Multi-hop/Direct Forwarding (MDF) for Static Wireless Sensor Networks

By: Jing Deng

J. Deng, "Multi-hop/Direct Forwarding (MDF) for Static Wireless Sensor Networks," *ACM Transactions on Sensor Networks*, vol. 5, no. 4, pp. 1-25, November 2009.

Abstract:
The success of Wireless Sensor Networks (WSNs) depends largely on efficient information delivery from target areas toward data sinks. The problem of data forwarding is complicated by the severe energy constraints of sensors in WSNs. In this work, we propose and analyze a data forwarding scheme, termed Multi-hop/Direct Forwarding (MDF), for WSNs where sensor nodes forward data traffic toward a common data sink. In the MDF scheme, a node splits out-going traffic into at most two branches: one is sent to a node that is h units away, the other is sent directly to the data sink. The value of h is chosen to minimize the overall energy consumption of the network. The direct transmission is employed to balance the energy consumption of nodes at different locations and to avoid the so-called “hot spot” problem in data forwarding. In order to calculate its traffic splitting ratio, a node only needs to know the distance toward the common data sink and that of the farthest node. Our analytical and simulation results show that the MDF scheme performs close to, in terms of energy efficiency and network lifetime, the optimum data forwarding rules, which are more complex and computation intensive.
Categories and Subject Descriptors: C.2.2.d [Communication/Networking and Information Technology]: Network Protocols
General Terms: Algorithms, Performance
Additional Key Words and Phrases: Energy efficient, multi-hop forwarding, wireless sensor networks

Article:

1. INTRODUCTION

Wireless communication has become ubiquitous with the development of miniature wireless devices. The substantially reduced size of wireless devices makes it possible to deploy wireless networks with large populations. One such example is Wireless Sensor Networks (WSNs) [Akyildiz et al.]. In WSNs, a large number of small sensors are deployed to a field to accomplish the goal of autonomous event detection and data gathering. The sensed results should be transmitted toward data sinks, which serve as the interface between the network and human. Depending on the nature of WSN applications, single or multiple data sinks are possible.

Large wireless networks, such as WSNs, are expected to have profound technological and economic impacts on our society. Their success is nonetheless determined by whether they can efficiently deliver information from target areas toward data sinks. This task is further complicated by severe energy constraints of sensors in WSNs. In the tiny sensors, the size and energy reserve of the sensor battery are extremely limited. Battery replacement or recharge is expected to be difficult. Therefore, use of the limited energy should be planned carefully [Sankar and Liu].

In this work, we focus on the data forwarding problem of sensor nodes toward the common data sink. We propose a data forwarding technique, termed the Multihop/Direct Forwarding (MDF) scheme. In the MDF scheme, the traffic that needs to be forwarded by a node is split into at most two branches: one is sent to the sensor node that is h units away; the other is sent directly to the data sink. The value of h is chosen to minimize the overall energy consumption of the network. The direct transmission is employed to balance the energy consumption of nodes at different locations and to avoid the so-called "hot spot" problem, which arises when some nodes need to forward much more traffic than others.

Our study is based on a network model with evenly distributed nodes on a straight line. This one-dimensional network model exists in many WSN applications such as highway surveillance, border-line surveillance, power-line monitoring, and street light surveillance (though the sensors may not be evenly distributed). Extensions of the MDF scheme into other network scenarios are provided in Section 5 including: two-dimensional networks; unevenly distributed nodes; and limited maximum transmission range. Note that data aggregation and data fusion do affect the data forwarding model, but the fundamentals of the problem remain. Also note that the focus of this work is not on collision domain or transmission scheduling. Rather, we focus on how to split the traffic of each node so that the lifetime of the network will be extended (in light-traffic WSNs).

We summarize our major contributions as follows:

—We have proposed a new data forwarding scheme termed MDF. The MDF scheme splits the traffic of each node into at most two branches toward the common data sink. The uniqueness of the MDF scheme lies in the simplicity of the forwarding rules and the small computation cost for each node. Each node only needs to know its index/distance toward the common data sink and that of the farthest node in order to calculate its traffic forwarding rule. We have provided closed-form equations for traffic splitting rules for one-dimensional networks. In two-dimensional networks, the relatively simple recursive calculations can be performed at individual nodes or at the data sink for broadcast.

—We have developed a framework to derive the data forwarding rules for the MDF scheme in a chain network. We then extend this framework to WSNs with different network properties such as two-dimensional topology, random node distribution, and limited maximum transmission range; and

—We have performed extensive simulations to evaluate the MDF scheme and to compare it with several related schemes.

The rest of the paper is organized as follows: In Section 2, we present the problem formulation and preliminaries. We then propose the MDF scheme in Section 3. Performance evaluation of the MDF scheme are presented in Section 4. We discuss related issues of the MDF scheme under various network conditions in Section 5. We summarize related work in Section 6. We then conclude our work in Section 7 and state some future research directions.

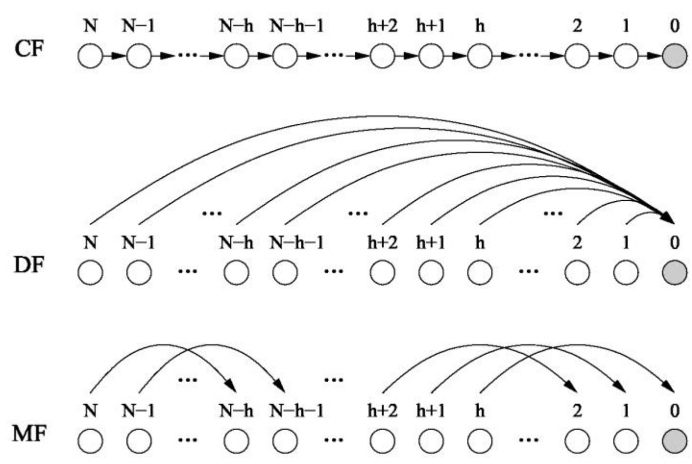

Fig. 1. Illustrations of the CF, the DF, and the MF schemes in a one-dimensional network. Nodes are assumed to be evenly distributed. In the CF scheme, nodes send to closest neighbor. Direct transmissions toward the data sink are employed in the DF scheme. In the MF scheme, multi-hop forwarding is used with hop distance of h units.

2. PROBLEM FORMULATION AND PRELIMINARY

We study a WSN in which the sensor nodes are evenly distributed on a straight line (i.e., the nodes form a chain as illustrated in Fig. 1.) The observer, or data sink, of the sensor network locates at one end of the line. All nodes have the same data generation rate, one packet per unit time (e.g., one second) to be forwarded to the data sink.

We establish the following notations:

—N: total number of nodes, excluding the data sink, in the chain network. These nodes are indexed as $i = 1, 2, \cdots, N$, starting from the node closest to the data sink. Therefore, node index also represents the *normalized distance* toward the data sink, which is indexed as node 0 (cf. Fig. 1);

—$T_{i,j}$: the traffic rate sent from node i to node j, $i, j = 0, 1, 2, \cdots, N$. $T_{i,i} = 0$. The traffic rates may not be integers;

—Energy consumption of sending one packet from node i to node j is:[1]

$$E = k_0 + (i - j)^w , \qquad (1)$$

where w is the path loss exponent and assumed to be 2 in this work [Feeney and Nilsson, Heinzelman et al., Zhao and Guibas]. We call k_0 the energy constant, which includes all energy consumption, such as receiving/idling energy consumption, that are unrelated to transmission distance. k_0 captures the effects of different ratios of energy spent on transmission, reception, and circuitry processing;

—E_i : the energy consumption of node i (to forward traffic), i.e.,

$$E_i = \sum_{j=0}^{N} T_{i,j}[k_0 + (i - j)^{\omega}] \; ; \qquad (2)$$

—$\mathcal{E}$: the maximum node energy consumption among all nodes;

—$\bar{E}$: the average energy consumption of all nodes.

