Multihop Cluster-based Architecture for Sparse Wireless Sensor Networks

C. Cano\textsuperscript{1}, B. Bellalta\textsuperscript{1}, P. Villalonga\textsuperscript{2}, J. Perelló \textsuperscript{2}

\textsuperscript{(1)} Universitat Pompeu Fabra, Barcelona, Spain
\{cristina.cano,boris.bellalta\}@upf.edu

\textsuperscript{(2)} Centre de Recerca i Investigació de Catalunya (CRIC)
\{pancras.villalonga, josep.perello\}@cric.cat

\textbf{Abstract}—Cluster-based architectures are one of the most practical solutions in order to cope with the requirements of large-scale WSNs. In this work a new cluster-based architecture for sparse WSNs is introduced, the key issue of the solution presented is the multihop approach that is followed both in the communication from sensors to cluster-heads and from cluster-heads to the sink. Intercluster multihop communication allows the deployment of sparse WSNs in which cluster-heads can not communicate directly to the sink or to other cluster-heads. Results show that the total information forwarded decreases if compared to a flat topology allowing to increase either the size of the network or the battery duration.

\textbf{Keywords:} Wireless Sensor Networks, Cluster, Multihop

\section{I. INTRODUCTION}

Wireless Sensor Networks (WSNs) are formed by a set of nodes that gather information and forward it to a sink. They are formed by small, inexpensive and resource limited devices that can interact with the environment (sensing or actuating) and communicate in a wireless manner with other devices [1]. WSNs present a new challenge research problem due to their high flexibility to support several real-world applications making the definition of a global technical solution for all of them difficult [2].

Cluster-based architectures improve the resource allocation and reduce the energy consumption by grouping sensor nodes in clusters. Nodes in the same cluster send the information to a cluster-head that is the responsible for forwarding the information to the sink [3]. There are different architectures proposed in order to design cluster-based WSNs. LEACH [3] is one of the most studied protocols for WSNs, in these networks the sensor nodes communicate directly (single-hop) to the cluster-head, and the cluster-heads route the intercluster data to the sink. LEACH considers that all nodes have the same capabilities but they assume different roles in order to maximize the network lifetime. M-LEACH [4] modifies LEACH allowing sensor nodes to use multihop communication within the cluster in order to increase the energy efficiency of the protocol. Other works define special nodes (called gateways) that are able to send the information generated inside the cluster directly to the sink [5].

This work extends the existing solutions by allowing multihop intercluster communication in sparse WSNs in which the direct communication between cluster-heads or the sink is not possible due to the distance between them. Thus, the main innovation of the solution proposed here is that the multihop approach is followed inside the cluster (messages from sensor nodes to the cluster-head) and outside the cluster (from cluster-heads to the sink using intermediate sensor nodes). Cluster-heads can also perform data fusion to the data received allowing a reduction in the total transmitted and forwarded data in the network.

This paper is organized as follows, Section II describes the architecture presented, in Section III the simulation results are analyzed and finally in the last section some conclusions are provided.

\section{II. MULTIHOP CLUSTER-BASED PROTOCOL}

\subsection{A. Scenario}

The scenario of interest is a WSN with static, homogeneous and energy-constrained sensor nodes. The network is assumed to be sparse, nodes are only capable of communicating with their direct neighbours, therefore multihop communication is needed in order to reach the rest of nodes in the network.

The intended application could be a monitoring application in which data fusion can be applied to periodical messages, data centric queries can also be sent by the sink.

\subsection{B. Network Elements and Architecture}

All nodes in the network act as a sensor nodes collecting information from the environment, apart from that they can act as a cluster-head, forwarding the information to the sink, or as bridges, interconnecting different clusters.

Each cluster is formed by a set of sensor nodes, one of them will assume the role of cluster-head and a set of them will work as bridges (the number of bridges depends on the number of cluster neighbours). Notice that this architecture allows homogeneous nodes but with different roles that could change during the network lifetime. The assignment of roles is done in the network formation phase and is allowed to change periodically.

Sensor nodes communicate to the cluster-head in a multihop way. The cluster-head stores the information it receives and performs aggregation tasks sending periodical messages to the sink. The cluster-head can inform about some parameters of
the covered area for that cluster or notify emergency messages (some value has exceeded a threshold), it can also respond to data-centric queries coming from the sink.

The network is formed by a set of clusters consisting of a group of sensor nodes and a cluster-head. All intercluster messages should reach the sink of the network. Cluster-heads will send the information to the sink by using multihop communication through the sensor nodes inside the clusters. Bridges are in charge of interconnecting two different clusters while cluster-heads should route the messages to the sink. In Fig. 1 a scheme of the general architecture is depicted.

