Optimal Number of Clusters in Dense Wireless Sensor Networks: A Cross-layer Approach

Li-Chun Wang, Senior Member, IEEE, Chung-Wei Wang, Student Member, IEEE, and Chuan-Ming Liu, Member, IEEE

Abstract — Cluster-based sensor networks have the advantages of reducing energy consumption and link-maintenance cost. One fundamental issue in cluster-based sensor networks is to determine the optimal number of clusters. In this paper, we suggest a physical (PHY)/medium access control (MAC)/network (NET) cross-layer analytical approach to determine the optimal number of clusters with an objective of minimizing energy consumption in a high density sensor network. Our cross-layer design can incorporate many effects, including lognormal shadowing and two-slope path loss model in the physical layer, various MAC scheduling and multi-hop routing schemes. Compared with the base-line case with one cluster per observation area, a sensor network with the proposed optimal number of the clusters can reduce energy consumption by over 80% in some cases. We also verify by simulations that the analytical optimal cluster number can still function effectively regardless of different density of sensors in various observation areas.

Index Terms — Wireless sensor networks; Optimal number of clusters; Cross-layer design.

I. INTRODUCTION

Cluster-based sensor networks have many advantages. For example, with clustering, energy consumption can be improved because only one representative node per cluster is required to be active and other nodes can enter the dormant mode [2]–[4]. Clustering architecture has important applications for high-density sensor networks because it is much easier to manage a set of cluster representatives from each cluster instead of managing whole sensor nodes. Environment measurement [5], target tracking [6], [7], intrusion detection [8], and pursuit-evasion games [9] are typical applications for high-density sensor networks due to the fault tolerance requirement.

One of key challenges for deploying a high-density cluster-based sensor network is to determine the optimal number of clusters. For a high-density sensor network, the coverage area can be partitioned into many disjoint spatial coherence regions [10], where the sensed information is highly spatial-correlated. In this paper, we call such a spatial coherence region a basic observation area. Intuitively, each basic observation area with highly correlated sensed information requires only one cluster representative [11]–[14]. In general, the number of clusters in a basic observation area can be determined from different aspects of protocol layers:

• From the physical (PHY) layer aspect, using more clusters can save more energy because the transmission distance between cluster representatives can be shortened.
• From the medium access control (MAC) layer aspect, having fewer clusters (or equivalently more nodes per cluster) decreases the average possibility of being a cluster representative for each sensor node and thus reduces energy consumption.
• From the multi-hop routing aspect in the network (NET) layer, fewer clusters yield fewer hop counts to the data sink, and result in less energy consumption.

Hence, to optimize the number of clusters in a sensor network becomes a cross-layer tradeoff design issue among the required transmission power in the PHY layer, the possibility of being a cluster representative in the MAC layer and the hop counts in the relay path in the NET layer.

A. Related Work

The issue of optimizing the number of clusters in a wireless sensor network has been addressed by many researchers from different aspects:

• First, from the viewpoint of propagation distance in the physical layer, many authors [15]–[20] discussed how to design the number of clusters in a sensor network. The authors in [15] concluded that the number of clusters should be as large as possible because the distance between the cluster head and its members can be shortened. However, in [16], [17] it was shown that a larger number of clusters also leads to more one-hop transmissions from the heads to the sink. Thus, an optimal number of clusters exists [18], [19]. Furthermore, the authors in [20] discussed the effects of shadowing and path loss exponents on the optimal number of clusters.
• From the MAC layer aspect, the authors in [21] suggested that the design of the optimal number of
clusters should include the impact of the contention mechanism. However, an explicit MAC protocol to achieve this goal was not presented.

• From the routing aspect, the authors in [22] assumed that each node sends data to the corresponding cluster heads using multihop routing. They found that more clusters will result in more one-hop transmissions from the heads to the sink although the total routing traffic within each cluster is reduced because of fewer members. Furthermore, [23] determined the optimal number of clusters from the viewpoint of minimizing the routing overhead using an information-theoretical approach.

Current methods to determine the optimal number of clusters are based on each individual protocol layer aspect. To our knowledge, a PHY/MAC/NET cross-layer analytical approach to determine the optimal number of clusters in a dense sensor network has received little notice in the literature.

B. Contributions and Organization of This Paper

The objective of this paper is to develop a PHY/MAC/NET cross-layer analytical method to calculate the optimal number of clusters in each basic observation area for a high-density sensor network. Based on the proposed analytical approach, an optimal number of clusters can be obtained without a time-consuming search and labor-intensive field trials. Therefore, the deployment of the energy-efficient cluster-based sensor network becomes easier.

The contributions of this paper can be summarized as follows:

• Firstly, we propose a cross-layer analytical approach to determine the optimal number of clusters. In the cluster number design problem, it is a concept in the average sense. Hence, we first suggest adopting the criterion of the minimal total average energy (instead of the lifetime in the extreme case) to calculate the optimal number of clusters per observation area. The PHY/MAC/NET cross-layer analytical design approach for the optimal number of clusters in high-density sensor networks has not been seen in the literature.

