The Time-Controlled Clustering Algorithm for Optimized Data Dissemination in Wireless Sensor Networks

S. SELVAKENNEDY# AND S. SINNAPPAN*

#School of Information Technologies, 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

As the communication task is a significant power consumer, there are many attempts to improve energy efficiency. Node clustering, to reduce direct transmission to the base station, is one such attempt to control data dissemination. Here, we derived the optimal number of clusters for TCCA clustering algorithm based on a realistic energy model using results in stochastic geometry.

1. Introduction

Wireless sensor networks (WSNs) have gathered an immense research interest in recent years mainly due to their possible wide applicability, such as monitoring, surveillance and pre-warning purposes. A WSN provides a global view of the monitored area based on local observations of each node. Even though the nodes are powered by non-rechargeable battery, these nodes are expected to operate for a long time. Furthermore, the nodes are also expected to be simple and cheap.

Since these nodes may be deployed in physically harsh or inaccessible area but still need to communicate with the base station, direct transmission may not be effective and in certain circumstances infeasible. To enable communication between nodes not within each other’s range, they use multihop transmission. Since the cost of transmitting a data bit is higher than the computation process [1], it appears to be advantageous to organize nodes into clusters. In the clustered environment, data gathered by the nodes is transmitted to the base station through clusterheads (CHs). As the nodes communicate data over shorter distances in such an environment, the energy spent in the network is likely to be substantially lower.

Many clustering algorithms in various contexts have been proposed [2-4]. 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 [4] uses stochastic self-election, where each sensor has a probability $p$ of becoming a CH in each round. 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.

For sensor networks with a large number of energy-constrained nodes, it is crucial to design a fast distributed algorithm to organize nodes in clusters. Bandyopadhyay et al. derived expressions for computing the optimal $p$ and number of hops ($k$) based on a simplified model of the LEACH network using results in stochastic geometry to minimize the total energy spent [5]. It was however assumed that all operations of a node have unit energy usage.

The authors introduced the Time-Controlled Clustering Algorithm (TCCA) that uses message timestamp and time-to-live (TTL) to control the cluster formation [6]. Here, we present an analytical model to derive the optimal $p$ that minimizes the energy spent in data dissemination using results in stochastic geometry.

2. The Analytical Model

The operation of TCCA is divided into rounds to enable load distribution among the nodes. 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. The detail operation of TCCA is given in [6]. The energy used for the data dissemination process by the nodes to the base station will depend on the parameters $p$, distance between the transmitting and receiving nodes, and the cluster size. The overall idea of the derivation...
is to define a function for the energy used to disseminate information to the base station.

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 $A = M \times M$. Lets assume that for a particular realization of the process, there are $n$ nodes in this area. The probability of becoming a CH is $p$, on average there will be $np$ nodes becoming CHs. Also, the CHs and the non-CHs are distributed as per independent homogeneous spatial Poisson processes $P_1$ and $P_0$ of intensity $\lambda_1 = p\lambda$ and $\lambda_0 = (1-p)\lambda$, respectively.

Using the ideas in stochastic geometry, each node joins the closest CH to form a Voronoi tessellation. The plane divides into zones called the Voronoi cells, representing the number of $P_0$ process points in each Voronoi cell and $N_v$ is the random variable representing the number of $P_0$ process points in each Voronoi cell and $L_v$ is the total length of all segments connecting the $P_0$ process points to the nucleus in a Voronoi cell, then their expectations are as in [5].

Now, to derive the energy usage per any node type, both the free space $\varepsilon_0 (d^2$ power loss) and the multipath fading $\varepsilon_{mp} (d^4$ power loss) models are used [4]. Accordingly, transmit and receive energy are determined, and depends on message size ($l$) and distance ($d$). The clustering algorithm is designed to ensure the expected number of clusters per round is $np$. The dissipated energy by the nodes can be analytically estimated using the computation and communication energy models. Each CH dissipates energy receiving signals from its members, aggregating the signals and transmitting the aggregate signal to the base station. As for each non-CH node, it only needs to transmit its data to the CH once during a round. This allows us to determine the energy spent in a cluster during each round and could be generalized network-wide (Due to space limit, we could only show the final derivation).

$$p_{opt} = \frac{1}{2} \sqrt{\frac{E_{fs}}{\bar{\lambda}(\varepsilon_{mp} d_{LOS}^4 - E_{elec} - E_{DA})}} \tag{1}$$

3. Analytical Experimentation

For these experiments, we assumed $M = 100$ m. The communication energy parameters are set as in [6]. Figure 1 shows the total energy spent by the network for different CH probabilities with number of nodes 1500 and 2000. Both curves depict a higher value of total energy dissipated in the network for very small CH probabilities. This is due to the presence of small number of CHs with large cluster size requiring significant energy expense for data collection from their members. However, as the CH probability increases, we reach an optimal point where the total energy consumed is minimal. Any further increase in the CH probability resulted in higher energy usage.

![Fig. 1. Total energy spent per round vs. CH probability for $n = 1500$ (left y-axis) and 2000 (right y-axis).](image)

4. Conclusions

As energy-awareness is highly critical in the design of sensor networks, we derived the optimal number of clusters for TCCA. TCCA is also able to control a cluster’s diameter based on the message TTL. It was demonstrated that there is an optimal probability, which could be determined from the given expression and preconfigured into the nodes, to achieve an overall energy efficient operation. It was also found that there is a decreasing improvement on network lifetime, when more nodes are deployed within the same region.

5. References