University of Wollongong

Research Online

Sydney Business School - Papers

Faculty of Business and Law

1-1-2005

A configurable time-controlled clustering algorithm for wireless sensor networks

S. Selvakennedy
University of Sydney

Sukunesan Sinnappan
University of Wollongong, suku@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/gsbpapers

Part of the Business Commons

A configurable time-controlled clustering algorithm for wireless sensor networks

Abstract

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. These nodes are typically highly resource-constrained. Since the communication task is a significant power consumer, there are various attempts to introduce energy-awareness at different levels within the communication stack. Clustering is one such attempt to control energy dissipation for sensor data routing. Here, we propose the time-controlled clustering algorithm to realise a network-wide energy reduction by the rotation of clusterhead role, and the consideration of residual energy in its election. A realistic energy model is derived to accurately quantify the network's energy consumption using the proposed clustering algorithm.

Keywords

Configurable, Time, Controlled, Clustering, Algorithm, for, Wireless, Sensor, Networks

Disciplines

Business

Publication Details

A Configurable Time-Controlled Clustering Algorithm for Wireless Sensor Networks

S. SELVAKENNEDY[#] AND S. SINNAPPAN[*]

[#]*School of IT, Madsen Bldg. F09, University of Sydney, NSW 2006, Australia.*
skennedy@it.usyd.edu.au
[*]*Graduate School of Business & Professional Development,*
University of Wollongong, Wollongong NSW 2522 Australia
suku@uow.edu.au

Abstract

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. These nodes are typically highly resource-constrained. Since the communication task is a significant power consumer, there are various attempts to introduce energy-awareness at different levels within the communication stack. Clustering is one such attempt to control energy dissipation for sensor data routing. Here, we propose the Time-Controlled Clustering Algorithm to realise a network-wide energy reduction by the rotation of clusterhead role, and the consideration of residual energy in its election. A realistic energy model is derived to accurately quantify the network's energy consumption using the proposed clustering algorithm.

1. Introduction

Wireless Sensor network is a critical emerging area of mobile computing that presents unique wireless networking issues due to their unusual application needs, highly constrained resources and functionality, small packet size and dynamic multihop topologies. It has gathered a considerable research interest in recent years mainly due to its possible wide applicability, such as monitoring (habitat, medical, seismic), surveillance and pre-warning purposes [1]. These networks usually contain hundreds or thousands of sensors, which may be randomly or selectively deployed. The unique application behaviour in sensor networks leads to very different traffic characteristics from that found in current networks. The main function of a sensor network application is to sample the environment for sensory information, such as atmospheric pressure, and propagate this data back to the monitoring point, while perhaps performing some in-network pre-processing, such as data fusion. These nodes are expected to operate for a long time, possibly several years. Furthermore, sensors are also expected to be simple and cheap. The goal of many micro-sensor projects underway is to make cubic millimetre sensors [2], [3]. Thus, the small size of sensor nodes will severely limit the available energy for data processing and communication tasks [4].

Since these sensors may be deployed in physically harsh and inaccessible area but still need to communicate with the base station (i.e. the gateway or sink), direct communication may not be effective and in certain circumstances infeasible. The dominant energy consumer in a sensor is its radio transceiver. This places significant restrictions on the power, limiting both the transmission range and the data rate. Thus, to enable communication between sensors not within each other's range, multihop transmission is a more feasible alternative.

There are numerous proposals to reduce energy usage by the protocols within the proposed leaner communication stack. Since the cost of transmitting a data bit is higher than the computation process [3], it appears to be advantageous to organize sensors into clusters. In the clustered environment, data gathered by the sensors is transmitted to the base station through clusterheads (CHs). As the sensors communicate data over shorter distances in such an environment, the energy spent in the network is likely to be substantially lower.

Various clustering algorithms in different contexts have been proposed in the literature. Some algorithms also distinguish themselves by how the CHs are elected. The LEACH algorithm [5] and its related extensions [6] use probabilistic self-election, where each sensor 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

energy-intensive CH function aims to distribute the power usage for prolonged network life. However, LEACH allows only one-hop clusters. Another clustering algorithm proposed in [7] aims to maximize the network lifetime, but it assumes the sensors 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 sensor network, the sensors are quasi-stationary and the instability of clusters due to mobility of sensors may not be an issue.