• Secondly, we show the existence of the optimal cluster number regardless of different density of sensors in various observation areas by simulations and analysis. We also take account of other randomness factors in our simulation platform (Fig. 6), including log-normal shadowing and more realistic two-slope path loss model. The simulation results are shown to match the proposed analytical results quite well. To our knowledge, this interesting finding of the optimal cluster number regardless of different density of sensors in various observation areas has not been reported in the literature yet.

The rest of this paper is organized as follows. In Section II, we introduce the system model and formulate an optimization problem to find the best number of clusters in high-density sensor networks. Section III discusses how to determine the optimal number of clusters from a PHY/MAC cross-layer perspective. In Section IV, we further consider the network layer aspect. The numerical results are shown in Section V. Finally, we give our concluding remarks in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

Consider a sensor network with \( N \) nodes in a grid of \( M \) basic observation areas, each of which contains \( K \) square-shaped clusters. A basic observation area (OA) represents a spatial coherence region with highly correlated information. Define \( d_{OA} \) as the minimal distance that results in uncorrelated information [10]. Figure 1 shows an example of a sensor network with \( M = 16 \) observation areas and four clusters per observation area (i.e., \( K = 4 \)). In any instant only one sensor node per cluster is active to sense the surrounding information, while the other nodes enter the dormant mode. Furthermore, we suppose that sensor nodes in each cluster can be synchronized by a certain clock synchronization mechanism [24].

To ease analysis, the grid model was used for sensor networks [25]–[29]. In general, the shape of clusters can be arbitrary (or close to circle) because of the randomness inherent in radio propagation. However, the grid model is still a good candidate to obtain insights for system design [30].

\( ^1 \)To design cluster-based sensor networks, we need some geometric shapes to tessellate the entire coverage area, such as triangle, square, and hexagon. Hexagon-shaped coverage is widely used for a base station in mobile cellular systems. Our approach can be also modified to the hexagon-shaped clusters, which is not included in this manuscript due to limitation on page length.
Now we briefly introduce the operations of the cluster-based sensor network considered in this paper. First, the clusters are formed according to some cluster formation mechanisms [31]. One cluster head is selected in each cluster to schedule the subsequent cluster representatives from its cluster members. When receiving the schedule broadcast from the cluster head, all sensor nodes wake up to be the cluster representatives in turn. The cluster representatives periodically report sensed information to the sink based on a multi-hop delivery. Note that the clusters are reformed periodically, and the reformation period of the clusters is called a round in the paper.

B. Problem Formulation

In the most of literature, the network lifetime is defined as the time elapsed from the start of the sensor network until the death of the first node [32]. In this case, the researches cared for the maximal energy consumption over all sensor nodes. Hence, the objective function of finding the optimal number of clusters can be expressed as follows:

\[ K_{opt} = \arg \min_{1 \leq K \leq N} \max_{1 \leq i \leq N} E_i(K) , \]

where \( E_i(K) \) is the energy consumption of the sensor \( i \) when the number of clusters in each basic observation area is \( K \).

For a dense sensor network, we believe that the lifetime in the extreme case may not be the only performance measure to be considered. It is very likely that the remaining sensor nodes in a dense sensor network can still form a network even if one sensor node is not functioning. Hence, our goal is to optimize the average energy consumption, instead of the lifetime in the extreme case (i.e., the duration for the first sensor running out its energy). In this paper, we formulate the Energy Minimizing Problem for sensor networks as follows: Given a set of system parameters, including the minimal uncorrelated distances \( d_{O/A} \), the number of sensor nodes \( N \), and the number of basic observation areas \( M \), find the optimal number of clusters (denoted by \( K_{opt} \)) in each basic observation area to minimize the average energy consumption (denoted by \( E(K) \)). Formally,

\[ K_{opt} = \arg \min_{1 \leq K \leq \frac{N}{M}} E(K) , \]

subject to:

\[
\begin{align*}
0 & \leq P_t \leq P_{\text{max}} ; \quad & \text{(Physical layer)} \\
0 & \leq \tau \leq \tau_{\text{max}} ; \quad & \text{(Physical/MAC layer)} \\
1 & \leq h \leq h_{\text{max}} ; \quad & \text{(Network layer)}
\end{align*}
\]

where \( P_t \), \( \tau \), and \( h \) are the transmission power, average retransmission times in the MAC layer, and average hop counts from the cluster representatives to the sink, respectively; and \( P_{\text{max}} \), \( \tau_{\text{max}} \), and \( h_{\text{max}} \) are the corresponding maximum values for \( P_t \), \( \tau \), and \( h \).

Clearly, the number of clusters \( K \) in a basic observation area affects the performance of a sensor network from the PHY/MAC/NET cross-layer perspectives. Note that \( P_t \), \( \tau \), and \( h \) are all the function of \( K \). Specifically, we consider the physical layer propagation model with different path loss exponents and shadowing components.

III. PHY/MAC CROSS-LAYER ASPECT

In this section, we discuss the impacts of the transmission power and the MAC scheduling policy on the number of clusters \( K \) from the PHY/MAC cross-layer perspectives.

