is

$$P_a = \int_{\mathcal{C}} \left(1 - Q\left(\frac{E_0 - \mu_{\mathbf{s}_1}}{\sigma_{\mathbf{s}_1}}\right)\right) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1. \tag{26}$$

In (65), $Q(\cdot)$ denotes the complementary cumulative distribution function of a normal distribution and $f_{\mathbf{S}_1}(\mathbf{s}_1)$ refers to the distribution of the position of sensor 1 over the cluster.

Proof. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2010.254. □

In the above analysis, according to (15), it is assumed that all nodes in the cluster serve as CH the same number of times. However, with death of one node, other sensors have to serve as CH more frequently. While this hurts the assumption made in (15), it does not directly affect the analysis. To see this, assume that the number of alive nodes at an arbitrary moment during the network operation is $N' < N$. In this case, each sensor becomes CH once after each N' rounds, but it forwards just the traffic of alive nodes. In single event reporting case, the traffic load generated by alive nodes is scaled by $\frac{N'}{N}$. Also, for multiple event reporting case, the ratio of the traffic loads forwarded by CH, before and after the nodes death, is

$$\frac{1 - e^{\frac{-\pi N' r_s^2}{\|\mathcal{C}\|}}}{1 - e^{\frac{-\pi N r_s^2}{\|\mathcal{C}\|}}}. \tag{27}$$

Using Taylor series, the ratio of traffic loads in (27) is accurately approximated by $\frac{N'}{N}$ when N and N' are not small. On the other hand, the number of times that a node serves as CH is scaled by $\frac{N}{N'}$, keeping the average number of packets transmitted by CH to the sink almost unchanged.

Death of nodes in the cluster has an indirect effect on the accuracy of our analysis. Usually, the first nodes to die are those farther from the sink. Thus, the effective cluster region $\mathcal{C}$ would change as nodes die. This influences the equations involving averaging over $\mathcal{C}$, e.g., (65). However, in the limit of the distance of the sink from the cluster and also in the limit of the lifetime, the effect is negligible and the analysis becomes exact.

Computer simulations presented in Section 8 verify the accuracy of the analysis for various setups.

6 Lifetime Analysis for Multihop Networks

In the previous section, we assumed that CH relays the data packets to the sink directly. This is not the case in all clustering protocols and data forwarding from CH to the sink may be completed through multihop communication [34], [35]. Evidently, the multihop communication between CH and data sink does not affect the energy consumption for intracluster transmissions among the sensors. Hence, we just need to modify E_c.

Suppose that a routing tree, rooted at the data sink, results from applying a routing protocol in the network. We study the lifetime of an arbitrary cluster in this routing tree. A CH is in level i of the tree if its shortest path to the sink has i hops. A CH in level i sends its data packets to the sink through another CH in level $i-1$. CHs in level 1 transmit their data directly to the sink.

To count the effect of multihop transmission into our analysis, it is required to consider two key points which differ from the analysis in the previous section. First, the multihop transmission changes the number of packets that a CH has to relay. Indeed, the CH should forward all data packets generated by all clusters located in its routing subtree. In addition, except the CHs in level 1 that send their data to the fixed data sink, the other CHs send their data to the CHs in the lower level whose position changes as the CH role rotates. In other words, CHs in level i, where $1 < i$, send data packets to a destination which has a random location over the neighboring cluster.

Assume that $\mathcal{C}_j$ is an arbitrary cluster in the mentioned routing tree which is located in level i. We study the amount of consumed energy in an arbitrary node in this cluster, called n_1 which is located at s_1. Also, $\mathcal{C}_j$ has N_j nodes and its CH compresses data (received from the sensors within its own cluster) with ratio γ_j. In addition, $\mathcal{C}_j$ forwards data to cluster $\mathcal{C}_k$ located in level $i-1$. As mentioned, $\mathcal{C}_j$ sends the data to the sink if it is located in level 1. If s' represents the position of CH in $\mathcal{C}_k$, the consumed energy for intercluster transmissions by n_1 is

$$\begin{aligned} E_c = &\left(\gamma_j \mathbf{P}_t + \sum_{i\in \mathrm{subtree}(\mathcal{C}_j)} \mathcal{Q}_i\right)\left(m_2\|\mathbf{s}_1-\mathbf{s}'\|^4 + c_2\right) \\ &+ e_r\left(\mathbf{P}_r + \sum_{i\in \mathrm{subtree}(\mathcal{C}_j)} \mathcal{Q}_i\right). \end{aligned} \tag{28}$$

In (28), $\mathcal{Q}_i$ denotes the number of packets sent to n_1 by CHs of $\mathcal{C}_j$'s neighbors in level $i+1$.

