
Mudasser Iqbal, Iqbal Gondal, Laurence Dooley
GSIT, Faculty of IT, MONASH University, Churchill, 3842 Australia

Abstract – This paper proposes a cross-layer protocol for energy-aware routing in wireless sensor networks. The protocol combines the energy depreciation rate, node distance and neighbourhood information from physical layer together with TDMA schedules from the MAC layer and also network life requirements from the application layer, to effectively determine the most efficient routes to the base station. This cross layer efficiency measure is then used to form dynamic clusters that adapt to changing traffic and energy conditions so assisting to both control the transmission power and schedule the sleep-wake cycles of nodes for better energy utilisation. The proposed protocol is a recursive aggregation scheme that transforms a network-wide routing dissemination problem into a single-hop query protocol that aids nodes in making multi-hop routing decisions based solely upon the information provided by their single-hop neighbours. Results confirm that the proposed technique balances the load on forwarding nodes, adapts the MAC layer precisely to the routing layer and minimizes data delivery time for increased traffic and large scale networks.

I. INTRODUCTION

The challenge in resource constrained ad hoc networks like dense wireless sensor networks is to disseminate network state information effectively, with the most desirable performance objective being to minimize the consumed energy while maximizing the network capacity. Due to the often hostile nature of such networks, a related challenge is to determine the most reliable routes so data from the sensor nodes can be reliably delivered to a base station.

Traditionally, these problems are addressed by splitting the network into clusters, with each cluster having a clusterhead (CH), with data being delivered to the CHs via either single or multi-hop communication [1, 2]. The CHs are given the role to aggregate the data received and then deliver it to the base station. As the network usually comprises of numerous tiny sensor nodes, the route from a sensor to a destination will inevitably involve in most cases multi-hop data transmission. In order to select a route to any available CH, some nodes take the responsibility of forwarding data to CHs. For energy-aware data delivery, these router nodes must know the load, link quality, energy and availability of each node on the route to the sink (clusterheads and base station). This information exchange has to be minimal, timely and meaningful so that an accurate network picture is maintained. This challenge is compounded with the many hop paths, hostile nature of traffic and resource constraints in sensor networks, where the traditional flooding (sending data to every node in the neighbourhood), gossiping (randomized path selection) [3], piggybacking or on demand beacon exchange schemes have been shown to be largely ineffective [4]. The two main deficiencies of these approaches that render them inadequate as a protocol for sensor networks are: i) Implosion – A node always sends data to its neighbours, regardless of whether the neighbour has already received the data from another source; and ii) Resource Blindness: Node activities are not modified based on the residual resources available.

A fundamental use of these core protocols and their variants is routing table dissemination. For example, nodes in OSPF [5] periodically broadcast their view of the network topology to their neighbours, with such protocols closely mimicking the classic flooding protocol. Using gossiping and broadcasting algorithms to disseminate information in distributed systems has been extensively explored as epidemic algorithms [3, 6] to maintain database consistency. Recently, such techniques have also been used for resource discovery in networks [7]. SPIN protocol [8] uses negotiation and resource-adaptation to overcome implosion and resource blindness deficiencies.

While clustering based techniques attempt to address the energy conservation at topology configuration level (usually as a middleware between MAC and Network layers), MAC level optimization and power control (physical layer) mechanisms [9, 10, 11] have been employed to control the sleep and wake schedules of nodes. Scheduled protocols such as TDMA are very attractive for applications in sensor networks because of their energy efficiency. Since slots are pre-allocated to individual nodes, they are collision- free so no energy wasted due to channel contention. Overhearing can also be easily avoided by turning off the radio during the free slots. The main strategy used in MAC level protocols to extend network life, has been to conserve energy by turning off those devices that either do not have scheduled packet transmissions or...
reception, which necessitates high synchronization between participating nodes.

Since current trends in ad hoc network designs are mainly based on a layered approach, their inflexibility and sub-optimality leads to overall poor performance in terms of energy conservation. The various approaches to minimize energy consumption at different layers such as data aggregation in clustering [2], power control approach at physical layer to minimize transmission energy, TDMA based sleep wake schedules at MAC layer [12] and application driven data acquisition methods, all ignore the fact that more than one layer usually participates in determining the overall energy drainage in the network. For instance, power control impacts on routing. Likewise, the power-control strategy needs connectivity information provided by the network layer while the amount of traffic routed by network layer defines the traffic load on a node, which in turn influences the MAC sleep-wake schedules.