A. Energy Consumption Model

To begin with, denote \( E_{i\rightarrow j} \) as the energy consumption of transmitting data from sensors \( i \) to \( j \) during each round where \( i \neq j \). Denote \( P_t \) and \( P_e \) as transmission power and electronics power consumption. Then, according to [18], [33], it follows that

\[ E_{i\rightarrow j} = a_{i\rightarrow j} \left[ (P_e + P_t) \cdot t + P_e \cdot t \right] = a_{i\rightarrow j} \cdot (2P_e + P_t)t , \]

where \( t \) and \( a_{i\rightarrow j} \) are the duration in each transmission, and the times of link \( i \rightarrow j \) being established during each round. Here, the energy consumption for sensing is ignored because it is much lower than the energy consumption for transmission [34], [35]. Taking the expectation of \( E_{i\rightarrow j} \) over all links from sensors \( i \) to \( j \), the average energy consumption of each link during each round can be expressed as follows:

\[ E_{i\rightarrow j} = \frac{\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} E_{i\rightarrow j}}{N(N-1)} = (2P_e + P_t)t \cdot \bar{a}_{i\rightarrow j} \]

where \( \bar{a}_{i\rightarrow j} = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} a_{i\rightarrow j}/(N(N-1)) \) is the average times of each link being established during each round. Usually, \( t \) will be a given system parameter. In Sections III-B and III-C, we discuss how to obtain \( P_t \), and \( \bar{a}_{i\rightarrow j} \).

B. Impact of Transmission Power

Consider the two-slope path loss model. Let \( n_1 \) and \( n_2 \) be the path loss exponents and \( X_1 \) and \( X_2 \) be the log-normally shadowing components, respectively. At the distance of \( d \), the received power \( P_r \) can be written as

\[ P_r = \begin{cases} 
\frac{P_t}{r_{a1} \left( \frac{d}{d_{a1}} \right)^{-n_1}} X_1 , & d \leq d_t \\
\frac{P_t}{L_{a2} \left( \frac{d}{d_{a2}} \right)^{-n_2}} X_2 , & d > d_t 
\end{cases} \]

where \( L_{a1} \) and \( L_{a2} \) are the path losses at the reference distances \( d_{a1} \) and \( d_{a2} \), respectively. In (4), the two different slopes \( (n_1 \) and \( n_2 \) are respectively adopted before and after distance threshold \( d_t \). In general, we have \( n_1 < n_2 \). This propagation model had been validated in outdoor sensor networks at 868 MHz [36].

Figure 2 shows a square-shaped basic observation area with 4 clusters \((K = 4)\). In order to guarantee any two cluster representatives in adjacent clusters can be
directly connected, the one-hop transmission distance of each sensor node (denoted by \( d \)) must be larger than or equal to \( \sqrt{\frac{s}{K}} d_{OA} \) [2]. In this paper, we assume that the cluster representatives are connected through the multi-hop communication method when the propagation distance is longer than \( \sqrt{\frac{s}{K}} d_{OA} \). Hence, given a value of \( K \), the value of \( d \) is either larger than \( d_t \) or smaller than \( d_t \). Then, \( P_t \) can be uniquely determined from one of the two particular cases in (4). Without loss of generality, we have

\[
P_r = \frac{P_t}{L_0} \left( \frac{d}{d_0} \right)^{-n} X. \tag{5}
\]

If \( d_{OA} \leq \sqrt{\frac{s}{K}} d_t \), \( L_0 = L_0, d_0 = d_0 \), \( n = n_1 \), and \( X = X_1 \). Otherwise, \( L_0 = L_{OA}, d_0 = d_{OA}, n = n_2 \), and \( X = X_2 \). In Section V-D, we will evaluate the impact of two-slope path loss model on the determination of \( K \).

From (5), the transmission power of a cluster representative needed to maintain the received power level of \( P_r \) at a distance of \( \sqrt{\frac{s}{K}} d_{OA} \) becomes

\[
P_t = P_r L_0 \left( \frac{\sqrt{5} d_{OA}}{\sqrt{K} d_0} \right)^n, \tag{6}
\]

where the shadowing effect is not considered here (i.e., \( X = 1 \)) until Section III-E. Hence, from the physical layer perspective, we know that with larger values of \( K \), the required transmission power of a cluster representative is lower. Note that because the cluster representatives in the adjacent clusters contend channel based on the IEEE 802.11 MAC protocol, only one cluster representative can transmit at any time instant. Thus, the inter-cluster interference does not occur, and thus we do not consider it in (6). The detail of channel contention is discussed in Section IV-C.

### C. Impact of MAC Scheduling

The MAC layer scheduling policy affects how often cluster representatives \( i \) and \( j \) establish a link, i.e., \( a_{i,j} \).

A time-driven sensor network usually specifies the number of reports from each cluster during each round, denoted by \( R \). For a sensor network with \( MK \) clusters,

\[
MKR = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} a_{i,j}. \tag{6}
\]

In (3) of Section III-A, we have

\[
\bar{a}_{i,j} = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} a_{i,j} / (N(N - 1)). \tag{7}
\]

From (7), once \( R \) is specified, \( \bar{a}_{i,j} \) in (3) can be obtained.