We assume that packet reception does not consume extra energy. This is justified by the fact that receiving nodes and idle nodes consume approximately the same amount of energy [Feeney and Nilsson]. We further assume that each node is able to adjust its transmission range by varying the transmission power. Each of these nodes can reach the data sink directly if necessary [Heinzelman et al.].

Note that, we do not consider transmission losses in our analysis. We argue that packet losses occurring in all transmissions will have a similar effect on all hops, which can be linearly mapped into data traffic. The packet loss over different transmission distances can also be compensated with different transmission powers, as shown in (1).

The problem of searching for the optimum data forwarding rules in a WSN is to look for a set of $T_{i,j}$, $0 \le i, j \le N$, such that the overall energy consumption is minimized and the network can run for a longer time. In the one-dimensional network as shown in Fig. 1, the problem can be formulated as:

Problem: *In a chain network with N nodes and one common data sink (cf. Fig. 1), each node generates one packet per unit time. How should the values of* $T_{i,j}$, $i, j \in \{0, 1, \cdots, N\}$ *be chosen such that the network lifetime is maximized?*

We first introduce our definition of network lifetime, which is measured as the time when the first node runs out of battery energy. Other network lifetime definitions are possible, such as the time when α, 0 < α < 1, of all nodes run out of battery energy. We will investigate these different network lifetimes in Section 4.

Based on the definition of network lifetime, the optimum values for $T_{i,j}$ should satisfy

$$\min_{T_{i,j}}[\max_i(E_i)] \ , \tag{3}$$

where E_i is given by (2) and $T_{i,j}$ should satisfy

$$1 + \sum_{a\in\{1,2,\cdots N\},a\neq i} T_{a,i} = \sum_{a\in\{0,1,\cdots N\},a\neq i} T_{i,a} \quad \text{for } i = 1,2,\cdots N \ . \tag{4}$$

This is a linear programming problem with constraints on each node's traffic flow, as pointed out by Perillo et al. [Perillo et al.]. It has $N(N + 1)/2$ variables and N constraints (these variables are needed in order to explore the entire possible transmission scheme domain). While this optimization may be computed at the data sink and the results be disseminated to the sensors, the large number of sensors make such computations and downstream information delivery costly.

Table I. Different forwarding schemes discussed in this paper

Scheme	Details
LPF	forwarding scheme based on linear programming
CF	the Closest Forwarding scheme
DF	the Direct Forwarding scheme
MF	the Multi-hop Forwarding scheme
GF	the Genie Forwarding scheme
MDF	the Multi-hop/Direct Forwarding scheme

2.1 Optimum Forwarding Distance

We look at an abstracted optimization problem of forwarding distance (transmission range) as follows [Gao]. Using an energy consumption model of (1) for wireless transmission, we further assume that the distance between the source and the destination is d units and there are enough nodes on the path between these two nodes. We shall derive the optimum transmission range, h, that the source should use to send the data packets toward the destination.

When a transmission range of h is used, the energy consumption of each hop is $k_0 + h^w$ and the number of hops that is needed to reach the destination is approximately $\frac{d}{h}$. The total energy consumption of sending one data packet toward the destination becomes

$$E \approx \frac{d}{h}(k_0 + h^\omega) \ .$$

Optimizing E with h, we obtain the optimum transmission range

$$h^* \approx \left(\frac{k_0}{\omega - 1}\right)^{\frac{1}{\omega}} \ .$$

In a chain network shown in Fig. 1, the hop distance must be positive integers therefore

$$h^* \approx \max\left(1, \left\langle \left(\frac{k_0}{\omega - 1}\right)^{\frac{1}{\omega}} \right\rangle \right) \ , \tag{5}$$

where $\langle \, . \, \rangle$ represents rounding to the closest integer.

Such an optimum transmission range may be interpreted as follows: when h is higher than h^*, the energy consumption of each hop increases due to the exponential path loss, leading to energy wastage. However, when h is lower than h^*, more hops are needed in order to reach the destination, consuming more energy because of the term k_0. We will exploit such an optimum transmission range in our proposed scheme (Section 3).

2.2 Some Special Forwarding Rules

Before presenting our proposed scheme, we study several special forwarding schemes summarized in Table I (cf. Fig. 1).

The *Linear Programming Forwarding* (LPF) scheme is the forwarding rule derived from linear programming. The energy consumption of the LPF scheme serves as a realistic lower bound for our study.

In the *Closest Forwarding* (CF) scheme, a node only forwards its traffic to its closest neighbor toward the data sink. Therefore, $T_{i,j} = 0$ except when $i - j = 1$. When k_0 is close to 0, the CF scheme consumes the least overall energy. The energy consumption of nodes, however, is unbalanced. For node i,

$$E_i^{(CF)} = (N - i + 1)(k_0 + 1^\omega) ,$$

where $1 \leq i \leq N$. The maximum node energy consumption occurs at node 1,

$$\mathcal{E}^{(CF)} = N(k_0 + 1) .$$

The average energy consumption is

$$\overline{E}^{(CF)} = \frac{\sum_{i=1}^{N} E_i^{(CF)}}{N} = \frac{(N+1)(k_0+1)}{2} .$$

In the *Direct Forwarding* (DF) scheme, all nodes forward their traffic directly to the data sink. Therefore, $T_{i,j} = 0$ except when $j = 0$. Similar to the CF scheme, the energy consumption of nodes in the DF scheme is unbalanced either. Node energy consumption increases with the distance from the data sink. For node i,

$$E_i^{(DF)} = k_0 + i^\omega ,$$

where $1 \leq i \leq N$. The maximum node energy consumption occurs at node N,

$$\mathcal{E}^{(DF)} = k_0 + N^\omega .$$

The average energy consumption is

$$\overline{E}^{(DF)} = \frac{\sum_{i=1}^{N} E_i^{(DF)}}{N} = k_0 + \frac{\sum_{i=1}^{N} i^\omega}{N} . \qquad (6)$$

In the *Multi-hop Forwarding* (MF) scheme, all traffic are sent through the optimum hop distance, h given by (5), toward the data sink. The MF scheme is similar to the CF scheme except that nodes do not send to their closest neighbors. Therefore, $T_{i,j} = 0$ except when $i - j = h$ or when $i < h$ and $j = 0$.

In the MF scheme, node energy consumption increases as node index decreases, depending on the number of nodes sending traffic through a node. Node i receives $\left\lfloor \frac{N-i}{h} \right\rfloor$ packets from nodes that are farther away from the data sink, where $\lfloor x \rfloor$ is the floor function returning the largest integer that is not larger than x. Therefore, node i will send out $\left\lfloor \frac{N-i}{h} \right\rfloor + 1$ packets. The energy consumption of node i is then

$$E_i^{(MF)} = [k_0 + \min(h,i)^\omega] \cdot \left(\left\lfloor \frac{N-i}{h} \right\rfloor + 1 \right) ,$$

where $1 \leq i \leq N$ and the min function represents those nodes that are less than h units from the data sink. The maximum node energy consumption occurs at node $1 \leq i \leq h$,

$$\mathcal{E}^{(MF)} = \max_{1 \leq i \leq h} E_i^{(MF)} .$$

The average energy consumption is simply $\overline{E}^{(MF)} = \sum_{i=1}^{N} E_i^{(MF)} / N$.

In the *Genie Forwarding* (GF) scheme, there exists a *genie* that is able to redistribute the energy among the N nodes without any extra cost. Therefore, the residual energy of the N nodes is always balanced. The GF scheme is similar to the MF scheme, except that the GF scheme has a cost-free energy-balancing genie. We can imagine this scheme to be operating in a network where all nodes share a virtual massive battery. Its node energy consumption serves as an unrealistic lower bound in our study.

