Title
Topology Control Protocols to Conserve Energy in Wireless Ad Hoc Networks

Permalink
https://escholarship.org/uc/item/5sd603sk

Authors
Xu, Ya
Bien, Solomon
Mori, Yutaka
et al.

Publication Date
2003-01-23
Abstract

In wireless ad hoc networks and sensor networks, energy use is in many cases the most important constraint since it corresponds directly to operational lifetime. This paper presents two topology control protocols that extend the lifetime of dense ad hoc networks while preserving connectivity, the ability for nodes to reach each other. Our protocols conserve energy by identifying redundant nodes and turning their radios off. Geographic Adaptive Fidelity (GAF) identifies redundant nodes by their physical location and a conservative estimate of radio range. Cluster-based Energy Conservation (CEC) directly observes radio connectivity to determine redundancy and so can be more aggressive at identifying duplication and more robust to radio fading. We evaluate these protocols through analysis, extensive simulations, and experimental results in two wireless testbeds, showing that the protocols are robust to variance in node mobility, radio propagation, node deployment density, and other factors.

Index terms: Wireless sensor networks, Adaptive topology, Topology control, Energy conservation

1 Introduction

Multihop, wireless, ad hoc networking has been the focus of many recent research and development efforts for its applications in military, commercial, and educational environments such as wireless LAN connections in the office, networks of appliances at home, and sensor networks.

A number of routing protocols have been proposed to provide multi-hop communication in wireless, ad hoc networks [21, 5, 22, 20]. Traditionally these protocols are evaluated in terms of packet loss rates, routing message overhead, and route...
length [6, 18, 11]. However, since ad hoc networks will often be composed of battery-powered nodes, energy consumption is also an important metric (as suggested for future work by some researchers [18]).

For unattended sensor networks energy consumption is the important metric, for it maps directly to network operational lifetime. In order understand energy use we examined existing ad hoc routing protocols using models of energy consumption [31] and radio propagation [6] for the Lucent WaveLAN direct sequence spread spectrum radio with IEEE 802.11-1997. We first only consider energy cost due to packet transmission or reception. We studied energy consumption of four ad hoc routing protocols (AODV, DSR, DSDV, and TORA) with a simple traffic model in which a few nodes send data over a multi-hop path [34] (Figure 1). With this energy model we found that on-demand protocols such as AODV and DSR consume much less energy than \textit{a priori} protocols such as DSDV (the dark bars in Figure 1). This makes sense since \textit{a priori} protocols are constantly expending energy pre-computing routes, even though there is no traffic passing on these routes.

In fact, a major source of extraneous energy consumption is from overhearing (as previously observed in PAMAS [29]). Radios have a relatively large broadcast range and all nodes in that range must receive each packet to determine if it is to be forwarded or received locally. Although most of these packets are immediately discarded, receiving them consumes energy.

Radios consume power not only when sending and receiving, but also when \textit{listening} or \textit{idle} (the radio electronics must be powered and decoding must occur to detect the presence of an incoming packet). Research [31, 19] shows that idle energy dissipation cannot be ignored in evaluating energy use. Stemm and Katz show idle:receive:transmit ratios of 1:1.05:1.4 by measurement [31], while more recent studies show ratios of 1:2:2.5 [19] and 1:1.2:1.7 [10]. With an energy model that takes into account energy use due to overhearing and idle listening, the ad hoc routing protocols we considered consume roughly the same amount of energy (within a few percent) as shown in the gray bars in Figure 1. In the scenario with modest traffic, idle time completely dominates system energy consumption.

The great energy cost associated with idle time and overhearing suggests that energy optimizations must turn off the radio, not simply reduce packet transmission and reception. We use information from above the MAC-layer to control radio power, since the application and routing layers provide better information about when the radio is not needed.

We observe that when there is significant node redundancy in an ad-hoc network, multiple paths exist between nodes. Thus we can power off some intermediate nodes while still maintaining connectivity. For example, in Figure 2(a), only one of nodes 2, 3, and 4 is required to forward data from 1 to 5: the other two are extraneous. Uninterrupted connectivity between nodes 1 and 5 can be maintained as long as any intermediate node is awake.

The contribution of this paper is to describe and analyze two protocols that exploit density to extend lifetime while preserving connectivity. The first, Geographic Adaptive Fidelity (GAF) [35], self-configures redundant nodes into small groups

![Figure 1: Comparison of energy consumed for four ad hoc routing protocols with different energy models (left, black bars are without considering energy consumed when listening; right, gray bars include this consumption). The simulation has 50 nodes in a 1500m*300m area. Nodes move according to the random way-point model. The energy model is based on Stemm and Katz [31].]
based on their locations and uses localized, distributed algorithms to control node duty cycle to extend network operational lifetime. The second, Cluster-based Energy Conservation (CEC), follows the same principle as GAF, but eliminates its dependency on location information and uniform radio propagation. The challenges of these protocols are how to identify network redundancy, how to control the duty cycle of redundant nodes to conserve energy, and how to maintain connectivity between communicating nodes in a dynamic network when redundant nodes are powered off.

Two additional requirements are needed for successful operation in ad hoc networks. First, the protocol must be self-configuring, meaning that it must actively measure the network state in order to react to network dynamics. Second, the protocol must find redundant nodes in a distributed and localized fashion, since it is prohibitively expensive to centralize or globally distribute state in rapidly changing ad hoc networks.