#### D. Optimal Number of Clusters \( (K_l) \)

To determine the optimal number of clusters in a basic observation area from both the MAC scheduling and the physical transmit power perspectives, we can substitute (6) and (7) into (3). Let \( \eta \) be the required receive power. Then the average energy consumption of each link during each round can be expressed as

\[
E_{i-j}(K) = \frac{MK}{N(N-1)} R \left[ 2P_e + \eta L_0 \left( \frac{\sqrt{5} d_{OA}}{\sqrt{K} d_0} \right)^n \right] t. \tag{8}
\]

From (8), one can find that \( K \) affects the energy consumption in two ways. On one hand, when \( K \) decreases, the number of sensor nodes in a cluster increases. Because a smaller value of \( K \) results in more members in a cluster and thus lower probability to be a cluster representative, energy consumption can be reduced. On the other hand, a smaller \( K \) can also increase the energy consumption of a sensor node due to the longer transmission distance in a larger coverage area. Because \( \frac{\partial^2 E_{i-j}(K)}{\partial K^2} > 0 \), (8) is a convex function. Then, let \( \frac{\partial E_{i-j}(K)}{\partial K} = 0 \), we can obtain the optimal number of clusters \( K_l \) to minimize \( E_{i-j} \) from both the physical and the MAC layer perspectives. Specifically, we can express \( K_l \) as

\[
K_l = \left\{ \begin{array}{ll}
\left( \frac{\left( \frac{n}{2} - 1 \right)}{2\alpha} \right)^{\frac{1}{2}}, & n > 2 \\
1, & n = 2
\end{array} \right.
\]

where \( \alpha = P_e \) and \( \beta = \eta L_0 \left( \frac{\sqrt{5} d_{OA}}{d_0} \right)^n \).

#### E. Shadowing Effect

Now we consider the shadowing effect in (5). Recall that the shadowing component \( X \) is modelled as a log-normal random variable with zero mean and standard deviation \( \sigma \). Thus, the probability density function of the received
signal power $P_r$ can be expressed as [37]

$$f_{P_r}(y) = \frac{1}{\sqrt{2\pi\sigma}} (\frac{\ln 10}{10})^y \exp \left\{- \frac{10\log_{10} y - 10\log_{10} \frac{P_r L_0}{L_0} \frac{5\sigma^{10}}{\sqrt{K d_0}}}{2\sigma^2}\right\}.$$  \hspace{1cm} (10)

Note that the logarithm of the log-normal random variable yields a normal random variable. Define

$$P_r(dBm) = 30 + 10\log_{10} P_r; \quad \eta(dBm) = 30 + 10\log_{10} \eta; \quad \xi(dBm) = 30 + 10\log_{10} \frac{P_t L_0}{P_t L_0} \frac{d}{d_0}^{-n},$$  \hspace{1cm} (11)

where dBm represents the dB value normalized to 1 milliwatt, $\eta(dBm)$ is the required received power level, and $\xi(dBm)$ is the mean received power without shadowing effect. Accordingly, the outage probability of $P_r$ can be calculated as follows:

$$\text{Prob}\{P_r < \eta\} = \text{Prob}\{P_r(dBm) < \eta(dBm)\} = \int_{-\infty}^{\eta(dBm)} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(z - \xi(dBm))^2}{2\sigma^2}\right\} dz = Q\left(\frac{\xi(dBm) - \eta(dBm)}{\sigma}\right),$$  \hspace{1cm} (12)

$$Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz.$$  \hspace{1cm} (13)

For the outage probability requirement $\theta$, i.e. $\text{Prob}\{P_r < \eta\} \leq \theta$, it follows that

$$Q^{-1}(\theta) \leq \frac{\xi(dBm) - \eta(dBm)}{\sigma}. $$  \hspace{1cm} (14)

Thus, substituting (12) and (13) into (14), we have

$$P_t \geq \eta L_0 \left(\frac{5\sigma^{10}}{\sqrt{K d_0}}\right)^n \left(10^{Q^{-1}(\theta)\sigma/10}\right)^{10^{-1}(\theta)\sigma/10}. $$  \hspace{1cm} (15)

Let $P_t = \eta L_0 \left(\frac{5\sigma^{10}}{\sqrt{K d_0}}\right)^n \left(10^{Q^{-1}(\theta)\sigma/10}\right)^{10^{-1}(\theta)\sigma/10}$ in (17) and substitute (7) into (3). By taking the differential against $K$, for outage probability $\theta$, we can express the optimal number of clusters subject to shadowing as

$$K_i = \left\{ \begin{array}{ll} \left(\frac{\gamma}{2P_e}\right)^{\frac{2}{n - 2}} & n > 2; \\
1 & n = 2, \end{array} \right. \hspace{1cm} (18)$$

where $\gamma = (\frac{n}{2} - 1) \eta L_0 \left(\frac{5\sigma^{10}}{\sqrt{K d_0}}\right)^n \left(10^{Q^{-1}(\theta)\sigma/10}\right)$.

In the above discussion, we consider the outdoor propagation model validated in [36]. In various environments [38]–[42], $K_i$ can also be evaluated by the similar approaches.

IV. PHY/MAC/NET LAYER ASPECT

Based on the PHY/MAC energy consumption model in the previous section, we now further include the effects of channel contention in the MAC layer and routing schemes in the network layer.