For sensor networks with a large number of energy-constrained sensors, it is crucial to design a fast distributed algorithm to organize sensors in clusters. Bandyopadhyay et al. derived simplified formulas for computing the optimal p based on a simplified energy model of the LEACH network and the optimal number of hops k using results in stochastic geometry to minimize the total energy spent [8]. However, it was assumed the sensors have unit energy consumption for each of a node's communication task. In [9], the authors presented a Hybrid Energy-Efficient Distributed clustering (HEED) protocol that periodically selects CHs according to a primary and secondary parameter, for example a node's residual energy and a node degree (or its proximity to neighbours), respectively. It capitalises on the availability of multiple power levels such as on the Berkeley motes. 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 level to ensure connectivity among the CHs. The configuration of these parameters requires the knowledge of the whole network.

In this paper, we introduce the Time-Controlled Clustering Algorithm (TCCA) that allows multihop clusters using message timestamp and time-to-live (TTL) to control the cluster formation. In the CH election, a node also considers its residual energy before volunteering. Subsequently, a numerical model to quantify its efficiency on energy usage is provided, which is derived using a realistic first-order radio energy dissipation model with the objective of minimizing the energy spent in communicating to the base station.

2. 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 collection, aggregation and transfer to the base station occurs.

In order to determine the eligibility to be a CH, a node's residual energy $E_{residual}$ is taken into consideration. Besides, 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. The threshold is computed as follows:

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

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

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 advertises itself as the CH to the neighbouring sensors within its radio range. This advertisement message (ADV) carries its node id, initial TTL, its residual energy and a timestamp. Upon receiving and processing, regular sensors 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 assumed that all nodes use the same fixed k value to simplify our mathematical model. Since the CH is able to calculate the first-hop successful transmission time based on its MAC layer feedback, it can use it to control the duration of the cluster setup phase. If the first-hop time is t, the clustering process time is $(2k-1)*t$ to ensure sufficient time for reply messages to reach the CH. To ensure that the network operation is stable, the steady-state phase should be significantly larger than $(2k-1)t$. To simplify the mathematical model representation, we will neglect the *marginal* effect of this setup phase in the overall computation of power dissipation, as the setup phase is substantially shorter than the transfer operation.

Any sensor that receives such an ADV message and is not a CH itself joins the cluster of the *nearest* CH. If there is a tie, the node could select the CH with higher residual energy. Once a sensor 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. The CH node also uses it to *learn* the appropriate setup phase time for future

rounds. Together with TTL, the CH could form a multihop view of its cluster, which could be used to create a collision-free transmission schedule. A transmission schedule is created by the CH based on its number of members and their relative distance to enable the reception of all sensed data in a collision free manner. At the end of the schedule, the CH communicates the aggregated information to the base station. The details of the transmission schedule formation are excluded here, as our current focus is on the clustering algorithm itself.

3. The TCCA Energy Usage Model

The energy used for the information gathered by the sensors to reach the base station 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 sensors in clusters to minimize overall energy consumptions, we need to determine the optimal value of the parameter k of our algorithm that would ensure minimization of energy usage. For the development of our model, the following assumptions are made:

a) The sensors are randomly scattered in a two-dimensional plane and have a homogeneous spatial Poisson process with λ intensity.

b) All nodes in the network are homogeneous. They transmit at the same power level and hence have the same radio range r. The communication from each sensor follows isotropic disk connectivity.

c) The base station is located at the centre of the field.

d) 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 derivation of the optimal system parameter value is to define a function for the energy used in the network to communicate information to the base station during the steady-state phase.

As per the assumptions, the sensors are distributed according to a homogeneous spatial Poisson process. The number of sensors in a square area of side M is a Poisson random variable, N with mean λA where $A = M \times M$. Let's assume that for a particular realization of the process, there are n sensors in this area. The probability of becoming a CH is $p = \dfrac{A}{n\pi(kr)^2}$. On average, there will be np sensors becoming CHs.

Now, to derive the energy usage, 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:

$$E_{Tx} = l\mathrm{E}_{elec} + l\varepsilon_{fs}d^2 \qquad (2)$$

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

$$E_{Rx} = l\mathrm{E}_{elec} \qquad (3)$$

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 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 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 CH (s_i), we simply multiply the area difference between the circle of ir and $(i-1)r$ radius and the mean node density, λ:

$$s_i = \lceil (2i-1)\pi r^2 \lambda \rceil \qquad (4)$$

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 routing path. The average number of all its 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:

$$c_i = \sum_{j=i}^{k} \left\lceil \frac{2k - (2j-1)}{2k - (2j+1)} \right\rceil \qquad (5)$$

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:

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

As for the CH energy usage computation, we need to include its average number of members, message aggregation cost (E_{DA}) and its communication (possibly multihop) cost to the base station. 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 base station is $\dfrac{0.765npM}{2}$ [8]. Thus, the average number of hops from a CH to the base station is $\dfrac{0.765M}{2r}$. The overall energy consumption of np CH nodes (C_2) could then be approximated as:

$$E[C_2 \mid N = n] =$$

$$np \times \left\{ \left\lceil \pi(kr)^2 \lambda - 1 \right\rceil E_{Rx} + E_{DA} l + \frac{0.765M}{2r}(E_{Tx} + E_{Rx}) \right\}$$

$$(7)$$

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

$$E[C \mid N = n] = E[C_1 \mid N = n] + E[C_2 \mid N = n] \quad (8)$$

Removing the conditioning on N yields:

$$
\begin{aligned}
E[C] \quad &= E[E[C \mid N = n]] \\
&= E[N] \times p(E[C_1] + E[C_2]) \\
&= \lambda A p(E[C_1] + E[C_2]) \quad (9)
\end{aligned}
$$

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 power 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:

$$L = \frac{B}{C/N} \times t = \frac{BNt}{C} \quad (10)$$

4. Experimentation and Discussions

This section discusses the numerical experimentation, which includes the description of the chosen parameters set and the adopted sensor network scenario. For these experiments, we assumed that there are N sensor nodes distributed randomly in a square $M \times M$ region with $M = 100$ m. The communication energy parameters are set as: $E_{elec} = 50$ nJ/bit and $\varepsilon_{fs} = 10$ pJ/bit/m^2. The energy for data aggregation is set to $E_{DA} = 5$ nJ/bit [6]. Initially, the radio range of each sensor node is taken as 30m. The message size of a sensor data item is fixed at 50 bits. Unless otherwise stated, all the following investigations adopt these values as their system parameters. The system input parameter being investigated here is the cluster size controlled by the hop parameter, k.

Figure 1 shows the total energy spent by the network against various cluster sizes for different number of sensor nodes, N. For the adopted scenario, it is evident that there exists an optimal value k that minimises the total energy consumption, and $k = 5$ is the most suitable size for all the tested cases. 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 base station. However, when the cluster is larger than the optimal size, the bigger number of members in a cluster results in 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 their cluster size. Another interesting observation to note is the larger clusters are significantly worse off than the smaller ones. The use of the optimal k is only marginally better than any smaller cluster size. This suggests that the proposals of LEACH [6] and HEED [9], which only allows one-hop clusters seems to be justified. Furthermore, the presence of only one-hop clusters would considerably simplify the generation of transmission schedule within a cluster, whose energy cost was omitted in our model.
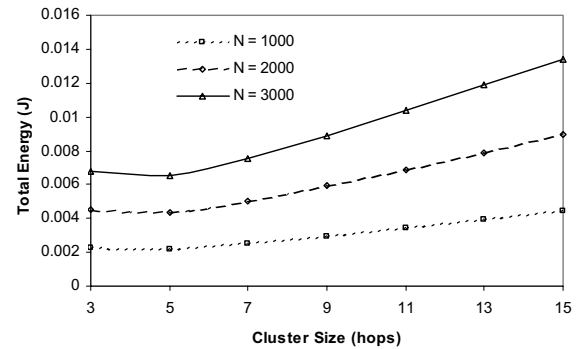

Fig. 1. Total energy usage against various TTL values for different sensor node density.

In Fig. 2, the impact of cluster size on the network lifetime is shown against different node density. As expected, a consistent result with the total energy usage behaviour given above is observed. When the cluster size is small as controlled by the TTL value, there are likely to be more CH nodes elected to communicate with the base station. This behaviour reduces to that of *direct transmission* albeit possibly using multihop links, which was shown to exhibit higher energy consumption than clustered-type communication [6]. However, when the TTL value is increased beyond the optimal value, the energy consumed per node increased substantially thereby reducing the overall network lifetime. As the cluster size increases, the number of clusters is smaller but the number of members in each cluster is larger. As such, there is a significant amount of intra-cluster communications 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 5 hops is inefficient. It may be surprising to the

reader that there is limited influence of density on the lifetime. It is likely due to our simplifying assumption of a collision- and error-free MAC protocol.
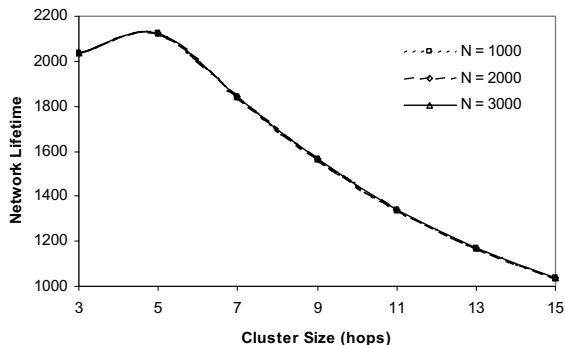

Fig. 2. Network lifetime against various TTL values for different sensor node density.