The average consumed energy by n_1, when it serves as CH, is conditioned on both s_1 and s'. Thus, to find μ_{s_1} and $\sigma^2_{\mathrm{s}_1}$, we need to first average over s' inside $\mathcal{C}_k$.

Mean and variance of $\mathcal{Q}_i$s are determined using the discussion provided in Section 4. The remaining steps of the analysis are the same as the previous section. It is worthy to note that for CHs located in level 1, the destination of the transmissions by CHs is fixed, thus, the analysis is similar to the previous section and we just need to adjust the packet transmission model of CH.

7 Some Remarks

Remark 1. According to the proposed analysis, one can determine whether clustering results in energy efficiency or not. To this aim, the average consumed energy by a single node in both situations can be studied and used as a measure of energy efficiency. As mentioned previously, the value derived in (20) is an accurate approximation of the average consumed energy by each sensor. We compare this value by the average consumed energy when no clustering is applied and propose a condition for the efficiency of the clustering algorithm.

The average value presented in (20) depends on the position of the node. To compare the energy efficiency of a clustered network with the nonclustered one, the unconditional value of the average consumed energy is required which can be obtained by averaging μ_{s_1} over the position of sensor 1. If E_d represents the average consumed energy by an arbitrary node in a nonclustered network where the sensors send their data directly to the sink, clustered network has a batter energy-efficiency when

$$\int_{\mathcal{C}} \mu_{\mathrm{s}_1} f_{\mathbf{S}_1}(\mathrm{s}_1)\, d\mathrm{s}_1 < E_d. \tag{29}$$

Notice that for multiple event reporting

$$E_d = T_\ell \lambda \pi r_s^2 \int_{\mathcal{C}} \left(m_2\|\mathbf{s}_1-\mathbf{s}_s\|^4 + c_2\right) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1. \tag{30}$$

Applying (29) and (30) results in the following condition on γ, we have

$$\begin{aligned} \gamma < &\frac{\pi r_s^2}{\|\mathcal{C}\|(1-e^{-\pi\delta})} \\ &\times \left[N - \frac{N e_r + (N-1)\displaystyle\int_{\mathcal{C}} E[z] f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1}{\displaystyle\int_{\mathcal{C}} (m_2\|\mathbf{s}_1-\mathbf{s}_s\|^4 + c_2) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1}\right], \end{aligned} \tag{31}$$

where z was defined in (24).

For single event reporting case, we have

$$E_d = \lambda T_\ell \frac{\|\mathcal{C}\|}{N} \int_{\mathcal{C}} (m_2\|\mathbf{s}_1-\mathbf{s}_s\|^4 + c_2) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1. \tag{32}$$

Thus, the following condition guarantees the preference of clustered networks over nonclustered ones

$$\gamma \leq 1 - \frac{e_r + \frac{N-1}{N}\displaystyle\int_{\mathcal{C}} E[z] f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1}{\displaystyle\int_{\mathcal{C}} (m_2\|\mathbf{s}_1-\mathbf{s}_s\|^4 + c_2 + e_r) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1}. \tag{33}$$

Remark 2. The effect of the idle power consumption can be considered in the analysis. Assume that P_i shows the idle power consumed to keep the radio part of the sensor on when no transmitting or receiving occurs. Since the transmission time of each packet, T_p, is very short compared to the desired lifetime, as a simple approach, one can ignore the effect of the transmission time intervals. Thus, a single node can achieve the lifetime when $E_t < E_0 - T_\ell P_i$. To accommodate the effect of the idle power in the network, we need to replace E_0 with $E_0 - T_\ell P_i$ in the proposed analysis.[1]

Remark 3. The MAC protocol affects the sensor energy consumption directly. As mentioned in [36], the energy consumption of TDMA is mainly due to the packet retransmissions. In fact, the consumed energy for transmitting acknowledgement packets is negligible. The packet retransmission scenario can be accommodated in our analysis as follows.