A. Impact of Multihop

The multihop delivery rely on number of middle cluster representatives to forward data to the destination. If a sensor node is far away from the sink, the multihop delivery becomes inevitable. The multihop routing scheme consumes less energy in each link than the one-hop delivery [43], but the total energy may also increase due to more relay links.

Assume that sensor $i$ has a $h_i$-hop routing path to the sink: $i \rightarrow \gamma_1 \rightarrow \gamma_2 \rightarrow \cdots \rightarrow \gamma_{h_i-1} \rightarrow sink$. In this routing path, the energy consumption of link $i \rightarrow j$ in each transmission is $\frac{E_{ij}}{P_t}$ because $a_{ij} - j$ is the times of link $i \rightarrow j$ being established and $E_{i\rightarrow j}$ is the sum of the energy consumption of link $i \rightarrow j$ during each round. Denote $E(i, sink)$ as the sum of the energy consumption from all the sensor nodes in this routing path. Then, we have

$$E(i, sink) = a_{i\rightarrow \gamma_1} \cdot \left\{ \frac{E_{i\rightarrow \gamma_1}}{a_{\gamma_1\rightarrow \gamma_2}} + \frac{E_{\gamma_1\rightarrow \gamma_2}}{a_{\gamma_1\rightarrow \gamma_2}} + \cdots + \frac{E_{\gamma_{h_i-2}\rightarrow \gamma_{h_i-1}}}{a_{\gamma_{h_i-2}\rightarrow \gamma_{h_i-1}}} + a_{\gamma_{h_i-1}\rightarrow sink} \right\}. $$  \hspace{1cm} (19)

Usually the energy consumption of receiving a packet at the sink can be ignored. Hence, we have $E_{\gamma_{h_i-1}\rightarrow sink} = a_{\gamma_{h_i-1}\rightarrow sink} \cdot (P_e + P_t)a_{i\rightarrow \gamma_1}$. Next, substituting (2) into (19), we can obtain

$$E(i, sink) = a_{i\rightarrow \gamma_1} \cdot [(h_i - 1)P_e + h_i(P_e + P_t)]t. $$  \hspace{1cm} (20)

Averaging over $N$ routing paths, the average energy consumption for each route can be expressed as

$$\bar{E}(i, sink) = \frac{1}{N} \sum_{i=1}^{N} E(i, sink) = \frac{1}{N} \sum_{i=1}^{N} a_{i\rightarrow \gamma_1} \cdot \left\{ \frac{\sum_{i=1}^{N} (h_i - 1)}{N} + (P_e + P_t) \cdot \frac{\sum_{i=1}^{N} h_i}{N} \right\}t. $$  \hspace{1cm} (21)

Recall $\overline{a_{i\rightarrow j}} = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} a_{i\rightarrow j} / (N(N-1))$ in (3). If the first relay node in the routing path from node $i$ to the sink is determined, sensor $i$ transmits to $\gamma_1$ only and $a_{i\rightarrow j} = 0$ for all $j \neq \gamma_1$. In this case, $\overline{a_{i\rightarrow j}}$ is simplified
Fig. 3. A sensor network with 64 clusters where the sink is at the center of the entire sensor network and the number in each cluster indicates the required hop counts from the associated cluster to the sink.

<table>
<thead>
<tr>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Fig. 4. The top-left portion of the considered sensor network in Fig 3, where $|V(j)|$ is proportional to the length of DE.

**Lemma 1:**

$$|V(h)| = \begin{cases} 4 \cdot h & , \quad 1 \leq h \leq \frac{\sqrt{K}M}{2} \\ 4 \cdot (\sqrt{K}M - h) & , \quad \frac{\sqrt{K}M}{2} \leq h \leq \sqrt{K}M - 1 \end{cases}$$

**Proof:** From Fig. 3, the pattern of the required hop counts in each cluster is symmetric around the center. Thus, we can consider only the top-left portion, as shown in Fig. 4. Consider the case of $1 \leq h \leq \frac{\sqrt{K}M}{2}$. Suppose that the set of the clusters associated with BC is $V'(1)$ and the set of those clusters associated with DE is $V'(j)$, respectively. Thus, we have

$$\frac{|V'(j)|}{|V'(1)|} = \frac{DE}{BC}.$$ 

Clearly, $|V'(1)| = 1$. Thus,

$$|V'(j)| = j.$$ 

Extending to include all four quadrants, we obtain $|V(h)| = 4h$ when $1 \leq h \leq \frac{\sqrt{K}M}{2}$.

When $\frac{\sqrt{K}M}{2} \leq h \leq \sqrt{K}M - 1$, one can also prove that $|V(h)| = 4(\sqrt{K}M - 1)$.