Although there are some cross-layer approaches to conserving energy [13, 14, 15], they ignore the crucial parameters from various layers that must be considered while designing a cross-layer solution. We observe that parameters such as the network life requirement from application layer, the TDMA slot allocation from MAC layer, the energy depreciation rate of nodes and neighbourhood nodes information from physical layer, all play a vital role in determining the load on each node and the routes traffic takes. In turn, a combination of these parameters determines the TDMA schedules that best suit a given traffic load and energy degradation profile. Moreover, the routing decision is dynamically based upon the historical pattern of load and energy depreciation, rather than an instantaneous information about these parameters [16].

These observations provided the motivation to investigate a cross-layer protocol for energy-aware routing, with the key features being:

1. It combines parameters from application, network, MAC and physical layers to determine the most efficient route to the base station.
2. It allows nodes to make local decisions for selecting best routes to many-hop away sink nodes. The network routing information is recursively aggregated and made available to the immediate neighbourhood of each node. This idea exploits the capability that any node can be used as an on-demand gateway to support data delivery to a many hops away sink.
3. It balances the load by exploiting the historical MAC schedule of nodes, the energy depreciation rate, the distance between nodes and availability of multiple sink nodes.

The rest of this paper is organized as follows: Section II details the proposed cross-layer protocol, while Section III presents the results to highlight the performance benefits of the protocol. Some conclusions are provided in Section IV.

II. THE PROPOSED PROTOCOL

A. The System Model

The network comprises randomly distributed sensor nodes in 2-dimensional space with a single base station. The nodes gather different types of data (temperature, humidity, motion and patient monitoring for example) to be routed to the base station. The nodes have the same maximum transmission power, which is dynamically adjustable. The objective of the proposed protocol is to organize this network in the form of decentralized clusters as shown in Fig. 1, with each CH establishing autonomous control over its proximity. In this clustered network, nodes are categorised on the basis of their communication capability. A node that is able to communicate directly with the base station is called a potential sink node, while a node which can communicate with any sink node either through single or multi-hop transmissions, is classed as a potential gateway. A potential sink node becomes a sink node (S-node) if it agrees to become a destination for traffic within the network. All the data from the network is delivered to a sink node, which in turn transfers it to the base station. It should be recalled that there can be multiple nodes able to directly communicate with the base station so there will be multiple sink nodes with sensor nodes having no priority to send the data to a particular sink node. The proposed protocol presented in this paper is an anycast routing protocol that has the flexibility to select the best sink node in various parts of the network, such that the transmission energy is minimized and traffic load is balanced to help in evenly deprecating energy amongst the sink nodes. Likewise, a potential gateway node becomes a gateway node (G-node) if it agrees to undertake this role. The responsibility of a gateway is to forward traffic from those nodes which are unable to communicate directly with a sink node. All nodes other than sinks and gateways are classified as client nodes (C-node).

The S/G-nodes form clusters. An S-node cluster comprises of the S-node acting as the S-clusterhead and some S/G/C-nodes as the S-cluster members. Similarly, a G-node cluster is formed by the G-node acting as the G-clusterhead and some G/C-nodes as their members. In this way the network is composed of distributed S/G-clusters. Each S/G-clusterhead regulates the data transmission from the cluster members by constructing the TDMA schedule and allocating transmission slots to the cluster members. The next section details CH selection and TDMA scheduling.

A basic anycast routing infrastructure is assumed to exist. If $r$ is the transmission range of a node $n_i$ at maximum power, and $\text{dist}(n_j, n_i)$ is its distance from another node $n_j$, then the complete neighbourhood $H(n_i)$ of a node $n_i$ is given by:

$$H(n_i) = \bigcup_{n_j \in N, \text{dist}(n_j, n_i) < r} \{n_j\}$$

(1)
Node Efficiency (NE): This measure combines parameters from the physical, MAC and application layers to determine the capability of a node to become a G-node, if the node has declared itself as a potential G-node. It is defined in terms of the remaining life of a node \( n_i \) and its ability to grant connections to G/C-nodes. Analytically, the efficiency measure \( \eta \) of a node \( n_i \) at time \( t \) is given by:

\[
\eta_t(n_i) = V_t(n_i) \times \xi_t(n_i)
\]  