Assume that the packet loss rate is ρ. To incorporate the influence of the packet loss rate, we need to adjust the mean and variance of $\mathcal{P}_e$, $\mathbf{P}$, and $\mathbf{P}'$. For this purpose, it can be shown that

$$E_n[\mathcal{P}_e] = \frac{1}{1-\rho} E[\mathcal{P}_e], \tag{34}$$

where $E_n[\mathcal{P}_e]$ is the modified mean of $\mathcal{P}_e$. Notice that $\frac{1}{1-\rho}$ is the average number of transmissions per data packet. Also,

$$\mathrm{VAR}_n[\mathcal{P}_e] = \frac{1+\rho}{(1-\rho)^2}\mathrm{VAR}[\mathcal{P}_e] + \frac{\rho}{(1-\rho)^2}E^2[\mathcal{P}_e], \tag{35}$$

where $\mathrm{VAR}_n[\mathcal{P}_e]$ denotes the modified variance of $\mathcal{P}_e$. The mean and variance of $\mathbf{P}$ and $\mathbf{P}'$ can be modified in a similar way.

Remark 4. Assume that the network is time-driven where sensors generate packet with rate λ'. In this case, the randomness of the traffic generation on the cluster lifetime is removed. Thus, $\mu_{\mathbf{s}_1}$ becomes

$$\mu_{\mathbf{s}_1} = \gamma\lambda' T_\ell (m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2 + e_r) + \frac{(N-1)T_\ell\lambda'}{N}E[z]. \tag{36}$$

Further, $\sigma^2_{\mathbf{s}_1}$ is

$$\begin{aligned}\sigma^2_{\mathbf{s}_1} &= \gamma^2\lambda'^2 T_\ell^2 (m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2 + e_r)^2 \\ &\quad + (N-1)\left(\frac{T_\ell\lambda'}{N}\right)^2\big(E[z^2] + \mathrm{VAR}[z]\big).\end{aligned} \tag{37}$$

The remaining part of the analysis is the same as event-driven networks.

Remark 5. The analysis is also applicable to multilevel transmission power models instead of the adjustable transmission power of (1), e.g., Mica Mote [17] and Intel Mote [37]. In this case, the transmitted power does not depend on the distance between the node and destination and $m_1\|\mathbf{s}_1 - \mathbf{s}_i\|^2 + c_1$ and $m_2\|\mathbf{s}_1 - \mathbf{s}_s\|^4 + c_2$ are replaced with two constant levels of the transmission energy. Thus, P_a is independent of the location of sensor. This removes the analysis step indicated in (64). If all the sensors in the cluster use a fixed power level for intracluster and another fixed power-level for intercluster transmissions, our analysis will not even be affected by the death of nodes and provides an exact description of the cluster lifetime. This is because, the location of nodes does not influence the energy consumption and consequently death of nodes occurs randomly over the area. Hence, the effective cluster region is not affected by the nodes death.

Remark 6. A reasonable definition of the network lifetime is the time duration over which all clusters are alive. In this case, the probability of achieving a desired lifetime by the network can be determined using the provided analysis for the clusters lifetime. In a single-hop network, lifetime of the clusters are independent from each other. Hence, if ℓ shows the number of clusters in the network, the probability of achieving T_ℓ by the network, P, is

$$P = \prod_{i=1}^{\ell} P_{\ell_i}, \tag{38}$$

where P_{ℓ_i}, $1 \le i \le \ell$, represents the probability of achieving T_ℓ by ith cluster in the network derived in Section 5.

In a multihop case, clusters in level 1 carry a higher traffic load compared to the clusters located in other levels and consequently they usually die earlier compared to the other clusters. Hence, to estimate the network lifetime, it is possible to just investigate the lifetime of the sensors in level 1 which are independent.

Remark 7. Our analysis can be used in the network design stage in order to optimize the network implementation cost. The network cost is a key factor in the deployment of a WSN and depends on the number of sensors and also the battery that they are equipped with. Here, we like each cluster to achieve T_ℓ with reliability (probability) R while the network cost is minimized. Assume that each sensor costs h and we have k different batteries with initial energy $\{E_1, E_2, \ldots, E_b\}$ which cost $\{g_1, g_2, \ldots, g_b\}$, respectively. Having the above assumptions, the network design problem can be formulated as follows:

$$\min_{N,i} N(h + g_i) \tag{39}$$

$$\text{s.t.}\ \ I_{P_a}(\lfloor(1-\beta)N\rfloor + 1, N - \lfloor(1-\beta)N\rfloor) \ge R, \tag{40}$$

$$P_a = \int_{\mathcal{C}} \left(1 - Q\left(\frac{E_i - \mu_{\mathbf{s}_1}}{\sigma_{\mathbf{s}_1}}\right)\right) f_{\mathbf{S}_1}(\mathbf{s}_1)\, d\mathbf{s}_1. \tag{41}$$

By solving the above optimization problem, the appropriate number of nodes within each cluster and the energy level of the sensor battery level can be found. Since b is typically a small number, one way to solve this optimization problem is to find the optimal N for each i. Notice that for a fixed i, the optimal N is the smallest N satisfying (40). Comparing all pairs of solutions using (39), the optimal pair of N and i is found.