**Theorem 1:** The average hop counts $\bar{h}$ from a cluster representative to the sink can be calculated as

$$\bar{h} = \begin{cases} \frac{\sqrt{K}M}{2} , \quad \text{when } \sqrt{K}M \text{ is an even integer} \\ \frac{\sqrt{K}M}{2} (KM - 1) + 1 \over KM , \quad \text{otherwise} \end{cases}$$

**Proof:** Recall that the total number of cluster representatives is equal to the total number of clusters, i.e., KM. Thus, in an area with KM clusters, the largest hop count
is equal to $\sqrt{KM} - 1$ as observed from Fig. 3. When $\sqrt{KM}$ is an even integer, we have

$$\sum_{h=1}^{\sqrt{KM}-1} |V(h)| = \sqrt{KM} .$$

In Appendix I, we derive the total hop counts among $KM$ clusters as follows:

$$\sum_{h=1}^{\sqrt{KM}-1} h \cdot |V(h)| = \frac{(\sqrt{KM})^3}{2} .$$

Combining (27) and (28), we can express the average hop counts for each cluster representative to the sink as

$$\bar{h} = \frac{\sqrt{KM} \cdot (KM - 1) + 1}{KM} .$$

C. Impact of Channel Contention in the MAC Layer

In this section, we extend (22) to incorporate the effect of channel contention at the MAC layer. Denote $\tau$ as the average retries for transmission in the contention procedures before successfully acquiring a channel. Then, (22) can be extended as

$$E(i, \text{sink}) = (N - 1)\bar{\pi}_{i \rightarrow j}(\tau + 1) \cdot [P_c \cdot (\bar{h} - 1) + (P_c + P_t) \cdot \bar{h} t] .$$

The MAC parameter $\bar{\tau}$ in (31) is related to the channel quality, the distribution of sensor nodes, the number of clusters, and the handshaking scheme. In this paper, the carrier sense multiple access with collision avoidance (CSMA/CA) MAC protocol is employed under an error-free channel. Based on the CSMA/CA protocol of the IEEE 802.11 standard, we derive the value of $\tau$ as follows.

Let $W$ be the minimum window size. For the maximum backoff stage $m$, the maximum window size is equal to $2^m W$. Denote $p_c$ as the collision probability when a packet is transmitted on the channel. Assume that each cluster representative always has data to send. According to [44], [45] in the context of the IEEE 802.11 MAC protocol, the probability that a cluster representative transmits data in a randomly chosen slot time can be expressed as:

$$\tau = \frac{2(1 - 2p_c)}{(1 - 2p_c)(W + 1) + p_c W (1 - (2p_c)^m)} .$$

If there are $q - 1$ other cluster representatives within the coverage area of a cluster representative, the collision probability $p_c$ for a particular cluster representative can be written as:

$$p_c = 1 - (1 - \tau)^{q-1} .$$

The value of $q$ can be estimated from the number of sensor nodes covered within the transmission range ($d$) of a cluster representative. Recall that $d = \sqrt{\frac{2}{K}d_{OA}}$ in Section III-B. Consider the example as shown in Fig. 5, where the cluster representative node $i$ located at $(x, y)$ is inside the center of the square-shaped cluster. There are 25 clusters in total within the transmission area of the cluster representative node $i$. We denote $A_j$ as the coverage area of the $j^{th}$ cluster, and $B_j$ as the area in the $j^{th}$ cluster interfered by the transmission of the cluster representative node $i$. Within $A_j$, only the cluster representative node in $B_j$ will be interfered with the cluster representative node $i$. Let $f_X(x)$ and $f_Y(y)$ be the probability density functions of the sensors’ locations in the $x$-axis and the $y$-axis, respectively. The total average number of contending nodes in 25 clusters can be approximated as follows:

$$q = \sum_{j=1}^{25} \frac{\int \int_{B_j} f_X(x) f_Y(y) dxdy}{\int \int_{A_j} f_X(x) f_Y(y) dxdy} .$$

In the special case when the distribution of sensor nodes’ locations ($f_X(x)$ and $f_Y(y)$) have a self-similar property [46] (e.g., uniform distribution), it follows that

$$q = 25 \cdot \frac{(\sqrt{\frac{2}{K}d_{OA}})^2}{(\sqrt{\frac{2}{K}d_{OA}})^2} = 5\pi .$$

Finally, the average retransmission times in the MAC layer can be obtained by substituting (33) into following
equation:
$$\tau = \sum_{r=0}^{\infty} r \cdot (1 - p_c) \cdot (p_c)^r.$$  
(36)

Note that $\tau$ and $p_c$ can be solved iteratively from (32) and (33).

In a more general case when the location of sensors are not self-similar, $q$ can be the function of $K$. In this case, $\tau$ is also a function of $K$ as described in the constraint in (1). Moreover, if the non-saturated traffic is considered, $\tau$ and $p_c$ must be re-evaluated according to the results in the current literature [47]–[51]. Hence, a new average retransmission times $\tau$ can be derived from (36).

D. Optimal Number of Clusters ($K_p$)

Without loss of generality, we consider the case that $\sqrt{KM}$ is an even integer. Substituting (6), (7), and (29) into (22), we have

$$E(i, \text{sink}) = \frac{MK}{N} R \cdot (\tau + 1) \left[ P_e \cdot \left( \frac{\sqrt{KM}}{2} - 1 \right) + \left( P_e + \eta L_0 \left( \frac{\sqrt{5d_{OA}}}{\sqrt{K}d_0} \right)^n \right) \cdot \frac{\sqrt{KM}}{2} \right] t,$$

(37)

where $\tau$ is defined in (36). Taking differential with respective to $K$ in (37), we can get

$$\frac{\partial E(i, \text{sink})}{\partial K} = \frac{M}{N} (\tau + 1) R \cdot \left[ \left( \frac{3\sqrt{KM}}{2} - 1 \right) P_e + \left( \frac{3 - n}{4} \right) \eta L_0 \left( \frac{\sqrt{5d_{OA}}}{\sqrt{K}d_0} \right)^n \right] t.$$  
(38)

Because $\frac{\partial^2 E(i, \text{sink})}{\partial K^2} > 0$, (37) is a convex function. Hence, we can obtain the optimal number of clusters in a basic observation area ($K_p$) by finding the root of $\frac{\partial E(i, \text{sink})}{\partial K} = 0$ in (38).

