WEIGHTED TREE BASED ROUTING AND CLUSTERING PROTOCOL FOR WSN

Riham S. Elhabyan, and Mustapha C.E. Yagoub, Senior Member, IEEE

School of Electrical Engineering and Computer Science, University of Ottawa
Ottawa, Ontario, Canada

ABSTRACT

In wireless sensor networks (WSN), clustering, a method of grouping sensor nodes, has numerous advantages. To ensure long-term operation of clustered WSN, two efficient routing protocols are proposed in this paper. The first is the Control Data Dissemination Protocol (CDDP) which forms a tree that connects all the nodes in the network, this tree will be used later to transfer control data to the base station. The second is a centralized weighted clustering protocol that uses the CDDP to transmit control data to the base station which will select the best cluster heads in the network according to a weighted criteria function. The effect of using a realistic energy consumption model in cluster-based communication for WSN will be investigated. The effectiveness of the proposed protocols are demonstrated by numerical simulations.

Index Terms— Clustering, Cluster Head, Routing, RSSI, WSN, Energy Model.

1. INTRODUCTION AND RELATED WORK

Wireless Sensor Network (WSN) technology is a key component for pervasive computing. It has become an important technology in realizing many applications including both simple phenomena monitoring applications and heavy-duty data streaming applications [1].

A WSN usually consists of tens to thousands of sensor nodes that can sense the environment, communicate with neighbouring nodes, and in many cases perform basic computations on the data being collected [2].

In order to achieve long-term operation of WSN, routing protocols based on clustering have been proposed. Clustering is a useful technique through which we can affect the network lifetime, scalability and load balancing. The main objective of a clustering protocol is to increase the WSN efficiency by organizing the network nodes into smaller clusters and cluster head (CH) for each cluster is elected. Sensor nodes in each cluster transmit their data to the respective CH and CH aggregates data and forward them to a central base station (BS).

Once the WSN has been divided into clusters, the communication between nodes can be either intra-cluster or inter-cluster. Intra-cluster communication comprises the message exchanges between the participating nodes and the CH. Inter-cluster communication includes the transmission of messages between the CHs or between the CH and the BS. The fact that only the CH is transmitting information out of the cluster helps avoid collisions between the sensors inside the cluster because they don't have to share the communication channel with the nodes in other clusters [3].

Many clustering protocols have been proposed for WSN in recent years. They can be divided into two main groups namely, distributed and centralized protocols. In the first, each sensor decides whether it will become a cluster head or not based on some criteria. In the second, a central node like the base station (BS) compares between the different nodes to select the best cluster heads among them.

Low energy adaptive clustering hierarchy (LEACH) [4] is one of the most popular distributed cluster-based routing algorithms in WSN. LEACH uses a time division multiple access (TDMA) principle to avoid collisions, and in order to maintain a balanced energy consumption, suggests that each node probabilistically become a cluster head. The cluster heads are selected without considering the residual energy or the other properties of the sensor nodes. This random mechanism of selecting the cluster heads doesn't guarantee even distribution of clusters over the network [3]. Furthermore, the protocol ignores the inter-cluster communication problem and assumes that each CH can send its information directly to the BS [2], which is not a realistic assumption due to the communication range restrictions of the sensor nodes.

Hybrid energy-efficient distributed Clustering (HEED) [5] is another distributed clustering protocol that is an extension of LEACH. It selects cluster heads according to a hybrid of the node residual energy and a secondary parameter, such as node proximity to its neighbours or node degree. Cluster formation is achieved with an iterative approach. The improvement over LEACH is that HEED can evenly distribute the cluster heads in the sensing area by local competition. HEED operates in multi-hop networks using an adaptive transmission power in the inter-clustering communication.

Both LEACH and HEED are fully distributed. However, both protocols may form a cluster consisting only of one node which will die much faster than other clusters and this will result in the hot-spot problem, in which nodes close to the BS die faster than the other nodes and the network lifetime cannot

978-1-4799-0033-6/13/$31.00

be extended.

