

# An Adaptive Data Dissemination Strategy for Wireless Sensor Networks\*

S. SELVAKENNEDY

*School of Information Technologies, J12 University of Sydney,  
NSW, Australia*

S. SINNAPPAN

*School of Economics and Information Systems, University of Wollongong,  
Wollongong, NSW, Australia*

*Future large-scale sensor networks may comprise thousands of wirelessly connected sensor nodes that could provide an unimaginable opportunity to interact with physical phenomena in real time. However, the nodes are typically highly resource-constrained. Since the communication task is a significant power consumer, various attempts have been made to introduce energy-awareness at different levels within the communication stack. Clustering is one such attempt to control energy dissipation for sensor data dissemination in a multihop fashion. The Time-Controlled Clustering Algorithm (TCCA) is proposed to realize a network-wide energy reduction. A realistic energy dissipation model is derived probabilistically to quantify the sensor network's energy consumption using the proposed clustering algorithm. A discrete-event simulator is developed to verify the mathematical model and to further investigate TCCA in other scenarios. The simulator is also extended to include the rest of the communication stack to allow a comprehensive evaluation of the proposed algorithm.*

**Keywords** Data Dissemination; Clustering; Routing; Energy-Efficiency; Lifetime; Simulation

## 1. Introduction

Even though physical sensing capability is a rather old technology, there is a renewed interest in it due to the recent technological advances in micro-electromechanical systems and the integration of wireless communication capability. The realization of wireless sensor networks (WSNs) has opened a whole set of new application domains, which were not possible with the traditional wired “dumb” sensors. New application areas such as forest fire detection, sophisticated structural monitoring, earthquake monitoring, and battlefield surveillance are beginning to benefit from this technology [1]. As many of these applications require non-intrusive participation of sensor nodes, they have to be small in size while remaining functional for a rather long period, possibly a couple of years. As these nodes are equipped with severely limited battery sources, this poses significant challenges in the design of network and communication stack to realize an energy efficient solution.

\*An earlier version of this paper appeared in HWISE-2005, co-located with ICPADS 2005.

Address correspondence to S. Selvakennedy, School of Information Technologies, J12 Madsen Bldg. F09, University of Sydney, NSW 2006, Australia. E-mail: skennedy@it.usyd.edu.au

Most possible WSN applications could either be classified as a periodic monitoring or event-triggered type application. Even the query-based applications such as TinyDB [2] and Cougar [3], may fall under one of the these types as they dynamically configure the network. In the periodic monitoring type applications, the task of data dissemination that happens at predefined intervals towards the network sink has to be managed efficiently to prolong network lifetime. In many cases, the sink node would not be in the radio range of most source nodes, therefore making direct transmission infeasible. Organizing these nodes into clusters is one of the popular mechanisms used to control data dissemination, as it could improve the scalability of a network [4]. The clustering process divides the network into subsets of nodes (i.e. a cluster) with each cluster being managed by a cluster-head (CH). Due to the spatial proximity of the nodes within a cluster, upon receiving sensory data from them, a CH could even aggregate them with minimal information loss as these readings may exhibit significant correlations. As such, large-scale multihop networks become feasible due to this trade-off [5]. For WSNs with a large number of energy-constrained nodes, it is crucial to design a fast distributed algorithm to organize the nodes into clusters.

Various clustering algorithms in different contexts have been proposed in the literature. Most algorithms are heuristic in nature and aim at generating the minimum number of clusters and transmission distance. These algorithms also distinguish themselves by how the CHs are elected. The LEACH algorithm [6] and its related extensions [7] use probabilistic self-election, where each sensor node has a probability  $p$  of becoming a CH in each round of monitoring. It guarantees that every node will be a CH only once in  $1/p$  rounds. This rotation of the energy-intensive CH function aims to distribute the power usage for prolonged network life. However, LEACH allows only one-hop clusters and assumes that all nodes are able to reach the sink directly. Another clustering algorithm proposed in [8] aims to maximize the network lifetime, but it assumes that the nodes are aware of the entire network topology. This assumption, however, may not be reasonable in many scenarios. Some of these algorithms were designed to generate stable clusters in environments with mobile nodes. In a typical WSN, the sensor nodes are stationary and the instability of clusters due to mobility of these nodes may not be an issue.

Bandyopadhyay et al. proposed a hierarchical multihop clustering algorithm, and then derived simplified formulas for computing the optimal number of clusters and number of hops  $k$  using results in stochastic geometry to minimize the total energy spent [9]. It was assumed that the sink is situated at the network center, and the nodes have unit energy consumption for each communication task. In [10], the authors presented a Hybrid Energy-Efficient Distributed clustering (HEED) protocol that periodically selects CHs according to some primary and secondary parameters, for example a node's residual energy and a node degree (or its proximity to neighbors), respectively. It exploited the availability of multiple power levels at a node such as on the Berkeley motes [11]. It was proven that this clustering process terminates in constant time, and achieves fairly uniform CH distribution across the network. However, HEED requires a number of parameters to be specified such as intra- and inter-cluster transmission power levels to ensure connectivity among the CHs. The configuration of these parameters requires the knowledge of the entire network. Also, this protocol only allows one-hop clusters. A fixed clustering algorithm that also involves an optimal power allocation was introduced in [12]. To our knowledge, this is the first paper to consider both clustering and its influence on the routing protocol in an integrated manner.

In this paper, we introduce the Time-Controlled Clustering Algorithm (TCCA) that allows multihop clusters, and uses message timestamp and time-to-live (TTL) to control the cluster formation. During the CH election, a node could also consider its residual

energy before volunteering. A probabilistic model to quantify its energy usage is then given, derived using a realistic first-order radio energy dissipation model with the objective of minimizing the energy spent in data dissemination. A discrete-event simulator is also developed to verify the mathematical results as well as to investigate the influence of sink placement on this algorithm. In order to perform a comprehensive evaluation of this algorithm, the simulator is further extended to incorporate the overhead of the rest of the communication stack to present an accurate energy usage pattern in an integrated study.

The rest of the paper is organized as follows. Section 2 presents the perspective of this area of research. Various clustering algorithms proposed in literature are discussed there. In Section 3, the details of the new clustering algorithm are described. Subsequently, its energy dissipation is mathematically derived. The simulator used to verify this mathematical model is described in Section 4. Section 5 presents and analyzes various experiments and their corresponding results for different scenarios. The main findings of this work are summarized in the final section with some directions for further work.

## 2. Related Work

Intense research in the field of sensor network technology in recent years has prompted further development in micro-sensor technology and low-power analog/digital electronics. Numerous issues have been addressed to tackle the challenge of sensor energy conservation. Some of the main issues are low-power signal processing architectures, low-power sensing interfaces, energy efficient wireless media access control (MAC), routing protocols, low-power security protocols, and essential key management architectures and localization systems. In order to address the energy-efficient data dissemination need, various clustering algorithms have been proposed in different context. Initially, these algorithms focused on the connectivity problem [13–15] but later energy-efficiency was more of interest in wireless ad hoc and sensor networks [6, 8, 9, 16–18]. However, almost all focus on reducing the number of clusters formed, which may not necessarily entail minimum energy dissipation.