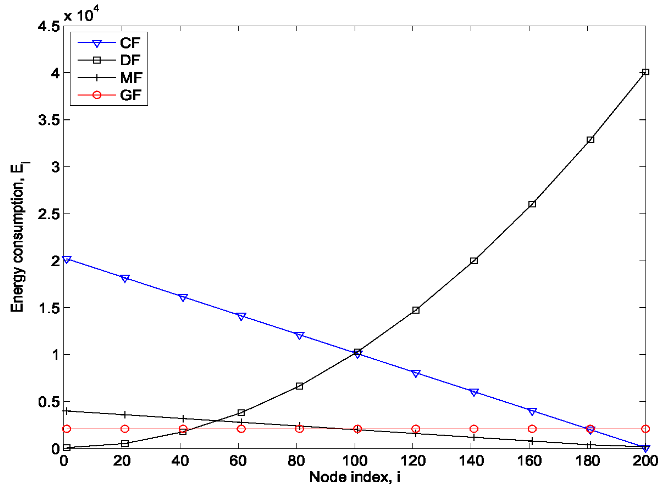

Fig. 2. Comparison of node energy consumption for the CF, the DF, the MF, and the GF schemes ($N = 200$, $k_0 = 100$).

Similar to the MF scheme, the pre-balanced energy consumption of node i is

$$E_i^{(GF)} = [k_0 + \min(h, i)^{\omega}] \cdot \left(\left\lfloor \frac{N-i}{h} \right\rfloor + 1 \right) ,$$

where $1 \leq i \leq N$. Due to the energy-balancing activity performed by the genie, the maximum and average node energy consumption are the same,

$$\mathcal{E}^{(GF)} = \overline{E}^{(GF)} = \sum_{i=1}^{N} E_i^{(GF)} / N .$$

We calculated energy consumption of different nodes in four of these special forward rules (the CF, the DF, the MF, and the GF schemes) and showed them in Fig. 2. Node energy consumption of the DF scheme increases exponentially with node index (distance from the data sink). The CF scheme exhibits a reversed trend, i.e., node energy consumption decreases as node index increases. Compared to the CF and the DF scheme, the MF scheme leads to a much more balanced energy consumption for the nodes. Still, a declining trend as i increases can be seen. The GF scheme results in an efficient and balanced energy consumption for all nodes. This, however, is achieved by the genie.

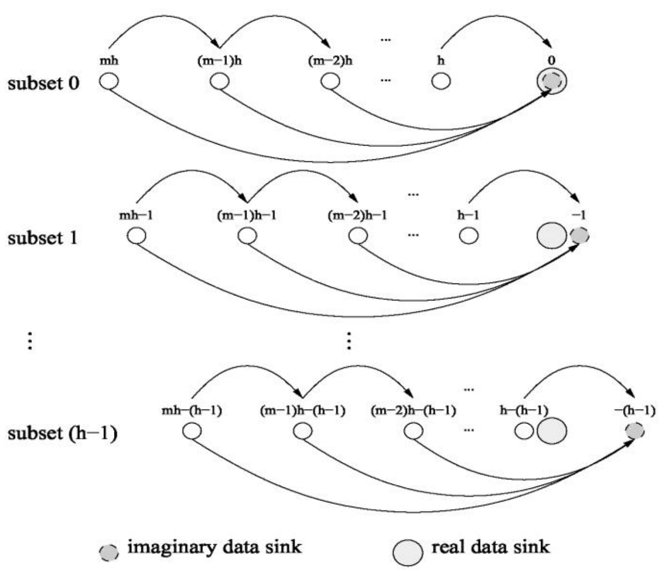

Fig. 3. Illustrations of the MDF scheme. There are h subsets. The dashed shaded circles represent the locations of imaginary data sinks. All traffic should be sent to the real data sink instead.

3. THE MULTI-HOP/DIRECT FORWARDING (MDF) SCHEME

In order to achieve simplicity and energy efficiency for data forwarding, we propose to restrict each node to split its data traffic into at most two branches, namely, $T_{i,(i-h)}$ and $T_{i,0}$ when $i > h$; and $T_{i,0}$ only when $i \leq h$. We term this scheme as the Multi-hop/Direct Forwarding (MDF) scheme, which exploits the optimum hop distance, *h*, given by (5).

Note that node *i*, $i > h$, does not have to send exactly $T_{i,\, i-h}$ packets before switching to the $T_{i,0}$ branch. Instead, it can spend a much longer time in each of the states of multi-hop transmission ($T_{i,\, i\text{-}h}$) and direct transmission ($T_{i,0}$). Such an implementation reduces the switching cost between different transmission states.

We have assumed that the nodes are able to switch to a transmission power that reaches the common data sink. We argue that WSNs with heterogeneous sensor deployments do have such capabilities. When the nodes cannot reach the common data sink, they will have to send the $T_{i,0}$ branch traffic to the node that is farthest away from itself (cf. Section 5.3).

The details of the MDF scheme can be explained with the help of Fig. 3. In the MDF scheme, the sensors are divided into *h* subsets, each of which sends their traffic independently toward the data sink. The distances of any two consecutive nodes in each of these subsets are *h*, except the last node toward the data sink, which might be closer than *h* units to the data sink. In order to simplify our discussions, we imagine there are *h* "data sinks", each of which serves as the evenly-spaced data sink of one subset. With the imaginary data sinks, all subsets are essentially the same.[2] Therefore, we only need to study one of these subsets, e.g., subset 0.

In subset 0, *m* nodes are sending information toward node 0, where

$$m = \left\lfloor \frac{N}{h} \right\rfloor . \tag{7}$$

Observing the traffic flow of node nh, we have:

$$T_{(n+1)h,nh} + 1 = T_{nh,(n-1)h} + T_{nh,0} , \tag{8}$$

where $1 < n < m$. When $n = 1$, we have:

$$T_{h,0} = T_{2h,h} + 1 . \tag{9}$$

The energy consumption of nodes $(n+1)h$ and nh is

$$\begin{aligned} E_{(n+1)h} &= T_{(n+1)h,nh} \cdot [k_0 + h^\omega] \\ &\quad + T_{(n+1)h,0} \cdot [k_0 + ((n+1)h)^\omega] , \end{aligned} \tag{10}$$

$$\begin{aligned} E_{nh} &= T_{nh,(n-1)h} \cdot [k_0 + h^\omega] \\ &\quad + T_{nh,0} \cdot [k_0 + (nh)^\omega] , \end{aligned} \tag{11}$$

respectively. Since our goal is to balance node energy consumption, we need to make sure that

$$\begin{aligned} & T_{(n+1)h,nh} \cdot [k_0 + h^\omega] + T_{(n+1)h,0} \cdot [k_0 + ((n+1)h)^\omega] \\ = {} & T_{nh,(n-1)h} \cdot [k_0 + h^\omega] + T_{nh,0} \cdot [k_0 + (nh)^\omega] . \end{aligned} \tag{12}$$

when $1 < n < m$.

Dividing (12) by $(k_0 + h^\omega)$ and combining it with (8), we have

$$\begin{aligned} & \frac{k_0 + ((n+1)h)^\omega}{k_0 + h^\omega} \cdot T_{(n+1)h,0} \\ = {} & \frac{(nh)^\omega - h^\omega}{(nh)^\omega + k_0} \cdot \frac{k_0 + (nh)^\omega}{k_0 + h^\omega} \cdot T_{nh,0} + 1 , \end{aligned} \tag{13}$$

when $1 < n < m$.

Defining

$$x_n \equiv \frac{k_0 + (nh)^\omega}{k_0 + h^\omega} \cdot T_{nh,0} , \tag{14}$$

and

$$k_n \equiv \frac{(nh)^\omega - h^\omega}{(nh)^\omega + k_0} , \tag{15}$$

for $n > 0$, we can rearrange (13) into

$$x_{n+1} = 1 + k_n \cdot x_n , \tag{16}$$

where $1 < n < m$.