LEACH-centralized (LEACH-C) [6] is a centralized version of LEACH. Unlike LEACH, where nodes self-configure themselves into clusters, LEACH-C uses the BS for cluster formation. Initially, each node sends its information (location and energy level) to the base station, which will use this information to find a predetermined number of cluster heads and configures the network into clusters. The clusters are chosen to minimize the energy required for non-cluster head nodes to transmit their data to their respective cluster heads.

Energy balanced unequal clustering protocol (EBUC) [7], the authors tried to solve the hot-spot problem by creating unequal clusters using a centralized particle swarm optimization (PSO) algorithm at the base station. The clusters are created such that the ones near the BS have fewer number of nodes and so increasing the number of clusters around the BS. For the Inter-cluster communication, the CH uses a greedy algorithm to choose a relay node based on the node's residual energy and distance to the BS.

LEACH-C and EBUC are both centralized, with the cluster head selection algorithm at the base station. Both also yield better results than those of distributed protocols in terms of network lifetime and packet delivery rate. Centralized approaches in clustering yield better results because the BS uses its global knowledge of the network to produce better clusters. However, both protocols ignored how the control packets will reach the BS. LEACH-C assumed the each node can transmit its control data directly to the sink, which is not applicable in realistic system due to the limited transmission range. In EBUC, the authors assumed that using flooding will deliver control packets to the BS. However, flooding waste valuable communication and energy resources by sending redundant information throughout the network [8]. The large amount of power required by flooding causes a prohibitively short network lifetime, which makes it difficult to apply basic flooding protocol to real WSNs [9].

In addition to the previously mentioned problems and up to our best knowledge, all the clustering protocols that were proposed so far use the energy consumption model suggested in [6], this energy model is fundamentally flawed for modelling radio power consumption in sensor networks. It ignores listening energy consumption, which is known to be the largest contributor to expended energy in WSN.

In this paper, a centralized clustering protocol is introduced. The best cluster heads will be selected based on number of neighbours, residual energy and number of times the node was selected as a cluster head before. Furthermore, a method to solve the problem of transferring control data to the BS is introduced. In this method, the BS starts to construct a tree that connects all the nodes in the network. This tree will be used for transferring the control data in two ways, from BS to network nodes and from the network nodes to the BS. This tree will also be used to construct the inter-cluster communication route. To minimize the hot-spot problem, the tree construction phase is triggered periodically by the BS and each node uses a greedy approach to select its next hop in the tree. In this paper, we also investigated using the realistic energy consumption model defined in [10] in our proposed protocols.

The remainder of this paper is organized as follows. The proposed system model is presented in section 2. Section 3 provides a detailed description of the proposed Control Data Dissemination Protocol (CDDP). Section 4 describes the proposed clustering protocol. Simulations' results will be illustrated in section 5. Finally, we concluded our work and highlighted a few future directions in Section 6.

2. THE SYSTEM MODEL

2.1. The WSN Model

We assume that a set of sensor nodes is randomly and densely spread throughout a two-dimensional square field. Sensor nodes are location-unaware and non-rechargeable.

2.2. The Energy Consumption Model

In our approach we used the realistic energy consumption model proposed in [10] which is based on the characteristics of the Chipcon CC2420 radio transceiver data sheet.
The total energy consumed by node i, E_i, is calculated as follows:

$$E_i = \sum_{statej} P_{statej} \times t_{statej} + \sum E_{transitions} \quad (1)$$

where the index $statej$ refers to the energy states of the sensor: sleep, reception, or transmission. Then, P_{statej} is the power consumed in each $statej$, and t_{statej} is the time spent in the corresponding state. Moreover, the energy spent in transitions between states, $E_{transitions}$, is also added to the node's total energy consumption.

3. CONTROL DATA DISSEMINATION PROTOCOL

The CDDP (Control Data Dissemination Protocol) consists of the following phases:

3.1. Neighbour Discovery

In this phase, each node in the network broadcast a HELLO packet that includes its ID, remaining energy and number of times this node was selected as a relay node in the inter-cluster route to the BS. Each node that receives this HELLO packet will update its neighbour table with the control data included in the packet a long with the Received Signal Strength Indicator (RSSI) value in the received packet.