(2)

where,

\[
V_t(n_i) = 1 - \frac{\sum_{j=1}^{T} x_j(n_i)}{k \times X_j(n_i)},
\]  

(3)

\( X_j \) is the information obtained from the MAC about the average number of total slots in a TDMA frame of node \( n_i \) at time \( j \), where \( x_j \) is the number of allocated slots at this time. \( V_t(n_i) \) is a measure of connection availability with node \( n_i \) and is not merely the snapshot of current TDMA allocation by node \( n_i \). Instead, it is calculated in terms of a one-step frame allocation prediction which is statistically averaged over the \( k \) past time instances to predict the next likely ratio of occupied slots to the total frame capacity. The higher \( V_t(n_i) \), the more empty slots a node has (lesser routing load) and thus becomes a stronger candidate to be a G-node.

The second term in equation (2) is given by:

\[
\xi_t(n_i) = \frac{k \cdot E_t(n_i)}{L_h \sum_{j=1}^{T} \sqrt{(E_j(n_i) - E_{j+1}(n_i))^2}},
\]  

(4)

where, \( E_t(.) \) is the physical layer parameter that gives the information about the residual energy of node \( n_i \) at a particular time. \( \xi_t(n_i) \) estimates the remaining life of node \( n_i \), defined in terms of the length of time for which the node can be sustained by the residual energy by calculating the normalized rate of energy depreciation over the past \( k \) time instances. It should be noted that the required network life \( L_h \) is an application layer parameter that is combined with the physical layer energy measure to give the estimated remaining life of a node. In calculating NE, the historical TDMA slot allocations and energy depreciation profile of a node helps to obtain the actual routing load picture and energy drainage.

Route Efficiency (RE): When there are multiple G-nodes available for a G/C-node, having similar efficiencies (NE), then deciding which one a node should forward its traffic to poses a conflict. For this purpose, another efficiency measure is proposed that utilizes the received signal strength (RSS) measure from the physical layer to determine the distance from each of the available G-nodes. The RE measure determines the efficiency of a route between two neighbouring nodes based on the NE of the G-node (destination) and its distance from the G/C-node (source) that intends sending some traffic. If \( n_i \) is the source and \( p \) is a neighbouring G-node then the RE between these two nodes at time \( T \) is given by:

\[
\lambda_r(n_i, p) = \eta_r(p) \frac{1}{c}
\]  

(5)

The RSS measure from the physical layer is used to represent the distance between two nodes because of its implementation simplicity. However, the signal strength measure is not converted into an actual distance using the underlying channel propagation model as only a comparative measure rather than the actual distance is required.

The capability of a node is defined as its ability to connect to a sink node, either via single hop or via many hops, and its consent to serve as a gateway node for an immediate neighbor. Capability analysis involves periodically updating the capability of a node, which is fixed to be every 10 frames. A process called on-demand propagation of capability is defined as: in response to a request from an immediate neighbor node \( i \), undergoing its own capability analysis, a node \( j \) sends its capability information to \( i \), if \( j \) decides to establish a trust with the requesting node \( i \) and can work as a reliable gateway for it.

B. S/G-clusters Formation

As the proposed protocol is anycast neighbourhood routing, this is formally defined as: Given a node \( n_i \), its neighbourhood \( H(n_i) \) and a set \( L \) of sink nodes. Node \( n_i \) should be able to get its data delivered to at least one S-node \( n_j \) in \( L \), merely by sending its data to only and at least one gateway node in \( H(n_i) \), irrespective of the number of hops between \( n_i \) and \( n_j \).

The core idea of our neighbourhood routing is to enable any node to have at least one reliable G-node in its one-hop neighbourhood referred to as immediate neighbourhood. This enables each node to query only its immediate neighbourhood and find out which node can be trusted as a gateway. Once a trusted gateway is found, the G/C-node in need of sending data to sink joins the G-cluster of the trusted G-node and hands over its data to the gateway. The G-node, having agreed to take the responsibility, considers the data sent by the G-cluster member node, as its own data. This process continues until the data reaches the immediate neighbourhood of an S-node, from where, the data is then delivered to sink. The G-node in the immediate vicinity of the S-node sends a request to the S-node for data transmission and if approved, the G-node joins the S-cluster. The formed S/G-clusters thus remain intact up to the next capability analysis cycle after which they are dynamically changed based upon the route efficiencies between the S/G-clusterheads and cluster members. The cluster updating process is explained in next section.