Generally, clustering algorithms segment a network into non-overlapping clusters comprising a CH each. Non-CHs transmit sensed data to CHs, where the sensed data could be aggregated and transmitted to the sink. Clustering algorithms maybe distinguished by the way the CHs are elected. The Linked Cluster Algorithm (LCA) [15] selects the CH based on the highest id among all nodes within one-hop. This was enhanced by LCA2 [19] that selects the node with the lowest id among all nodes that is neither a CH nor is one-hop of the previously selected CHs. In [20], the authors developed a similar distributed algorithm to LCA2, which identifies the CH by choosing the node with the highest degree. Other algorithms such as the Distributed Cluster Algorithm (DCA) [21] and Weighted Clustering Algorithm (WCA) [22] rely on weights to select CHs.

Load balancing heuristics were proposed in [17] and [23]. In [23], the proposed clustering algorithm was designed to ensure that each cluster has an equal number of nodes while keeping minimal distance between the nodes and their CH. However, most of the above algorithms have a restricted scope of application and are suitable for only a small number of nodes. They generate only one-hop clusters and require synchronized clocks, which makes them less favorable for practical usage. Moreover, the “load-balanced” algorithms focus mainly on balancing the intra-cluster traffic load without due consideration to the external traffic. The max-min  $d$ -cluster algorithm was proposed to achieve better load balancing among CHs as well as to reduce the number of CHs as compared to LCA or LCA2 [18]. It generates  $d$ -hop clusters with a run-time of  $O(d)$  rounds. In [8], clustering algorithms were proposed to maximize the network lifetime by varying the cluster size

and the duration of a node being nominated as a CH based on the assumption that the locations are known a priori. These algorithms ignore inter-cluster traffic and also need the recognition of the whole network topology, which may not be feasible in most cases.

LEACH [6] self-elects CHs using a nominated probability  $p$ . The algorithm ensures that every node will be nominated as a CH only once in  $1/p$  rounds for a certain fixed duration. Based on LEACH, the authors of [9] proposed a clustering algorithm and derived a simplified energy model for predicting the optimal  $p$  using the results in stochastic geometry. This model was also extended to optimize the cluster radius as well as to support a hierarchical sensor network. Hybrid Indirect Transmission (HIT) [24], a combination of LEACH and Power Efficient Gathering in Sensor Information Systems (PEGASIS) [25], was proposed to allow simultaneous transmissions both among clusters and within a cluster, without requiring any position knowledge. HIT also looks at the transmission delay issue with its chaining ability.

In [26], the authors proposed a scheme that incorporates the effect of inter-cluster traffic and the sink location to maximize the network topological lifetime. However, energy-load balancing was not considered. In [12], a fixed clustering algorithm that performs energy-balancing to improve network lifetime was proposed. It also takes into consideration the interaction between clustering and routing. Two schemes were introduced. The first scheme computes the optimal cluster size and the CH locations, whereas the second allows a CH to probabilistically choose to either relay the traffic to the next-hop or to deliver it directly to the sink. It assumes a heterogeneous network, where the CH nodes have bigger resources than the others.

In this paper, we expanded the work proposed in [6] and [9]. In [9], a node self-elects with probability  $p$  and advertises itself as a CH to the nodes within its radio range. This information is also passed on to  $k$ -hop neighbors. Such nodes are known as *volunteer* CH. Those nodes that do not receive any advertisements within a time duration of  $t$  become the *forced* CH. The duration is in the order of the time needed for the data to be sent by the leaf nodes from  $k$ -hop away to their CH. The non-CHs join the closest CH forming a Voronoi tessellation. The CH aggregates data from its cluster members and communicates the information to the sink. Based on these assumptions, the authors derived the optimal  $p$  as well as  $k$  values that minimize the total energy spent. However, a simplified unit energy model with the same dissipation behavior for the radio's transmitter and receiver was assumed. This may not present an accurate energy usage pattern, and it was reported in [27] that the transmitter consumes almost 2.5 times more energy than the receiver. In this work, the Time-Controlled Clustering Algorithm (TCCA) is introduced, where the cluster formation is controlled through the inclusion of timestamp and TTL fields in its messages. Furthermore, its energy dissipation is analytically derived using a realistic radio model. The resultant expression is further used to derive the network lifetime metric.

### 3. The TCCA Algorithm

The operation of TCCA is divided into rounds to enable load distribution among the nodes, similar to the LEACH algorithm. Each of these rounds comprises a cluster setup phase and a steady-state phase. During the setup phase, CHs are elected and the clusters are formed. During the steady-state phase, the cycle of periodic data dissemination that involves data collection, aggregation and transfer to the sink occurs.

In order to determine the eligibility to be a CH, each node  $i$  generates a random number between 0 and 1. If the number is less than a variable threshold  $T(i)$ , the node becomes a CH for the current round  $r$ . Besides, a node's residual energy  $E_{residual}$  could also be taken into consideration. Thus, the threshold is computed as follows:

$$T(i) = \max \left( \frac{p}{1-p \left( r \bmod \frac{1}{p} \right)} \times \frac{E_{residual}}{E_{max}}, T_{min} \right) \quad \forall i \in G$$

$$T(i) = 0 \quad \forall i \notin G \quad (1)$$

Where  $p$  is the desired CH probability,  $E_{max}$  is a reference maximum energy,  $T_{min}$  is a minimum threshold (to avoid a very unlikely possibility when  $E_{residual}$  is small), and  $G$  is the set of nodes that have not been CHs in the last  $1/p$  rounds. When a CH has been self-elected, it advertizes itself as the CH to the neighboring nodes within its radio range. This advertisement message (ADV) carries its node id, an initial TTL, the node's residual energy and a timestamp. Upon receiving, first-hop nodes forward the ADV message further as governed by its TTL value. The selection of the TTL value may be based on the current energy level of the CH and could be used to limit the diameter of the cluster to be formed. However, in this work, we assume that all nodes use the same fixed  $k$  value to simplify our mathematical derivation.

Any node that receives such an ADV message and is not a CH itself joins the cluster of the *nearest* CH from the TTL value. If there is a tie, this node could select the CH with higher residual energy. Once a node decides to be part of a cluster, it informs the corresponding CH by generating a join-request message (JOIN-REQ) consisting of the node's id, the CH's id, the original ADV timestamp and the remaining TTL value. The timestamp is included to assist the CH in approximating the relative distance of its members. Together with TTL, the CH could form a multihop view of its cluster, which could be used to create a collision-free transmission schedule in a time-division multiplexed scheme as in [6]. A transmission schedule is created by the CH based on its number of members and their relative distance to enable the reception of all sensory data in a collision free manner. At the end of the schedule, the CH communicates the aggregated information to the sink. For the nodes that do not receive any ADV message, they are forced to route their data, using the routing protocol, to the sink. Their dissemination process does not include any aggregation task but the transfer to the sink may involve multihop transmissions. To simplify the mathematical model representation, we will neglect the *marginal* effect of the cluster setup phase in the computation of energy dissipation, as the setup phase is substantially shorter than data transfer. However, the complete overhead of this algorithm is incorporated in the simulator.