1. One can involve the minor effect of transmission time in the above discussion as follows: First note that $E_t = E_n + E_c + E_i$, where $E_i = \big[T_\ell - T_p(\gamma\mathbf{P}_t - \mathbf{P}_r - \sum_{j=1}^{(N-1)}\mathcal{P}_k)\big]P_i$. Using the above equations, $\mu_{\mathbf{s}_1}$ and $\sigma^2_{\mathbf{s}_1}$ have to be modified accordingly. Notice that mean and variance of E_i are calculated easily using the mean and variance of $\mathbf{P}_t$, $\mathbf{P}_r$, and $\mathcal{P}_k$ which was discussed previously.

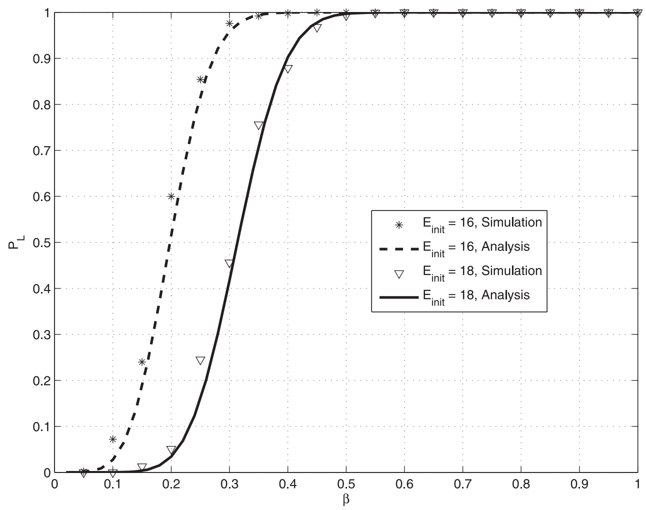

Fig. 1. Probability of achieving $T_\ell = 2{,}000$ time units by a cluster in a single-hop network for single event reporting scheme.

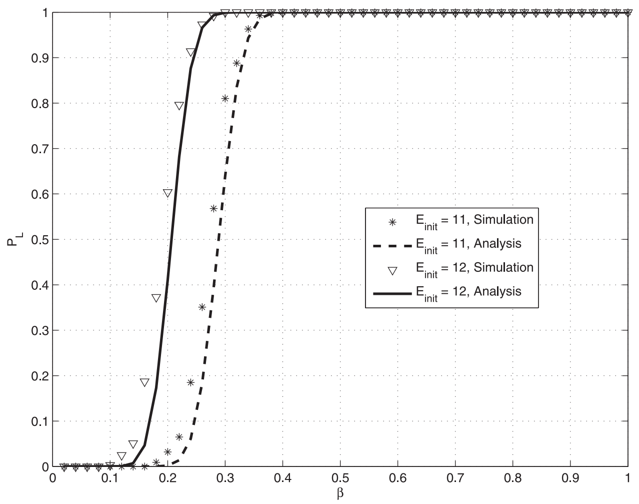

Fig. 2. Probability of achieving $T_\ell = 6{,}000$ time units by a cluster in a single-hop network for multiple event reporting scheme.

8 Numerical Results

In this section, the analysis is verified through computer simulations using MATLAB. Please notice that in our simulations, we use the accurate network model and do not involve the approximations made in the analytical results to find a simple lifetime model.

It is assumed that $\lambda = 0.1$ (packets/time unit/m^2), $T = 1$ time units, and $\gamma = 1$ except for the setup where the effect of γ is studied. Having data packets with length 100 bits, the loss coefficient in (1) is $m_1 = 0.13\ \mathrm{pJ/m}^4$ when $\alpha = 4$ and $m_2 = 1\ \mathrm{nJ/m}^2$ for $\alpha = 2$ [5]. Also, $e_r = 4\ \mu\mathrm{J}$ and $c_1 = c_2 = 1\ \mu\mathrm{J}$. For multiple event reporting case, $r_s = 3$ m. Here, the CH role is assigned to the sensors periodically based on a predetermined schedule. If the CH assigning schedule reaches a sensor which is already dead, the dead sensor will be removed from the schedule and the next sensor in the list will be assigned as CH. Simulations are conducted for both single event reporting and multiple event reporting schemes separately. Also, to consider the randomness of the sensors placement, simulation was run 100 times for each different setup.