3.2. Main route establishment

The BS starts the process of creating a main route that connect most of the nodes in the network. The BS goal is to choose the best next hop based on the following attributes:

A) Link Quality

The link quality of neighbour i is determined by the following function:

$$LQ_i = \frac{MaxRSSI}{RSSI_i} \quad (2)$$

where $MaxRSSI$ is the maximum RSSI received by the BS. $RSSI_i$ is the RSSI value of neighbour i. RSSI is a register in the CC2420 transceiver that is used to compute the receiver-side link quality [11].

B) Residual Energy

A neighbour node with a higher level of energy is a better candidate to include in the route. The following function is used to balance the energy consumption among all the network nodes:

$$EN_i = \frac{remEnergy_i}{MaxRemEnergy} \quad (3)$$

where $MaxRemEnergy$ is the maximum remaining energy among all the neighbours and $remEnergy_i$ is the remaining energy of neighbour i.

C) Relay node Frequency

This function is used to make sure we minimize the hot-spot problem and therefore maximize the network lifetime. It is expressed by the following function:

$$RN_i = \frac{1}{numOfSelAsRN_i} \quad (4)$$

$numOfSelAsRN_i$ is the number of times neighbour i was selected as a relay node in the inter-cluster communication route.

The BS uses all the criteria defined above and employs a greedy approach when selecting its next best hop using the following function:

$$wght_i = w_1 \times LQ_i + w_2 \times EN_i + w_3 \times RN_i \quad (5)$$

w_1, w_2 and w_3 are weight coefficients. In our simulations we used $w_1 = w_2 = w_3 = 0.33$.

The BS selects the node with the highest weight to include in the main route by sending a CONNECT packet to that node, this packet also contains the BS ID. A neighbour node which receives a CONNECT packet will perform the following steps:

1. Update its next Hop towards the BS (BS-nxtHop) using the CONNECT packet sender ID.

2. Choose its best neighbour according to (5). To avoid loops, the node should select the best neighbour that is not in the route constructed so far.

3. Update its next Hop towards the route end (RE-nxtHop) to be equal to its best neighbour determined in step 2.

4. Change its status to reflect its in route.

5. Update the CONNECT packet received from the BS with its ID and resend the packet to its next best neighbour determined in step 2.

These steps will be repeated by each node that received the CONNECT packet until the last node in the route finds that all of its neighbours are in the route, then it will mark itself as the route end.

3.3. Branch routes establishment

In this phase, each non-connected node will broadcast a request to connect (RTC) packet to the main route which was constructed in the previous phase and will mark itself as a branch node. After a non-connected node broadcast the request, it will wait for an amount of time enough to receive replies from other nodes. A node in the main route, who receives the RTC packet, will reply with a packet containing its own ID, remaining energy and number of times it was selected as a relay node. After the time expires at the sending node, it will perform the following steps:

1. Checks the different nodes who replied to its RTC, if any, and calculates their weights according to equation 2. Then, it will select the best node and send a CONNECT packet to that node.

2. A node who receives a CONNECT packet from a branch (non-connected node) will change its status to reflect it has a branch and will add the node ID to its branch vector. It will also confirm by sending a CONFIRM packet to the branch node.

3. The branch node upon receiving the CONFIRM will update its next hop to BS (BS-nxtHop) to the sender of the CONFIRM packet and will change its status to reflect it is in route now.

If the branch node didn't receive any reply from a node in the main route after the time expires, it will resend the RTC packet and wait again. A branch node who didn't receive a RTC packet declare itself as a leaf node. The protocol finishes when all the nodes are connected to the route. At the end of this phase, each node will have two next hop variables: BS-nxtHop is the next hop towards the BS and RE-nxtHop is the next hop towards the route end. Leaf nodes and route end node have no RE-nxtHop while BS has no BS-nxtHop.

3.4. Control Data Dissemination

In this section we will describe how the control data are transferred to and from the BS using the tree constructed in phases 1-3.