#### 4. The TCCA Energy Dissipation Model

The energy used for the data dissemination towards the sink will depend on the cluster size controlled through  $k$  (i.e. TTL) and distance between the transmitting and receiving nodes. Since the goal of our work is to organize nodes into clusters with minimal overall energy consumptions, we need to determine the optimal value of  $k$ . For the development of our model, the following assumptions are made:

- a. The nodes are randomly scattered in a two-dimensional plane and follow a homogeneous spatial Poisson process with  $\lambda$  intensity.
- b. All nodes in the network are homogeneous. They transmit at the same power level and hence have the same radio range  $r$ . Each transmission follows the isotropic propagation model.
- c. The sink is located at the center of the field. (For other sink locations, we will rely on the simulator to investigate its effect.)
- d. Intermediate nodes forward data exchanged between two nodes not within each other's range.
- e. All nodes are able to reach the sink, at least in a multihop fashion. (This may not be true in simulation as there may be isolated nodes during random deployment.)
- f. The energy needed for the transmission of one bit of data from node  $u$  to node  $v$  is the same as to transmit from  $v$  to  $u$  (i.e. only symmetric propagation channels).
- g. A routing and MAC infrastructure is in place. The link-level communication using the MAC is collision- and error-free.

The overall idea of the optimal value derivation is to define a function for the energy used to disseminate information to the sink during the steady-state phase.

As per the assumptions, the nodes are distributed according to a homogeneous spatial Poisson process. The number of nodes in a square area of side  $M$  is a Poisson random variable,  $N$  with mean  $\lambda A$  where  $A = M \times M$ . Let us assume that for a particular realization of the process, there are  $n$  sensors in this area. The probability of becoming a CH is  $p$ . On average, there will be  $np$  nodes becoming CHs.

Now, to derive the energy usage, we need to approximate the energy required to drive the transceiver. For the transmit energy estimation, the free space ( $d^2$  power loss) channel model is used [6]. Power control is used to invert this loss by suitably configuring the power amplifier. Thus, to transmit an  $l$ -bit packet a distance  $d$ , the radio expends:

$$E_{Tx} = lE_{\text{elec}} + l\epsilon_{\text{fs}}d^2 \quad (2)$$

Where  $E_{\text{elec}}$  is the electronic energy that depends on factors like digital coding, modulation, filtering, and spreading of the signal, and  $\epsilon_{\text{fs}}d^2$  is the amplifier energy that depends on the distance to the receiver and the acceptable bit-error rate. As to receive a packet, the radio expends:

$$E_{Rx} = lE_{\text{elec}} \quad (3)$$

To estimate the energy consumption, we need to compute the average energy dissipation per cluster and multiply against the average number of clusters. If we assume the maximum number of hops is  $k$ , the average hop for a CH to reach each of its members is  $k/2$ . Any communication between a CH and its member not in direct radio range requires multihop transmission with intermediate nodes acting as the relay nodes. Thus, each non-CH node dissipates energy not only due to the transmission of its own message, but mainly due to its relay function, except for the leaf nodes. To estimate the average number of nodes at a certain hop from the CH, we represent a cluster as concentric circles with radius as multiple of  $r$  (i.e.  $r, 2r, 3r$  etc.). For example, to obtain the average number of

nodes at  $i$ -hop from the CH ( $s_i$ ), we simply multiply the area difference between the circle of  $ir$  and  $(i-1)r$  radius (i.e. an annulus area) and the mean node density,  $\lambda$ :

$$s_i = (2i - 1)\pi r^2 \lambda \quad (4)$$

Each upstream node towards the CH has to transmit its message as well as to route messages from all its downstream children as part of the cluster. Its average number of downstream nodes ( $c_i$ ) is given by the sum of the ratio of number of nodes in level- $(i + 1)$  and level- $i$  repeated till  $k$ -hop:

$$c_i = \sum_{j=i}^{k-1} \frac{2k - (2j - 1)}{2k(2j + 1)} \quad (5)$$

Thus, the total energy consumption by all the non-CH nodes ( $C_1$ ) is obtained by iteratively adding each hop- $i$  contribution for  $np$  clusters as follows:

$$E[C_1 | N = n] = np \sum_{h=1}^{k-1} s_h [c_h E_{Rx} + (c_h + 1)E_{Tx}] \quad (6)$$

As for the CH energy usage computation, we need to include its average number of members, the message aggregation cost ( $E_{DA}$ ), and its communication (possibly multihop) cost to the sink. Since there are on average  $np$  CHs and the location of any CH is independent of the locations of the other CHs, the total length of the segments from all these CHs to the sink is  $\frac{0.765npM}{2}$  [9]. Thus, the average number of hops from a CH to the

sink is  $\left\lceil \frac{0.765M}{2r} \right\rceil$ . The overall energy consumption of  $np$  CH nodes ( $C_2$ ) could then be approximated as:

$$E[C_2 | N = n] = np \times \left\{ \left[ \pi (kr)^2 \lambda - 1 \right] E_{Rx} + E_{DA} + \left\lceil \frac{0.765M}{2r} \right\rceil (E_{Tx} + E_{Rx}) \right\} \quad (7)$$

It is possible that some nodes may not receive any ADV messages at all for some random topologies but could still reach the sink (c.f. our model assumptions). As such, these nodes will need to transmit their own message to the sink given by  $C_3$  as follows:

$$E[C_3 | N = n] = n(1 - p\pi(kr)^2 \lambda) \times \left\{ \left\lceil \frac{0.765M}{2r} \right\rceil (E_{Tx} + E_{Rx}) \right\} \quad (8)$$

Therefore, the total energy consumption (C) for each *round* of sensing and transfer is:

$$E[C|N = n] = E[C_1|N = n] + E[C_2|N = n] + E[C_3|N = n] \quad (9)$$

Removing the conditioning on  $N$  yields:

$$\begin{aligned} E[C] &= E[E[C|N = n]] \\ &= E[N] \times p(E[C_1] + E[C_2] + E[C_3]) \\ &= \lambda A p(E[C_1] + E[C_2] + E[C_3]) \end{aligned} \quad (10)$$

It is difficult to simplify  $E[C]$  further to determine the optimal cluster size  $k$  analytically. However, it is amenable to numerical evaluation for the computation of the total energy dissipation for various cluster size.