First, we focus on the single-hop network. To this end, we consider a network with just one cluster. Considering more number of clusters will give similar results. When the sensors follow a single event reporting scheme, Fig. 1 presents the analysis and simulation results for $T_\ell = 2{,}000$ time units when $N = 50$ nodes are spread over the cluster area. In this figure, the probability of achieving the desired lifetime by a single cluster is plotted versus the ratio of the dead nodes over the area, β. In Fig. 1, it is assumed that the cluster has a square shape with side length of $L = 50$ m and centered at the origin. Sensors are spread uniformly over the area. In addition, the sink is located on $(0, -175)$ m.

Fig. 1 shows the simulation and analysis results for two different values of E_0. As it can be seen, the proposed analysis can predict the network performance accurately.

Simulation and analysis results for the case where the sensors follow a multiple reporting scheme are shown in Fig. 2. In this scenario, $N = 200$ nodes are deployed over the cluster area. L and d have the same value as the previous simulation. Here, we like to study the probability of achieving $T_\ell = 6{,}000$ time units.

The results of our analysis is also applicable to predict the quality of the network performance after a desired lifetime. Quality of the network performance, e.g., network coverage and connectivity, can be measured by β. For example, in a single event reporting case, assume that one is interested to find the network quality of performance with reliability $P_L = 0.9$ after $T_\ell = 1{,}500$ time units. With $N = 50$, $E_0 = 12$, and $L = 50$ m for a single event reporting single-hop network, our analysis gives $\beta = 0.4$. This means that 60 percent of the nodes reach the desired lifetime and can operate satisfyingly. Computer simulation shows that with the abovementioned network setup $\beta = 0.39$, meaning that 61 percent of nodes reach the desired lifetime.

Rate of data compression, γ, plays an important role in the network lifetime. Due to the similarity of the results for single event reporting and multiple event reporting schemes, we just present the results for single event reporting scenario in the following. Fig. 3 shows the simulation and analysis results for the probability of achieving $T_\ell = 2{,}000$ time units for different data compression ratios. $N = 50$ nodes with

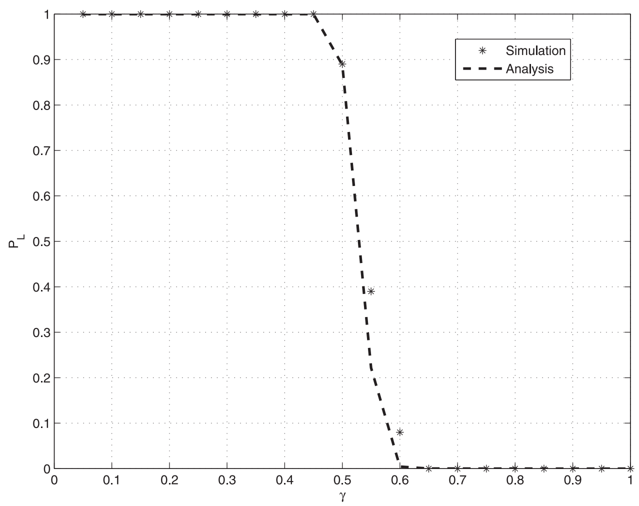

Fig. 3. Probability of achieving $T_\ell = 2{,}000$ time units by a cluster for different data aggregation ratios.

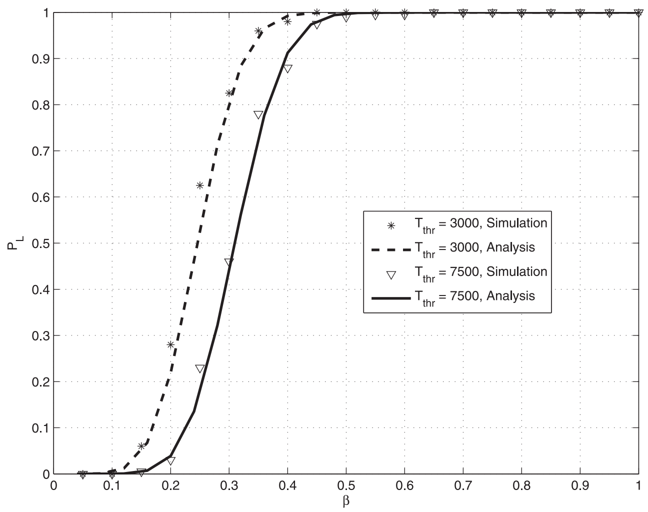

Fig. 4. Probability of achieving $T_\ell = 3{,}000$ time units and $T_\ell = 7{,}500$ time units by a cluster in a multihop network for single event reporting scheme.

