EAR: An Energy and Activity Aware Routing Protocol for Wireless Sensor Networks in Smart Environments

DEBRAJ DE[†1], WEN-ZHAN SONG[‡1], SHAOJIE TANG[♭2] AND DIANE COOK[*3]

[1]*Sensorweb Research Laboratory, Department of Computer Science, Georgia State University, Atlanta, GA, USA.*
[2]*Department of Computer Science, Illinois Institute of Technology*
[3]*Artificial Intelligence Laboratory, School of Electrical Engineering and Computer Science, Washington State University*
Email: [†]*dde1@student.gsu.edu,* [‡]*wsong@gsu.edu,* [♭]*stang7@iit.edu* [*]*djcook@wsu.edu*

Sensor network, unlike traditional communication network, is deeply embedded in physical environments and its operation is mainly driven by the event *activities* **in the environment. In long-term operations, the event** *activities* **usually show certain patterns which can be learned and exploited to optimize network design. However, this has been underexplored in the literature. One work related to this is using** $ATPG$ **for radio duty cycling ([1]). In this paper we present a novel Energy and Activity aware Routing (**EAR**) protocol for sensor networks. As a case study, we have evaluated** EAR **with the data trace of real Smart Environments. In** EAR **an** ***Activity Transition Probability Graph*** **(**$ATPG$**) is learned and built from the event activity patterns.** EAR **is an online routing protocol, which chooses the next-hop relay node by utilizing:** *activity* **pattern information in the** $ATPG$ **graph and a novel index of energy balance in the network.** EAR **extends network lifetime by maintaining an energy balance across the nodes in the network, while meeting application performance with desired throughput and low data delivery latency. We theoretically prove that: (a) the network throughput with** EAR **achieves a competitive ratio (i.e., the ratio of the performance of any offline algorithm that has knowledge of all past and future packet arrivals to the performance of our online algorithm) which is asymptotically optimal, and (b)** EAR **achieves a lower bound in network lifetime. Extensive experimental results from: (a) 82 node** ***Motelab*** **sensor network testbed [2] and (b) varying size network (20-100) in sensor network simulator TOSSIM, validate that** EAR **outperforms the existing methods both in terms of network performance (network lifetime, network energy consumption) and application performance (low latency, desired throughput) for an energy-constrained sensor network.**

Keywords: Routing; activity-aware; energy balance; Atkinson Index; competitive ratio; wireless sensor networks; smart environments

1. INTRODUCTION

Unlike traditional communication networks, sensor networks are deeply embedded in physical environments to provide high degree of visibility into environmental physical processes. Its data generation operation is mainly driven by the physical event activities in the environments. Here, by *activity* we mean the events those are sensed and reported by the sensor nodes to base station (sink) node. For example, in a smart home environment, the detected motion events of the residents constitute activities.

Figure 1 shows the event activities detected and reported by a node (in a 30 node smart workplace sensor network deployed across a floor) with motion sensor and corresponding node energy consumption. Through observation, it is clear that node energy consumption (thus node operations) is strongly correlated to the event activities. In long-term operations, these activities usually show certain periodic patterns (e.g. Figure 2), which can be learned and exploited to optimize network design. However, this has been underexplored in the literature. In this paper we present a novel Energy and Activity aware online Routing (EAR) protocol for sensor networks. EAR can adapt its network operations based on learned activity

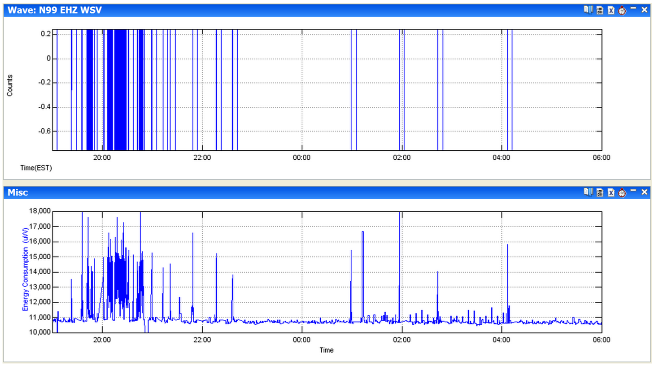

FIGURE 1. Event activity detected and reported by node with motion sensor in smart workplace, and corresponding energy consumption pattern of sensor node

patterns and an index of balance in node remaining energy in the network.

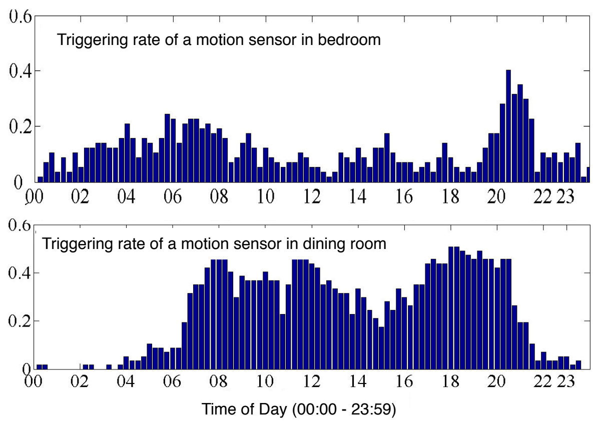

FIGURE 2. 24-hour activity pattern in bedroom and dining room of CASAS Smart Home testbed [3]

In this paper, we use Smart Environments (such as Smart Home [4] [3], Smart Workplace etc.) as an application case study. Figure 2 shows the daylong average triggering rate of two motion sensors (in different rooms) in CASAS Smart Home testbed [3]. The triggering rates for two motion sensors, one in bedroom and another in dining room, are calculated using 24 hour data of 57 days. It can be observed that the amount of sensing data generated in the network is driven by event activities and their patterns of occurrence. There are distinct activity patterns that vary in time (say, throughout a day) and space (say, bedroom vs. living room). This means at different time, the activity patterns in the same space are different, and at the same time the activity patterns at different space are also different. With such pattern the set of active data sources in the network changes in context of time and space. This leads to significantly non-uniform energy consumption of nodes across the network, resulting in severe imbalance in remaining energy of the nodes across the network. But if the activity patterns are utilized in form of activity-awareness, then it is possible to optimize network operations and thus significantly improving performance.

Motivated by this observation, we design an innovative activity-aware and energy-balanced online routing protocol. EAR protocol aims to maximize the network lifetime (defined as the time till first node depletes its energy), while meeting application data throughput requirement with low data delivery latency, without making any assumptions on future message arrivals. EAR tends to route packets through nodes with: larger remaining energy, neighborhood with larger balance in remaining energy, and relatively less sensing and data processing.

We theoretically prove that EAR achieves a lower bound in network lifetime. Also, the network throughput with EAR achieves a competitive ratio (i.e., the ratio of the performance of any offline algorithm that has knowledge of all past and future packet arrivals to the performance of our online algorithm), which is asymptotically optimal. Extensive experimental results on large scale 82 node Motelab real testbed and sensor network simulator $TOSSIM$ (with CASAS smart home [3] event activity data trace) validate that EAR outperforms the existing methods both in terms of network performance (network lifetime, network energy consumption) and application performance (low latency, desired throughput) for an energy-constrained sensor network.

The rest of the paper is organized as follows. In section 2 we have discussed the related works. In section 3 we have presented our EAR routing protocol in details. In section 4 we have presented theoretical analysis of EAR performance. Then in section 5 we have evaluated the system performance of EAR on the basis of experimental results. Finally we conclude in section 6 with discussion on possible future work.

2. RELATED WORK

To best of our knowledge, activity-aware routing is almost an unexplored area in the literature. Some relevant works on energy-balanced or lifetime-maximized routing design issues include [5], [6], [7]. The work in [5] has formulated the lifetime maximization problem as a linear program and has solved it using distributed heuristics technique. But this work assumes that the message generation rate at nodes are fixed and known. In [6] the observation has been made that the linear program is equivalent to that of a maximum concurrent flow problem. The algorithms proposed in [5] and [6] are able to determine how the traffic (generated at a constant rate) should be split among the different routes in order to maximize the network lifetime. Since the traffic generation rates are assumed to be constant and known in these works, the network can solve the energy aware routing problem off-line. The work in [8] converts the maximum network lifetime problem into a utility-based nonlinear optimization problem and proposes a distributed routing algorithm to solve the problem. But this work also assumes that the data generation rate at each node is fixed and known in advance. But sensor networks are majorly driven by activities. Therefore the data generation rates at nodes are usually non-uniform and not known exactly in advance. Our proposed protocol EAR uses this more realistic view.