Effectively, this neighbourhood-clustering approach

---

---
transforms the problem of routing to the sink node using global knowledge of the network routes to discovering reliable gateways only in the immediate neighbourhood. Consider a wireless sensor network as a graph \( G = (V, E, L) \), whereby \( L \) represents a set of sink nodes, \( V \) is set of and an edge \((u, v)\) is in set \( E \) if nodes \( n_u \) and \( n_v \) are immediate neighbours of each other such that \( n_v \in H(n_u) \) and \( n_u \in H(n_v) \). A typical single hop and multihop routing scenario is shown in Fig. 2. In this case, sensor nodes \( L \) and \( M \) are in the immediate neighbourhood of sink node \( S \), while nodes \( N \) and \( O \) are two and three hops away from \( S \) respectively. If it is assumed that \( S \) is the only sink node in the network, then obviously, node \( O \) needs to know the availability of both nodes \( N \) and \( M \), along with the status of routes \( y \) and \( x \). The situation is further complicated with the larger number of hops between \( O \) and \( S \).

In Fig. 3, a methodology by which this multiple hop routing scenario is transformed into a single hop by recursive aggregation of the network state information, to enable \( O \) to decide which route to send its data. If nodes are represented by the vertices of the graph and edges represent the route efficiency between two nodes, then:

\[
\begin{align*}
\text{Path}_{(M,S)} &= x \\
\text{Path}_{(N,S)} &= y + \text{Path}_{(M,S)} \\
\text{Path}_{(O,S)} &= z + \text{Path}_{(N,S)} \\
\text{Path}_{(P,S)} &= q + \text{Path}_{(O,S)} \\
\end{align*}
\]

where \( \text{Path}_{(i,j)} \) defines the route from \( i \) to \( j \). These route equations are then translated into \( RE \) measures.

(6) aggregates the efficiency measures from the sink node to the immediate neighbourhood of a node, which then gives the clear picture of the route and nodes from source node to sink, irrespective of the number of hops. The next sections detail how this information is propagated and updated in an energy efficient manner using only single hop ‘hello’ protocol.

C. S/G-clusters Maintenance

The protocol works as follows: Periodically, each node carries out its capability analysis. For this purpose, each node maintains a Capability Descriptor Routing Table (CDR), as shown in Table I, which contains information about the immediate neighbourhood nodes which can be trusted as reliable G-nodes. A node willing to transmit data, or just for periodic cluster maintenance, adjusts its transmission power to maximum and requests the capability of each immediate neighbour by broadcasting a ‘hello’ message. In response, each node willing to work as a gateway sends a packet comprising fields other than \( RSS \) and Actual Route Efficiency (ARE) in \( CDR \).

The first field is the node \( id \) which is unique within the neighbourhood. Hop to sink (HTS) is the number of hops that a G-node in the immediate neighbour is away from an S-node. The G-node does not send its efficiency measure, as defined in equation (2), instead it send the two bounds that govern the minimum and maximum level of quality of service in terms of \( RE \), that the requesting node can expect. In order to define the limits, consider a node with communication radius \( R \). It is clear that the \( ARE \) can only be calculated by the requesting node, using (5), upon receiving the response and estimating the distance in terms of \( RSS \). However, \( ARE \) always lies between the two \( RE \) limits which are effectively defined by \( R \). Analytically, if \( n_i \) is the node undergoing capability analysis and node \( n_j \) is doing on-demand propagation of capability, both separated by a distance \( c \), then \( n_i \) is guaranteed minimum routing service if \( c < R \). On the other hand, a \( c = 0 \) results in maximum guarantee of service. The two limits on \( RE \) are given as:

\[
\lim_{c \to R} \lambda_c(n_i, n_j) = \frac{\eta_c(n_i)}{R},
\]

and

\[
\lim_{c \to 0} \lambda_c(n_i, n_j) = \eta_c(n_j),
\]

\[
\Rightarrow \frac{\eta_c(n_i)}{R} \leq \lambda_c(n_i, n_j) \leq \eta_c(n_j)
\]

Since \( ARE \) always lies between these limits, propagating limits on \( RE \) instead of absolute \( RE \) makes the system robust and self correcting, in case there is some error in RSS calculation. Upon receiving the neighbourhood nodes response, the requesting node updates its \( CDR \) table.