For the following experiment, the cluster size is fixed at $k = 5$. In Fig. 3, the impact of the sensor transmission range on the total energy usage is depicted. When the transmission range is very short, the energy consumption is significantly higher. For example, at $N = 2000$, when comparing the energy consumption for range of 15m to 30m, there is almost 70% 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 for the shorter range. When the range is increased, the total energy usage reduced initially, but later increased albeit slowly. Thus, there is an optimal transmission range that achieves the lowest energy dissipation. Any further increase to the range do not result in further saving mainly due to a fixed cluster size controlled by the TTL value and the fixed number of sensor nodes in the network. Interference, which could have been yet another factor here, is not represented in our model as we have assumed that the MAC has a perfect schedule.
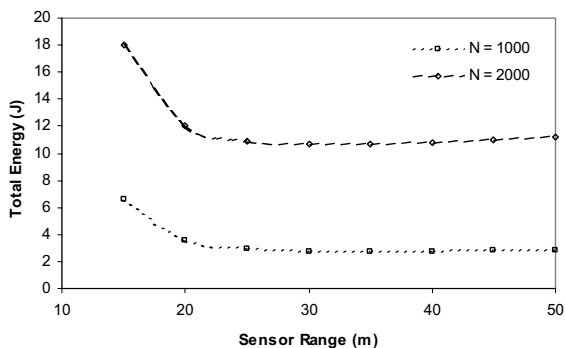

Fig. 3. Total energy usage against transmission range for different sensor node density.

5. Conclusions

As energy-awareness is highly critical in the design of sensor networks, we proposed the Time-Controlled Clustering Algorithm (TCCA). The objective of TCCA is to minimise the total energy dissipation by using non-monitored rotating clusterhead election with residual energy level consideration. TCCA is also able to control the cluster diameter using an appropriate TTL value. It was numerically demonstrated that there is an optimal cluster size, which could be determined from the given model, and then used to pre-configure the nodes to achieve an overall energy efficient operation. It is found that smaller cluster sizes, including one-hop clusters, have almost similar performance level as the optimal *k*-hop clusters. Furthermore, the generation of the transmission schedule for such clusters has only *O*(1) complexity, which should make such smaller sizes to be more attractive for sensor networks.

6. References

[1]   I. F. Akyildiz *et al*, "Wireless sensor networks: a survey," *Journal of Computer Networks,* Elsevier*,* Vol. 38 No. 4, 2002, pp. 393–422.

[2]   B. Warneke *et al*, "Smart Dust: Communicating with a Cubic-Millimeter Computer," *IEEE Computer Magazine*, Vol. 34 No. 1, Jan. 2001, pp. 44- 51.

[3]   G. J. Pottie and W. J. Kaiser, "Wireless Integrated Network Sensors", *Communications of the ACM*, Vol. 43, No. 5, May 2000, pp 51-58.

[4]   J. M. Kahn, R. H. Katz and K. S. J. Pister, "Next Century Challenges: Mobile Networking for Smart Dust,"*Proc. of 5th ACM/IEEE Int. Conf. on Mob. Comp. and Net.*, Seattle, US, Aug. 1999, pp. 271-278.

[5]   C. F. Chiasserini *et al*, "Energy Efficient design of Wireless Ad Hoc Networks," *Proc. of Eur. Wireless*, Pisa, Italy, Feb. 2002.

[6]   W. Heinzelman, A. Chandrakasan and H. Balakrishnan, Energy-Efficient Communication Protocol for Wireless Microsensor Networks, in: *Proc. of IEEE Proc. Of the Hawaii International Conf. on System Sciences*, Hawaii, US, Jan. 2000, pp. 1-10.

[7]   M. J. Handy, M. Haase and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," *Proc. of IEEE Int. Conf. on Mobile and Wireless Communications Networks*, Sept. 2002, pp. 368 – 372.

[8]   S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *Proc. of the 22nd Joint Conf. of the IEEE Computer and Communications Societies*, June 2003, pp. 1713-1723.

[9]   O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. on Mob. Comp.*, vol. 3, Oct 2004, pp. 366-379.