For many practical applications (for example smart environment activity detection sensor networks) the message generation rate at different nodes are non-uniform and also dynamic. Therefore there is no complete advance knowledge of future message generation process. This problem needs to be solved with online routing protocol which does not need to know the message generation rates. Our proposed EAR is an online routing protocol. An online routing algorithm max-min zP_{min} is proposed in [7] for network lifetime maximization and it provides a competitive analysis. $CMAX$ [9] proposes an algorithm that tries to maximize the network capacity using shortest path computation with routing metric based on node remaining energy. The work in [10] proposes $E-WME$ online routing algorithm for the scenario of energy harvesting sensor nodes. Most of these energy aware online routing algorithms are based on remaining energy of relaying nodes. Unlike our proposed online routing protocol EAR, these works don't try to maintain energy balance in the network as a whole, and don't learn from activity patterns. In this paper we have considered the issue of energy balance across the network. An energy-welfare index (using Atkinson Inequality Index) is utilized in [11] to keep energy balance in network. But the forwarder selection procedure is complex and expensive (computing the index for each forwarder for each packet). Also there is no theoretical analysis of its proposed routing. The energy balance index in EAR is localized and simple to compute. We have also provided theoretical performance analysis of EAR. Also to note that, for the goal of maximizing network lifetime one possible solution may be to route the messages along the path with the maximal minimal fraction of remaining energy (the max-min routing). The performance of max-min path can degrade in situations (as described with some example in [7]). Another issue with the max-min routing is that following route with max-min energy node may be expensive compared to other possible paths. For large number of data streams there can be significant energy consumption for common nodes on max-min routing paths. This creates bottleneck nodes with high energy consumption and thus degrades the network lifetime quickly.

There are some works on activity-aware or context-aware networking. The work of *ActSee* ([1]) proposes an activity-aware radio duty cycling protocol that utilizes the learned event activity pattern information to intelligently adjust radio duty cycles in wireless sensor networks. The work in [12] presents a proactive communication algorithm for context aware sensor network. A framework for integrated unicast and multicast routing in context-aware ordered meshes is presented in [13]. But utilizing activity awareness for energy efficient and resource optimizing networking is not explored in these works, while EAR attempts these issues in depth.

TABLE 1. List of symbols used

E_i	Initial energy of node i
$E_i(k)$	residual energy of node i before routing message k
s_k, d_k, l_k	Source, destination and size of message k
c_{ij}	energy required by node i to send unit size data to j
$p_{tr}(x, y)$	probability of activity transition from node x to y
$t_{tr}(x, y)$	predicted activity transition delay from node x to y
$CL(i)$	activity cluster of node i

3. ENERGY AND ACTIVITY AWARE ROUTING

We first introduce system models and formal problem definitions, then we describe EAR protocol in details.

3.1. Preliminaries

The symbols used in EAR are listed in Table 1.

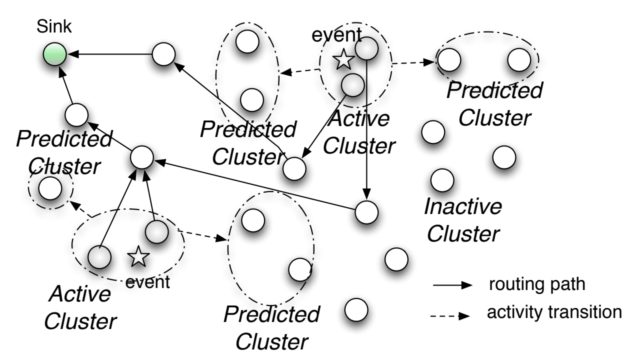

FIGURE 3. Illustration of activity patterns and data generation in network.

Learning Activity Patterns: Figure 3 illustrates the event activity patterns and data generation in a sensor network in smart environments. Based on the context of event activities, the nodes in the network at any moment belong to one of the three types of clusters (set of sensor nodes): *Active Cluster* (where the event activities are occurring in current time period), *Predicted Cluster* (predicted to be in active cluster in next time period), *Inactive Cluster* (with no activity in current period and no predicted activity in next time period). The membership of nodes being in clusters changes with time according to an *Activity Transition Probability Graph* ($ATPG$). In such a graph, the edge from node x to node y denotes the transition tuple $< p_{tr}(x,y), t_{tr}(x,y) >$, where $p_{tr}(x,y)$ is the predicted activity transition probability and $t_{tr}(x,y)$ is the predicted activity transition delay, for transition of activity from x-th cluster to y-th cluster. To note that the nodes in *Active Clusters* will be involved in activity detection and computation, followed by sending the data to base station node. Therefore it will be better to avoid involving nodes in Active or Predicted clusters as data forwarder. The nodes with higher probability of being in inactive clusters can be more involved in data forwarding task. This will save some energy (of data forwarding) for active nodes, and will also save some MCU computation resource for the node's own tasks (such as activity detection, processing, communication etc.).

System Model: The energy cost of sleep state is much lower than that of transmitting/receiving state. The energy consumption is considered only for transmitting/receiving state in our system model. The node overhear energy consumption model is used in various earlier works (e.g. in [14]). The sensor network is considered as a graph G=(V, E), where V is the set of nodes and E is the set of edges. Let n=$|V|$ be the number of nodes. Each node i starts with initial energy E_i. The source and destination of message k (of size l_k) are denoted as s_k and d_k respectively. In data collection scenario, d_k is always the base station. Now suppose in multi-hop routing, node i decides to forward message k to node j through link ij. Then node i consumes c_{ij} energy per unit length of data, therefore consuming a total $l_k.c_{ij}$ amount of energy for transmitting message k.

Objective: The design objective of *EAR* is to meet application performance requirement (e.g., throughput and delay) while maximizing network lifetime, by utilizing the activity pattern information in *ATPG* graph.

3.2. Algorithm and Protocol Design

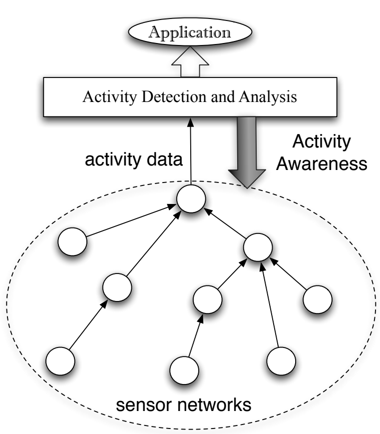

FIGURE 4. Activity-awareness for sensor network.

Distribution of Computation in EAR: The activity-awareness in sensor network is used in *EAR* as shown in Figure 4. The event detection data in network is collected at base station node for application purpose, and also used for constructing activity patterns in form of *ATPG*. Then the *ATPG* information is disseminated back into the network once. Now until a node dies, or a new node is added, or the activity pattern changes (that can happen only in long time period, typically at least several days, in smart environments), the *ATPG* information stored in the nodes is not changed. Therefore, the networkwide dissemination of *ATPG* information is performed rarely. Thus there is minimum communication overhead. The activity pattern analysis is only done in base station. All other calculations involved in *EAR* are distributed and localized. So all computations, except activity pattern analysis are distributed and localized in the network.

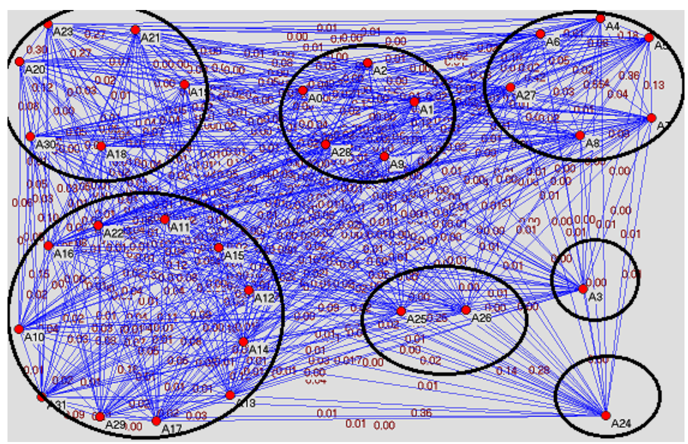

FIGURE 5. Activity Transition Probability Graph (including both significant and negligible transition probabilities) generated from the CASAS Smart Home testbed with layout shown in Figure 6.

Building and Maintaining *ATPG*: As a case study, we calculated an *ATPG* graph based on sensed events in CASAS [3] smart home testbed. The probability of transition between two sensor nodes x and y is based on the relative frequency of events at sensor node x followed by events at sensor node y. In *ATPG*, a node is generated for each sensor node that exists in the environment. The probability associated with edge x, y is estimated using the formula in equation 1. The example *ATPG* with both significant and negligible transition probabilities is shown in Figure 5. The revised *ATPG* with only significant activity transitions is shown in Figure 6.

$$p(x,y) = \frac{|\ events\ for\ sensor\ x\ followed\ by\ sensor\ y\ |}{|\ events\ for\ sensor\ x\ |} \tag{1}$$

Activity-Aware Routing Metric: Now we describe the notion of *activity-awareness* in *EAR*.