Using (16) recursively, we obtain

$$\begin{aligned} x_n &= 1 + k_{n-1}x_{n-1} \\ &= 1 + k_{n-1}(1 + k_{n-2}x_{n-2}) \\ &\cdots \\ &= 1 + \sum_{i=3}^{n-1}\left(\prod_{j=i}^{n-1} k_j\right) + \left(\prod_{j=2}^{n-1} k_j\right) x_2 \ , \end{aligned} \tag{17}$$

where $1 < n \leq m$. $k_0 + h^w$

We need to find $x_2 = \frac{k0 + (2h)w}{k0 + hw} \cdot T_{2h,0}$. In fact, the energy consumption of node $2h$ and node h should be the same,

$$\begin{aligned} & T_{2h,h} \cdot [k_0 + h^\omega] + T_{2h,0} \cdot [k_0 + (2h)^\omega] \\ = \ & T_{h,0} \cdot [k_0 + h^\omega] \ . \end{aligned} \tag{18}$$

Combining (9) and (18), we have

$$T_{2h,0} = \frac{k_0 + h^\omega}{k_0 + (2h)^\omega} \tag{19}$$

and

$$x_2 = 1 \ .$$

Thus, x_n can be expressed as

$$x_n = 1 + \sum_{i=2}^{n-1}\left(\prod_{j=i}^{n-1} k_j\right) = 1 + \sum_{i=2}^{n-1}\left(\prod_{j=i}^{n-1} \frac{(jh)^\omega - h^\omega}{(jh)^\omega + k_0}\right) \ .$$

$T_{nh,0}$ is then

$$T_{nh,0} = \frac{k_0 + h^\omega}{k_0 + (nh)^\omega} \cdot \left[1 + \sum_{i=2}^{n-1}\left(\prod_{j=i}^{n-1} \frac{(jh)^\omega - h^\omega}{(jh)^\omega + k_0}\right)\right] \ , \tag{20}$$

where $1 < n \leq m$.

In order to calculate $T_{nh,(n-1)h}$ for $n > 1$, we need to find the energy consumption of node nh. Since the energy consumption of every node is the same, we firstly study node mh, which is farthest away from the data sink in this subset.

$$\begin{aligned} E_{mh} &= T_{mh,(m-1)h} \cdot [k_0 + h^\omega] + T_{mh,0} \cdot [k_0 + (mh)^\omega] \\ &= [k_0 + h^\omega] + T_{mh,0} \cdot [(mh)^\omega - h^\omega] \ , \end{aligned}$$

where we have used $T_{mh,(m-1)h} + T_{mh,0} = 1$, because it receives no traffic from other nodes. On the other hand, the energy consumption of node nh is

$$E_{nh} = T_{nh,(n-1)h} \cdot [k_0 + h^\omega] + T_{nh,0} \cdot [k_0 + (nh)^\omega] \ .$$

Since $E_{nh} = E_{mh}$, we have

$$\begin{aligned} & T_{nh,(n-1)h} \\ = \ & \frac{T_{mh,0}[(mh)^\omega - h^\omega] - T_{nh,0}[k_0 + (nh)^\omega]}{k_0 + h^\omega} + 1, \end{aligned} \tag{21}$$

where $T_{nh,0}$ and $T_{mh,0}$ are given by (20), and $1 < n \leq m$.

In order for node i to find out its data forwarding rule, it only needs to know its index (distance) toward the data sink and the total number of nodes in the network.[3] It calculates the values of n and m in its subset as

$$n = \left\lceil \frac{i}{h} \right\rceil \tag{22}$$

$$m = n + \left\lfloor \frac{N-i}{h} \right\rfloor \ , \tag{23}$$

where $\lceil x \rceil$ is the ceiling function returning the smallest integer that is not smaller than x, and apply them to (20), (21), and (9). Note that the value of m varies slightly in different subsets because of border effects.

Neglecting border effects and assuming that the real data sink locates at the position of the imaginary data sinks, we would have all nodes in the network with the same energy consumption. The actual energy consumption is slightly different. We express the energy consumption of node i as follows:
When $i > h$,

$$E_i = T_{nh,0}\left(k_0 + i^{\omega}\right) + T_{nh,(n-1)h}\left(k_0 + h^{\omega}\right), \tag{24}$$

where $T_{nh,0}$, $T_{nh,(n-1)h}$, n, and m are given by (20), (21), (22), and (23), respectively.

When $i \leq h$,

$$E_i = T_{h,0}\left(k_0 + i^{\omega}\right), \tag{25}$$

where $T_{h,0}$ is given by (9) and m is given by (23).

4. PERFORMANCE EVALUATION

We used Matlab to perform simulations in order to evaluate the MDF scheme and to compare it with related schemes. The methodology of our simulations is the following: we place N nodes evenly on a straight line and use different data forwarding rules including the MF, the LPF, and the MDF schemes and measure node energy consumption and network lifetime. Each node is assumed to generate 1 data packet per unit time. The distance between any two consecutive sensors is 1 unit.

Figure 4 shows energy consumption of different hop distance, h, when the MDF scheme is employed and k_0 = 100. Different N values have been chosen. In order to show the optimality more clearly in Fig. 4, we present normalized energy consumption, which is calculated as the average energy consumption divided by the minimum value of energy consumption along all possible h, i.e., E/E_{min}. It can be seen that both small values of h and large values of h lead to higher energy consumption. The optimum h matches (5) very well. Such a trend is insensitive to the change of N. The saw-shape curves shown in Fig. 4, especially when h is large and N is small, are caused by border effects.

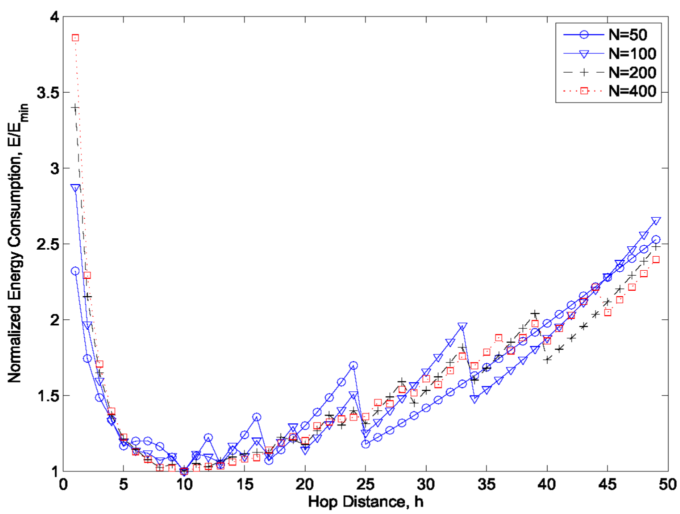

Fig. 4. Normalized energy consumption of different hop distance, h, when the MDF scheme is employed and $k_0 = 100$. Normalized energy consumption is calculated as the average energy consumption divided by the minimum value of energy consumption along all possible h.

Another interesting observation from Fig. 4 is the relatively flat shape of the curves in the neighborhood of the optimum h. This suggests that an h value derived from a slightly inaccurate k_0 may still achieve relatively good performance in energy conservation.

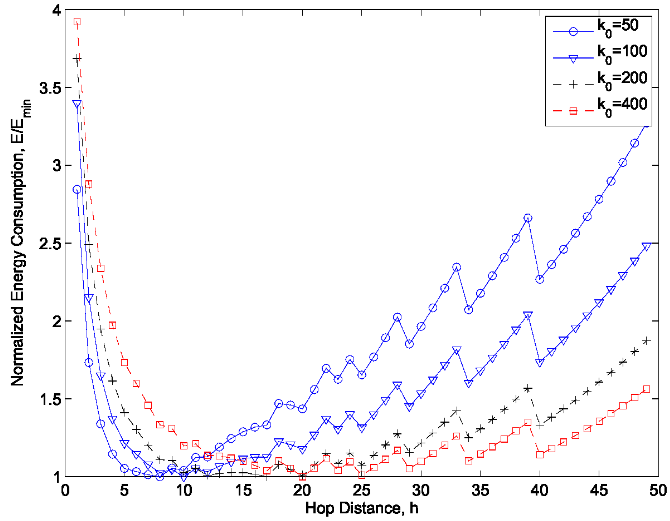

Fig. 5. Normalized energy consumption of different hop distance, h ($N = 200$). Optimum h matches (5) well.