In addition to demonstration of the effectiveness of these protocols, we show that topology control protocols can be designed independent of the underlying ad hoc routing protocol provided that protocol can recover from topology changes quickly. We also show the relationship between node deployment density and robustness. Finally, we evaluate GAF and CEC with simple analysis, extensive simulation, and laboratory experiments. To our knowledge, this paper is the first to present experimental results of topology control over real hardware.

2 Related Work

Reducing energy consumption has been a recent focus of wireless ad hoc network research. One approach has been to adaptively control the transmit power of the radio. LINT/LILT [25] adjusts transmit power in order to form a desired network topology, while the “lazy” scheduling of Prabhakar et al. [24] transmits packets with the lowest possible transmit power for the longest possible time such that delay constraints are still met.

However, based on the energy use study presented in the Introduction, we argue that the most
energy savings will come from turning off unused radios rather than by dynamically adjusting power (both classes of approaches can, of course, be combined for further energy savings). Such protocols can be classified into groups based on the network protocol level from which information is received to identify radios that can be turned off. GAF and CEC can receive information at or above the routing layer. We believe that it is important to get this information from a high level so that our protocols have access to useful node mobility predictions and a good understanding of node redundancy from the routing perspective. We now describe protocols that get information from various levels of the network protocol stack.

By doing energy conservation with application level information it is possible to save much more energy, yet the sacrifice is having a network with application-specific characteristics. ASCENT [8] measures local connectivity based on neighbor threshold and packet loss threshold to decide which nodes should join the routing infrastructure based on application requirements. PicoNet [3] goes a step farther by designing a system with application-specific hardware and protocols so that energy can be conserved.

Because MAC level protocols have a very small view of the network, the main approach followed by such energy-efficient protocols has been to turn off radios that are not actively transmitting or receiving packets. Because it takes generally takes time to turn the radios back on when they are needed, MAC protocols typically trade-off network delay for energy conservation. Energy-efficient MAC and routing protocols can be used together to increase energy conservation.

With TDMA protocols [23], sets of nodes take turns having their radios on and off. PAMAS [29, 30] reduces energy consumption due to overhearing by using a second radio channel to detect activity on its neighboring nodes and turning on its main radio in response to such activity. Sparse Topology and Energy Management (STEM) [28] accepts delays in path-setup time in exchange for energy savings. It uses a second radio (operating at a lower duty cycle) as a paging channel. When a node needs to send a packet, it pages the next node in the routing path. This node then turns on its main radio so that it can receive the packet. Sensor-MAC (S-MAC) [36] treats per-node fairness and latency as secondary to energy conservation. It periodically turns off the radios of idle nodes and uses in-channel signalling to turn off radios that are not taking part in the current communication.

Like energy-efficient MAC protocols, energy-efficient routing protocols periodically power off nodes; however, such routing protocols do not necessarily cause longer latency. The major difference is that energy-efficient routing only powers off the redundant nodes while energy-efficient MAC protocols periodically power off all of the nodes. This difference is due to the fact that unlike MAC protocols, energy-efficient routing protocols have access to network topology information. In addition, routing-level energy conservation protocols can use routing information to ensure that connectivity will be maintained when nodes are turned off.

AFECA [34] seeks to maintain a constant density of active nodes by periodically turning radios off for an amount of time proportional to the measured number of neighbors of a node. By following this approach, as the density increases more energy can be conserved. While AFECA must be conservative in its local density measurement so that network connectivity is not hurt, MIT’s Span [10] adaptively measures local connectivity in order to makes its decisions. If a node detects that two of its neighbors are not connected by either one or two hops over a network backbone of nodes, then the node joins the backbone itself; otherwise it goes to sleep.

Finally, there are examples of routing protocols that themselves seek to conserve energy and extend network operational lifetime. Chang and Tassiulas [9], Pottie et al. [23], and LEACH [15] suggest selection of routes based on available energy, so that energy is consumed evenly among nodes and network lifetime is extended. Our approach can complement such efforts.

Finally, our work is related to adaptive fidelity [13] and RTCP [27] adaptive frequency techniques. Other examples include beacon density for localization [7] and route fidelity under high mobility [1]. Currently our work maintains a constant fidelity (as do [7, 1]), but future work may explore adaptive fidelity.
3 Geography-informed Energy conservation protocol

With GAF, Geographic Adaptive Fidelity [35], nodes that are redundant for communication as determined by geographical position turn off their radios in order to save energy. Nodes alternate having their radios on in order to accomplish load balancing.

3.1 Determining node equivalence

GAF uses location information and an idealized radio model to determine node equivalence. Location information may be provided by GPS or other location systems under development (for example [2, 7, 12]). For our initial discussion, we assume that there is no error in the location information.

Even with location information it is not trivial to find equivalent nodes in an ad hoc network. Nodes that are equivalent for communication between one pair of nodes may not be equivalent for communication between a different pair of nodes. GAF addresses this problem by dividing the whole area where nodes are distributed into small “virtual grids”. A virtual grid is defined as follows: for two adjacent virtual grids A and B, all nodes in A can communicate with all nodes in B and vice versa. Thus, in each grid all nodes are equivalent for routing. For example, Figure 2(b) overlays virtual grids on Figure 2(a), creating three virtual grids, A, B, and C. According to our definition of virtual grids, node 1 can reach any of nodes 2, 3, or 4, and nodes 2, 3, and 4 can all reach node 5. Therefore nodes 2, 3, and 4 are equivalent and two of them can go to sleep.