Another crucial metric of a sensor network is the *system lifetime*. Here, lifetime is defined as the time period from the instant the network is deployed to the moment when the first sensor node runs out of energy. Once the total energy dissipation is determined (C), we can determine the average energy dissipated per sensor in each round of transmission. Assuming each node initially has  $B$  joule of battery energy, and there is a single transmission of sensed data to the CH per round of  $t$  period, we could approximate lifetime,  $L$  in seconds, through:

$$L = \frac{B}{C/N} \times t = \frac{BNt}{C} \quad (11)$$

## 5. The Simulation Model

In order to validate the above analysis, we initially considered the use of some existing simulators. As evident in the literature, TinyOS [28] is the most widely adopted operating platform for WSNs. Thus, we naturally opted for TOSSIM [29], which is included as part of the TinyOS distribution. However, TOSSIM and its extension PowerTOSSIM [30] have a significant shortcoming. Both simulators do not readily have the notion of location and field dimension. As such, we would not be able to control the field size as well as to determine the distance to neighbors easily. Even though the message structure used by TOSSIM has a *strength* attribute to represent the received signal strength indicator (RSSI), the current implementation does not include a realistic radio propagation model to govern the attribute's behavior. As the notions of field dimension and location are crucial for our purpose, we resorted to develop our own discrete event simulator.

In the description of the simulator, we assume that each sensor node is aware of:

- i. its position, as some nodes could be equipped with a GPS receiver while the others could infer their location by means of any localization technique [31];
- ii. the position of its neighbors (even if it changes) due to the occasional beaconing by the sink and the cluster setup phase of TCCA; and
- iii. the network is synchronized (at least among the neighbors) by means of any time synchronization protocol [32].

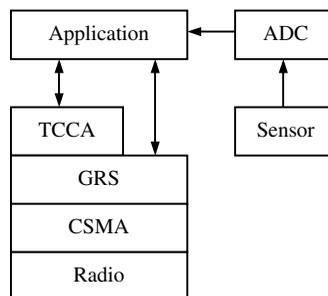
An investigation on the impact of positioning and synchronization errors is left for future work.

The simulator is developed in C++ and adopts the object-oriented approach to allow natural mapping to a real sensor network. The limitations of the above mathematical model due to its simplifying assumptions (see previous section) to ensure tractability are mostly relaxed for this simulation purpose. The radio model adopted here is however retained to represent isotropic propagation. We implemented the CSMA [33] protocol to govern the channel access. When a node becomes active after its sleep time, it listens to the channel if it needs to access. If the channel is sensed idle, it then transmits. Otherwise, it backs off randomly and repeats the above. At the network layer, we implemented a simple routing mechanism in Greedy Routing Scheme (GRS) [34]. The forwarding objective is to minimize the number of hops between the sink and the other nodes. Whenever a node is not in the overlay clustered-structure, it uses this routing protocol to get to the sink.

To establish this minimum hop routing tree, the sink periodically broadcasts a *beacon* message with a hop count field, which is initialized to zero. To overcome the problem of asymmetrical links that may be prevalent in WSNs [35], the sink broadcasts at a power level lower than the maximum level of a regular node, a move similar to He's proposal [36]. Upon receiving the beacon, each node records the sender's id, increments the hop count by one, and then rebroadcasts at a power level below its maximum level. A node only rebroadcasts if the new hop count is smaller than its stored value. Since we are focusing on a stationary type of application, the sink node only needs to perform occasional beaconing to avoid a significant overhead. This forwarding rule establishes a minimum hop tree rooted at the sink. TCCA is implemented between the application and the network layer, and thus, the overall system architecture is as shown in Fig. 1.

Since most current applications require or even demand unstructured sensor node deployment, we have adopted a random uniform distribution strategy. However, while running the simulator for smaller node density scenarios (<400 nodes per km<sup>2</sup>), we found that the above strategy makes some nodes isolated from the rest, and especially of interest is the isolation from the sink. This is mainly due to the underlying assumptions of the built-in *rand()* function. As such, a straightforward application of this function will not suffice. Instead, before the nodes' locations are generated, the field is divided into four equal quadrants. Then, an equal number of nodes are uniformly randomly distributed within each quadrant, while ensuring a node degree of at least two is obeyed to reduce the isolation probability.

Another *interesting* topological phenomenon was observed when running the simulator for certain scenarios. According to the TCCA operation, during the cluster setup phase, each regular node waits for a fixed join timer to expire to allow it to receive most ADV



**FIGURE 1** The system simulator architecture.

messages from nearby CHs before deciding to join a cluster. This flexibility occasionally causes a child and its parent nodes to join different clusters based on their local decision on the *best* CH node. We chose not to permit such phenomenon to occur as this would lead to more state maintenance at the parent nodes, and could pose a scalability problem in larger networks. When a parent receives a JOIN message for further forwarding, if it finds the child's choice of CH differs from itself, the parent modifies and reroutes this message to its CH instead. The child is also informed of this decision. Since the greedy selection of the best CH is a local decision, it may not be the globally optimal choice. Therefore, by forcing this change, the parent makes a minor trade-off between more state maintenance to the choice of a sub-optimal CH, which may be acceptable due to the severe resource constraints.

Finally, for the purpose of the mathematical result verification, we did not include the overhead of the rest of the stack as per the analysis assumptions. For all other simulation results, the complete overhead is incorporated as discussed.

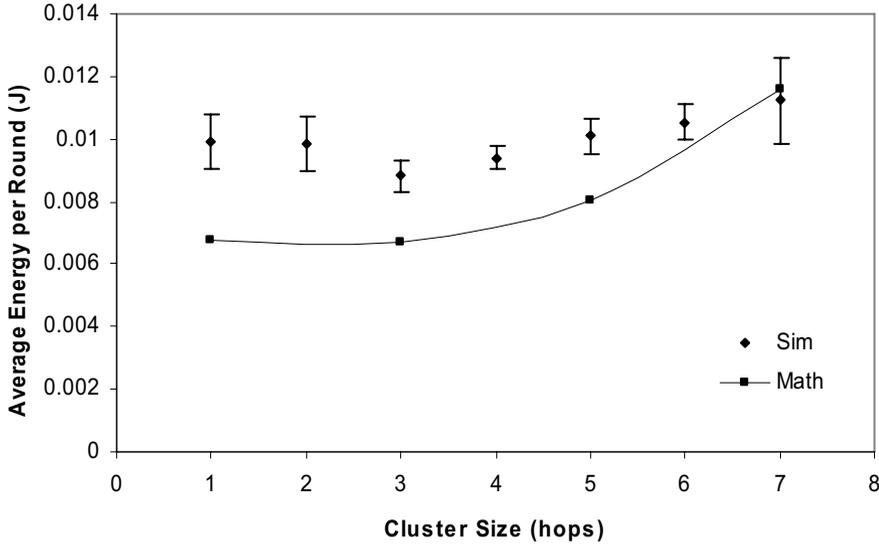
## 6. Results and Discussions

To investigate the performance behavior of TCCA, we initially used both the mathematical model and the simulator. For these experiments, we assumed that there are  $N$  sensor nodes distributed uniformly at random in an  $M \times M$  square terrain with  $M = 500$  m. The transceiver energy parameters are set as:  $E_{\text{elec}} = 50$  nJ/bit and  $\epsilon_{\text{fs}} = 10$  pJ/bit/m<sup>2</sup>. The energy for data aggregation is set to  $E_{\text{DA}} = 5$  nJ/bit per signal [6]. Initially, the radio range of each sensor node is taken as 80 m. The message size of a sensor data item is fixed at 30 bytes and sensory data is generated at 1-second interval. A CH node retains its CH status for 20 seconds. The CH probability is fixed at 5% for the adopted scenario. The optimal value was computed using the analytical expression reported in [37]. Unless otherwise stated, all the following investigations adopt these values as their system parameters as depicted in Table 1. The system input parameter being investigated here is the cluster size controlled through the hop parameter,  $k$ . To make both the mathematical and simulation results comparable, we did not include the energy cost during the cluster setup phase of TCCA as well as other layers' costs. The complete cost is however determined when considering only the simulation results later. For all simulation results, each experiment is repeated 5–10 times and a 95% confidence interval is obtained (shown as error bars on the graphs).