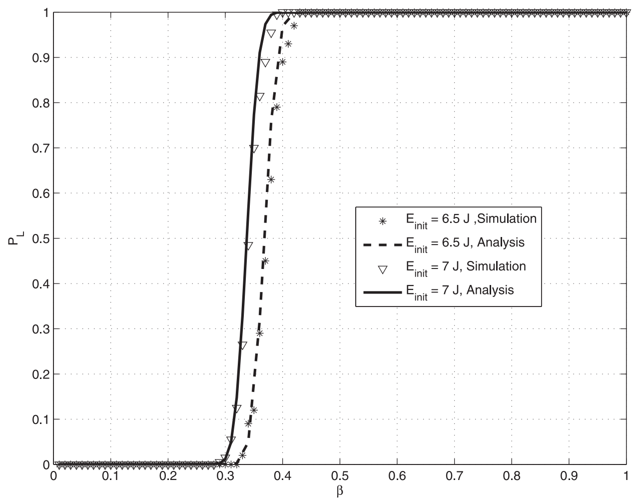

Fig. 5. Probability of achieving $T_\ell = 12{,}000$ time units by a cluster in a multihop network for multiple event reporting scheme.

$E_0 = 2$ J are spread uniformly over a square with side length $L = 50$.

The probability of achieving the lifetime threshold by a cluster in a multihop network is also studied by computer simulation. For this purpose, it is assumed that the network has two clusters with square shape and the intercluster transmissions occur between neighboring CHs. Fig. 4 shows results for the cluster in the second level of the routing tree (a leaf of the tree) for single event reporting case where $N = 50$ sensors are spread over a square-shaped clusters with side $L = 50$ m. Also, for $T_\ell = 3{,}000$ time units and $T_\ell = 7{,}500$ time units, the initial sensors energy are $E_0 = 1.5$ J and $E_0 = 3.5$ J, respectively.

Results of the multihop transmission associated with multiple event reporting scheme are shown in Fig. 5 where the studied cluster is a leaf of the routing tree. Here, $N = 200$ nodes are spread over a square with side $L = 100$. For $T_\ell = 12{,}000$ time units, two different values for E_0 have been used for simulation. As it can be seen, our analysis provides an accurate estimation of the network lifetime behavior.

9 Conclusion

In this paper, we proposed an approach to determine the probability of achieving a lifetime by the clusters of a WSN. The reason for a probabilistic analysis comes from the fact that the network lifetime has a random essence due to the randomness in the network deployment and traffic generation. We studied the lifetime at the sensor level by characterizing an appropriate traffic model for each sensor and considering the sensor random location within the cluster. Moreover, we showed that depending on the data compression ratio, clustered networks may or may not outlive nonclustered ones. Our analysis can predict the quality of the network performance after a desired lifetime, represented by the number of live nodes in the network. As a possible application of this work, one may use our proposed analysis to adjust the network setup such that while the desired lifetime is achievable, the network implementation cost, which is associated with number of nodes and their initial energy, is kept as low as possible.

Acknowledgments

The authors would like to thank Natural Sciences and Engineering Research Council of Canada (NSERC) and Informatics Circle of Research Excellence (*i*CORE) for supporting our research. Some of the results of this paper have been presented in *IEEE* Wireless Communications and Networking Conference (WCNC) 2008.