We size our virtual grid based on the nominal radio range $r$, the farthest possible distance between two nodes in adjacent grids (since they must be able to communicate). If a virtual grid is a square with $r$ units on a side, then the longest possible distance between nodes in adjacent grids is the length of the long diagonal connecting the two grids. Therefore, we get that $r^2 + (2r)^2 \leq R^2$ and thus $r \leq \frac{R}{\sqrt{5}}$.

3.2 GAF state transitions

In GAF, nodes are in one of three states: sleeping, discovery, active. A state transition diagram is shown in Figure 3. Initially a node starts out in the discovery state with its radio turned on and exchanges discovery messages with its neighbors in order to find other nodes within the same grid. The discovery message is a tuple of node ID, grid ID, estimated node active time ($enat$), and node state. If, after waiting for $T_d$ seconds, the node has not already determined that it should not be the active node in its grid, the node moves to the active state. The node remains active for $T_a$ seconds, then returns to discovery state. A node in discovery or active states can change state to sleeping when it can determine some other equivalent node will handle routing. Nodes negotiate which node will handle routing through an application-dependent ranking procedure described in Section 4.2 (CEC uses the same ranking system). Transitioning to the sleeping state, causes a node to cancel all pending timers and power down its radio. A node in the sleeping state wakes up after an application-dependent sleep time $T_s$ and transitions back to discovery.

4 Cluster-based Energy Conservation (CEC) Algorithm

In many settings, such as indoors or under trees where GPS does not work, location information is not available. The dependency on global location information thus limits GAF’s usefulness. In addition, geographic proximity does not always lead to
network connectivity. As we show in Section 7.1, GAF must make very conservative connectivity assumptions because it guesses at connectivity (based on a radio model) instead of directly measuring it. Being conservative requires more nodes to stay active than necessary, leading to less energy conservation.

This motivates Cluster-based Energy Conservation (CEC), which, unlike GAF, does not rely on location information. Further, CEC itself directly and adaptively measures network connectivity and thus can find network redundancy more accurately so that more energy can be conserved.

4.1 Determining network redundancy

CEC organizes nodes into overlapping clusters that are interconnected to each other as shown in Figure 4. A cluster is defined as a subset of nodes that are mutually “reachable” in at most 2 hops. As shown in Figure 4, a cluster can be viewed as a circle around the cluster-head with the radius equal to the radio transmission range of the cluster-head. Each cluster is identified by one cluster-head, a node that can reach all nodes in the cluster in 1 hop.

A gateway is a node that is a member of more than one cluster. The gateway nodes connect all clusters together to ensure overall network connectivity. A node is ordinary if it is neither a cluster-head nor a gateway node and is thus redundant.

4.2 Distributed Cluster Formation

In order to elect cluster-heads and gateway nodes, each node periodically broadcasts a discovery message that contains its node ID, its cluster ID, and its estimated lifetime. A node’s estimated lifetime can be conservatively set by assuming the node will constantly consume energy at a maximum rate until it runs out of energy.

While forming clusters, CEC first elects cluster-heads, then elects gateways to connect clusters.

1. Cluster-head Selection – A node selects itself as a cluster-head if it has the longest lifetime of all its neighbor nodes, breaking ties by node ID. Each node can independently make this decision based on exchanged discovery messages. Each node sets its cluster ID to be the node ID of its cluster-head.

2. Gateway Node Selection – Among the gateway nodes, those nodes that can hear multiple cluster-heads are primary gateway nodes and those that can hear a combination of cluster-heads and primary gateway nodes are secondary gateway nodes.

When multiple gateway nodes exist between two adjacent clusters, CEC suppresses some of them in order to conserve energy since these gateway nodes are redundant. Gateway selection is determined by several rules. First, primary gateway nodes have higher priority than secondary gateway nodes since at least two secondary gateway nodes, instead of just one primary gateway node, are needed to connect adjacent clusters. Second, gateway nodes with more cluster-head neighbors have higher priority, since this will require fewer nodes to be kept awake. Third, gateway nodes with longer lifetimes have higher priority in order to balance node energy. Note that the gateway selection algorithm does not guarantee that only one or one pair of gateway nodes exist between adjacent clusters. In order to support gateway selection, CEC extends the basic discovery message to include the IDs of the clusters that a gateway node can connect.
Figure 4 shows an example of CEC cluster formation in which all nodes have the same estimated network operational lifetime. Nodes 1 and 10 can directly decide they are the cluster-heads because they have the lowest ID of all of their neighbors. Node 7 becomes a cluster-head after nodes 2 and 3 choose node 1 as their cluster-head. Nodes 2 and 3 are primary gateway nodes because they are neighbors of two cluster-heads: nodes 1 and 7. Note that one of nodes 2 and 3 is redundant. Nodes 9 and 11 are secondary gateway nodes between clusters 7 and 10.