Figure 2 shows the average energy spent per round by the network against different cluster sizes for  $N = 100$  nodes. The mathematical result follows a similar pattern as the simulation and is able to reach within 7% of the simulation results. For this scenario, it is evident that there exists an optimal value  $k$  that minimizes the total energy dissipation, and

**TABLE 1** The System Parameters

Parameter	Value
Field dimension, $M$	500 m
Radio range, $r$	80 m
Message size	30 bytes
Sensor timer	1 sec
CH setup timer	20 sec
CH probability, $p$	5%
Electronics energy, $E_{\text{elec}}$	50 nJ/bit
Amplifier energy, $\epsilon_{\text{fs}}$	10 pJ/bit/m <sup>2</sup>
Data aggregation energy, $E_{\text{DA}}$	5 nJ/bit/signal

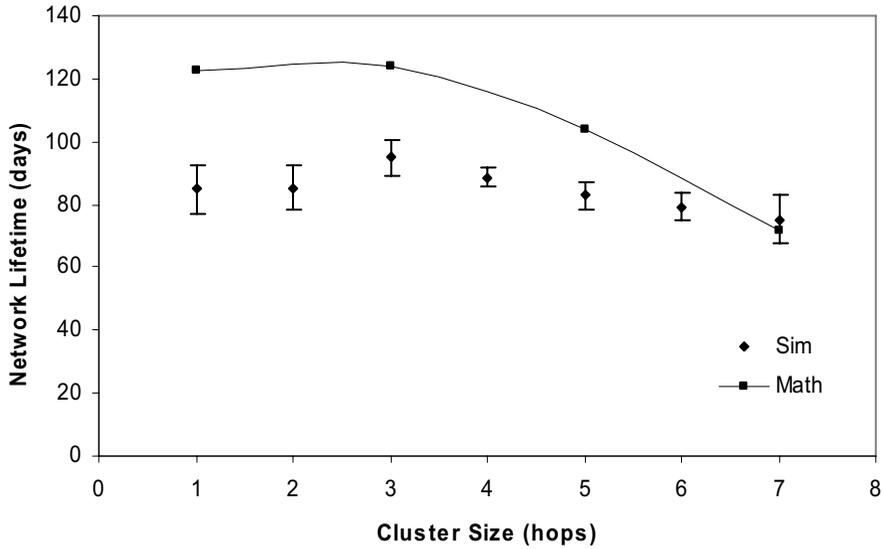


**FIGURE 2** Average energy dissipated against cluster size through the mathematical (continuous line) and simulation results (dots).

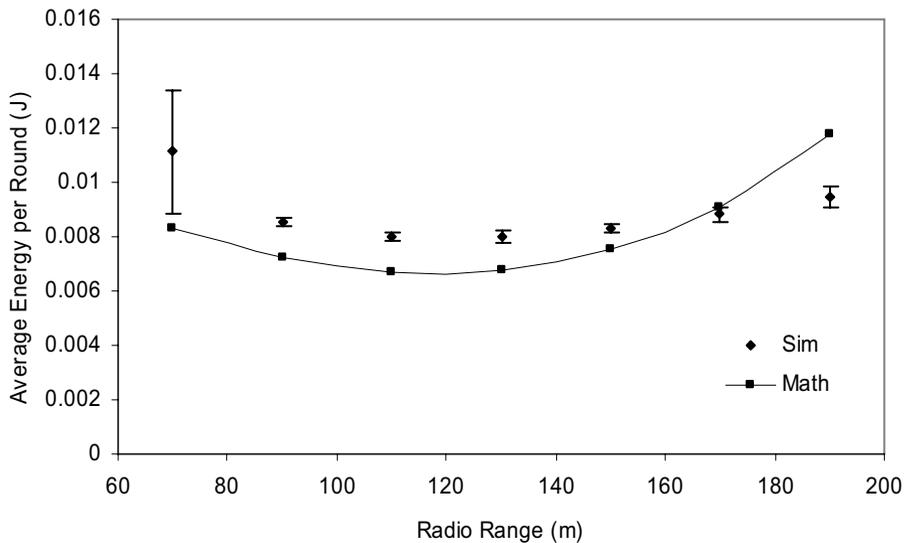
$k = 3$  is the optimal size. Any smaller or larger  $k$  results in higher energy consumption. A smaller cluster size implies the likely existence of many clusters, and the need for many CH nodes to communicate with the sink. However, when the cluster is larger than the optimal size, the bigger number of members in a cluster results in a higher intra-cluster communication cost, consequently increasing the overall energy dissipation. Thus, the CHs could use this optimal value to set the TTL field in their ADV messages to control their memberships and indirectly the cluster size.

In Fig. 3, the impact of the cluster size on the network lifetime is depicted. For this study, we chose to use a one-hour data collection interval rather than 1-sec interval. This does not however change the results pattern. As expected, a consistent result with the total energy usage behavior given above is observed. When the cluster size is small as controlled through TTL, there are likely to be many nodes forced to communicate straight to the sink. This behavior reduces to that of *direct transmission* albeit possibly using multi-hop links, which was shown to exhibit a higher energy consumption than clustered-type communication [6]. However, when TTL is increased beyond the optimal value, the energy consumption increased substantially, resulting in a reduced network lifetime. As the cluster size increases, for the same number of clusters, there are more members in each cluster. As such, there is a significant amount of intra-cluster communication required per round with many nodes acting as relay nodes to forward their downstream nodes' messages towards their CH. Thus, for the chosen network scenario, a cluster size larger than 3 hops is inefficient.

For the following experiment, the cluster size is fixed at  $k = 2$ . In Fig. 4, the impact of the sensor transmission range on energy dissipation is depicted. There is a rather close match between the mathematical and simulation results with at most 20% deviation. For the case of 70-m radio range, the simulation result exhibits high variance (c.f. long error bar) due to the higher probability of node isolation resulting in a disconnected graph. This probability is reduced at longer radio ranges. From Fig. 4, it is evident that for a short transmission range, the energy consumption is significantly higher. For example, when



**FIGURE 3** Network lifetime against cluster size through the mathematical (continuous line) and simulation results (dots).