References

[1] Y. Zhou, M. Hart, S. Vadgama, and A. Rouz, "A Hierarchical Clustering Method in Wireless Ad Hoc Sensor Networks," *Proc. IEEE Int'l Conf. Comm. (ICC '07),* pp. 3503-3509, June 2007.
[2] C. Liu, K. Wu, and J. Pei, "An Energy-Efficient Data Collection Framework for Wireless Sensor Networks by Exploiting Spatiotemporal Correlation," *IEEE Trans. Parallel and Distributed Systems,* vol. 18, no. 7, pp. 1010-1023, July 2007.
[3] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Trans. Wireless Comm.,* vol. 1, no. 4, pp. 660-670, Oct. 2002.
[4] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," *IEEE Trans. Mobile Computing,* vol. 3, no. 4, pp. 366-379, Oct.-Dec. 2004.
[5] W.B. Heinzelman, "Application-Specific Protocol for Wireless Networks," PhD dissertation, Massachusetts Inst. of Technology, 2000.
[6] D.M. Blough and P. Santi, "Investigating Upper Bounds on Network Lifetime Extension for Cell-Based Energy Conservation Techniques in Stationary Ad Hoc Networks," *Proc. ACM MobiCom,* pp. 183-192, 2002.
[7] H. Zhang and J. Hou, "On Deriving the Upper Bound of α-Lifetime for Large Sensor Networks," *Proc. ACM MobiHoc,* pp. 121-132, 2004.
[8] E.J. Duarte-Melo and M. Liu, "Analysis of Energy Consumption and Lifetime of Heterogeneous Wireless Sensor Networks," *Proc. IEEE Global Telecomm. Conf. (GlobeCom '02),* vol. 1, pp. 21-25, Nov. 2002.
[9] G. Ferrari and M. Martalò, "Extending the Lifetime of Sensor Networks through Adaptive Reclustering," *EURASIP J. Wireless Comm. and Networking,* vol. 2007, 2007.
[10] S.K. Prasad and A. Dhawan, "Distributed Algorithms for Lifetime of Wireless Sensor Networks Based on Dependencies among Cover Sets," *Proc. 14th Int'l Conf. High Performance Computing (HiPC '07),* pp. 381-392, 2007.

[11] A. Fapojuwo and A. Cano-Tinoco, "Energy Consumption and Message Delay Analysis of QoS Enhanced Base Station Controlled Dynamic Clustering Protocol for Wireless Sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 8, no. 10, pp. 5366-5374, Oct. 2009.

[12] M. Peng, Y. Xiao, and P. Wang, "Error Analysis and Kernel Density Approach of Scheduling Sleeping Nodes in Cluster-Based Wireless Sensor Networks," *IEEE Trans. Vehicular Technology*, vol. 58, no. 9, pp. 5105-5114, Nov. 2009.

[13] Y. Zhuang, J. Pan, and L. Cai, "Minimizing Energy Consumption with Probabilistic Distance Models in Wireless Sensor Networks," *Proc. IEEE INFOCOM*, pp. 1-9, 2010.

[14] N. Vlajic and D. Xia, "Wireless Sensor Networks: To Cluster or Not to Cluster?" *Proc. Int'l Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM '06)*, June 2006.

[15] S. Soro and W.B. Heinzelman, "Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering," *Proc. IEEE 19th Int'l Parallel and Distributed Processing Symp.*, Apr. 2005.

[16] O. Younis, M. Krunz, and S. Ramasubramanian, "Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges," *IEEE Networks*, vol. 20, no. 3, pp. 20-25, May/June 2006.

[17] J.L. Hill and D.E. Culler, "Mica: A Wireless Platform for Deeply Embedded Networks," *IEEE Micro*, vol. 22, no. 6, pp. 12-24, Nov./Dec. 2002.

[18] A. Goldsmith, *Wireless Communications.* Cambridge Univ., 2005.

[19] J.N. Al-Karaki and A.E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey," *IEEE Wireless Comm.*, vol. 11, no. 6, pp. 6-28, Dec. 2004.

[20] S.G.M. Kramer, F.P. Leon, and B. Appert, "Fiber Optic Sensor Network for Lightning Impact Localization and Classification in Wind Turbines," *Proc. IEEE Int'l Conf. Multisensor Fusion and Integration for Intelligent Systems*, pp. 173-178, Sept. 2006.

[21] F. Schäfer and R. Janovsky, "Impact Sensor Network for Detection of Hypervelocity Impacts on Spacecraft," *Acta Astronautica*, vol. 61, no. 10, pp. 901-911, 2007.

[22] Y. Ma and J.H. Aylor, "System Lifetime Optimization for Heterogeneous Sensor Networks with a Hub-Spoke Topology," *IEEE Trans. Mobile Computing*, vol. 3, no. 3, pp. 286-294, July/Aug. 2004.

[23] C. Hua and T.-S.P. Yum, "Asynchronous Random Sleeping for Sensor Networks," *ACM Trans. Sensor Networks*, vol. 3, no. 3, p. 15, 2007.

[24] Y. Ling and C.-M. Chen, "Energy Saving via Power-Aware Buffering in Wireless Sensor Networks," *Proc. IEEE INFOCOM*, pp. 2411-2415, 2007.