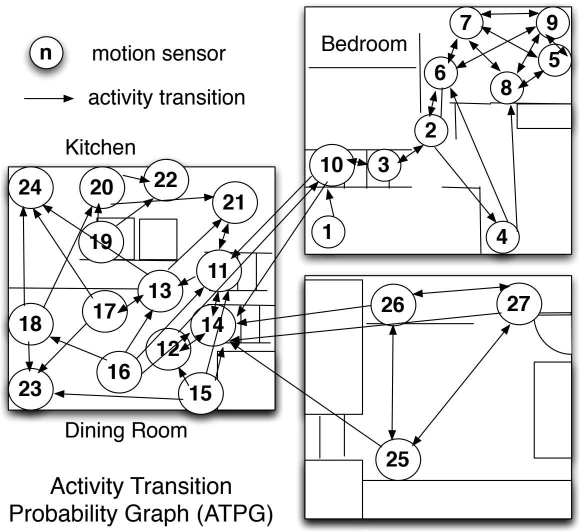

FIGURE 6. Activity Transition Probability Graph (pruned for significant activity transition) learnt from the CASAS Smart Home testbed. The significant transition probability for example from node 27 to nodes 14, 25 and 26 are 12%, 45% and 40% respectively.

Suppose a node i is trying to forward message k whose source is node s_k, which is in activity cluster $CL(s_k)$. Then node i tries to forward the data to some node j with (a) less computed probability $p(CL(s_k), CL(j))$ of being active (given $CL(s_k)$ is active) and (b) less duration of being active $(t_{active}(j))\cdot p(CL(s_k), CL(j))$ can be calculated by summing up the computed probability along all the paths from node $CL(s_k)$ to node $CL(j)$ in the $ATPG$. $t_{active}(j)$ can be calculated from weighted (based on transition probability) activity transition delay from $CL(j)$ to the next clusters. Let the period of activity pattern be T_P (which is 24 hours for smart home environments). Then activity-awareness metric for node j when routing of message k is $a(j,k) = p(CL(s_k), CL(j))\cdot t_{active}(j)$.

$$p(CL(s_k), CL(j)) = \sum_{P \in CL(s_k) \rightsquigarrow CL(j)} \prod_{(xy) \in path\ P} p_{tr}(x,y) \quad (2)$$

$$t_{active}(j) = \frac{\sum_{q \in N(CL(j))} p_{tr}(CL(j),q).t_{tr}(CL(j),q)}{T_P} \quad (3)$$

Low computation overhead for activity metric: In real application scenario, the nodes don't need to compute the parameters $p(CL(s_k), CL(j))$ and $t_{active}(j)$ each time. The transition graph information (transition probability and duration) can be stored (and updated if necessary in long time duration) in the nodes in an MxM vector, where M is the number of clusters in the network. This will indicate the values of $p(CL(s_k), CL(j))$. Based on that matrix, the nodes can save calculated $t_{active}(j)$ in an 1xM matrix. So the nodes can directly access the routing metric $a(j,k)$. This indicates that Activity Awareness metric doesn't incur much computation overhead.

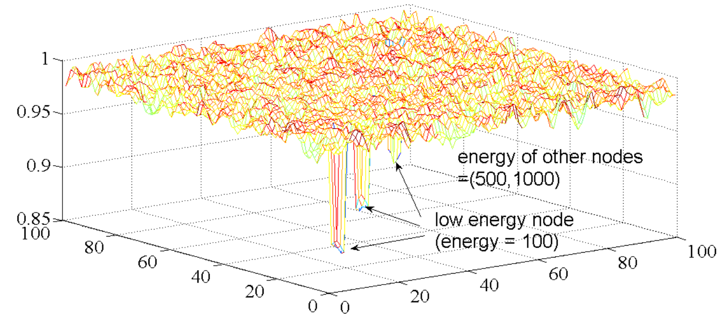

FIGURE 7. Distribution of Energy Balance index (B)

Energy-Aware Routing Metric: Now we describe the notion of *energy balance* in *EAR*. In order to reach a balance in energy consumption rate across the network we use Atkinson's Inequality Index [15]. It is a measure of economic income inequality in a society. The index can be turned into a normative measure by imposing a coefficient ε to weight incomes. Greater weight can be placed on changes in a given portion of the income distribution by choosing ε, the level of *Inequality Aversion*. The Atkinson index becomes more sensitive to changes at the lower end of the income distribution as ε approaches 1. Conversely, as the level of inequality aversion falls (ε approaching 0) the Atkinson Index becomes more sensitive to changes in the upper end of the income distribution. Atkinson index A is defined as in equation 4.

$$A = 1 - \frac{1}{\mu}\left[\frac{1}{N}\sum_{i=1}^{N} y_i^{(1-\varepsilon)}\right]^{1/(1-\varepsilon)} \quad (4)$$

Where $0 \leq \varepsilon < 1$, y_i is the individual income of i-th entity (i = 1, 2, ..., N) and μ is the mean income of total N entities. We calculate B=$(1 - A)$ as the energy balance index which is computed locally in 1-hop neighborhood. So the index of energy balance $B_i(k)$ computed by each node i (before routing of some message k) is shown in equation 5. The term $e_i(k)$ = $E_i(k)/E_i$ denotes the normalized remaining energy before routing message k. $E_i(k)$ is the residual energy of node i before routing message k. The neighbor set of node i is denotes as $N(i)$. Δ (=$\{N(i) \cup i\}$) is the set of 1-hop neighbors of node i and the node itself. So the index B_i is calculated using remaining energy information of the neighbors and the node i itself. In Figure 7 the effectiveness of B metric is shown. In a simulated environment in MATLAB with 100x100 sensor network grid, each node has maximum 8 neighbors. Only three nodes in the network have energy value of 100. But other nodes have energy between 500 to 1000. Then the distribution of locally computed B (with ε=0.8) across the network is shown in Figure 7. It can be observed that B is high enough (close to 1) everywhere, except in the neighboring region of the nodes with low energy. Therefore in a distributed network B is a meaningful indicator of region with significant energy imbalance. Lower B indicates higher degree of energy imbalance.

$$B_i(k) = \frac{|\Delta|}{\sum_{j\in\Delta} e_j(k)}[\frac{1}{|\Delta|}\sum_{j\in\Delta} e_j(k)^{(1-\varepsilon)}]^{1/(1-\varepsilon)} \quad (5)$$

The nodes in network maintain the hopcount from base station based on the default transmission power level. Now for purpose of routing convergence with delay control, (i) data is forwarded to node with same or less hopcount, (ii) if data is carried by at most H forwarders with same hopcount, it has to be forwarded next to a node with strictly lesser hopcount.

Routing Policy: Let messages are indexed in the order they are generated. Let s_k, d_k and l_k be the source, destination and length of message k. Suppose $E_i(k)$ is the residual energy of node i when the message k is generated but not routed. So $E_i(1)$ is the starting energy E_i for node i. Let the variable $\alpha_i(k)$ = 1-($E_i(k)/E_i$). Variable $B_i(k)$ is the computed energy balance index of node i as described before. $B_i(k)$ denotes the degree of energy balance around the neighborhood of node i before message k is routed. For activity awareness, the previously introduced parameter $a(j,k)$ (= $p(CL(s_k), CL(j))\cdot t_{active}(j)$) is used where $j \in N(i)$. In the description of the protocol there are two constant parameters λ and σ.

Algorithm 1 *EAR* algorithm

1. Set the weight or routing metric w_{ij} for the link ij as
w_{ij}=$c_{ij}.a(j,k).(\lambda^{\alpha_i(k)} - 1)/B_i(k)$.
2. Find the best path from s_k to d_k in the graph with the forwarder node selection method described. If node i has data packet to forward, select node j as its forwarder node as follows:

if *hop_travel* $< H$ **then**
  $j = \arg\min\{w_{iq}, q : q \in N(i)\ AND\ hopcnt(q) \le hopcnt(i)\ AND\ c_{iq} < E_i(k)/l_k\}$; *if* $hopcnt(j) = hopcnt(i)$ then *hop_travel*$+= 1$;
else
  $j = \arg\min\{w_{iq}, q : q \in N(i)\ AND\ hopcnt(q) < hopcnt(i)\ AND\ c_{iq} < E_i(k)/l_k\}$; *hop_travel* $= 0$;
end if

3. Let γ_k be the cost of the best path found for message k. Now if $\gamma_k \le \sigma$, then route the message k along the computed path, otherwise reject it. To note that γ_k=∞ if no such path is found.