In Fig. 5, we present normalized energy consumption of the MDF scheme as a function of h for different k_0. The number of nodes is fixed at $N = 200$. As k_0 increases, the optimum h increases, matching (5) well. Note that the identical occurrences of saw-shapes in the curves for different N are caused by border effects.

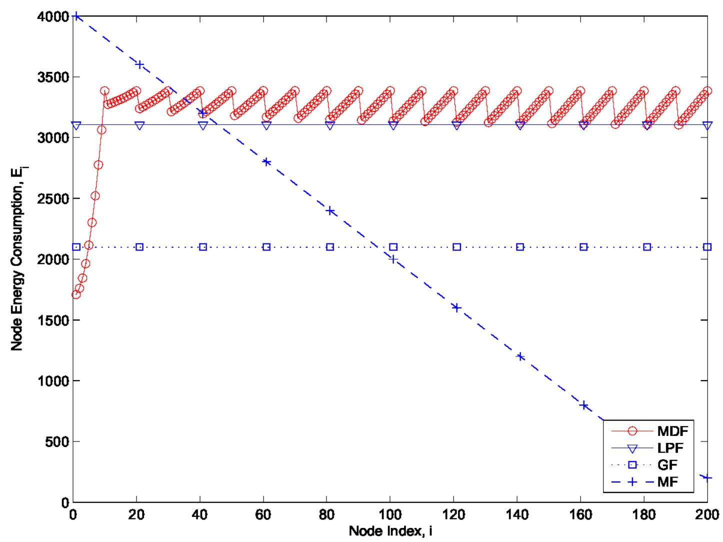

Fig. 6. Node energy consumption for the MDF, LPF, GF, and MF schemes ($N = 200$, $k_0 = 100$).

In Fig. 6, we compare node energy consumption when the MDF, the LPF, the GF, and the MF schemes are employed. When the MDF scheme is used, node energy consumption is rather balanced for all nodes, except for those nodes that are less than h units from the data sink. The node energy consumption of the MDF scheme is very close to that of the LPF scheme. Both of the MDF and the LPF schemes are outperformed by the unrealistic GF scheme. When the MF scheme is employed, node energy consumption decreases as node index (distance from the data sink) increases. Note that the other two schemes, the CF and the DF schemes, are outperformed by the schemes shown in Fig. 6 (cf. Fig. 2).

We compare the average node energy consumption of the MDF and the LPF schemes as a function of k_0 for different N in Fig. 7. Naturally, energy consumption increases with k_0. Note that, in LPF, all nodes have the same energy consumption. In MDF, however, sensors have different energy consumption due to border effects and their different distances toward the real data sink. We can observe that the energy consumption of the MDF scheme is at most 10% higher than that of the LPF scheme.

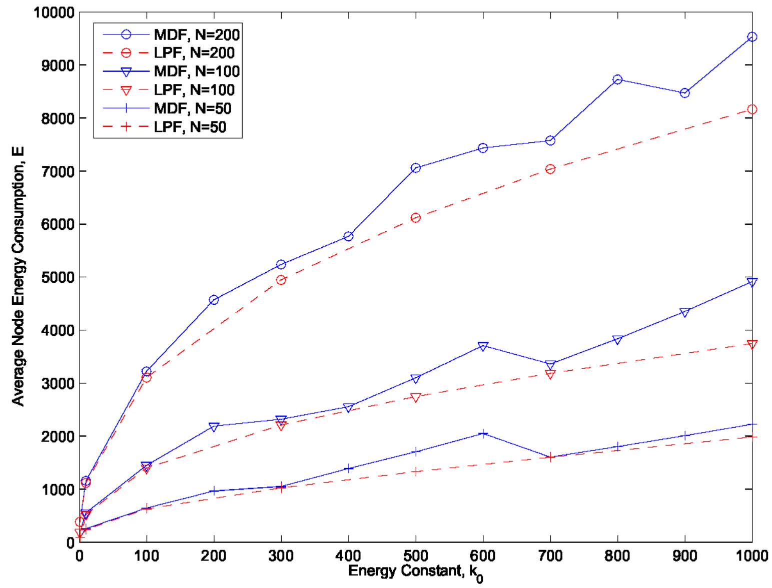

Fig. 7. Average node energy consumption of the MDF and the LPF schemes as a function of k_0.

Based on Figs. 6 and 7, we can see that the MDF scheme has a nice performance that is very close to the LPF scheme. This is mainly due to our design of the MDF scheme that takes into the consideration of the traffic forward from all of the outer nodes to the data sink. The use of multi-hop forwarding ensures the optimality of the MDF scheme, while direct forwarding balances the energy consumption among different nodes. Yet, the MDF scheme has much lower computation complexity than the LPF scheme.

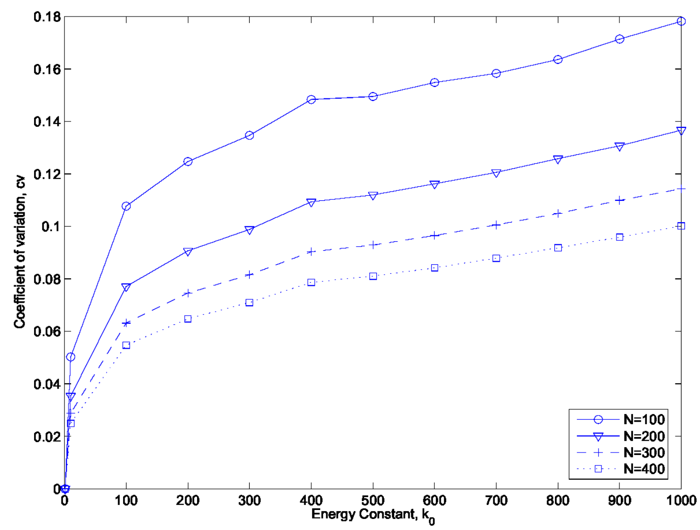

Fig. 8. Coefficient of variation of energy consumption when the MDF scheme is used. Coefficient of variation is calculated as standard deviation divided by mean.

Since the actual node energy consumption of the MDF scheme is not balanced, we should investigate the difference of energy consumption on different nodes. Figure 8 serves this purpose by showing the coefficient of variation, *cv*, of node energy consumption when the MDF scheme is employed. Coefficient of variation is calculated as standard deviation divided by the value of mean. The value of *cv* increases with k_0, caused by the larger difference in each packet transmission. As *N* increases, border effects diminish, leading to lower *cv*. Most of these differences in energy consumption can be attributed to the lower energy consumption of those nodes that are less than h units from the data sink (cf. Fig. 6).

In Fig. 9, we compare the network lifetime for different forwarding schemes: MDF, LPF, GF, MF, CF, and DF. The network lifetime shown in Fig. 9 is defined as the time when the first node in the network runs out of battery energy.[4] As we can observe in Fig. 9, the lifetime of the CF and the DF schemes is significantly shorter than all other schemes. This is because of the imbalance energy consumption of nodes in the network. The network lifetime of the MDF scheme approaches that of the LPF scheme, both of which are better than the MF scheme. The GF scheme serves as the unrealistic upper limit. As can be seen in Fig. 9, network lifetime decreases as k_0 increases because of the increased energy consumption in all transmissions.

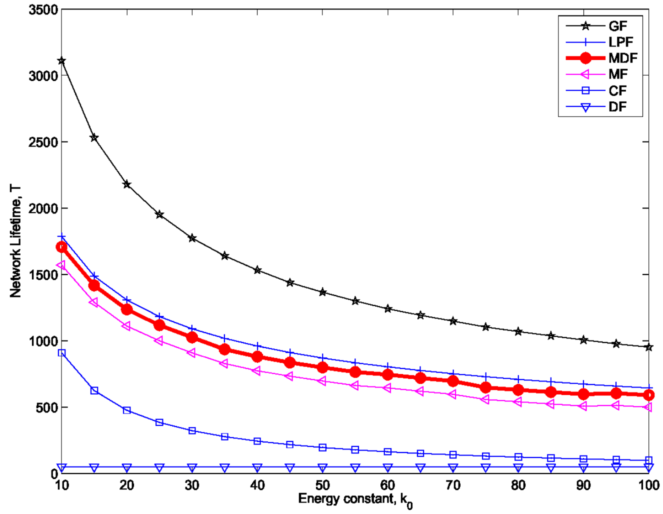

Fig. 9. Network lifetime when the MDF, LPF, GF, MF, CF, and DF schemes are employed, respectively ($N = 200$). Network lifetime in this figure is defined as the time when the first node runs out of battery energy.