4.3 Controlling the Duty Cycle of CEC Nodes

After the selection of cluster-heads and gateway nodes, the remaining redundant nodes are powered off to conserve energy.

Whenever a cluster is formed, each redundant node sets a wake-up timer that will wake it up in time \( T_s \), \( T_s \) is set to some fraction of the estimated node lifetime (enlt) of the cluster-head. In our CEC implementation, we normally set \( T_s \) to be \( enlt/2 \). In order to avoid thrashing, we set \( T_s \) to be \( enlt \) when it becomes less than a threshold (say 30s).

All nodes in the same cluster will thus be powered on to re-form the cluster before the cluster-head runs out of energy. While re-forming clusters, it is more likely that the last cluster-head has less remaining energy than the other nodes in the cluster since most have been in a sleeping state and conserving energy. CEC therefore achieves the goal of balanced energy use.

When a node’s radio is powered off, its forwarding role can be replaced by other nodes. An interesting question is how a sleeping node handles traffic originating from it or destined to it. In the former case, if the node has data to send it can simply power on its radio and send out data. In the latter case, the situation can be addressed as follows. First, for some applications (such as sensor nets), packets are usually not addressed to a particular node, but to a group of nodes with similar properties [17]. Thus, when a node is powered off, other nodes can stay alive to pick up the traffic. Second, some MAC protocols, such as 802.11, support a power-saving mode in which active nodes (typically base-stations) can temporarily buffer data for sleeping nodes. We can follow a similar strategy in our schemes.

4.4 Adapting to Network Mobility

With only a subset of the nodes active, it is possible that network mobility could cause a loss of connectivity. If a cluster-head moves then it might no longer be able to serve as a cluster-head. CEC uses mobility prediction in order to maintain network connectivity.

By estimating how soon a cluster-head will leave its current cluster and informing all nodes in the cluster of that time, the clustered nodes can power themselves on before the cluster-head leaves its cluster. This time is estimated as \( R/s \) where \( s \) is the cluster-head’s current speed and \( R \) is its radio transmission range.

Note that if the \( R/s \) estimate is too large, the connectivity between the moving cluster-head and some nodes might be lost before this time. However, if this estimate is too small, CEC will not be able to conserve any energy. In our CEC implementation, we set the estimate as \( R/4s \) to balance energy conservation and connectivity.

We extend the basic discovery message to include the predicted cluster-leaving time. All nodes in a cluster should wake up to reconfigure clusters before the shorter of \( T_s \) and the cluster-leaving time of its current cluster-head. The cluster-leaving time estimate is used analogously in the gateway node selection process (gateway nodes roughly estimate their cluster-leaving time as \( R/2s \)).

Although GAF uses a similar method for dealing with mobility, it anticipates hand-offs by using location information, while CEC uses only local measurements. With such global information, GAF may have more accurate mobility predications, but CEC is more practical and localized in nature.

5 Analysis of energy conservation protocols

In order to get an upper bound on how much GAF may extend network lifetime, we consider \( n \) nodes that are evenly distributed in an area of size \( A \). The
radio of each node has a nominal range of $R$. As we saw in Section 3.1, the length of a virtual grid square is at most $\frac{R}{\sqrt{5}}$. The minimum number of virtual grid squares, $m$, would then be $A/(\frac{R}{\sqrt{5}})^2$. Because our nodes are evenly distributed, each grid would have at most $n/m$ nodes, which is equal to $\frac{nR^2}{A}$. At best (assuming stationary nodes and no GAF overhead), only one node in each grid will be active while the rest sleep. Based on the maximum number of nodes in each grid, the network lifetime will be extended by at most $\frac{nR^2}{(5A)}$ times.

The formula reflects the fact that with GAF both more nodes and fewer virtual grids will lead to longer network operational lifetime. The number of virtual grids depends on the nominal radio transmission range and the size of the deployment area.

In order to get the same upper bound for CEC, we again consider $n$ nodes distributed in an area of size $A$. Again, each node has a radio with a nominal range of $R$. A cluster area can be viewed as a circle of radius $R$ around the cluster-head. The minimum number of clusters, $m$, to cover the whole area is equal to $\frac{A}{\pi R^2}$. In each cluster, the cluster-head and a few gateway nodes must stay alive in order to maintain network connectivity. If the average number of adjacent clusters is $k$, we need at least $m(k+1)$ nodes to cover the whole area. At best (assuming stationary nodes and no CEC overhead), the network lifetime can be extended by $\frac{n}{m(k+1)}$ times, or $\frac{\pi R^2}{(k+1)A}$ times.

Since GAF can extend network operational lifetime by at most $(nR^2)/(5A)$ times, these equations show that CEC can extend network operational lifetime longer than GAF when $k < 16$. This is a reasonable value for $k$ in most scenarios. One reason for this difference is that GAF conservatively uses smaller grid sizes to group redundant nodes. CEC uses connectivity measurements to discover network redundancy and thus does not have the same constraint.

6 Simulation of Topology Control

Because it is difficult to capture the details of GAF/CEC performance in an analytical model, we implemented GAF/CEC in the ns-2.1b6 snapshot of the ns-2 simulator [4] and used AODV and DSR to route packets.

We ran GAF/AODV, CEC/AODV, AODV, GAF/DSR, CEC/DSR, and DSR on the same simulated scenarios to compare the effects of variations in node movement, traffic patterns, and energy models on the performance of the protocols, as measured by energy use and data delivery quality.