Rationale behind routing policy: Here we explain the rationale behind the routing metric chosen. (a) For the link ij the weight w_{ij} increases with increase in c_{ij} (energy spent by node i for routing message k over link ij). So routing avoids the links with high message transmission cost. (b) w_{ij} increases with increase in both $p(CL(s_k), CL(j))$ (i.e. for nodes with more probability of being active) and $t_{active}(j)$ (i.e. for nodes with more expected active duration). Thus routing tries to select node (as forwarder) with less probability of being active and with less activity duration. This property makes it activity-aware. (c) w_{ij} increases with increase in the energy utilization $\alpha_i(k)$ of node i. So routing avoids nodes with low normalized residual energy. (d) Then w_{ij} increases with decrease in the energy balance $B_i(k)$ of node i. So routing avoids nodes whose neighborhood is relatively out of balance in residual energy. In addition, to note that in admission control, setting the value of σ to infinity (then the only reasons for rejecting a message is insufficient energy for routing) has shown results with good performance.

Competitive Ratio in Data Delivery: We now describe the calculated competitive bound for *EAR*. Let $c_{max} = max_{ij\in E} c_{ij}$, $c_{min} = min_{ij\in E} c_{ij}$, $a_{min} = min_{ij\in E} a(j,k)$ and ρ=$\frac{c_{max}}{c_{min}a_{min}}$. Let $L(k)$ be the total size of messages that is successfully routed by *EAR* till the arrival of message k. Let $L_{opt}(k)$ be the total size of messages that is successfully routed by optimal offline algorithm till the arrival of message k. Then the obtained competitive ratio result for *EAR* is as shown below. The detailed proof is provided in section 4.

THEOREM 3.1. *Suppose* λ=$2(n\rho + 1)$, ρ=$\frac{c_{max}}{c_{min}a_{min}}$, $\sigma = nc_{max}$ *and Q is a positive constant. For all message k, let*

$$l_k \le \frac{min_{i\in V} E_i}{c_{max} lg(\lambda)} \quad (6)$$

then, $\frac{L(k)}{L_{opt}(k)} \ge \frac{1}{1+Qlg(\lambda)}\ \forall k$

Delivery latency: It can be proved that *EAR* is a *H*-hop spanner. The H factor assures that the routed data, carried through less active nodes and energy balanced neighborhood, is converged to the base station. It is important to note that *EAR* actually reduces the data delivery latency to base station, by routing them through less active nodes (nodes less busy with sensing, data processing and forwarding). This is also supported from the experimental results (described later).

Network Lifetime: *EAR* is also proved in following theorem to provide sub-optimal network lifetime.

THEOREM 3.2. *Let* T_{ear} *and* T_{ML} *are network lifetime (time till first node dies) for EAR and optimal network lifetime algorithm (algorithm for maximum lifetime) respectively. Then*

$$T_{ear} > \frac{T_{ML}\sum_{k=1}^{S} P_{Min}(s_{m_k})}{\sum_{k=1}^{S} P_{atp}(s_{m_k})} + \frac{\delta(\sum_{i=1}^{n} E_i^{ML} - \sum_{i=1}^{n} E_i^{ear})}{\sum_{k=1}^{s} P_{atp}(s_{m_k})}$$

S is the number of message generated in the period T_P. $\sum_{k=1}^{S} P_{Min}(s_{m_k})$ is the total energy consumption for routing S message in T_P, when minimum energy path routing scheme is used. $\sum_{k=1}^{S} P_{atp}(s_{m_k})$ is the total energy consumption for routing S message in T_P, when purely activity-aware routing scheme is used. $(\sum_{i=1}^{n} E_i^{ML} - \sum_{i=1}^{n} E_i^{ear})$ denotes the difference between total remaining energy in network after time T_{ML} and T_{ear}. The detailed proof is provided in section 4.

Reliable Data Delivery: EAR follows routing metric based on energy and activity index. But EAR is not affected by link failure rate in lossy wireless medium. It has been observed through a number of experimental works (e.g. in [16]) that for any link, Packet Reception Rate (PRR) saturates to sufficiently high (almost 100%) when the link $RSSI$ is at least -90 dBm, or when the Link Quality Indicator LQI is 100. In system implementation of EAR, a node eliminates its neighbor node from routing table, to whom it's $RSSI$ is < -90dBm or it's LQI is < 100. So EAR can achieve gains in overall energy and resource usage, while not suffering data delivery guarantee because of failure rate of the links. This makes it practically applicable in any kind of harsh application environment.

Network Energy Balance: Through localized energy balance, EAR tries to keep a balance in remaining energy of nodes across network. This is crucial both for networks with uniform and non-uniform (e.g. heterogeneous network) starting energy. This is also useful for energy harvesting sensor networks. Maintenance of energy balance across network inherently increases lifetime, also gives the opportunity to intelligently utilize dynamically available energy sources. According to [17], Atkinson index measurement of inequality remains unchanged if there is an equi-proportionate change of all levels of income. Now EAR ensures messages are not forwarded by nodes with low energy, or not overheard by nodes with very low energy. Thus from the property mentioned, EAR tries to thwart the degradation in energy balance in local neighborhood due to routing of messages generating from nodes in non-uniform rate. In this way EAR tries to keep better energy balance in the network.

4. THEORETICAL ANALYSIS

We now present the theoretical proof of Theorem 1 and Theorem 2 described in earlier section.

■ **Proof of Theorem 1**: We associate a cost f_i for each node $i \in V$. Now the cost $f_i(k)$ for node i before the arrival of message k is as described in equation 7.

$$f_i(k) = E_i(\lambda^{\alpha_i(k)} - 1)/B_i(k) \tag{7}$$

Let $S(k)$ be the set of messages those are successfully routed by EAR until the arrival of message k. Now to prove the competitive ratio, we first find the lower bound of total message length successfully routed by EAR, in terms of node cost f_i.

Lemma 4.1. $\sum_{i\in V} f_i(k) \le 2qMP.lg(\lambda).\sigma L(k)$

Proof. Considering any message $k' \in S(k)$, from equation 7, for any node $i \in V$:

$$\begin{aligned}
&f_i(k'+1) - f_i(k') \\
&\le \frac{E_i.(\lambda^{\alpha_i(k'+1)} - 1)}{B_i(k'+1)} - \frac{E_i.(\lambda^{\alpha_i(k')} - 1)}{B_i(k')} \\
&= \frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)}.\lambda^{\alpha_i(k'+1)-\alpha_i(k')} - 1) \\
&\quad - \frac{E_i}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)} - 1) \\
&\le \frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)}.\lambda^{l_{k'}e_{ij}/E_i} - 1) \\
&\quad - \frac{E_i}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)} - 1)
\end{aligned}$$

Δ (=$\{N(i) \cup i\}$) is the set of node i and its neighbors. We define two terms $X(k')$=$\sum_{p\in\Delta} E_p(k')$ and $Y(k')$=$\sum_{p\in\Delta-i} E_p(k')^{(1-\varepsilon)}$. Then due to cost of routing message k' for node i and cost of overhearing message k' by awake neighbors of i:

$$\sum_{p\in\Delta} E_p(k'+1) = X(k'+1) = X(k') - l_{k'}e_{ij} - \beta_1 \tag{8}$$

$$\sum_{p\in\Delta} E_p(k')^{1-\varepsilon} = Y(k') + E_i(k')^{1-\varepsilon} \tag{9}$$

$$\sum_{p\in\Delta} E_p(k'+1)^{1-\varepsilon} = Y(k') + (E_i(k') - l_{k'}e_{ij})^{1-\varepsilon} - \beta_2 \tag{10}$$

β_1 and β_2 are energy cost due to message overhearing, and they vary with every message k'. Now we compute the expression $\frac{B_i(k')}{B_i(k'+1)}$.

$$\begin{aligned}
&\frac{B_i(k')}{B_i(k'+1)} \\
&= \frac{\frac{|\Delta|}{\sum_{p\in\Delta} E_p(k')}.(\frac{1}{|\Delta|}\sum_{p\in\Delta} E_p(k')^{(1-\varepsilon)})^{1/(1-\varepsilon)}}{\frac{|\Delta|}{\sum_{p\in\Delta} E_p(k'+1)}.(\frac{1}{|\Delta|}\sum_{p\in\Delta} E_p(k'+1)^{(1-\varepsilon)})^{1/(1-\varepsilon)}} \\
&= \frac{(X(k') - l_{k'}e_{ij} - \beta_1).(\frac{1}{|\Delta|}(Y(k') + E_i(k')^{(1-\varepsilon)}))^{1/(1-\varepsilon)}}{X(k').(\frac{1}{|\Delta|}(Y(k') + (E_i(k') - l_{k'}e_{ij})^{(1-\varepsilon)} - \beta_2))^{1/(1-\varepsilon)}} \\
&= \frac{(X(k') - l_{k'}e_{ij} - \beta_1).2^{1/(1-\varepsilon).(lg(Y(k')+E_i(k')^{(1-\varepsilon)})-lg(|\Delta|))}}{X(k').2^{1/(1-\varepsilon).(lg(Y(k')+(E_i(k')-l_{k'}e_{ij})^{(1-\varepsilon)}-\beta_2)-lg(|\Delta|))}} \\
&= \frac{(X(k') - l_{k'}e_{ij} - \beta_1)}{X(k')}.2^{1/(1-\varepsilon).lg(\frac{Y(k')+E_i(k')^{(1-\varepsilon)}}{Y(k')+(E_i(k')-l_{k'}e_{ij})^{(1-\varepsilon)}-\beta_2})} \\
&= \frac{(X(k') - l_{k'}e_{ij} - \beta_1)}{X(k')} \\
&\quad .(\frac{Y(k') + E_i(k')^{(1-\varepsilon)}}{Y(k') + (E_i(k') - l_{k'}e_{ij})^{(1-\varepsilon)} - \beta_2})^{1/(1-\varepsilon)}
\end{aligned}$$