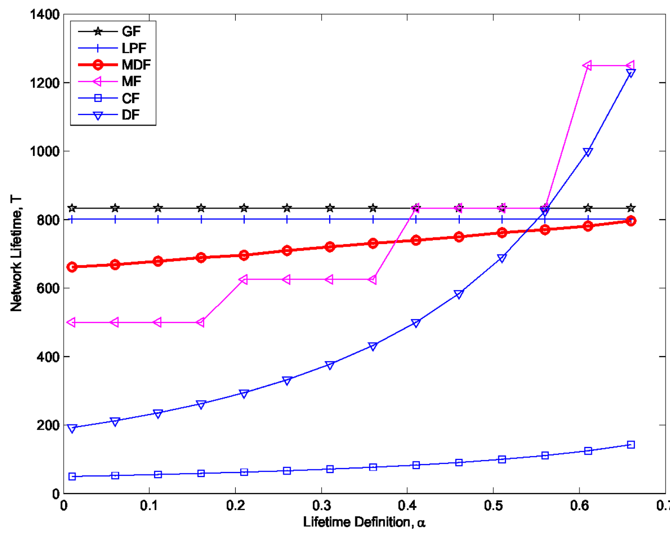

Fig. 10. Network lifetime when the MDF, LPF, GF, MF, CF, and DF schemes are employed, respectively ($N = 100$, $k_0 = 400$). Network lifetime has been defined as the time when α of all sensors run out of battery energy.

We study network lifetime in Fig. 10. Instead of defining network lifetime as the time when the first sensor runs out of energy, we investigate a network lifetime measured as the time when α, $0 < \alpha < 1$, of all sensors run out of energy. Therefore, network lifetime should be a non-decreasing function of α, confirmed by Fig. 10. It can be seen clearly that the network lifetime of the MDF scheme approaches that of the LPF scheme, both of which are lower than that of the GF scheme. The network lifetimes of the LPF and the GF schemes are horizontal lines because node energy consumption is balanced. The MF, CF, and DF schemes, on the other hand, have an unbalanced node energy consumption, leading to the rising curve in Fig. 10.

5. DISCUSSIONS OF THE MDF SCHEME

We discuss the MDF scheme under different network scenarios in this section. These include two-dimensional networks, random node distribution, and limited maximum transmission distance.

5.1 Two-Dimensional WSNs

Thus far, we have only considered one-dimensional WSNs. When sensors are deployed to two-dimensional fields, our system model needs to be revised. In fact, the power assignment problem of two-dimensional wireless networks have been shown to be NP-hard [Cagalj et al.]. In this work, however, we use an approach similar to that of [Perillo et al.] to identify the optimum transmission schedule of these nodes. We group the nodes that are approximately *i* units away from the data sink into a single virtual node *i*. The traffic generated by these virtual nodes is termed λ_i and the energy reserve is assumed to be b_i.

Based on random node distribution on a circular disk, the number of nodes that are roughly *i* units away from the data sink is proportional to $i^2 — (i — 1)^2 \approx 2i$ [Chang and Liu, Perillo et al.]. The traffic generated from the virtual node *i* is $\lambda_i \approx 2_i$. The energy reserve of the virtual node *i* is $b_i \approx 2_i$. Therefore, our system model becomes heterogeneous in residual energy as well as traffic generation. Note that, when $\lambda_i = 1$ and b_i is a constant, the system degenerates to the chain network that we investigated in Section 3.

In Fig. 11, we illustrate a two-dimensional network where the common data sink sits at the center. The nodes with the same distance toward the data sink are grouped into a virtual node with b_n overall battery energy and λ_n generated traffic.

Observing the traffic flow of node *nh*, we have the following equation:

$$T_{(n+1)h,nh} + \lambda_{nh} = T_{nh,(n-1)h} + T_{nh,0} \ , \tag{26}$$

when $1 < n < m$.

In order to balance the energy consumption of nodes with different distances toward the common data sink, we should make sure (instead of (12))

$$\begin{aligned} & \frac{T_{(n+1)h,nh} \cdot [k_0 + h^\omega] + T_{(n+1)h,0} \cdot [k_0 + ((n+1)h)^\omega]}{b_{(n+1)h}} \\ = & \frac{T_{nh,(n-1)h} \cdot [k_0 + h^\omega] + T_{nh,0} \cdot [k_0 + (nh)^\omega]}{b_{nh}} \ . \end{aligned} \tag{27}$$

when $1 < n < m$.

In order to derive the form for $T_{nh,(n-1)h}$ and $T_{nh,0}$ for the nodes with *nh* distance from the common data sink, we need to know the form of λ_{nh} and b_{nh}. In the following, we demonstrate the derivation of $T_{nh,(n-1)h}$ and $T_{nh,0}$ based on random two-dimensional node distribution where $\lambda_{nh} = 2nh$ and $b_{nh} = 2nh$.

Defining k_n, R_n and S_n as

$$\begin{aligned} k_n &= k_0 + (nh)^\omega \quad n = 0, 1, \ldots, m \\ R_n &= \frac{T_{nh,(n-1)h} + hn(n-1)}{n} \quad n = 2, 3, \ldots, m \\ S_n &= \frac{T_{nh,0} \cdot \frac{k_n}{k_1} - hn(n-1)}{n} \quad n = 2, 3, \ldots, m \ , \end{aligned} \tag{28}$$

we can rewrite (27) and (26) into the following recursive equations:

$$\begin{aligned} R_n + S_n &= R_{n-1} + S_{n-1} = \cdots = R_2 + S_2 \\ R_n &= \frac{n-1}{n} R_{n-1} + T_{(n-1)h,0} \ , \end{aligned} \tag{29}$$

where $n = 3, \ldots, m$. Therefore,

$$S_n = R_2 + S_2 - R_n = T_{h,0} - R_n \ , \tag{30}$$

where $n = 3, \ldots, m$ and we have used the results of $R_2 = S_2 = \frac{T_{h,0}}{2}$ in Appendix.

Once the values of S_n is known, $T_{nh,0}$ can be calculated as

$$T_{nh,0} = [nS_n + hn(n-1)] \cdot \frac{k_1}{k_n} \ , \tag{31}$$

Therefore, $T_{nh,0}$, n = 2, 3, ... , m, can be shown as a linear function $T_{h,0}$, i.e.,

$$T_{nh,0} = a_n T_{h,0} + c_n \ , \tag{32}$$

where n = 2, 3, ... , m.