C. Network Formation

The network formation is based on the one defined in [3] where the network lifetime is divided in rounds in order to balance the energy consumption of the nodes. The sink broadcasts a net_formation message at the beginning of each round. Each sensor node that receives the message will forward it, so all nodes in the network will receive the message. Upon a reception of the net_formation message nodes decide if they will act as a cluster-head for that round, the probability to become cluster-head depends on the number of times a node has been a cluster-head and the remaining energy [3]. Cluster-heads send a message to the sink in order to receive the cluster identifier that will serve them to configure their address and the addresses of the nodes of their cluster. Nodes that do not become cluster-heads wait for the reception of multiple cluster_formation messages and choose the cluster-head to communicate with based on the number of hops and the total energy required to reach them. If a node does not receive any cluster_formation message it will become cluster-head for that round. A node that decides to join a cluster will send a nodeid_request message to that cluster-head by specifying its hardware address in the payload. The cluster-head will reply to that message with a nodeid_response message in which the address of that node will be included.

The flow diagrams of the sink and the nodes are depicted in Fig. 2 and Fig. 3.

After the network formation, all nodes will send a neighbour_notification message to inform their neighbours about their address and cluster identifier. Nodes that receive a message from a neighbour with different cluster identifier will configure themselves as bridges and will notify the cluster-head the clusters they could reach with a bridge_notification message including this information in the payload.

D. Addressing

Addressing remains an important problem in large WSNs in which a global identifier can not be assigned to each one of the nodes in the network. Local addressing is more appropriate for cluster-based WSNs with the cluster-head responsible for assigning local identifiers to the cluster member nodes [6]. In the architecture proposed the cluster-head node assigns identifiers to the nodes inside the cluster based on the cluster identifier that is provided by the sink during the network formation. Notice that as the architecture will change during the network lifetime the address of the nodes will also change. The cluster identifier and the node identifier have been defined to be 1 byte long in order to obtain 256 possible clusters and 256 nodes inside a cluster, that leads to a maximum network size of 65,635 nodes.
E. Routing Algorithm

The routing algorithm of this architecture is based on the AODV (Adhoc On-demand Distance Vector) protocol. Each time a node wants to send a message inside the cluster it has to look up its routing table. In case the route has not been learnt before it has to broadcast a route request message that will be replied by the destination. Once the reply traverse the path intermediate nodes and the originator will store the next hop to arrive to the destination. Notice that sensor nodes will send messages to their cluster-head but not intercluster messages.

In order to establish the intercluster routes an extension of the same protocol has been designed. Once a cluster-head or the sink wants to send a message it will broadcast at intercluster level (to all the clusters in the network but not to all the sensor nodes) a route request message by sending it to all the bridges in the cluster. The destination will send a route reply and all the intermediate cluster-heads and the originator will store the next cluster identifier to arrive to the destination.

F. Packet Format

The structure of the packet format is depicted in Fig. 4, there are two extra headers, the network header is represented as the intracluster header and the intercluster header that is only present in intercluster messages.

In the intracluster header the following fields are included:
- Message ID (1 byte) (Intra or Inter Cluster message)
- Message Type (1 byte)
- Sequence Number (1 byte)
- Node Destination Address (2 bytes)
- Node Source Address (2 bytes)

Only intercluster messages contain the intercluster header with the following items:
- Message Type (1 byte)
- Cluster ID Destination (1 bytes)
- Cluster ID Source (1 bytes)

G. Medium Access Control Protocol

As all nodes communicate only to their one-hop neighbours it not necessary to use different transceivers or communication mechanisms for intra and intercluster communications. All messages could be sent using a CSMA technique with periodic listening (between sleep intervals) [7]. In order to achieve a lower delay nodes in active paths (those that have forwarded a route response message) could remain awake in order to forward messages. To simplify, the results presented in the following section have been performed using CSMA without periodic listening (nodes do not sleep).

The hidden terminal problem is specially important in sparse wireless networks as it is more probable that the neighbours of the recipient could not hear each other due to the distance between them. In this solution a mechanism to reduce this type of collisions should be used, the RTS/CTS (Request To Send / Clear To Send) procedure could be a good alternative but the extra energy consumption of using this technique should be taken into consideration.

H. Energy Consumption

The proposed architecture distributes the consumption of sending intercluster messages between several nodes in the network (not only cluster-heads) due to the multi-hop communication from cluster-heads to the sink. One could argue that nodes closer to the main cluster will deplete their battery...
before due to the high forwarded traffic. This fact could be solved by alternating the sink role during the network lifetime, therefore, different nodes in the network have to be capable of communicating the information to the remote control center. Nodes that act as sinks should have higher battery or an additional source of energy because this communication is more energy-consuming than the rest. The overhead of creating and maintaining the cluster architecture has to be considered in terms of energy consumption and message loss or delay during role changes.