Now, the term $T1$=$\frac{(X(k')-l_{k'}e_{ij}-\beta_1)}{X(k')}$ is slightly lower than 1, the term $T2$= $(\frac{Y(k')+E_i(k')^{(1-\varepsilon)}}{Y(k')+(E_i(k')-l_{k'}e_{ij})^{(1-\varepsilon)}-\beta_2})^{1/(1-\varepsilon)}$ is slightly higher than 1. This is due to relatively small amount of energy consumption in each routing step (with respect to the remaining energy of nodes). Then it can be proved that $T1.T2 \leq M$, where M is a relatively high positive constant. Then, $\frac{B_i(k')}{B_i(k'+1)} \leq M$.

Now from expression of $f_i(k'+1) - f_i(k')$: $f_i(k'+1) - f_i(k') \leq 2\frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)}.\lambda^{l_{k'}e_{ij}/E_i} - 1)$, and since value of λ is high. Therefore:

$$\begin{aligned}
&f_i(k'+1) - f_i(k') \\
&\quad \leq 2\frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(\frac{B_i(k')}{B_i(k'+1)}.\lambda^{l_{k'}e_{ij}/E_i} - 1) \\
&\quad \leq 2\frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(M\lambda^{l_{k'}e_{ij}/E_i} - 1) \\
&\quad = 2\frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(M2^{l_{k'}e_{ij}lg(\lambda)/E_i} - 1)
\end{aligned}$$

Since $l_k \leq \frac{min_{i\in V} E_i}{c_{max} lg\lambda}$, therefore $l_{k'}c_{ij}lg(\lambda)/E_i \leq 1$. For $0 \leq x \leq 1$, $2^x \leq (x+1)$. Therefore:

$$\begin{aligned}
&f_i(k'+1) - f_i(k') \\
&\quad \leq 2\frac{E_i.\lambda^{\alpha_i(k')}}{B_i(k')}.(Ml_{k'}c_{ij}lg(\lambda)/E_i + M - 1) \\
&\quad \leq \frac{2qM.l_{k'}c_{ij}lg(\lambda)\lambda^{\alpha_i(k')}}{B_i(k')}
\end{aligned}$$

This is because λ is very high and q is a relatively large positive constant. Now let $P(k')$ be the path over which the message k' was successfully routed. Therefore $\sum_{ij\in P(k')} c_{ij}a(j,k')(\lambda^{\alpha_i(k')} - 1)/B_i(k') \leq \sigma$.

$$\begin{aligned}
&\sum_{i\in V}(f_i(k'+1) - f_i(k')) \\
&\quad = \sum_{ij\in P(k')}(f_i(k'+1) - f_i(k')) \\
&\quad \leq \sum_{ij\in P(k')} \frac{2qM.l_{k'}c_{ij}lg(\lambda)\lambda^{\alpha_i(k')}}{B_i(k')} \\
&\quad = 2qM.lg(\lambda)l_{k'} \sum_{ij\in P(k')} \frac{c_{ij}(\lambda^{\alpha_i(k')} - 1)}{B_i(k')} \\
&\qquad + 2qM.lg(\lambda)l_{k'} \sum_{ij\in P(k')} \frac{c_{ij}}{B_i(k')} \\
&\quad \leq 4qM.lg(\lambda)l_{k'}\sigma
\end{aligned}$$

To note that $|P(k')| < n$. For $k' \notin S(k)$, $f_i(k'+1) - f_i(k') = 0$, $f_i(1) = 0\ \forall i \in V$. Then:

$$\begin{aligned}
&\sum_{i\in V} f_i(k) \\
&\quad = \sum_{k'\in S(k)} \sum_{i\in V}(f_i(k'+1) - f_i(k')) \\
&\quad \leq \sum_{k'\in S(k)} 4qM.lg(\lambda)l_{k'}\sigma \\
&\quad = 4qM.lg(\lambda)\sigma L(k)
\end{aligned}$$

□

Let $NS(k)$ be the set of messages successfully routed by the optimal off-line algorithm but rejected by EAR, until arrival of message k. Now we show that: $\forall k' \in NS(k)$, $\sum_{ij\in P(k')} c_{ij}a(j,k')(\lambda^{\alpha_i(k')} - 1)/B_i(k') > \sigma$.

A message $k' \in NS(k)$ is rejected by EAR if: (i) there is not sufficient energy on some node to forward the message, or (ii) $\gamma'_k > \sigma$. Now the lemma holds true for situation (ii). We have to prove the lemma for situation (i). Let message k' is rejected due to situation (i) in the protocol. That message k' is successfully routed by optimal offline algorithm through path say $P_{opt}(k')$. But for EAR, there is at least a link $i'j' \in P_{opt}(k')$, for which $E_{i'}(k') < l'_k c_{i'j'}$. Therefore $\alpha_{i'}(k') = 1 - E_{i'}(k')/E_{i'} \geq 1 - (1/lg\lambda)$ (using equation 6). Therefore:

$$\begin{aligned}
&\sum_{ij\in P_{opt}(k')} c_{ij}a(j,k')(\lambda^{\alpha_i(k')} - 1)/B_i(k') \\
&\quad \geq c_{i'j'}a(j',k')(\lambda^{\alpha_{i'}(k')} - 1)/B_{i'}(k') \\
&\quad > c_{i'j'}a(j',k')(\lambda^{1-(1/lg\lambda)} - 1)/B_{i'}(k') \\
&\quad = c_{i'j'}a(j',k')(\lambda/2 - 1)/B_{i'}(k') \\
&\quad \geq c_{min}a_{min}(\lambda/2 - 1) = nc_{max} = \sigma
\end{aligned}$$

Finally we show that:

$$nc_{max}(L_{opt}(k) - L(k)) \leq \sum_{i\in V} f_i(k) \qquad (11)$$

$$\begin{aligned}
&nc_{max}(L_{opt}(k) - L(k)) \\
&\quad \leq \sum_{k'\in NS(k)} nc_{max}l_{k'} \\
&\quad < \sum_{k'\in NS(k)} \sum_{ij\in P(k')} l_{k'}c_{ij}a(j,k')(\lambda^{\alpha_i(k')} - 1)/B_i(k') \\
&\quad \leq \sum_{k'\in NS(k)} \sum_{ij\in P(k')} l_{k'}c_{ij}f_i(k')/E_i \\
&\quad \leq \sum_{i\in V} f_i(k) \sum_{k'\in NS(k), ij\in P(k')} l_{k'}c_{ij}/E_i \\
&\quad \leq \sum_{i\in V} f_i(k)
\end{aligned}$$

The second last step uses the fact that the node cost f_i is non-decreasing. Last step uses the fact that the total energy spent for routing the messages at a node cannot exceed its initial energy. Finally from Lemma 1 and equation 11, we can prove the following expression, thus proving Theorem 1. ($Q = 4qM$ is a positive value.)

$$\frac{L(k)}{L_{opt}(k)} \geq \frac{1}{1 + Q.lg(\lambda)} \tag{12}$$

■ **Proof of Theorem 2**: Competitive ratio analysis implicitly proves the sub-optimality of EAR in lifetime w.r.t application point of view. To note that the competitive ratio analysis for EAR used no previous knowledge of message arrival. Now for analysis of another definition of network lifetime (time till the first node dies), we have utilized a property that is common to Smart Environment applications. The nodes in sensor networks in such scenario generate same amount of data in each time period, although in each period the data generation sequence may be different. The time period can be short or long. This is actually common to a lots of sensor network applications, for example Smart Home sensor networks, where the daily activity patterns are same, thus message generation is roughly periodic. This is validated through collected motion detection sensor network data in real experiments. So we have assumed here that in each time period $[t, t+\delta)$, the message distributions on the nodes in the network are the same. Then its possible to schedule the message routing with the same policy in each time period of δ.