D. Best Clusterhead Selection

The number of entries in \( CDR \) reflects the number of reliable G-nodes available in the immediate neighbourhood of the node. Handling over data to anyone of them guarantees the delivery of data to at least one of the sink nodes available in the network. The key objective however, is to extend the network life by selecting the most efficient routes. As alluded earlier, \( RE \) combines metrics from various layers: application (expected network life), MAC (TDMA allocated slots ratio) and physical layer (energy depreciation rate, neighbourhood, \( RSS \)); so any efficient route ensures an even drainage of energy and minimizes the maximum load on all nodes from source to sink, including the sink node. It is thus necessary that out of all the S/G-nodes available, the one that affords the most efficient route and is closest to an S-node should be selected.
To achieve this, a node $n_i$ undergoing capability analysis, after receiving capability information from prospective S/G-nodes sequentially sorts two fields in the $CDR$ table: $ARE$ and $HTS$ count. Firstly, the S/G-nodes are sorted in descending $ARE$ order to bring those nodes across the most efficient routes and closest to $n_i$ to the top in the table. Subsequently, the nodes are sorted in the ascending order of $HTS$ to prioritize nodes that are close to an S-node, with the net result being that nodes are ranked so having the maximum route efficiency, minimum distance from $n_i$ and minimum hop counts to a sink appear highest. Having sorted the table, CH selection simply involves selecting the top node in the table. Analytically, if $Q$ is the set of S/G-nodes for a node $n_i$, with neighbourhood $H(n_i)$, such that $Q \subseteq H(n_i)$, then the energy efficient route from $n_i$ to an S/G-node in the neighbourhood defines the cluster membership of node $n_i$. It is given by:

$$\text{Route}(n_i, S / G) = HTS_{\text{min}} \left( ARE_{\text{max}} \{Q\} \right)$$  \hspace{1cm} (8)

If the node $n_i$ is in the immediate neighbourhood of one or more S-nodes, the entries in the $CDR$ may include one or more S-nodes along with some G-nodes. In this case, the sorting process always brings S-nodes highest in the table, which stops further hoping of traffic, and node $n_i$ joins S-cluster to deliver the data, once the data has arrived in the immediate neighbourhood of at least one S-node.

CH selection involves the calculation of best outbound route using measures from various layers simultaneously. As a result, once a node joins an S/G-cluster, it no longer transmits data at its maximum power. Instead, the power level is adjusted according to the $RSS$ distance measure to the selected S/G-node. As the selection process prefers the most efficient S/G-node which is at minimum distance from node $n_i$, the transmission power at physical layer is adjusted at a minimum, thereby reducing the energy required for data transmission. This helps in minimizing redundant broadcasts, contention and collisions in the neighbourhood. Traditionally, minimizing the transmission power between every two hops on the route to the sink node is achieved at the expense of maximizing the number of hops to the sink node [13]. In contrast, our approach utilizes parameters from multiple layers to simultaneously minimize both the hop counts to an S-node and the transmission power between every two hops on the route to a sink node.

**E. Sink node to Base Station**

Once the data has been delivered to S-nodes, it is aggregated and transmitted to the base station. It has been shown [1] that a multi-level hierarchy of clusters conserves energy more than a two level hierarchy. Fig. 4 illustrates a multi-level hierarchy in which the S-clusterheads form the so called super-clusters. Like a normal cluster, a super cluster also has a super-CH which is responsible for data aggregation and transmission to the next higher level – the base station in our case.

While in contemporary algorithms [1, 2], the decision to have super-clusters in the network is made at the design level with the clustering protocol tuned accordingly, the new approach is more dynamic and flexible. As every node in the network is treated identically, the clusters are formed based on the route and power efficiency gains. Unless a node joins a cluster, the decision about where to forward the traffic is left to the node, which makes decision in favour of the most reliable next hop S/G-node.