[25] N. Rao, M. Shankar, J.-C. Chin, D. Yau, S. Srivathsan, S. Iyengar, Y. Yang, and J. Hou, "Identification of Low-Level Point Radiation Sources Using a Sensor Network," *Proc. Int'l Conf. Information Processing in Sensor Networks (IPSN '08)*, pp. 493-504, 2008.

[26] M. Bhardwaj, T. Garnett, and A.P. Chandrakasan, "Upper Bounds on the Lifetime of Sensor Networks," *Proc. IEEE Int'l Conf. Comm. (ICC '01)*, vol. 3, pp. 785-790, 2001.

[27] W. Ye, J. Heidemann, and D. Estrin, "An Energy-Efficient MAC Protocol for Wireless Sensor Networks," *Proc. IEEE INFOCOM*, vol. 3, pp. 1567-1576, 2002.

[28] R. Madan and S. Lall, "Distributed Algorithms for Maximum Lifetime Routing in Wireless Sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 5, no. 8, pp. 2185-2193, Aug. 2006.

[29] J.-H. Chang and L. Tassiulas, "Energy Conserving Routing in Wireless Ad-Hoc Networks," *Proc. IEEE INFOCOM*, pp. 22-31, Mar. 2000.

[30] S. Adlakha and M. Srivastava, "Critical Density Thresholds for Coverage in Wireless Sensor Networks," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC '03)* vol. 3, pp. 1615-1620, Mar. 2003.

[31] W. Ye and J. Heidemann, *Medium Access Control in Wireless Sensor Networks.* Kluwer Academic, 2004.

[32] B.A. Bash and P.J. Desnoyers, "Exact Distributed Voronoi Cell Computation in Sensor Networks," *Proc. Sixth Int'l Conf. Information Processing in Sensor Networks (IPSN '07)*, pp. 236-243, 2007.

[33] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, D. Zwillinger and A. Jeffery, eds., sixth ed. Academic Press, 2000.

[34] Y.-C. Chang, Z.-S. Lin, and J.-L. Chen, "Cluster Based Self-Organization Management Protocols for Wireless Sensor Networks," *IEEE Trans. Consumer Electronics*, vol. 52, no. 1, pp. 75-80, Feb. 2006.

[35] X. Du and F. Lin, "Improving Routing in Sensor Networks with Heterogeneous Sensor Nodes," *Proc. IEEE 61st Vehicular Technology Conf. (VTC '05)*, vol. 4, pp. 2528-2532, May/June 2005.

[36] L. Shi and A. Fapojuwo, "TDMA Scheduling with Optimized Energy Efficiency and Minimum Delay in Clustered Wireless Sensor Networks," *IEEE Trans. Mobile Computing*, vol. 9, no. 7, pp. 927-940, July 2010.

[37] R.M. Kling, "Intel Motes: Advanced Sensor Network Platforms and Applications," *Proc. IEEE MTT-S Int'l Microwave Symp. Digest*, June 2005.

[38] P. Hall, *Introduction to the Theory of Coverage Process.* John Wiley and Sons, 1988.

[39] J.-S. Ferenc and Z. Néda, "On the Size Distribution of Poisson Voronoi Cells," *Physica A: Statistical Mechanics and Its Applications*, vol. 385, no. 2, pp. 518-526, 2007.

Moslem Noori received the BSc degree in electrical engineering and applied mathematics from the Amirkabir University of Technology in 2005 and 2006, respectively. Then, he began a graduate studies program in the Department of Electrical and Computer Engineering at the University of Alberta, where he obtained the MSc degree in 2008 and is currently working toward the PhD degree. His research interests include performance analysis of random networks and network coding. He is a recipient of several awards and scholarships including NSERC Vaniver CGS, the *i*CORE PhD Scholarship in ICT, and the Connaught Scholarship. He is a student member of the IEEE.

Masoud Ardakani received the BSc degree from the Isfahan University of Technology in 1994, the MSc degree from Tehran University in 1997, and the PhD degree from the University of Toronto in 2004, all in electrical engineering. He was a postdoctoral fellow at the University of Toronto from 2004 to 2005. Currently, he is working as an associate professor in electrical and computer engineering and is an Alberta Ingenuity New Faculty member at the University of Alberta. His research interests include the general area of digital communications, codes defined on graphs, and iterative decoding techniques. He serves as an associate editor for the *IEEE Transactions on Wireless Communications* and the *IEEE Communications Letters*. He is a senior member of the IEEE.