Now let the network starts at time $t = 0$, network lifetime on optimal routing algorithm (for maximum lifetime) is say T_{ML}, network lifetime on EAR routing algorithm is T_{ear}. The initial energy content of each node $i \in V$ is E_i, remaining energy of each node $i \in V$ after time T_{ML} is $E_i(T_{ML})$, remaining energy of each node $i \in V$ after time T_{ear} is $E_i(T_{ear})$. Let the message sequence in any time period is m_1, m_2, ..., m_{S-1}, m_S.

$$\sum_{i=1}^{n} E_i = \sum_{i=1}^{n} E_i(T_{ML}) + \sum_{k=1}^{M(T_{ML})} P_{m_k}^{ML} \tag{13}$$

$$\sum_{i=1}^{n} E_i = \sum_{i=1}^{n} E_i(T_{ear}) + \sum_{k=1}^{M(T_{ear})} P_{m_k}^{ear} \tag{14}$$

Where $M(T_{ML})$ and $M(T_{ear})$ are the number of messages routed from time 0 to T_{ML} and from time 0 to T_{ear} respectively. $P_{m_k}^{ML}$ and $P_{m_k}^{ear}$ are the power consumption of the k-th message m_k by running optimal algorithm for maximum lifetime and EAR algorithm respectively. The messages are same in any two periods, without considering the sequence. Therefore it is possible to schedule the messages so that the message rates along the same route are the same in any two periods. Therefore:

$$\sum_{k=1}^{M(T_{ML})} P_{m_k}^{ML} = \frac{M(T_{ML})}{S} \sum_{k=1}^{S} P_{m_k}^{ML} = \frac{T_{ML}}{\delta} \sum_{k=1}^{S} P_{m_k}^{ML} \tag{15}$$

$$\sum_{k=1}^{M(T_{ear})} P_{m_k}^{ear} = \frac{T_{ear}}{\delta} \sum_{k=1}^{S} P_{m_k}^{ear} \tag{16}$$

$P_{m_k}^{ear}$ is the energy consumption of the message m_k in a period by running algorithm EAR. Now EAR considers remaining energy, energy balance and activity-awareness. Thus the total energy consumption $\sum_{k=1}^{S} P_{m_k}^{ear}$ will be less than that of ($\sum_{k=1}^{S} P_{m_k}^{act}$) a purely activity-aware routing algorithm (say act) that uses routing metric $a(j, m_k)$ for each node j. So, $\sum_{k=1}^{S} P_{m_k}^{ear} < \sum_{k=1}^{S} P_{m_k}^{act'}$. Now, for each message m_k, it is possible to construct the Network Activity Transition Probability graph for the sensor network G. $ATP(s_{m_k})$ is the constructed ATPG graph where the data source is s_{m_k}, the weight for each node j is $a(j, m_k)$, and $P_{atp}(s_{m_k})$ is the computed energy consumption of greedily selected path from s_{m_k} to base station using the node weight $a(j, m_k)$. So $\sum_{k=1}^{S} P_{atp}(s_{m_k})$ can be computed from G and ATP. Now, $\sum_{k=1}^{S} P_{m_k}^{ear}$ $< \sum_{k=1}^{S} P_{m_k}^{act} = \sum_{k=1}^{S} P_{atp}(s_{m_k})$. On the other hand $\sum_{k=1}^{S} P_{m_k}^{ML} > \sum_{k=1}^{S} P_{Min}(s_{m_k})$, where $P_{Min}(s_{m_k})$ is that of the the minimum energy consumption path in G from s_{m_k} to base station. $P_{Min}(s_{m_k})$ can be computed from G. Therefore:

$$\sum_{i=1}^{n} E_i^{ear} + \frac{T_{ear}}{\delta} \sum_{k=1}^{S} P_{atp}(s_{m_k}) > \sum_{i=1}^{n} E_i^{ML} + \frac{T_{ML}}{\delta} \sum_{k=1}^{S} P_{Min}(s_{m_k}) \tag{17}$$

$$T_{ear} > \frac{T_{ML} \sum_{k=1}^{S} P_{Min}(s_{m_k})}{\sum_{k=1}^{S} P_{atp}(s_{m_k})} + \frac{\delta(\sum_{i=1}^{n} E_i^{ML} - \sum_{i=1}^{n} E_i^{ear})}{\sum_{k=1}^{s} P_{atp}(s_{m_k})} \tag{18}$$

5. IMPLEMENTATION AND PERFORMANCE EVALUATION

In this section we have described the implementation, experiments and the analysis of results in detail.

5.1. Implementation of EAR

In EAR protocol with admission control, the data source node needs to have some knowledge about the network topology and the current energy of nodes. However, in practical network the topology and energy level of the nodes change frequently. It may work for small networks using information dissemination, but will be difficult to maintain for large networks. In this aspect, in our implementation, EAR is locally applied to each one hop neighborhood in the network. We have implemented EAR and other comparing routing

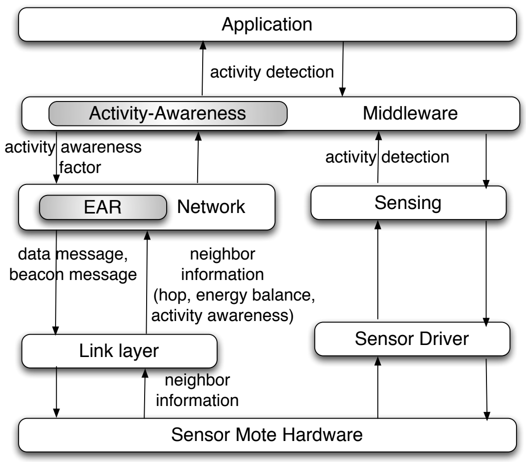

FIGURE 8. TinyOS node software stack included with activity-awareness and energy-balance design for *EAR*.

protocols in *TinyOS-2.x*, one of the most popular event based operating system environment for wireless sensor networks. The TinyOS node software stack with activity-awareness and energy-balance design support for *EAR* is shown in Figure 8. Regarding activity awareness in the experiments, the set of paths of activity in network is used for probabilistic path selection, and the network nodes are injected with intelligence of corresponding probabilistic transition information (*ATPG* graph).

5.2. Evaluation in *Motelab* Tested

Evaluation environment: We have evaluated our proposed *EAR* protocol in large scale 82 node network of *TelosB* motes (physically distributed in three floors, as shown in Figure 9(a), 9(b) and 9(c)) in Harvard *Motelab* sensor network testbed [2]. The experiments are conducted in 82 node network physically distributed across three floors.

Activity transition and data generation: From a separately deployed motion sensor network testbed we have learned the activity transition patterns and have validated the construction of activity transition graph *ATPG* (presented in section 3). The activity transition patterns are modified to be scalable for a 82 node Motelab testbed, and is injected in the testbed for activity event generation and activity transition. The activity transition decides the order with which nodes will be active.

The activity event generation makes node(s) active, letting it send data to sink node (base station) at a high rate (we used data sending rate of 480 Bytes/second). We have emulated the activity events by generating three independent sequences of active nodes (indicating motion trails) each in one of the three floors. From a remote server, periodic serial

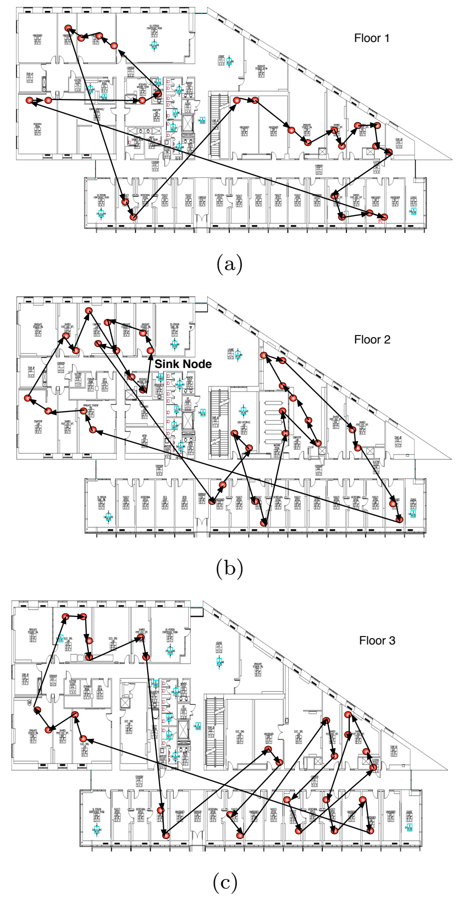

FIGURE 9. Most frequent activity sequences (order of active nodes) occurred in each floor of Motelab testbed during experiment

message (containing new active node numbers) is sent to the sensor motes in the testbed to generate the activity sequences. The sensor nodes receiving the serial message with it's ID start generating sensor data. Other nodes act as relay only. This periodic activation of nodes through serial message follow the activity transitions defined in the corresponding *ATPG*. In this way the activity transition experiments are performed with networkwide data collection. In addition each node periodically sends one local status data packet (containing information of remaining energy, hop count etc.) to sink every 30 seconds.