A base station is considered one of the sink nodes (S-node) from the network. It is evaluated on the basis of node and route efficiencies in the same manner as the other sink, gateway and client nodes. An S-node intending to transmit data to the base station updates its $CDR$ table that might include other S-nodes and G-nodes, along with the base station itself. The clusterhead (next hop in the route) selection process (Section II-D) prioritizes those S/G-nodes that minimize required transmission power and maximize the balance of traffic and even energy drainage, an S-node may select another S-node as the next-hop to form a cluster (super-cluster). Therefore, it is not a static design level decision in our protocol that an S-node must always form a super-cluster, or must always transmit data directly to the base station. The decision is adaptive and is made dynamically based on the network condition. Fig. 5 illustrates this phenomenon where in Fig. 5(a) an S-node $p$ forms a super cluster with S-node $q$, which in turn acts like a gateway node for $p$ to deliver data from $p$ to the base station. Fig. 5(b) in contrast shows that after the next capability analysis cycle, the same S-node $p$ decides to abandon S-node $q$ super cluster and transmits data directly to the base station since this is far more energy efficient.

**F. Recursive Aggregation:**

As yet we have not discussed how our protocol recursively aggregates route information to bring in the immediate neighbourhood of each node. We have discussed about on-demand propagation of capability, which occurs in response to a request. To propagate each node’s capability, they need to maintain their own capability field updated which includes:

a. Calculating $RE$ limits in (7) using energy depreciation rate and pattern of load.

b. Updating own $HTS$ and $RE$ limits by recursive aggregation.

Step a) involves each node monitoring energy depreciation rates and its load profile which does not require any communication with other nodes. Step b) involves updating the $RE$ limits from step a) in such a way that the most efficient path to the sink node is reflected in the $HTS$ and $RE$ of the node.
With the low-duty-cycle operation, nodes must delay sending a packet until the next listen period of a destination, which increases latency. In addition, by limiting the opportunity to content for the channel, throughput is reduced to one message per frame. These delay costs accumulate at each hop of a multi-hop network, so to optimally reduce this delay an adaptive listening strategy is proposed whereby rather than letting the C-nodes wait until their next scheduled transmission interval, the G/S-clusterheads do not sleep immediately once the allocated slots in a frame have been handled. Instead, a CH remains active for one extra slot to allow any node that has already utilized its allocated time slot and needs to transmit data urgently which is referred to as an extended-active slot (EAS). If a G/S-clusterhead receives an RTS packet while in EAS, the requesting C-node is served. The EAS is extended to the next time slot and the clusterhead remains available for any urgent data transmission request. This process continues until an EAS is passed without receiving any RTS request, after which the clusterhead turns off its radio and switches to sleep mode. This allows immediate contention for the channel, either by another node with data to send or for the next hop in a multi-hop path. With adaptive listening, the overall multi-hop latency is reduced by at least 50%.

It should be noted that the MAC layer is affected by the cluster membership decisions by individual nodes. Previous sections have elaborated that cluster membership decision are based on parameters taken from the application, MAC and physical layers. In turn, the MAC schedules formed as a result of cluster memberships affect the energy depreciation rate (physical layer), the neighbourhood of a node (nodes may die due to loss of energy), and the routing load on each S/G-node (network layer).

### III. RESULTS AND ANALYSIS

Extensive numerical results were obtained to evaluate the performance of the proposed model. The model was implemented in a distributed manner for different load profiles and node densities. The radio consumes 330 mW for idle listening, 346 mW for receiving and 462 mW for sending. Each node has limited power and we initialize the amount of energy for each node with 100J. The channel data rate is 150 kbps and the radio transmission range is 50 meters. The simulation area is 1000x1000 meters square. We vary the density of the network by changing the total number of nodes within the simulation area. For ease of comparison, we name our protocol as ENR (for energy-aware neighbourhood routing). We compare the ENR protocol with the distributed coordination function (DCF) of IEEE 802.11, with respect to their capability to adapt the MAC layer according to the changes in routing load. DCF has been a popular MAC protocol for ad hoc wireless network research community due to its simplicity, commercial availability, and robustness to the well-known hidden terminal problem. We also evaluated the performance of ENR in terms of the load balancing in the
network and packet delivery time by varying network size and also varying the amount of traffic. The results were compared with the traditional source routing, limited flooding, Closest Node First (CNF) clustering and Maximum Energy First (MEF) clustering approaches. The protocol was coded in the Microsoft C# language as a discrete event simulator.