A) Control data transfer from nodes to BS

The leaf nodes start transferring their control packets first, each leaf node sends its control data to its BS-nxtHop, which in turn will forward the control data until it reaches a node in the main route which will aggregate all the received packets from its branches and its own control packet into one packet. Many leaf nodes may share the same BS-nxtHop, which will cause collisions if they transmit their control data at the same time. To solve this problem, the leaf nodes transmit their control data at random times.

After a reasonable amount of time, enough to send the control packets from branches to main route, the node at the route end will start sending its control packet towards its BS-nxtHop. When this packet reaches the BS-nxtHop, the BS-nxtHop will append its control packet and this process continue until the aggregated control packet that contains all the control data packets reaches the BS.

As the control packets traverse the route, the number of hops from each node to the BS is updated, this information will be used later to construct the inter-cluster communication route.

B) Control data transfer BS to all other nodes
When the BS finishes its centralized processing, it aggregate all the control packets it wants to send to the network into one packet. This packet will be sent to its next hop towards the route end using the RE-nxtHop determined in phase 2. After the control packet reaches the route end, the nodes in the main route which have branches will broadcast this control packet to its branches which in turn will broadcast it again if they are not leaf nodes.

4. THE PROPOSED CLUSTERING PROTOCOL

In the proposed clustering protocol, the network operating time is divided into rounds. Each round consists of the following two phases:

4.1. The Set-up Phase

The network is configured in this phase. The best CHs and relay nodes will be chosen by the BS. The set-up phase consists of the following steps:

A)Control Data Transfer: The CDDP protocol described in section 4 is executed to form a route to transfer the control data to the BS. While the aggregated control packet transfers the network to the BS, the number of hops to the BS is updated for each node in the network.

B)Cluster Head Election: The BS goal is to choose the best CHs that maximize the following functions:

1. **Remaining energy level:**
The remaining energy of a sensor node could be a criterion for selecting CHs since a node with a better battery life is a better candidate for the cluster management and the data aggregation. The BS uses the following function:

$$ENR_i = \frac{energy_i}{MaxEnergy} \quad (6)$$

where $energy_i$ is the residual energy of node i. $MaxEnergy$ is the maximum residual energy among all the nodes.

2. **Number of neighbours:**
The more neighbours a sensor node has, the better is its connectivity and so the more it can be an appropriate candidate as a CH. The following function is used to express a node degree of connectivity:

$$NBR_i = \frac{numNbrs_i}{MaxNumNbrs} \quad (7)$$

where $numNbrs_i$ is the number of neighbours of node i. $MaxNumNbrs$ is the maximum number of neighbours among all the nodes.

3. **CH frequency:**
To balance the load between different nodes, the role of CH should be rotated among all the network nodes. To ensure this, the BS uses the following function:

$$CHF_i = \frac{1}{CHFrq_i} \quad (8)$$

where $CHFrq_i$ is the number of times node i acted as a CH.

The BS then calculates the final weight for each node i using the following equation:

$$wght_i = w_1 \times ENR_i + w_2 \times NBR_i + w_3 \times CHF_i \quad (9)$$

where w_1, w_2 and w_3 are weight coefficients. In our simulations, we used $w_1 = w_2 = w_3 = 0.33$ to get a good balance between the different parameters.
The CH choose the highest weight node as the first CH. When choosing the next best CH, the BS chooses the next best CH under the constraint that it is not a neighbour of the first CH, this is to make sure that the selected CHs are evenly distributed in the network. This process of adding the best CH that is not a neighbour of the previous CH continues until the predetermined number of CHs are elected.

C)Inter-cluster route construction: For each selected CH, the BS searches among the selected CH neighbours for the neighbour with the minimum number of hops to the BS. The node with the minimum number of hops becomes a relay node for that CH. For the selected relay node, the process of building the route is repeated recursively until a complete route is formed from the CH to the BS.

D)Configuration announcement: The BS then forms one aggregated control packet that contain all the CHs and relay nodes along with their next hop toward the BS for the inter-cluster communication route. This packet is sent to the network using the technique described in section 4. When a node that will become a CH or a relay node receives this packet, it changes its status accordingly.