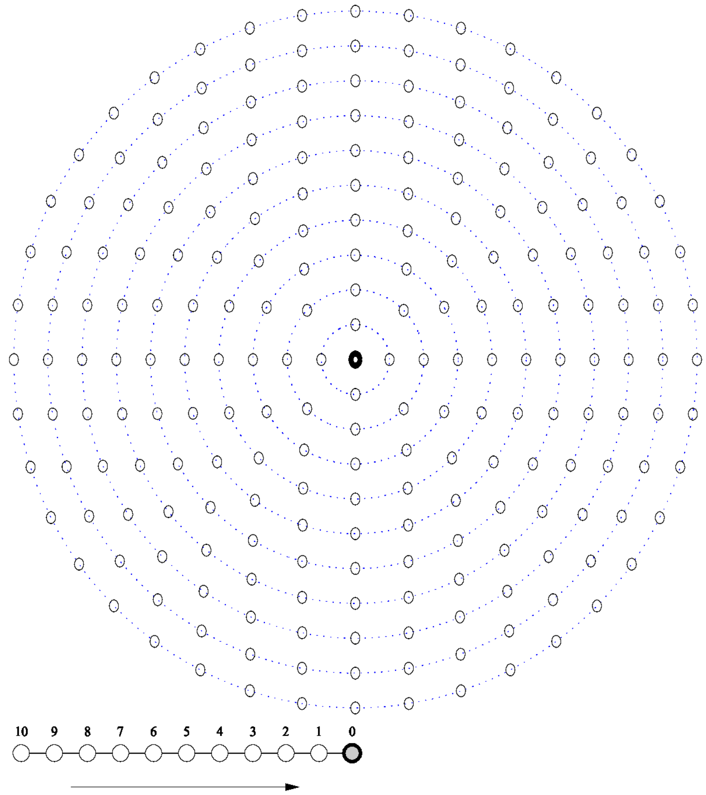

Fig. 11. Illustration of a two-dimensional network and virtual chain network where all nodes send their traffic toward the common data sink (darker node). For example, node i in the chain network represents all those nodes in the two-dimensional network that are i units away from the data sink, i.e., those nodes on the same dotted line circles.

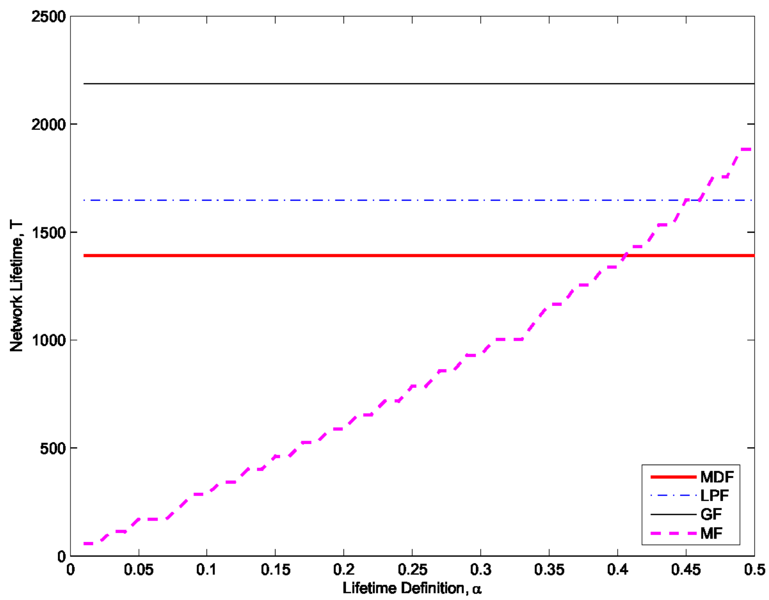

Fig. 12. Network lifetime when the MDF, LPF, GF, and MF schemes are employed in two-dimensional networks ($N = 50$, $k_0 = 100$).

The boundary condition may be obtained through traffic generation of all nodes

$$\sum_{n=1}^{m} T_{nh,0} = \sum_{n=1}^{m} 2hn = hm(m-1) \ , \tag{33}$$

which can be used to solve $T_{h,0}$, i.e.,

$$\sum_{n=1}^{m} T_{nh,0} = T_{h,0} + \sum_{n=2}^{m} [a_n T_{h,0} + c_n] = hm(m-1) \ .$$

Therefore, $T_{h,0}$ can be calculated as

$$T_{h,0} = \frac{hm(m-1) - \sum_{n=2}^{m} c_n}{1 + \sum_{n=2}^{m} a_n} \ , \tag{34}$$

and $T_{nh,0}$ are given by (32) and (34). $T_{nh,(n-1)h}$ can be calculated through R_n:

$$T_{nh,(n-1)h} = nR_n - hn(n-1) \ , \tag{35}$$

where $R_n = R_2 + S_2 — S_n = T_{h,0} — S_n$ and S_n is given by (28).

Note that these calculations may be performed by each of the nodes individually or by the data sink before broadcasting to all nodes.

In Fig. 12, we present network lifetime measured as the time when a of all sensors run out of battery energy. In the two-dimensional network, the "hot spot" problem is more severe with more nodes that are farther away from the data sink [Chang and Liu], as can be seen from the network lifetime of the MF scheme. The MDF scheme, however, offers a network life that is close to that of the LPF scheme.

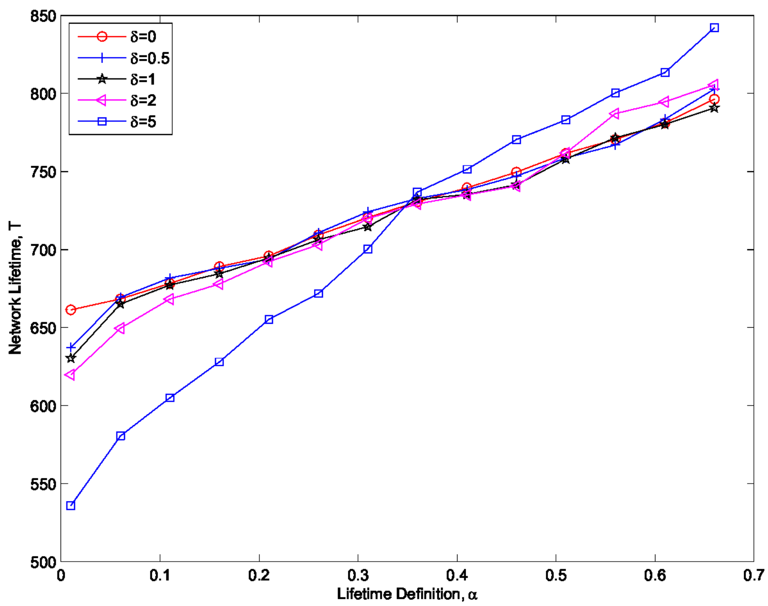

Fig. 13. Network lifetime of the MDF scheme when nodes are randomly shifted from their original locations by at most $\pm\delta$ units ($N = 100$, $k_0 = 400$). The $\delta = 0$ curve represents chain networks as in Fig. 1. The steeper network lifetime curve as δ increases shows that node energy consumption is more imbalance.

5.2 Random Node Distribution

In our analysis, we have assumed that nodes are evenly distributed on a straight line. In this section, we demonstrate the effectiveness of our scheme in scenarios where the nodes are not evenly distributed. Our evaluation is based on the following model: the nodes on the chain network are assumed to be randomly shifted from their original locations. Such a shift is bounded by δ. Therefore, a node with index of n may be physically at any location in (n - δ, n + δ). We measure the network lifetime of the MDF scheme as δ changes from 0 to 5.

The results are presented in Fig. 13, where we show the network lifetime of the MDF scheme with different δ. Based on the curves in Fig. 13, we can see that as δ increases, the energy consumption of nodes becomes more imbalance (hence steeper increase in the network lifetime curve with α). On the other hand, even if δ = 2, with nodes being randomly shifted as much as 2 units away, does not affect the network lifetime (energy balance) of the MDF scheme significantly.

5.3 Limited Transmission Range

It is expected that sensors may have limited transmission range. Therefore, some nodes may not be able to reach the common data sink directly. In this subsection, we investigate the effect of such limited transmission range of sensors on the performance of the MDF scheme.