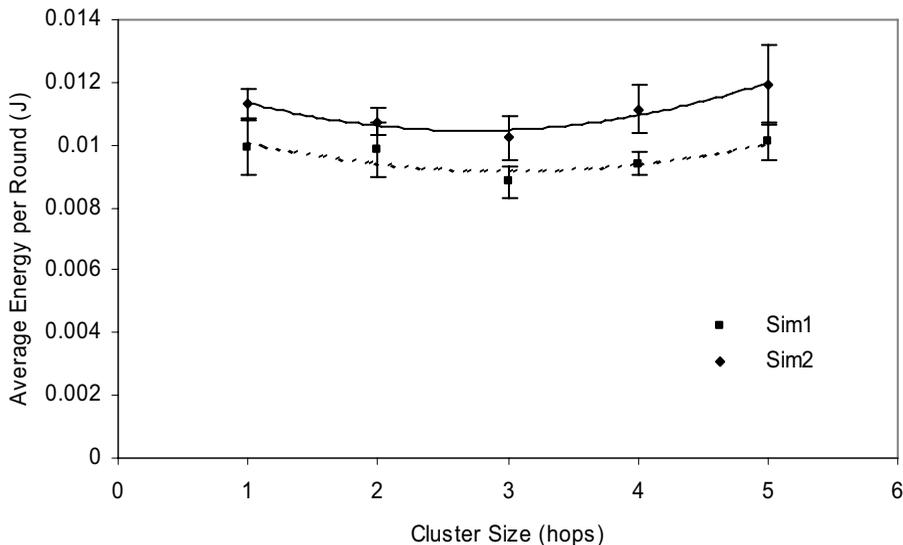
**FIGURE 4** Average energy dissipated against radio range through the mathematical (continuous line) and simulation results (dots).

comparing the energy consumption for a range of 70 to 110 m, there is almost 40% more energy used in the former for the same monitoring scenario. This is mainly due to the likely increased average number of hops required to reach a node. When the range is increased, the energy usage reduced initially, but later increased. Even though the average hops between nodes tend to decrease with a longer radio range, the energy usage will be dominated by the higher power required to drive the amplifier rather than the number of transmissions. Thus, there is an optimal transmission range at 110 m for this scenario that

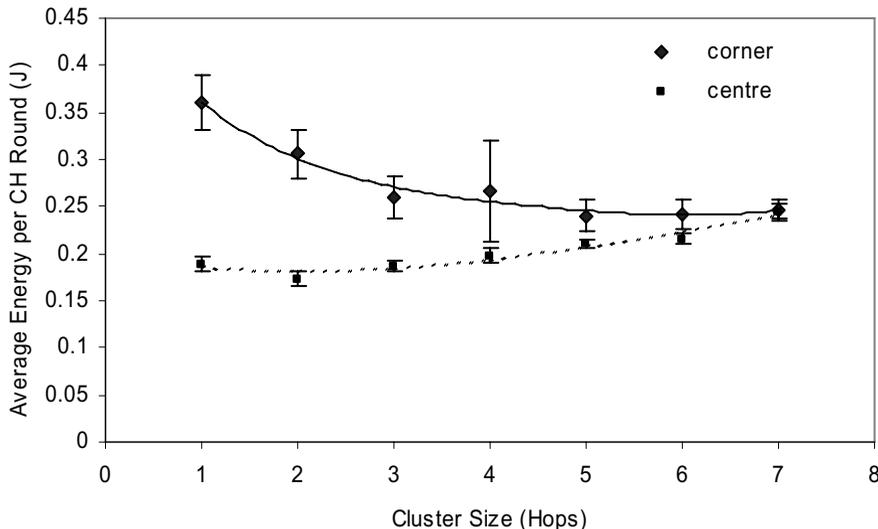
achieves the lowest energy dissipation. From Figs. 2–4, it is evident that the mathematical results mainly underestimated the energy usage. This is mainly due to the perfect uniform distribution assumption in the analysis that may not be achieved in every run of the simulation. As such, for these random topologies, the energy consumption is not at the optimal behavior resulting in higher energy estimation in simulation.

For the following investigations, only the simulator is used. However, we now consider the results with the full communication and computation overhead integrated according to the architecture shown in Fig. 1. In order to compute the computation cost, we assume a 4-MHz mote processor and a 3-V power source are used. On average, the Mica2 mote processor uses 8 mA per instruction. We just need to determine the average number of instructions to determine the processing cost [30]. For the first study, we repeated the same experiment of average energy dissipation against the cluster size for two cases. For the first case, the same assumptions as in Fig. 2 are made. For the second case, we assume a complete overhead integration into the simulator. Figure 5 depicts these cases with the curves *Sim1* and *Sim2* representing the above cases, respectively. As expected, the average energy dissipated for the second case is about 15% higher than the first case. More interestingly though, even with the complete overhead consideration, the optimal cluster size is still three. This clearly exhibits TCCA's ability to assure optimal performance for the specified values even in the comprehensive communication cost evaluation. It is an indication that this algorithm would work equally well against other types of routing and MAC protocols as well.

One of the assumptions made for the mathematical analysis was that the sink node is located at the center of the field. To investigate how the sink placement affects the network energy behavior, we repositioned the sink to the top right corner of the field. Any other corner position could be representative of the other corner locations as well. Figure 6 shows the impact of the sink placement on the energy dissipation. As expected, the sink location at the center is more energy efficient than the corner location. This is mainly due to the reduced average hops between the CH nodes and the sink resulting in lower forwarding energy cost. As the cluster size increases, the energy cost for sensory data collection and



**FIGURE 5** Average energy dissipated against cluster size for some overhead (*Sim1*) and complete overhead (*Sim2*).

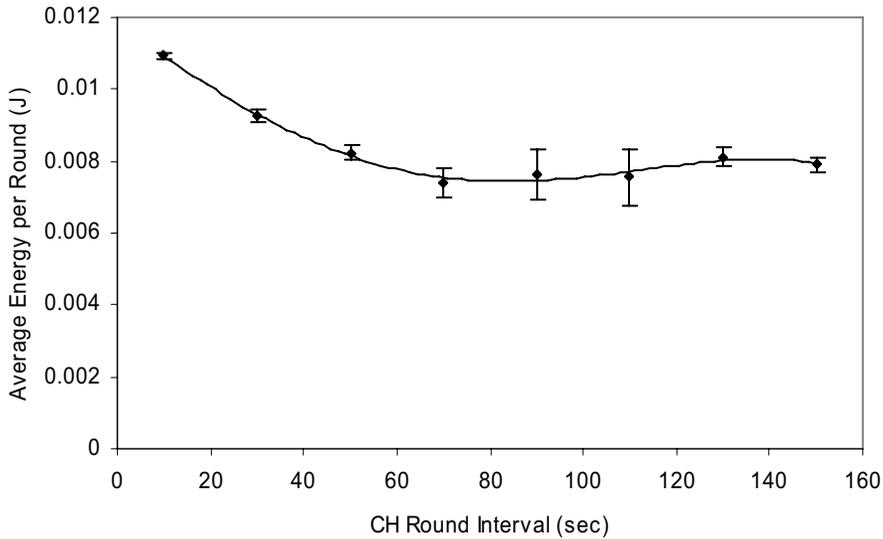


**FIGURE 6** Average energy dissipated per CH round against cluster size for sink located at the centre (square) and at a corner (diamond).