Traffic, mobility, and radio models: Nodes in the simulation move according to the random waypoint model used in [6]. Nodes pause and then move to a randomly chosen location at a fixed speed. We consider seven pause times: 0, 30, 60, 120, 300, 600, and 900 seconds and for each we generate 10 sets of initial placements and random way-points. Nodes move at two different speeds: from uniform distributions between 0 and 20m/s and 0 and 1m/s. Nodes move in a 1500m by 300m area.

In most scenarios we use 50 transit nodes that route data and 10 traffic nodes that act as sources and sinks. When we vary node density we use 100 and 200 nodes, while keeping the area constant.

Traffic was generated by continuous bit rate (CBR) sources spreading the traffic randomly among the 10 traffic nodes. The packet sizes were 512 and 1024 bytes and the packet rate was set to four different values: 1 pkt/s, 10 pkts/s, 20 pkts/s and 200 pkts/s. Note that when packet size is 1024 bytes and packet rate is 200 pkts/s, the traffic reaches the maximum link bandwidth of 2Mb/s.

We model a radio with a nominal range of 250 meters both with the two-ray-ground propagation model [6] and a non-deterministic shadowing model [26].

Energy model: Our energy consumption model is based on Stemm and Katz’s measurements of a

---

1CMU contributed an extended version of DSR [32] and a validated 802.11 (2M) MAC layer with the simulation package. Our AODV implementation was an improved version from the AODV designers [11]. We have verified that our integration of CMU’s ad hoc routing reproduces their results [6], and that our simulation results of unmodified ad hoc protocols are consistent with other published results [6, 11, 18].
1995 AT&T 2Mb/s WaveLAN (pre-802.11) wireless LAN [31]. They measured costs of 1.6W for transmitting, 1.2W for receiving, and 1.0W for listening. To this we add a cost of 0.025W when sleeping. Newer evaluations of more recent versions of the WaveLAN card and compatible hardware by other vendors show very similar costs [19, 10].

Since it is impossible to evaluate the behavior of the network if the traffic nodes run out of energy before the transit nodes, we give traffic nodes infinite energy. Traffic nodes follow the same mobility model as transit nodes, but they do not run GAF or forward traffic. Because we treat traffic nodes specially, we do not count them when reporting the number of nodes in the simulation.

We give each transit node enough energy so that it can listen for about 450 seconds.

In our GAF simulations we model GPS as consuming 0.033W, the amount of power necessary for reporting location every 8 seconds, since GAF does not require constant position information. We do not turn off GPS when we turn off the radio in order to avoid modeling satellite acquisition time, and because the GPS cost is quite small (about equal to radio sleep cost).

6.1 Energy Conservation

In order to quantify energy consumption, we define the mean energy consumption per node (meen) as follows. At the start of the simulation the \( n \) nodes have a total initial energy, \( E_0 \). After time \( t \), the remaining total energy of the \( n \) nodes is \( E_t \). The (meen) equals \( \frac{E_0 - E_t}{n} \).

Our results show that CEC uses almost half the energy of GAF except in the scenario where nodes move at high speed (20m/s) constantly (zero pause time). GAF typically can save 30-40% more energy than plain AODV, while CEC can save about 60-70% more energy than plain AODV. When nodes move at high speed and with constant movement, CEC adjusts by turning off nodes for shorter times, thus leading to more frequent cluster formations. Such overhead causes CEC to use more energy than GAF in this scenario, though still about 30% less than plain AODV. With the help of global location information, GAF is more energy efficient when dealing with high mobility.

Varying the traffic load does not affect the energy conservation results.

6.2 Extending network operational lifetime

We now examine how these energy savings extend network operational lifetime.

Figure 5 shows the fraction of the network with remaining energy over time when nodes move at 20m/s (When nodes move at low speed, 1m/s, the CEC plot is close to that of the 900s pause time CEC curve, regardless of actual pause time). For both CEC and GAF we plot a zero pause time, representing constant node movement, and a 900s pause time, representing almost no node movement.

All nodes running plain AODV run out of energy at the same time, around 430s. Since AODV does nothing to conserve energy, this result reflects the cost of continuously listening.

We can see that CEC balances energy use more evenly among nodes than GAF. For example, at time 900s at least 80% of CEC nodes are still alive while at most 40% of GAF nodes are alive except the scenario with a pause time of 0. CEC is more

---

**Summary:** We conducted our comparison in two phases. In the first phase, we simulated 50 nodes for 900s. Our goal in this phase was to show that our schemes do not reduce the quality of routing, but do in fact conserve energy and extend network operational lifetime. In the second phase we do the same comparison for 3600s while varying the number of nodes in order to see how long network operational lifetime is extended for different node deployment densities.

In each phase we consider 1680 simulations: all combinations of 6 protocols, 7 movement patterns, 10 initial placements, 3 traffic loads, and 2 movement speeds. Based on our results, the difference when running AODV and DSR is not noticeable, so in this section we present results only for the AODV simulations.
Figure 5: Comparison of non-zero energy node fraction over time: CEC, GAF and plain AODV under different network mobility. Traffic load is 20pkts/s. Different traffic loads do not affect the result. In the legend, “CEC,x” means running CEC with pause time x, so is “GAF,x”.

effective at balancing energy because of its connectivity measurement based approach.