V. Numerical Results

In this section, we present numerical results to illustrate the relation of the optimal number of clusters and the related parameters in the PHY, MAC, and NET layers. Furthermore, the analytical results are validated through some simulation tests.

A. Simulation Setup

In the experiment, we consider a 16 meters $\times$ 16 meters sensor network with 16 basic observation areas ($M = 16$). The sensor nodes are uniformly distributed with four kinds of density as shown in Fig. 6. In this figure, the density of sensor nodes are 90, 270, 450, and 630 nodes per observation area in regions 1, 2, 3, and 4, respectively. Hence, we have $N = 5760$. Furthermore, we first assume only one set of path loss parameters is needed to obtain the relation between the energy consumption and the number of clusters $K$. We will discuss the impact of the two-slope path loss model in Section V-D. Referring to [52], we adopt the following system parameters: $R = 5.76 \times 10^6$, $t = 10$ ms, $P_e = 5$ mW, $\eta = 12.43$ pW, $d_0 = 0.2$ meters, and $L_0 = 52$ dB in (8) and (37).

B. PHY/MAC Layer

Figure 7 illustrates the energy consumption per sensor node against the number of clusters per basic observation area according to (8) and simulation results. As shown, the proposed analytical model can match the simulation results quite well. Furthermore, the optimal number of clusters (denoted by circle in figure) in a basic observation area are $K_1 = 5$, 30, and 81 for $n = 3$, 4, and 5, respectively. Thus, a sensor network in an environment with a smaller path loss exponent prefers less clusters in a basic observation area. Second, when $K < K_1$, $E_{\text{link}}$ decreases as $K$ increases because a smaller $K$ makes the distance between the adjacent cluster representatives...
longer. For \( n = 4 \), the energy consumption per sensor node for the optimal \( K_1 = 30 \) is reduced to 16.4 mJ compared to 242.6 mJ for \( K = 1 \). In this example, the optimal \( K_1 \) can reduce energy consumption by 93%. Last, when \( K > K_1 \), the energy consumption of the sensor node is proportional to \( K \) because the number of times being the cluster representative increases. Compared to \( K = 60 \), the optimal \( K_1 = 30 \) can decrease the energy consumption from 20.7 mJ to 16.4 mJ, i.e., an energy reduction 20%.

Figure 8 shows the impact of the edge length of the basic observation area (\( d_{OA} \)) on the optimal number of clusters \( K_1 \) with \( \sigma = 2 \) and 4 dB. In this example, \( \theta = 0.1 \). We find that as \( \sigma \) increases, \( K_1 \) increases. To overcome more serious shadowing, a shorter transmission distance is preferred, thereby requiring more clusters. Furthermore, the network with fewer basic observation areas (or larger \( d_{OA} \)) prefers to have more clusters in a basic observation area because more clusters can reduce the transmission distance. Finally, the impact of the shadowing effect and \( d_{OA} \) on \( K_1 \) is more significant for a larger \( n \) than that for a smaller \( n \). One can explain this phenomenon from (18). A larger \( n \) may amplify the shadowing effect on the optimal number of clusters.

Figure 9 shows the percentage of energy reduction of the system using the optimal \( K_1 \) compared to the system using \( K = 1 \) for different shadowing standard deviations. As shown in the figure, the energy saving percentage \((1 - \frac{E_{path}(K=1)}{E_{path}(K=K_1)})\) increases as the path loss exponent \( n \) and the shadowing standard deviation (\( \sigma \)) increases. For \( \sigma = 3 \) dB, the optimal \( K_1 \) can provide energy savings of 48%, 89%, and 96% for path loss exponents \( n = 3, 3.5, \) and 4, respectively.

C. PHY/MAC/NET Layer

Figure 10 depicts the relation between the number of clusters (\( K \)) and the total energy consumption of all the sensor nodes in the routing path from a PHY/MAC/NET cross-layer perspective. From this figure, the proposed analytical model can also match the simulation results quite well. Furthermore, when the cluster representatives share the channel without contention, (i.e., \( \tau = 0 \)), the optimal number of clusters (\( K_p \)) in a basic observation area from the PHY/MAC/NET cross-layer perspectives are 1, 17, and 59 for \( n = 3, 4, \) and 5, respectively. Comparing Figs. 7 and 10, we notice that \( K_p < K_1 \). This is because \( K_p \) considers all the PHY/MAC/NET effects. Specifically, fewer clusters yield fewer hop counts to the data sink, and result in less energy consumption in the NET layer aspect. Moreover, for a channel with higher collision probability, e.g., \( \tau = 0.6545 \), the total energy consumption increases.