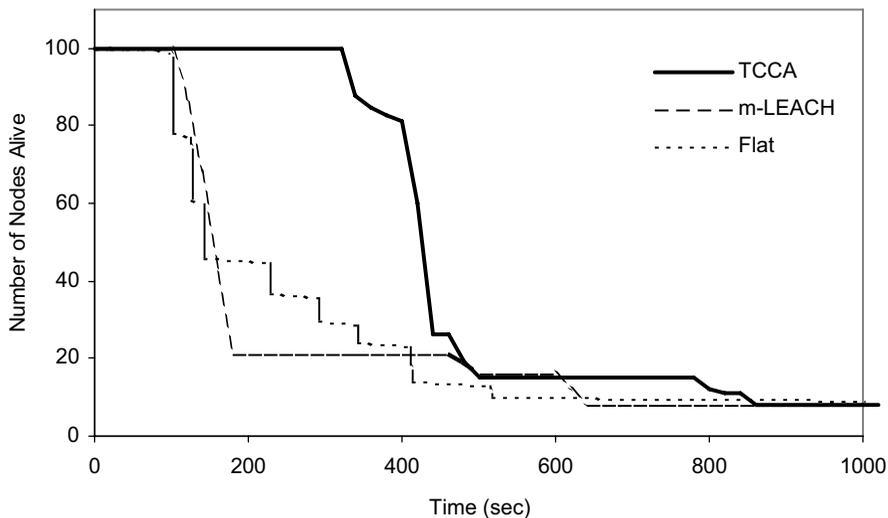
aggregation outweighs the forwarding cost. This effect is more pronounced for the corner placement. By having a larger cluster size, more nodes are likely to be members of some cluster and fewer nodes are likely to be forced to communicate directly to the sink. Thus, when we are faced with a deployment where the sink is located at the edge, we should adopt a larger TTL value to ensure a more energy efficient TCCA operation.

For most of the considered scenarios above, we have assumed a rather intensive data collection scenario where the sensor generated data at 1-second interval with CH round timer is fixed at 20-second interval. To see the impact of the CH round duration, this period is varied between 10 and 150 seconds for the same fixed 1-second data collection interval. Its impact on the average energy dissipation is depicted in Fig. 7. As per expectation, the energy usage decreased with the increase in the CH round duration. Since the cluster setup phase is not repeated as frequently for the longer round durations, energy dissipation naturally reduces. However, any further increase beyond 70 seconds, results in almost negligible changes. The curve appears to saturate at a specific energy level at these longer durations. As the round duration is increased, the influence of the cluster setup phase gets spread over more rounds. This results in its marginal effect at longer durations. The energy saturates rather than decreases any further due to the existence of a fixed cost in data collection and dissemination to the sink. It is however undesirable to adopt large durations as this would promote more unbalanced energy distribution and premature death of some nodes, mainly the CH nodes.

To investigate TCCA's performance against other clustering algorithms in the literature, we compared it against LEACH and a flat strategy using GRS only. Since LEACH cannot be applied directly to a multihop network, we modified this algorithm to use the GRS routing protocol to forward messages whenever the destination is not within the radio range. We termed this modified algorithm as multihop-LEACH (or m-LEACH). In Fig. 8, the improvement gained through TCCA is further exemplified by the network lifetime graph. TCCA exhibits the longest lifetime with all the nodes remaining fully functional. It is found that TCCA achieves more than twice the lifetime of m-LEACH and the flat approach. When the first node dies in the network, almost 80% of nodes die subsequently



**FIGURE 7** Average energy dissipated against the CH round duration.



**FIGURE 8** Network lifetime against simulation time of TCCA, m-LEACH and the flat strategy.

within the period of a third of the lifetime. This indicates that TCCA is able to enforce load balancing among the nodes.

It is clearly demonstrated that TCCA is able to realize significant energy savings when the parameters are properly configured guided by the given expression. Even with different sink locations, TCCA could be reconfigured to achieve optimal performance. The overhead due to the dynamic cluster setup is only marginal and could be leveraged against a longer configurable round duration. Since the CH nodes only need to maintain minimal state, we expect that the TCCA protocol to readily scale with network size.

## **7. Conclusions and Future Work**

As energy-awareness is highly critical in the design of sensor networks, we introduced the Time-Controlled Clustering Algorithm (TCCA) to control data dissemination towards the network sink. The objective of TCCA is to minimize the total energy dissipation by using a non-monitored rotating clusterhead election. TCCA is also able to control the cluster radius using an appropriate TTL value. A probabilistic mathematical model of this algorithm is developed to compute its realistic energy dissipation and network lifetime behavior. This model is then verified through a simulator and found to predict its behavior rather accurately. The simulator is further extended to incorporate other overheads due to the rest of the communication stack to enable a comprehensive TCCA evaluation. It was demonstrated that there is an optimal cluster size, which could be determined from the given model, and could be used to pre-configure the nodes to achieve an overall energy efficient operation. There is also an optimal radio range that should be adopted to control the transmit power level. The sink node placement plays an important role in the overall network energy usage. If the network center could not be accessed due to the nature of the application or some physical constraints, the sink could be deployed at the network edge with TCCA adopting a larger TTL.

As part of our current research we have assumed that the environment is error-free. We feel that this aspect could be investigated using a realistic radio propagation model through simulation. Currently, we have not considered the application data semantics in the TCCA behavior. We believe data correlation and the resulting data redundancy may be common and should be considered. This is left for future work.

## **Acknowledgement**

This work was supported in part by grants from the ARC Discovery Project DP0664782 and USyd R&D L2844 U3230.

## **About the Authors**

S. Selvakennedy obtained his PhD in computer communications in 1999 from the University of Putra, Malaysia. Earlier in 1996, he obtained his first-class honors degree in computer science also from the same university. He is currently serving as a Senior Lecturer at the University of Sydney, Australia. He has published more than 50 articles in journals and conferences mainly in the area of communications protocols for all-optical networking and wireless networking. His current research interest lies in developing algorithms for media access, routing, localization and topology control issues in wireless sensor networks, wireless ad hoc networks, and wireless mesh networks.

Dr. Selvakennedy is a member of ACM and IEEE. He has served on many conference program committees.