III. RESULTS

The proposed architecture has been evaluated using the Component Oriented Simulation Toolkit (COST) [8] and has been compared with the results obtained in a flat topology in which all messages should reach the sink. The flat topology has been used for evaluation purposes only because the use of other architectures like LEACH or M-LEACH would not be possible in sparse WSNs. Two multihop cluster-based topologies have been considered, in the former the number of nodes inside a cluster has been fixed to 9 (named MHC-9) and in the second to 25 (called MHC-25) including the cluster-head. The topology is not changed during all the simulation and in order to simplify the evaluation, nodes are not energy-constrained and periodic listening is not used.

A typical application has been simulated in which nodes send messages with \( \lambda_{Node} = 1/5 \) packets/s and packet length \( L_{Node} = 480 \) bits. Cluster-heads send an aggregated packet of \( L_{CH} = 560 \) bits with \( \lambda_{CH} = 1/5 \) packets/s, then they should store only the information received in this interval and perform data fusion.

In Fig. 5 the total information transmitted in the network is shown. Each node sends data packets with the same frequency \( (L_{Node} = L_{CH}) \) but in the cluster-based topology the cluster-head payload is higher than the payload of sensor nodes, this is the cause of the slightly higher load in the cluster-based topology. The total information sent in bits/s could be expressed as shown in Eq. 1 where \( N_{Node} \) is the number of sensor nodes and \( N_{CH} \) is the number of cluster-heads. Notice that in the flat topology all nodes are sensor-nodes while in the cluster-based architecture some of them act as a cluster-heads.

\[
T_p = (\lambda_{Node} \times L_{Node} \times N_{Node}) + (\lambda_{CH} \times L_{CH} \times N_{CH}) \quad (1)
\]

The total information received at sink is depicted in Fig. 6. It could be seen that with the cluster-based topology there is a reduction in the number of messages that arrive to the sink, only one message per cluster-head (then, it decreases with the cluster size) compared to one message per node in the flat topology. Messages that arrive to the sink in the cluster-based topology are longer than the ones in the flat topology but the total information received is still noticeably smaller leading to a reduction of the total energy spent by the nodes in the network. The saturation point could also be seen in this figure, notice that in a flat topology the maximum number of nodes without losses is 200 while in the multihop cluster-based topology it increases to 700 and 2000 with 9 and 25 nodes per cluster respectively.

The total number of bits forwarded in the network is plotted in Fig. 7. In the flat topology all messages generated by the sensor nodes should reach the sink in a multihop way but in the cluster-based topology these messages will arrive to the cluster-head and only one message per cluster should be routed to the sink. The reduction of the total bits forwarded becomes more noticeable with the number of nodes in the network. The MHC-25 topology shows more information forwarded than the MHC-9 topology due to the increase of the number of hops inside the clusters.

The number of lost packets due to full queue for the different architectures are depicted in Fig. 8. It could be seen that when there is saturation in the network the number of lost packets begin to increase. Notice that the number of intercluster messages lost is higher than the number of intracluster messages because, unlike intracluster messages, all intercluster messages should traverse the most congested clusters in the network.
It is important to consider that in the flat topology the messages lost contain only information of an individual node but in the cluster-based topology the intercluster messages lost contain information of the fusion of several individual node measurements, therefore to lose an intercluster message is more important than the loss of an intracluster message as the amount of information lost is higher.

In spite of slightly higher load, the total saved number of bits received at sink and forwarded leads to a reduction of the total energy spent in communication if the multihop cluster-based architecture is used instead of a flat topology. As a weak point the loss of intercluster messages should be considered as the information lost in these messages is higher than the information that could be lost in the flat topology in which all messages contain the information of only one node.

IV. CONCLUSIONS

In this work a new architecture for WSNs has been proposed. The network is organized into clusters and the communication inside and outside the cluster is made in a multihop way in which sensor nodes are also responsible for forwarding intercluster messages. This architecture is suitable for sparse WSNs in which the direct communication between cluster-heads is not possible. The proposed solution could be useful for monitoring and event-based applications in which not all the individual messages have to be received by the sink, cluster-heads perform data fusion and could respond to data centric queries.

The results show how the total load of the network could be reduced decreasing the delay of intercluster messages if compared to a flat topology. The total information received at sink and forwarded in the network is reduced.

V. ACKNOWLEDGEMENTS

This work has been partially supported by the European Commission, contract number COOP-CT-2006-032094 (CityBee). The authors would like to thank Steve Lean from Triteq Ltd for his useful comments.

REFERENCES