CEC also shows a different trend from GAF in regard to mobility. With CEC more nodes survive under low mobility (900s pause time), while in GAF, more nodes survive under high mobility (zero pause time). With CEC, high mobility causes more frequent cluster formations and more overhead. With GAF, high mobility helps balance energy use because changes in node location cause active node re-election within grids. In addition, GAF is more efficient in predicting mobility due to its access to global location information.

In Figure 6 we plot the time at which only 20% of the nodes remain alive against varying degrees of mobility. From this, we can see that CEC extends network operational lifetime at most two times longer than GAF and five times longer than AODV. With CEC, network operational lifetime increases with the pause time. As explained above, this is the effect of the adjustments of CEC for high mobility.

6.3 Network Connectivity

It is easy to conserve energy if one does not care about connectivity: in the extreme, one could turn off the whole network. It is thus important to evaluate the network connectivity produced by our protocols.

We define data delivery ratio as the ratio of the number of packets received to the total number of packets sent and the average data transfer delay as the mean delay for those received packets. Under varying traffic loads, these metrics truly reflect the effect of our protocols on network capacity as well as on connectivity.

Within the normal AODV lifetime, we found that CEC performs almost the same as AODV. With low network mobility, CEC even performs better than AODV for the following reason. Under heavy traffic, energy use in AODV is unbalanced: those nodes on the routing path or close to the routing path consume more energy and run out of energy sooner than other nodes. The premature loss of these nodes leads to a worse data delivery ratio. This does not happen with CEC since redundant nodes are powered off and node energy use is balanced. This effect does not stand out in high mobility scenarios since different nodes will move in and out of the heavy traffic region— mobility thus leads to balanced energy use.

We also noticed that CEC can maintain almost the same data delivery ratio at extended network operational lifetime. This means that the amount of data carried by the network is doubled.

In order to further understand how CEC and GAF perform at extended network lifetime, we
compare CEC with GAF at extended network operational lifetime in Figure 7. We also plot the connectivity with plain AODV at its normal network operational lifetime as the ideal value to see how connectivity changes in the extended network operational lifetime.

Under high mobility (pause time less than 120s), both CEC and GAF can maintain the same data delivery ratio as the ideal value (with reasonable standard deviation as shown in Figure 7(a)). However, as the pause time increases (larger than 120s), the GAF data delivery ratio becomes worse. At the worst case (pause time 900s), the GAF packet delivery ratio dramatically decreases to only 60% of the ideal level. However, CEC still follows the trend of the ideal data delivery ratio: the lower the mobility, the better the data delivery ratio. The difference between the CEC data delivery ratio and the ideal ratio remains below 5%.

The bad performance of GAF at lower mobility is due to its static gridding mechanism. When node density is high enough to keep at least one node in each grid, GAF works fine. However, when the node density decreases in the extended lifetime, connectivity is affected. With high mobility the situation is not very severe because the movement can help change the uneven distribution.

The same trend is reflected in the delay time as shown in Figure 7(b). CEC follows the trend of the ideal delay time: the lower the mobility, the lower the delay time. GAF performs better than CEC under high mobility but performs worse at low mobility.

In summary, with a high enough network density, both CEC and GAF can keep good network connectivity. As the network density decreases, CEC can maintain better connectivity than GAF, especially under low mobility.

### 6.4 Sensitivity to network density

Because our protocols exploit network redundancy, they should extend network operational lifetime farther for a more dense (and more redundant) network. We quantify density as the number of nodes in nominal radio range (ninra) [35], so that our results are independent of the number of nodes and the size of the topology.

![Figure 7: Data Delivery Quality as a function of pause time comparison: CEC vs. GAF under moving speed 20m/s at extended network operational lifetime. Traffic load is 20pkts/s. Other loads do not change the result. The AODV performance under normal network operational lifetime is plotted as the ideal value.](image-url)
Figure 8 shows that under high network density (ninra larger than 20) both CEC and GAF extend network operational lifetime in proportion to the increase of node density while network operational lifetime under plain AODV remains constant. Under high mobility, CEC and GAF perform about the same; however, under low mobility CEC extends network operational lifetime consistently longer than GAF. With a 4-fold increase in node density (ninra 88), CEC extends network operational lifetime 12 times longer than plain AODV and 3 times longer than GAF.

In a low density (ninra less than 20) network, a network without redundancy, both CEC and GAF have the same network operational lifetime as plain AODV.

Another observation about the effect of mobility is that GAF extends network operational lifetime more under high mobility, while the larger extensions for CEC are present under low mobility. This trend agrees with our previous analysis regarding the effects of mobility on CEC and GAF.

6.5 CEC protocol overhead

We measured the energy used by CEC control messages and computed the percentage of energy used by these over the total system energy usage. We find that protocol overhead is always less than 0.4% but is higher with higher mobility. The reason for this is that under high mobility CEC turns on more nodes more frequently. The relationship between mobility and protocol overhead is available in more detail in [33].