E)CH Announcement and Cluster formation: A CH will announce its role by broadcasting an announcement (ANNC) packet. a non-CH node who receives this announcement will store the CH ID and the RSSI value received from that CH. Each non-CH node will select the CH with the highest RSSI to become its CH and send that CH a JOIN packet. A CH which receives the JOIN packet adds the source of that packet into its cluster members. After a CH receives all the JOIN packets, it creates a TDMA schedule to each node in its cluster and broadcast this schedule.

4.2. The Steady-state Phase

In this phase, each non-CH node uses its TDMA schedule to transmit its data to its respective CH. When a CH receives this data it uses its next relay node to forward the data to the BS. When a non-CH node finishes its data transmission slot, it enters the sleep state to save its energy.

5. SIMULATIONS AND RESULTS

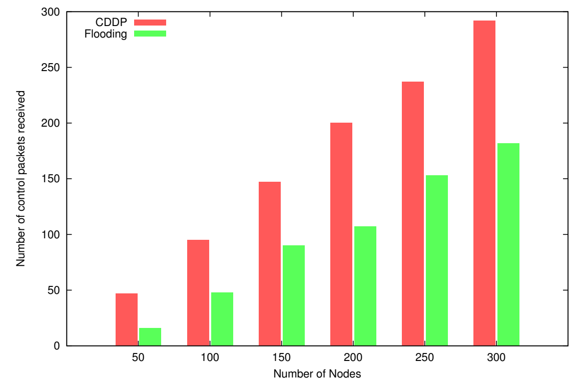
Simulations were carried on Castalia, which is based on the OMNeT++ platform and can be used to test WSN protocols in realistic wireless channel and radio models, with a realistic node behaviour. The objective of our simulations is to verify the performance of our proposed protocols.

5.1. CDDP vs. Flooding

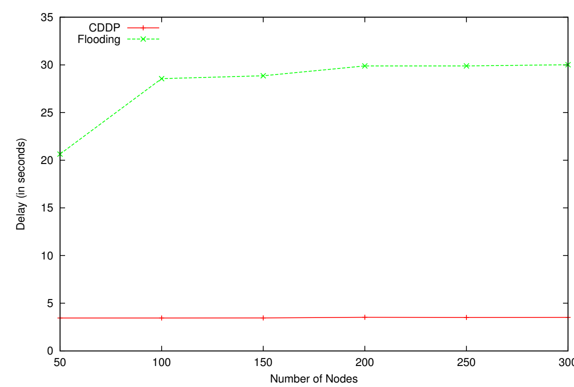
In this simulation, we compared the proposed Control Data Dissemination Protocol (CDDP) to Flooding. We tested the total number of control packets received by the BS and the delay it takes from the start of executing the two protocols until the BS receives the last control packet. The simulation used six different networks with 50, 100, 150, 200, 250 and 300 nodes in an area of 70 meters by 70 meters. The BS was located at (0,0) which is in the top left corner of the network grid. Every node sends one control packet. Figure 1-a shows that the total number of received control packets is much higher than Flooding. In the same time, in CDDP less delay is encountered in transferring the control data as illustrated in figure 1-b.

5.2. Weighted Clustering Protocol vs. LEACH

In this simulation, the Weighted Clustering Protocol was compared with LEACH. A network of 100 nodes was used in an area of 70 meters by 70 meters. The round length is 40 seconds and the number of cluster heads is 5%. Each sensor initial energy is 18720 joules. The data rate is one data packet per second. The objective of the simulation is to test the total number of data packets received by the BS and the average energy consumed per node in both protocols using variable number of rounds (from 1 to 10). Figure 2-a shows that the



(a) Total number of control packets received by the BS



(b) Delay (in seconds)

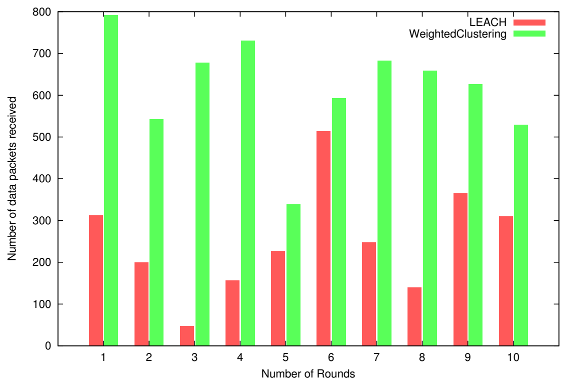
Fig. 1. CDDP vs.Flooding

weighted clustering protocol has a much higher number of received data packets by the BS and at the same time lower average energy consumption per node as illustrated in figure 2-b.