Source routing suggests promiscuous listening to learn new routes. In order to be used in case of many-hop routing scenarios, each node needs to inquire multiple nodes on the various possible routes, before a best route is selected. On the other hand, restricted flooding restricts the amount of broadcast needed to maintain routing information, based on some predefined criteria, like geography. In the CNF approach, nodes form both static and dynamic clusters with the closest CH while the MEF approach lets nodes become members of the cluster that has maximum residual energy.

Figures 6 and 7 analyse the theoretical performance of ENR, with the source and restricted flooding routing methodologies. Fig. 6 shows time taken for a packet to be delivered to any sink node, while varying the average number of hops to any sink node. It can be seen that because of many-hop snooping, source routing requires much longer time for getting the feedback from various nodes and then selecting the route. The packet delivery time is observed to be linearly decreasing with the reduction in the number of hops to sink nodes. While restricted flooding performs better than source routing, and it converges to reduced delivery as the number of hops reduce, however, ENR protocol because of its single-hop snooping requires the least amount of inquires. This results in quick discovery of most efficient route, which minimizes packet delivery to sink nodes. It should also be observed that the ENR converges faster to reduced packet delivery time, as the network becomes denser in sink nodes. This effect is attributed to lesser recursive aggregation needed and reduced HTS, that maximizes the chance of getting a reliable gateway in the immediate neighbourhood, closer to sink node.

Fig. 7 plots the packet delivery time against the amount of traffic in the network. Given a particular density of sink and sensor nodes, the amount of traffic defines the chances of nodes agreeing to work as a gateway node as well. As expected, it can be analysed that in case of source routing, the packet delivery time is almost double than that of ENR. Though, restricted flooding attempts to cope with the increasing traffic, its performance entirely depend upon the criteria of restricting the control traffic. ENR performs better than both because each node is affected only by the traffic in its immediate neighbourhood. Whereas, in case of other protocols where multi-hop away nodes are also inquired for making routing decisions, the traffic load on multiple nodes causes a longer delay in response. Although generally better than the other two, the recursive aggregation process of ENR also suffers from rapid increase in packet delivery time when the traffic load reaches the request handling capacity of nodes.

We conducted experiments in order to analyse the capability of ENR to adjust the MAC sleep-wake schedules of nodes according to the amount of traffic. Fig. 8 compares the performance of IEEE DCF protocol with our proposed ENR protocol. We calculated the average frame capacity in the network and analysed the change in sleep-wake up ratio as the traffic load changed. It was found that the proposed protocol can very precisely adapt the MAC schedules to the routing load on the S/G-nodes, and turns the nodes off when the traffic is not expected to arrive. On the other hand, the standard IEEE
DCF protocol keeps the node awake for 50% time longer on the average than required.

The balance of load in the network was analysed using the standard deviation of maximum load on each CH, using the ENF, CNF and MEF methodologies. Fig. 9 shows the number of undelivered packets and the corresponding situation of the balance of load in the network. It is observed that where all the three techniques could maintain acceptable balance of load for low traffic, CNF and MEF fail to keep the balance in the event of high network traffic. This is attributed to their routing criteria, biased to short distance and maximum energy respectively. Conversely, ENR continues to maintain control over redistribution of traffic by taking into account the historical load, energy drainage and the distance while redirecting the high traffic to better routes.

IV. CONCLUSIONS

This paper has proposed a cross-layer data dissemination and routing protocol for dynamically adaptive clustering under stringent energy constraints. The model combines parameters from the physical, MAC and application layers to calculate the efficiency of outbound routes, which helps in adapting the TDMA schedules according to the routing load. Through cross-layer operation, the protocol forms energy-aware dynamic clusters that minimize transmission power and balances the traffic such that the average energy depletion is distributed throughout the network, which maximizes network life. The network-wide routing information dissemination problem is addressed through single-hop exchange of node capabilities which helps in minimizing packet delivery time due to minimal query exchange (only single-hop), and selecting the shortest and most energy efficient route as indicated by the results. The results prove the superior performance of ENR over traditional DCF, source routing, restricted flooding routing; CNF and MEF based clustering approaches, in terms of precisely scheduling the sleep-wake cycles of nodes adapting to the routing load, balancing the load across the network and minimizing packet delivery time for variable traffic loads and network sizes.

REFERENCES