6.6 Result sensitivity to propagation model

The simulation studies so far have considered a deterministic radio propagation model (the two-ray-ground model). In reality though, radio propagation is strongly affected by multi-path effects (fading). In addition, observations of radio communication in the field show that the shadowing model cannot completely reflect the characteristics of radio propagation. Zhao et al. [37, 38] found that the quality of radio communication between nodes varies dramatically, leading to the belief that there is time-varying interference affecting radio communication in the field.

We therefore extend the shadowing model to a time-varying shadowing model by adding a statistical factor to the path loss model so that the attenuation of radio reception changes probabilistically.

We repeated our simulations using our time-varying shadowing model. We chose a value in the range of 3.0 to 4.0 for the path loss exponent in order to reflect a typical outdoor environment and used a shadowing deviation of 4.0 in the simulation. Our time-varying shadowing model controls how often the path loss exponents should be changed, following an exponential distribution. We observed that 10% to 20% of the links were asymmetric over the simulation time with this model.

As shown in Figure 9(a), under high mobility (pause time less than 120s), AODV, both with and without CEC, has a packet delivery ratio 20% worse with the time-varying propagation model than with the two-ray-ground model. However, the time-varying model does not change the relative performance of AODV with and without CEC under high mobility.

Under low mobility (pause time larger than 300s), in contrast, AODV with CEC shows almost a 30% worse packet delivery ratio than that of plain AODV. The reason for the bad performance of CEC under low mobility is that CEC does not sense network topology changes quickly enough. The time-varying shadowing leads to a more frequent change in radio conditions, which affects the performance of CEC.
of network topology and has an effect similar to that of high mobility.

CEC performs well under time-varying propagation model when network mobility is high because CEC’s mobility prediction algorithm forces CEC to sense the network connectivity more frequently. This frequent measurement ensures that CEC can adapt to any network topology changes, whether a result of mobility or time varying propagation.

The above observations suggest that we can make CEC work more robustly under the time-varying propagation model by having the nodes measure network connectivity more frequently. We can achieve this by adjusting CEC’s $T_s$ parameter. Without mobility prediction we set $T_s$ to be $\text{enlt}/2$, so when a cluster-head had a large amount of energy (a large $\text{enlt}$) all nodes in the cluster would check the network connectivity at a very low frequency. To make nodes measure network connectivity more frequently, we can set $T_s$ to be the minimum value of $\text{enlt}/2$ and a threshold, $T_h$, decided based on measurement. By doing this, nodes will take measurements at a higher frequency.

We change our CEC implementation by forcing each node to sense network connectivity on average every 10 secs (equivalent to a 20m/s node movement speed under high mobility). As shown in Figure 9(b), this change eliminates the difference in performance when using and not using CEC.

In summary, the time-varying propagation model introduces the issue of frequent network topology changes even under low mobility. Depending on the application, CEC might need to more frequently measure network connectivity in order to maintain the same data delivery quality as plain AODV.

7 Implementation and Experimentation of Topology Control

Although many wireless sensor network schemes seem to work well in simulation, such results do not necessarily guarantee good performance in the real world (which has obstacles, e.g.) or on nodes with real radios (the propagation of which is very difficult to model). GAF is an example of such a protocol. Therefore, in this section we present
the results of running implementations of GAF and CEC on experimental testbeds with real radios.

7.1 GAF with Experimental Radios

GAF’s off-line construction of virtual grids based on radio range is a cause for concern, since it assumes a circular radio propagation model, in which the radius of the circle is the range of the radio. In order to test the validity of making such use of radio ranges, we did an experiment in which we measured the connectivity between pairs of nodes. Based on this measure and some chosen thresholds, pairs of nodes are classified as well-connected (95% connectivity between the two nodes) or poorly-connected (10% connectivity between the two nodes). When comparing the connectivity measure to the actual distance between a pair of nodes, we see very high variability. Figure 10 shows an approximation of the radio range of each node. For each node, the left side of the bar shows the shortest distance to a poorly-connected node and the right side of the bar shows the longest distance to well-connected node. In order to trust radio ranges, we would want these two distances to have close to the same value. However, for only a few nodes (nodes 15, 16, 26, 32, and 38) is this actually the case. Such variance in radio range is also observed in [14]. Because radio ranges cannot be accurately estimated by distance, GAF is difficult to configure—a conservative setting of $R$ results in grids too small to provide energy conservation.

We evaluated GAF in a network of 15 PC/104s equipped with Radiometrix packet radios, running directed diffusion as the routing protocol [17]. We considered three configurations: small network size with fixed $R$, larger network with fixed $R$, and manually chosen non-uniform $R$. Given our observations about variation in radio range, we chose a conservative $R$ for the fixed scenarios. As a result, the larger network with fixed $R$ had several empty grid cells, though the network could still function. Due to space constraints we do not present detailed results here, but in general energy savings was proportional to the number of nodes per grid cell (densities were about 1.5 for fixed $R$ and 2.7 for manual configuration). Thus configuration with non-uniform grids provided better energy savings. While GAF was effective, the difficulty at selecting good values for $R$ suggests that CEC can provide better performance in practical settings.

We also observed an interaction between GAF and the routing protocol. Since they operate independently, in some cases GAF would turn off a node that was actively routing packets. Thus in some cases we observed periods of interrupted communication and very high latency. These problems were not observed in simulations of AODV and DSR because those routing protocols include a local repair mechanism for broken routes. We are currently adding such a mechanism to directed diffusion.