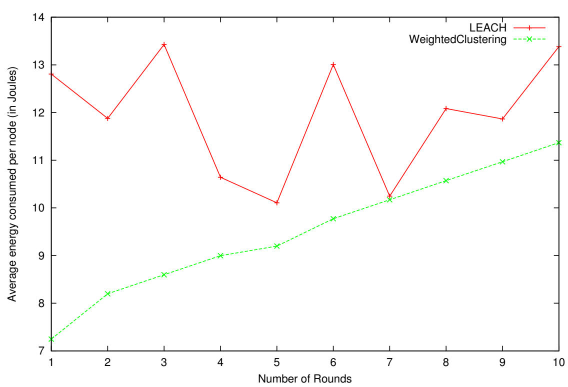
6. CONCLUSION AND FUTURE WORK

In conclusion, the proposed CDDP outperforms Flooding protocol in terms of total control data received by BS and lower delays. This is due to using a determined route for transferring the control data rather than just broadcasting as in the Flooding technique. As for the weighted clustering protocol, simulation shows a much higher performance in terms of the total data received by the BS mainly due to the use of inter-cluster communication route. Also, lower average consumed energy by the network nodes because of the node sleep scheduling.

Clustering in WSN involves many parameters to optimize



(a) Total number of data packets received by the BS



(b) Average energy consumption per node

Fig. 2. Weighted Clustering vs. LEACH

which makes swarm intelligent techniques a good candidate for finding the best cluster heads. In the future, we will extend our work by using swarm intelligent techniques in the clustering process.

7. REFERENCES

[1] Y. Yu, V.K. Prasanna, and B. Krishnamachari, *Information Processing and Routing in Wireless Sensor Networks*, chapter 1, World Scientific Pub Co Inc, 2006.

[2] A.M. Zungeru, L. Ang, and K.P. Seng, "Classical and swarm intelligence based routing protocols for wireless sensor networks: A survey and comparison," *Journal of Network and Computer Applications*, vol. 35, no. 5, pp. 1508–1536, September 2012.

[3] M.C. Arboleda and N. Nasser, "Comparison of clustering algorithms and protocols for wireless sensor networks," in *Canadian Conference on Electrical and Computer Engineering, CCECE '06'*.

[4] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*.

[5] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, 2004.

[6] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, 2002.

[7] C.J. Jiang, W. Shi, M. Xiang, and X. Tang, "Energy-balanced unequal clustering protocol for wireless sensor networks," *Journal of China Universities of Posts and Telecommunications*, 2010.

[8] V.Vasanthi, P.Nagarajan, B.Bharathi, and M.Hemalatha, "A perspective analysis of routing protocols in wireless sensor network," *International Journal on Computer Science and Engineering*, 2010.

[9] L. Zhao, G. Liu, J. Chen, and Z. Zhang, "Flooding and directed diffusion routing algorithm in wireless sensor networks," in *Ninth International Conference on Hybrid Intelligent Systems*.

[10] A. Barberis, L. Barboni, and M. Valle, "Evaluating energy consumption in wireless sensor networks applications," in *Digital System Design Architectures, Methods and Tools, 2007. DSD 2007. 10th Euromicro Conference on*, aug. 2007, pp. 455 –462.

[11] N. Baccour, A. Koubaa, L. Mottola, M.A. Zuniga, H. Youssef, C.A. Boano, and M. Alves, "Radio link quality estimation in wireless sensor networks: A survey," *ACM Transactions on Sensor Networks*, 2012.