D. Impact of Two-Slope Path Loss Models

Now, we further relax the assumption in Figs. 7 ~ 10 when only one set of path loss parameters is needed to obtain the relation between the energy consumption and number of clusters \( K \). Now we consider when two sets of path loss parameters are required to obtain \( K \) based on
consumption. With its flexibility, the proposed cross-layer optimal number of clusters in a basic observation area with the network layer. The closed-form expressions for the perspectives. Specifically, this cross-layer analytical model observation area from the PHY/MAC/NET cross-layer to determine the optimal number of clusters in a basic smaller path loss exponent value of $K$.

Let $Z = \sqrt{KM}$, then we have
\[
\sum_{h=1}^{Z} h \cdot v(h) = \sum_{h=1}^{2Z-1} h \cdot [h \cdot 4(2Z - h)]
\]
\[
= \sum_{h=1}^{2Z-1} h^2 - 4Z \sum_{h=1}^{Z-1} h^2 + 8Z \sum_{h=1}^{Z-1} h.
\]
\[
= 8 \sum_{h=1}^{Z-1} h^2 + 8Z \sum_{h=1}^{Z-1} h - 4 \sum_{h=1}^{2Z-1} h^2.
\]

Now we apply the formula of sum of series and can derive
\[
\sum_{h=1}^{\sqrt{KM} - 1} h \cdot v(h) = \frac{8}{6} Z(1 + Z)(2Z + 1) + 8Z[(1 + Z)(2Z - 1)\cdot (Z - 1)]
\]
\[
+ \frac{4}{6}(2Z - 1)(2Z - 1 + 1)[2(2Z - 1) + 1]
\]
\[
= 4Z^3.
\]

VI. Conclusions

In this paper, we presented an analytical approach to determine the optimal number of clusters in a basic observation area from the PHY/MAC/NET cross-layer perspectives. Specifically, this cross-layer analytical model integrates the effects of the transmission distance, power, and shadowing in the physical layer, the possibility of being a cluster representative and the retransmission times in the MAC layer, as well as the number of hops in the network layer. The closed-form expressions for the optimal number of clusters in a basic observation area with various shadowing and path loss conditions are presented. We demonstrate that the suggested analytical optimal number of clusters can significantly improve the energy consumption. With its flexibility, the proposed cross-layer analytical model can facilitate the design of the optimal number of clusters for a sensor network in different radio environments. Furthermore, our simulation results show the existence of the optimal cluster number regardless of different density of sensors in various observation areas.

One interesting research topic that can be extended from this work is to apply the similar methodology to determine the optimal number of clusters based on other MAC and routing strategies with different values of $\tau$ and $\tilde{n}$. Furthermore, it is also worthwhile to investigate how to determine the optimal number of clusters when data aggregation and random traffic model in upper protocol layers are considered.

Appendix I

Proof of (28) in Theorem 1

Let $Z = \sqrt{KM}$, then we have
\[
\sum_{h=1}^{\sqrt{KM} - 1} h \cdot v(h) = \frac{8}{6} Z(1 + Z)(2Z + 1)
\]
\[
+ 8Z[(1 + Z)(2Z - 1)\cdot (Z - 1)]
\]
\[
+ \frac{4}{6}(2Z - 1)(2Z - 1 + 1)[2(2Z - 1) + 1]
\]
\[
= 4Z^3.
\]

REFERENCES

Li-Chun Wang (S’92-M’96-SM’06) received the B.S. degree in electrical engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 1986, the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1988, and the M.Sc. and Ph.D. degrees in electrical engineering from Georgia Institute of Technology, Atlanta, in 1995 and 1996, respectively.

In 1995, he was affiliated with Northern Telecom in Richardson, Texas. From 1996 to 2000, he was with AT&T Laboratories, where he was a Senior Technical Staff Member in the Wireless Communications Research Department. Since August 2000, he has joined the Department of Communication Engineering of National Chiao-Tung University in Taiwan as an Associate Professor and has been promoted to a full professor since August 2005. Dr. Wang was a corecipient of the Jack Neubauer Best Paper Award from the IEEE Vehicular Technology Society in 1997. His current research interests are in the areas of cellular architectures, radio network resource management, cross-layer optimization for cooperative and cognitive wireless networks. He is the holder of three U.S. patents with three more pending.

Chung-Wei Wang (S’07) received the B.S. degree in electrical engineering from the Tamkang University, Taipei, Taiwan, and then started the M.S. program in the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2003. Then, he has directly worked toward the Ph.D. degree in that since 2004. His current research interests are in the areas of cross-layer optimization and MAC protocols design in wireless sensor networks, ad hoc networks, and cognitive radio networks.

Chuan-Ming Liu is an associate professor in the Department of Computer Science and Information Engineering, National Taipei University of Technology (NTUT), TAIWAN. He received his Ph. D. in Computer Sciences from Purdue University in 2002 and B.S. and M.S. degrees both in Applied Mathematics from National Chung-Hsing University, Taiwan, in 1992 and 1994, respectively. He is a member of Upsilon Pi Epsilon Honor Society in Computer Science since 1998. His research interests include parallel and distributed computation, data management and data dissemination in wireless environments, ad hoc and sensor networks, and analysis and design of algorithms.