7.2 CEC Implementation and Experimental Setup

By measuring connectivity instead of assuming the correlation of radio range to distance, CEC provides a solution to the problem of variable radio ranges.

We ran experiments with CEC on a testbed of 21 iPAQs, each equipped with a UCB mote [16] as a radio interface and an 802.11 card for experimental control and logging purposes. The nodes were arranged in a square grid, with nodes at most of the vertices of the grid. Each node has an average of seven neighbors.

Data packets are routed through the use of flooding: each node keeps a cache of the data packets it has received and, with each subsequent arrival of a packet, forwards (broadcasts) that packet only if
it is not already present in the cache. All nodes in the network generate traffic according to the following rule: every five seconds, each node probabilistically decides whether to generate a new data packet. The probabilities used generate an expected one new packet per five seconds over the whole network.

The energy usage of each node is modeled in the same way as in our simulations. Again, we only model the energy usage of the radio, since the energy usage of other components is assumed to be negligible in comparison. Each node is given enough energy to remain in idle listening mode for 450 seconds.

7.2.1 CEC Extension of Network Operational Lifetime

Figure 11(a) depicts the extension of network lifetime by showing the number of nodes with remaining energy over time. Without CEC, there is a sharp drop at time 450s to zero, when all nodes run out of energy. With CEC, though, we see that network operational lifetime (20% of nodes remaining) is extended until time 1000s. The curve is qualitatively similar to what we observed in Figure 5, validating those results. Further, the fact that the curve is smooth indicates that CEC successfully balances energy usage among equivalent nodes.

7.2.2 CEC Data Delivery Ratio

As we pointed out in the discussion of the simulations, it is important to look both at energy savings and quality of data delivery, since optimizing for one hurts the other. There are two elements to this metric. First, we must judge the connectivity of the active nodes. Second, we must see how well the active nodes cover the network.

Figure 11(b) shows the data delivery ratio among CEC active nodes over time as well as the percentage of the network that is active. The first thing to notice is that the plot of the number of active nodes receiving packets and the plot of the number of active nodes are about the same. This means that about 100% of the active nodes are receiving the data packets and that the set of active nodes is indeed connected. The second thing to notice is

Figure 11: CEC experimental results.
the fluctuations in the bottom two plots. The first reason for these is the fact that nodes in discovery phase are counted as active: during discovery phase (e.g. just after time 200s) a larger number of nodes will temporarily be active. While the periodic discovery phases amplify the fluctuations in this plot, the fluctuations are present even if one does not count nodes in discovery state as active. CEC chooses active nodes based on two factors: a node’s remaining energy and a node’s number of neighboring nodes. In the first iteration of CEC, all nodes have approximately equal amounts of energy. As a result, CEC chooses active nodes primarily based on the number of neighboring nodes of a node. In the first iteration, therefore, the number of active nodes is close to the smallest possible for the given network topology. However, in the second iteration these same nodes (which make up something close to the minimum dominating set of nodes) cannot be picked (since they have used up half of their energy in the first iteration). Thus, a different and larger set of active nodes must be chosen in the second iteration.

Figure 11(c) shows the second part of the data delivery metric, the total number of nodes in the network that receive packets over time. The number of nodes depicted in this graph includes nodes that are asleep: the number of nodes on the y-axis is the sum of the number of active nodes and the number of nodes that are both asleep and one hop away from an active node. This metric shows us how much of the network is covered by (i.e. within one hop of) the active nodes. We can see that although CEC turns off a subset of nodes, data is still effectively delivered throughout the network.

Two interesting features of the plot are the dips in the number of nodes receiving packets around times 200s and 400s. By looking at Figure 11(b), one can see that the dips occur during discovery phases. We attribute these dips to the fact that all nodes are taking part in the flooding of data packets. There is thus a higher probability of collisions with all of the nodes forwarding packets than there is with only the active nodes forwarding packets.

Starting just before time 800s, we see increased variance in the data plotted. This is to be expected as the nodes in the network run out of energy, since the network at this point might be left disconnected.

In a disconnected network, a packet could get to vastly different amounts of the network depending on the topological location of the node that generates the packet.

Something that we noticed is that packets often reached either every node in the network or none of the nodes in the network. In Figure 11(c) we have ignored the packets that reach none of the network. Such packets are not even successfully transmitted over the first link. We suspect that perhaps by lowering the packet generation rate, we can eliminate such packets.

In our experiments with real radios we have seen that CEC is an effective scheme for saving energy without sacrificing the usefulness of networks for delivering data. CEC eliminates the redundancy that can be safely given up.

8 Conclusions

We have contributed two protocols, GAF [35] and CEC. GAF determines redundant nodes and controls node duty cycle to extend network operational lifetime while maintaining network connectivity, independent of the involvement of ad hoc routing protocols. GAF can substantially conserve energy (40% to 60% less energy than an unmodified ad hoc routing protocol), allowing network operational lifetime to increase in proportion to node density. CEC eliminates the dependency of GAF on global location information and its assumption about radio range. CEC measures local connectivity with low overhead and is thus able to dynamically adapt to a changing network.

In addition to an analysis and simulations results, we have presented measurements from an experimental implementation of our protocols. These results show that our protocols truly are effective with real radios in real systems.

References