Comparison: For performance comparison we have compared *EAR* with standard existing routing schemes to show performance improvement. Following relevant routing protocols are used: *PMin* (shortest path routing), *CTP* [18] (very commonly used data collection protocol for sensor networks, that uses link and path quality), and *CMAX* (an energy aware protocol [9] where data is forwarded preferably to neighbor with higher remaining energy in the neighborhood).

The 82 node network formed a 9 hop routing tree with -5 dBm transmission power of CC2420 radio of TelosB motes. The sink node is in middle of the three floors. In

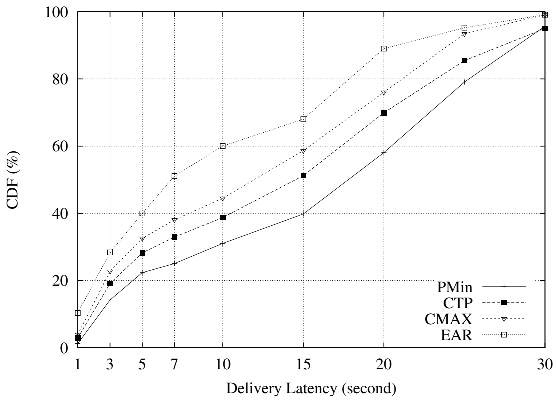

FIGURE 10. Distribution of data delivery latency.

this network data collection scenario we have evaluated following parameters: (i) data delivery latency, (ii) data throughput, (iii) minimum node energy in the network through time (indicating network lifetime). Now we describe the performance analysis of our proposed *EAR* protocol compared to existing protocols.

Latency: Figure 10 represents the distribution of data delivery latency of packets in the 82 node network. It can be observed that *EAR* provides much lower latency than each of the comparing protocols *PMin*, *CTP* and *CMAX*. In *PMin*, *CTP*, *CMAX*, 80% of the packets are delivered with latency between 22 seconds to 25 seconds. But in *EAR* the 80% of the packets are delivered within latency around 18 seconds. Therefore *EAR* provides much lower delivery latency, providing better performance to the application. *EAR* achieves this improvement in latency by avoiding selection of currently active nodes (which are busy with sensing and sending own data) as relays.

Data Throughput: Figure 11 shows the data throughput for each node at sink. More throughput indicates more event data successfully delivered and reported at sink. It is observed that for each of the 82 nodes, *EAR* provides much improved data throughput than others. For all the 82 nodes *EAR* provides a data throughput improvement ranging from 6% to 13%. This advantage in *EAR* comes from avoiding selection of currently active nodes (which are busy with sensing and sending own data) as relays.

Lifetime: Figure 12 represents the minimum energy of any node in the entire network through time. This property decides the network lifetime. The energy consumption is configured in such a way that all nodes start with energy 2200 mAh. For indicating the rate of drop in minimum remaining node energy in network, Figure 12 represents the energy drop with respect to a chosen value of 0.05806 mAh. This value is chosen for purpose of analysis because the minimum node energy in $PMin$ reduces by an amount of nearly 0.05806 mAh energy in experiment run time of 1800 seconds (i.e. 30 minutes). This chosen value for analysis doesn't

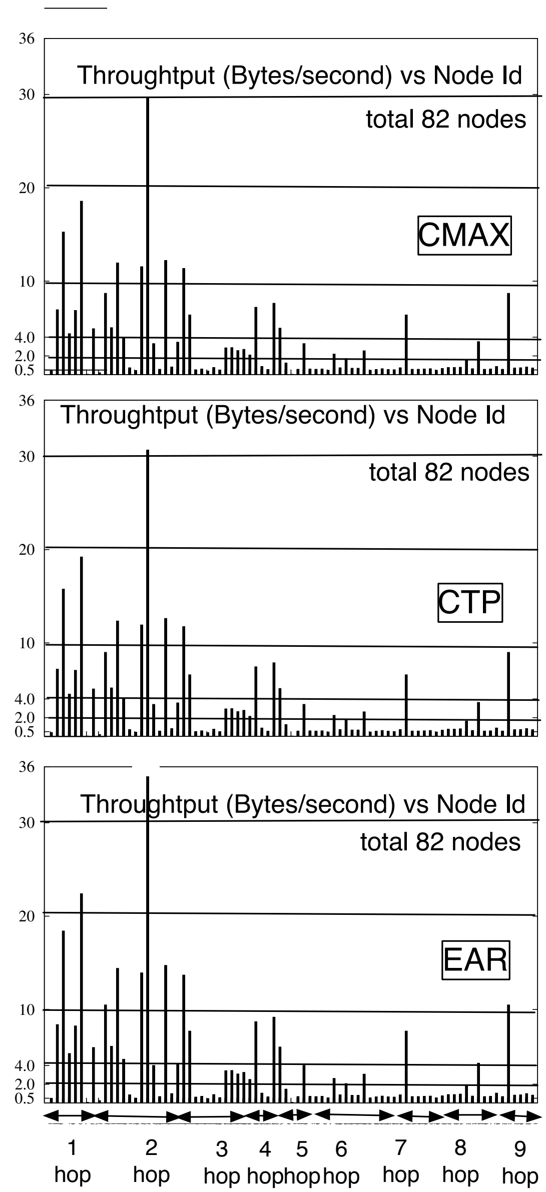

FIGURE 11. Data throughput at Sink

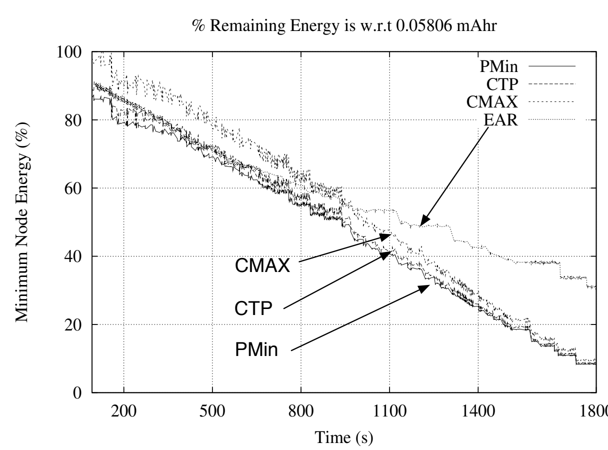

FIGURE 12. Minimum node energy in network.

affect the nature of energy consumption of network. Now it can be observed that the minimum energy of any node in *PMin*, *CTP* and *CMAX* depletes faster than the one in *EAR*. Therefore in protocols other than *EAR* network node depletes almost all it's energy within time 1800 seconds, had the network start with 0.05806 mAh for all nodes. But in same scenario in same time the minimum energy of any node in *EAR* would have still around 32% energy left. Therefore it is clear that network lifetime for *EAR* will also be much higher than others. *EAR* achieves this advantage because of energy balanced relay node selection.

5.3. Evaluation in TOSSIM Simulator

To validate the scalability of *EAR* we have used network size containing 20, 40, 60, 80 and 100 nodes (all with lossy wireless channel). The topology of 20 node network closely follows the node distribution in kitchen, dining room and bedroom of *CASAS* testbed, as shown in Figure 6. The topology of nodes in the other networks also follow the layout in Figure 6, but modified according to the network size. The activities are probabilistic and follow the activity transition patterns. We have generated *ATPG* with similar activity patterns for larger networks containing 40, 60, 80 and 100 nodes. When activities occur in a node, it performs some processing and then sends out bunch of data packets containing activity detection data.

The energy consumption is calculated using the relevant model of: CC2420 radio parameters (19.7 mA current consumption in receive mode, 17.4 mA current consumption in transmit mode, 250 kbps data rate with 48 kByte data packet size), and the MSP430 MCU parameters (3 mA current consumption in active mode due to sensing and computation). To note that due to timer and ADC read operations, sensor nodes can consume as high as 3mA current (as observed in [19]). The remaining node energy is updated accordingly.

EAR is compared with following relevant routing protocols: *PMin*, *CTP*, *MaxEn* (data forwarded to node with maximum remaining energy among the relay nodes) and *CMAX*. Each experiment with a network size is conducted for 2 hours. This generates multiple possibility paths of activity due to probabilistic activity transition in *ATPG*.

Data delivery latency: Despite preferring activity-aware and energy-balanced path, *EAR* also provides better data delivery latency. This is because of the activity aware property of *EAR*, which prefers less active nodes as forwarder (i.e. relay) node. This is validated from experimental results in Figure 13. For different network sizes, *EAR* provides from 6.8% to 19.1% less data delivery latency over others. *CMAX* and *MaxEn* are only energy-aware, so in non-uniform data generating network (leading to non-uniform energy nodes) the routed data packets sometimes deviate and follow a longer path. This leads to high data delivery latency. *PMin* has better data delivery latency by following shorter path, but suffers from retransmissions and from processing delay when being forwarded through active nodes (busy in sensing, processing and sending own data). *CTP* provides better delivery latency, but still performs worse than *EAR* because it doesn't learn from activity patterns.