Sukunesan Sinnappan got his doctorate from the University of Newcastle, Australia in 2004. He is currently with the School of Economics and Information Systems, University of Wollongong, Australia. He has published in numerous fields such as online consumer behavior, website analysis, and wireless sensor networks. His current research interests include online self-profiling, location-based services, online privacy and trust apart from wireless sensor networks.

## References

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, pp. 102–114, Aug 2002.
2. S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TinyDB: An acquisitional query processing system for sensor networks," *ACM Transactions on Database Systems*, vol. 30, pp. 122–173, Mar 2005.
3. Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *ACM SIGMOD Record*, vol. 31, pp. 9–18, 2002.
4. M. Lotfinezhad and B. Liang, "Effect of partially correlated data on clustering in wireless sensor networks," *Proc. of the First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, Santa Clara, CA, USA, Oct 2004, pp. 172–181.
5. A. Scaglione, "Routing and data compression in sensor networks: stochastic models for sensor data that guarantee scalability," *Proc. of IEEE International Symposium on Information Theory*, Yokohama, Japan, Jun 2003, pp. 174.
6. W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660–670, Oct 2002.
7. M. J. Handy, M. Haase, and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," *Proc. of 4th International Workshop on Mobile and Wireless Communications Network*, Sept 2002, pp. 368–372.
8. C.-F. Chiasserini, I. Chlamtac, P. Monti, and A. Nucci, "An energy-efficient method for nodes assignment in cluster-based ad hoc networks," *Wireless Networks*, vol. 10, pp. 223–231, May 2004.
9. S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *Proc. of 22nd IEEE Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, San Francisco, CA, USA, Mar 2003, pp. 1713–1723.
10. O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, pp. 366–379, Oct 2004.
11. Crossbow Technology, "<http://www.xbow.com/>".
12. T. Shu, M. Krunz, and S. Vruthula, "Power balanced coverage-time optimization for clustered wireless sensor networks," *Proc. of the Sixth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Urbana-Champaign, IL, USA, May 2005, pp. 111–120.
13. M. Gerla and J. T.-C. Tsai, "Multicluster, mobile, multimedia radio network," *Wireless Networks*, vol. 1, pp. 255–265, Aug 1995.
14. F. Kuhn, T. Moscibroda, and R. Wattenhofer, "Initializing newly deployed ad hoc and sensor networks," *Proc. of the 10th Annual International Conference on Mobile Computing and Networking*, Philadelphia, PA, USA, Sept 2004, pp. 260–274.
15. D. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm," *IEEE Transactions on Communications*, vol. 29, pp. 1694–1701, Nov 1981.
16. G. Gupta and M. Younis, "Load-balanced clustering of wireless sensor networks," *Proc. of IEEE International Conference on Communications*, Anchorage, AL, USA, May 2003, pp. 1848–1852.
17. A. D. Amis and R. Prakash, "Load-balancing clusters in wireless ad hoc networks," *Proc. of the Third IEEE Symposium on Application-Specific Systems and Software Engineering Technology*, Richardson, TX, USA, Mar 2000, pp. 25–32.
18. A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh, "Max-min d-cluster formation in wireless ad hoc networks," *Proc. of the Nineteenth Joint Conference of the IEEE Computer and Communications*, Tel Aviv, Israel, Mar 2000, pp. 32–41.
19. A. Ephremides, J. Wieselthier, and D. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proc. of the IEEE*, vol. 75, pp. 56–73, Jan 1987.

20. C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 1265–1275, Sept 1997.
21. S. Basagni, "Distributed clustering for ad hoc networks," *Proc. of the Fourth International Symposium on Parallel Architectures, Algorithms, and Networks*, Perth, Australia, Jun 1999, pp. 310–315.
22. M. Chatterjee, S. K. Sas, and D. Turgut, "An on-demand weighted clustering algorithm (WCA) for ad hoc networks," *Proc. of the IEEE Global Telecommunications Conference*, San Francisco, CA, USA, Dec 2000, pp. 1697–1701.
23. S. Ghiasi, A. Srivastava, X. Yang, and M. Sarrafzadeh, "Optimal Energy Aware Clustering in Sensor Networks," *Sensors*, vol. 2, pp. 258–269, Jul 2002.
24. B. J. Culpepper, L. Dung, and M. Moh, "Design and analysis of Hybrid Indirect Transmissions (HIT) for data gathering in wireless micro sensor networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 8, pp. 61–83, Jan 2004.
25. S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, pp. 924–935, Sept 2002.
26. J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen, "Topology control for wireless sensor networks," *Proc. of the Ninth International Conference on Mobile Computing and Networking*, San Diego, CA, USA, Sept 2003, pp. 286–299.
27. W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," *Proc. of the Twenty-First Joint Conference of the IEEE Computer and Communications Societies*, New York, USA, June 2002, pp. 1567–1576.
28. J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister, "System architecture directions for networked sensors," *Proc. of the Ninth International Conference on Architectural Support for Programming Languages and Operating Systems*, Cambridge, MS, USA, Nov 2000, pp. 93–104.
29. P. Levis, N. Lee, M. Welsh, and D. Culler, "TOSSIM: Accurate and scalable simulation of entire tinyOS applications," *Proc. of the First ACM International Conference on Embedded Networked Sensor Systems*, Los Angeles, USA, Nov 2003, pp. 126–137.
30. V. Shnayder, M. Hempstead, B.-r. Chen, G. W. Allen, and M. Welsh, "Simulating the power consumption of large-scale sensor network applications," *Proc. of the Second International Conference on Embedded Networked Sensor Systems*, Baltimore, MD, USA, Nov 2004, pp. 188–200.
31. J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *IEEE Computer*, vol. 34, pp. 57–66, Aug 2001.
32. B. Sundararaman, U. Buy, and A. D. Kshemkalyani, "Clock synchronization for wireless sensor networks: a survey," *Elsevier Ad Hoc Networks*, vol. 3, pp. 281–323, May 2005.
33. A. Woo and D. Culler, "A transmission control scheme for media access in sensor networks," *Proc. of the Seventh Annual International Conference on Mobile Computing and Networking*, Rome, Italy, Jul 2001, pp. 221–235.
34. G. Finn, "Routing and addressing problems in large metropolitan-scale internetworks," *ISI, ISU/RR* 87–180, Mar 1987.
35. D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker, "Complex Behavior at Scale: An Experimental Study of Low-Power Wireless Sensor Networks," *UCLA, LA, UCLA CS Technical Report UCLA/CSD-TR 02–0013*, 2002.
36. T. He, S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, J. Hui, and B. Krogh, "Wide-area monitoring of mobile objects: Energy-efficient surveillance system using wireless sensor networks," *Proc. of the Second ACM International Conference on Mobile Systems, Applications, and Services*, Boston, USA, June 2004, pp. 270–283.
37. S. Selvakennedy and S. Sinnappan, "The Time-Controlled Clustering Algorithm for Optimized Data Dissemination in Wireless Sensor Networks," *Proc. of the IEEE Conference on Local Computer Networks*, Sydney, Australia, Nov 2005, pp. 509–510.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