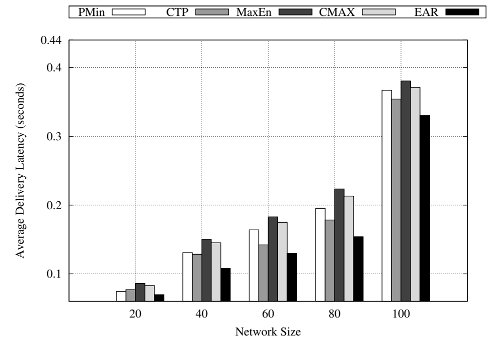

FIGURE 13. Mean data delivery latency (seconds) with varying network size.

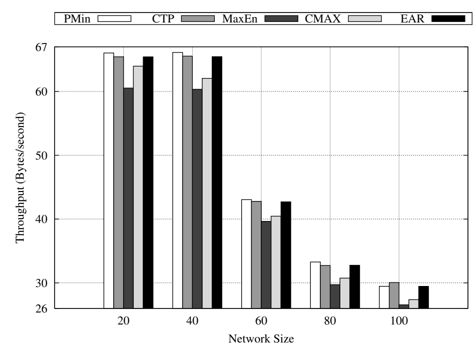

FIGURE 14. Data throughput at base station (successfully delivered message per unit time) in Bytes/second with varying network size.

Data throughput: *EAR* tries to minimize the network lifetime, with maintenance of network energy balance. Despite providing these advantages, *EAR* doesn't degrade the throughput (successful message received at sink per unit time) much. This is validated through results in Figure 14. *PMin* and *CTP* provide better throughput. But throughput performance of *EAR* closely follow (within upto 2% lesser) that of *PMin* and *CTP*. Energy-aware only protocols *CMAX* and *MaxEn* suffer worse throughput for lack of activity-awareness and lack of faster convergence in non-uniform activity generation network. To note that despite following activity and energy awareness, *EAR* make sure faster convergence by using hop spanner property discussed earlier.

Network lifetime: From experimental results in Figure 15 it can be observed that *EAR* achieves the maximum network lifetime for all the network sizes. For different network sizes, *EAR* achieves an improvement

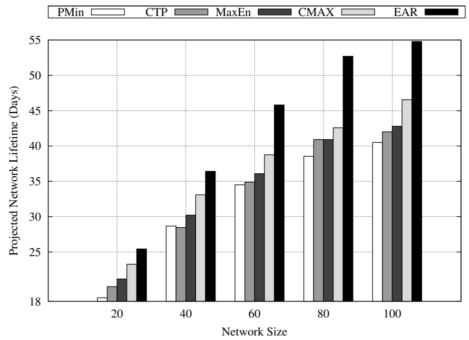

FIGURE 15. Projected network lifetime (days) with varying network size.

in lifetime over others from 9.31% to 23.77%. Figure 15 $CMAX$ and $MaxEn$ have better network lifetime than CTP and $PMin$, because $CMAX$ and $MaxEn$ are energy aware protocols. But their performance is worse than EAR due to inability to keep energy balance and to be activity-aware, CTP and $PMin$ suffer because they are not activity-aware or energy balancing. This is interesting observation for networks where data generation is a non-uniform and dynamic process, but have some patterns.

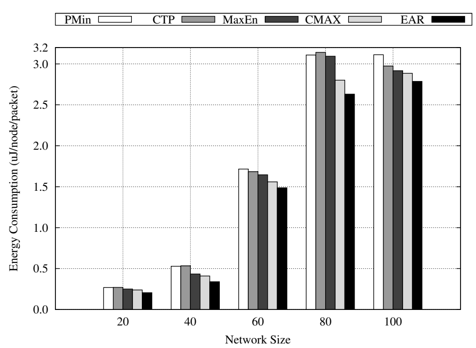

FIGURE 16. Energy consumption per successfully delivered message per node (uJ/packet/node) with varying network size

Network energy consumption: EAR also has minimum network energy consumption for all network sizes. In Figure 16 we have used the parameter for indicating effective network energy consumption. This is represented by the parameter: the total energy consumption per delivered packet per node. This indicates the average amount of energy spent by a node to enable one successful routing and collection of a data packet from network to sink. It can be observed that EAR has the minimum observed network energy consumption. For different network sizes, EAR provides from 3.4% to 17.2% improvement in network energy consumption over others. This proves the effectiveness of activity-awareness and energy balance of EAR. Due to energy balance property and avoiding active nodes for forwarding, the network as a whole spends less amount of energy for delivering data packets.

Scalability: All the advantages of EAR are achieved for network size varying from 20 to 100. This proves the scalability, thus its real-world applicability for pervasive environments.

6. CONCLUSION

Unlike traditional networks, the operation of sensor networks is driven by activities in the embedded environment. These activities show certain patterns in long-term operation. But the existing sensor network design seldom learns and exploits the activity patterns to optimize network operations. In this paper we have proposed EAR for activity-aware and energy-balanced routing. As a case study EAR is evaluated with Smart Environment data trace. The experimental results have demonstrated its efficiency both with respect to application and network performance, as well as its scalability. The possible future work includes adaptive joint radio scheduling and routing, with delay control.

REFERENCES

[1] Tang, S., De, D., Song, W.-Z., Cook, D., and Das, S. (2011) ActSee: Activity-Aware Radio Duty-Cycling for Sensor Networks in Smart Environments. *Eighth IEEE International Conference on Networked Sensing Systems (IEEE INSS).*

[2] Werner-Allen, G., Swieskowski, P., and Welsh, M. (2005) MoteLab: a wireless sensor network testbed. *Fourth International Symposium on Information Processing in Sensor Networks (IPSN).*

[3] CASAS Smart Home Project. http://ailab.wsu.edu/casas/.

[4] Smart Home-based Health Platform for Behavioral Monitoring and Alteration for Diabetic and Obese Individuals. http://www.icta.ufl.edu/projects_nih.htm.

[5] Chang, J. H. and Tassiulas, L. (2000) Energy Conserving Routing in Wireless Ad-hoc Networks. *INFOCOM*, pp. 22–31.

[6] Chang, J. H. and Tassiulas, L. (2000) Fast Approximate Algorithms for Maximum Lifetime Routing in Wireless Ad-hoc Networks. *Networking.*

[7] Li, Q., Aslam, J., and Rus, D. (2001) Online power-aware routing in wireless Ad-hoc networks. *Mobicom.*

[8] Xue, Y., Cui, Y., and Nahrstedt, K. (2005) A Utility-based Distributed Maximum Lifetime Routing Algorithm forWirelessNetworks. *QShine.*

[9] Kar, K., Kodialam, M., Lakshman, T. V., Tassiulas, L., and Tassiulas, R. (2003) Routing for Network Capacity Maximization in Energy-constrained Ad-hoc Networks. *INFOCOM.*

[10] Lin, L., Shroff, N. B., and Srikant, R. (2007) Asymptotically Optimal Energy-Aware Routing for Multihop Wireless Networks With Renewable Energy Sources. *IEEE/ACM Transactions on Networking*, **15**.

[11] Ok, C., Mitra, P., Lee, S., and Kumara, S. (2007) Maximum Energy Welfare Routing in Wireless Sensor Networks. *NETWORKING 2007.*

[12] Ahn, S. and Kim, D. (2006) Proactive Context-Aware Sensor Networks. *EWSN*.

[13] Menchaca-Mendez, R. and Garcia-Luna-Aceves, J. J. (2010) Robust and Scalable Integrated Routing inMANETs Using Context-Aware Ordered Meshes. *INFOCOM*.

[14] Basu, P. and Redi, J. (2004) Effect of Overhearing Transmissions on Energy Efficiency in Dense Sensor Networks. *IPSN*.

[15] Atkinson, A. B. (1970) On the measurement of inequality. *Journal of Economics Theory* , **?**

[16] Lin, S., Zhang, J., Zhou, G., Gu, L., Stankovic, J. A., and He, T. (2006) ATPC: adaptive transmission power control for wireless sensor networks. *SenSys '06: Proceedings of the 4th international conference on Embedded networked sensor systems*, New York, NY, USA, pp. 223–236. ACM Press.

[17] Das, T. and Parikh, A. (1982) Statistical interpretation, decomposition and properties of Atkinson's inequality index. *Decisions in Economics and Finance* , **?**

[18] Gnawali, O., Fonseca, R., Jamieson, K., Moss, D., and Levis, P. (2009) Collection Tree Protocol. *Proc. of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys)*.

[19] Lu, G., De, D., Xu, M., Song, W.-Z., and Shirazi, B. (2010) TelosW: Enabling Ultra-Low Power Wake-On Sensor Network. *INSS 2010*, June.