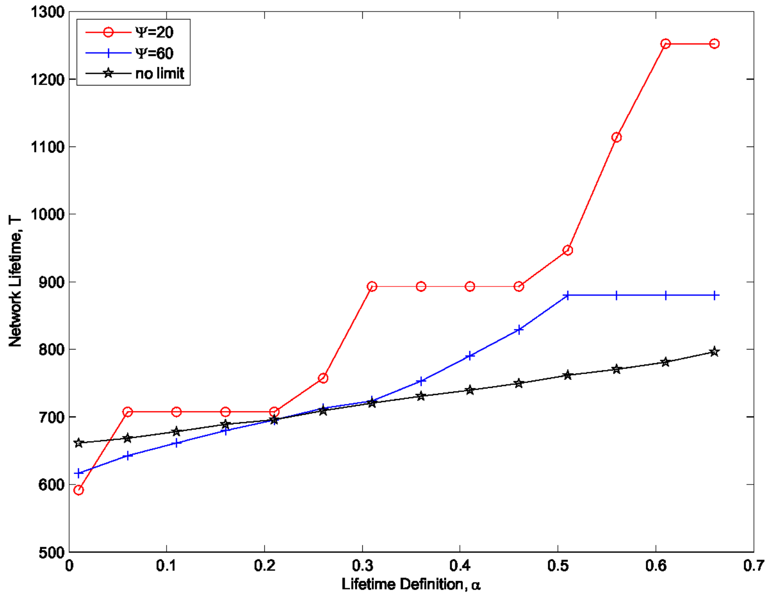

Fig. 14. Network lifetime of the MDF scheme when nodes have a limited maximum transmission range of Ψ units ($N = 100$, $k_0 = 400$). In all other figures in this work, we have assumed $\Psi = \infty$. As Ψ decreases, node energy consumption is more imbalance.

When sensors cannot reach the common data sink directly, they will send the traffic of $T_{nh,0}$ with maximum power. Assume the limit of the transmission distance is Ψ units, node $i > \Psi$ will send its $T_{i,0}$ traffic toward node $i - \Psi$ instead of the common data sink.

We present the network lifetime of the MDF scheme with different Ψ in Fig. 14. When Ψ = ∞, the MDF scheme is not affected. From the figure, we can see that a lowering Ψ shortens the network lifetime with smaller α, but the network lifetime with larger α actually increases. Basically, smaller Ψ increases the imbalance of energy consumption among nodes (in a way similar to the random node distribution does).

6. RELATED WORK

Sensor networks have been an active research field in recent years. Many researches were focused on efficient information forwarding with different constraints [Sankar and Liu, Borghini et al.]. Kulik et al. [Kulik et al.] proposed a Sensor Protocols for Information via Negotiation (SPIN) scheme to disseminate a sensor's observation to all the sensors in a network. In SPIN, meta-data negotiation is exploited to eliminate redundant data transmission. Yu et al. [Yu et al.] proposed a Geographic and Energy Aware Routing (GEAR) scheme to route packets toward target regions. The strategy attempts to balance energy consumption and thereby increase network lifetime. In this work, however, we focus on the problem of data forwarding from sensors toward the data sink [Zhao and Guibas].

Heinzelman et al. proposed the Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol for energy efficient delivery in microsensor networks [Heinzelman et al.]. LEACH uses localized coordination to improve scalability and robustness for dynamic networks, and incorporates data fusion into data forwarding to reduce the amount of information that must be transmitted to the base station (data sink). In the LEACH scheme, cluster heads are selected and used to send information to the base station instead of asking all nodes to do the same. Our MDF scheme applies to the transmission of information from such cluster heads toward the data sink.

Directed Diffusion [Intanagonwiwat et al.] employs initial data flooding and gradual reinforcement of better paths to deliver information from sensors toward observers. Individually compressed data are sent to the data sink and are aggregated through the transmission. Chang and Tassiulas [Chang and Tassiulas] combined two metrics, residual power and energy cost of sending packets, to evaluate and choose routes for data forwarding in sensor networks. The main ideas are to avoid nodes with low residual energy and to favor short hops. Different to their approach, we use the idea of branching the traffic to provide efficient and energy-balanced transmission.

Our work is closely related the works of Perillo et al. [Perillo et al., Perillo et al.]. The “hot spot” problem was investigated through the use of more intelligent transmission power control policy, such as longer transmission range for nodes that are farther away from the data sink [Perillo et al.]. It was shown that the benefit of employing such policy is rather limited. It was further suggested to utilize clustering hierarchy, where heterogeneous sensors are deployed, some of which can act as a data aggregator/compressor. This is different to our approach, which is a simple forwarding rule that balances the energy consumption of nodes and extends the network lifetime.

The problem focused in this work is also closely related to the problem of optimizing transmission range in wireless networks [Takagi and Kleinrock, Hou and Li, Rodoplu and Meng, Chen et al., Gao] . Chen et al. investigated the optimization of transmission range as a system design issue [Chen et al.]. The wireless network was assumed to have high node density, and consisting of nodes with relatively low mobility and short transmission range. As justified by the assumption of high node density, the authors further assumed that intermediate router nodes are always available at the desired location whenever they are needed.[5] Hou and Li suggested to use the lowest possible transmission power to the nearest neighbor in the forwarding direction [Hou and Li].

In [Jurdak et al.], Jurdak et al. investigated the battery lifetime of underwater acoustic sensor networks. Several optimization techniques were proposed and studied. Their approach focused on four areas for network lifetime extension: transmission frequency, update period, average transmission distance, and cluster size. In contrast, we focus on optimal transmission distance with multi-hop forwarding and energy balance with direct forwarding. Another difference between our work and [Jurdak et al.] is that we do not consider transmission loss while it was investigated in [Jurdak et al.].

7. CONCLUSIONS AND FUTURE WORK

We have proposed a Multi-hop/Direct Forwarding (MDF) scheme to forward data traffic toward the data sink in WSN. In the MDF scheme, the traffic forwarding rules are much simpler than the forwarding rules obtained from linear programming (LPF). We have also developed an analytical framework to derive the traffic splitting rules for the MDF scheme. The MDF scheme has been shown to approach the performance of LPF but with much lower computation complexity. The multi-hop forwarding in the MDF scheme achieves close to optimality (as of the LPF scheme), while direct forwarding balances the energy consumption among sensors with different distances toward the data sink.

While the MDF scheme specifies efficient data forwarding rules for WSNs, the actual data transmission needs to be supported by Medium Access Control (MAC) scheme design. With the simple two-branch data forwarding rule, we can design MAC schemes that are both simple and efficient. We leave this as our future research. In our study of two-dimensional networks, we have assumed that the common data sink sits at the center of the region and nodes form a perfect circle. In reality, the common data sink may be located at other positions and the nodes form an irregular shape. Interestingly, we can extend the MDF scheme by considering each node to forward traffic to the data sink in a straight line. The node only needs to consider the nodes on this extended line. An irregular shape network topology can be treated as nodes forward traffic with different maximum distance m toward the data sink. Also, experimental results from real sensor networks should be collected to ensure the performance of our proposed scheme. We leave further discussions of this to our future work.

Notes:

[1]Note that we have normalized this energy by the Euclidean distance of neighboring nodes.

[2]Some subsets may have one fewer node if N/h is not an integer.

[3]In fact, node i only needs to know its index or distance toward the common data sink if traffic generation is known *a priori*. With this information, it can calculate n and find out $T_{nh,0}$, which is the traffic to be sent to the common data sink directly. It should send the rest of the traffic to the node that is h-hop closer to the common data sink. The value of h can be coded into the nodes prior to deployment since it is hardware related.

[4]These results only have relative significance, as network lifetime depends largely on initial energy, k_0, N, and other system parameters. We assume an initial energy of 2×10^6 J for each sensor.

[5]By using an evenly-spaced distribution for nodes, we have made a similar assumption in this work.

APPENDIX: Proof of $S_2 = R_2 = \frac{T_{h,0}}{2}$

Related to Section 5.1, we prove that $S_2 = R_2 = \frac{T_{h,0}}{2}$ in this appendix.

Considering the traffic of node h, we have:

$$T_{2h,h} + 2h = T_{h,0} \tag{36}$$

Since nodes $2h$ and h should have the same energy consumption per node, we have

$$\frac{T_{2h,h} \cdot k_1 + T_{2h,0} \cdot k_2}{4h} = \frac{T_{h,0} \cdot k_1}{2h} ,$$

which can be re-arranged as:

$$T_{2h,0} \cdot \frac{k_2}{k_1} - 2h = T_{2h,h} + 2h = T_{h,0} .$$

Using the definition of R_n and S_n in (28), we have proven that

$$R_2 = S_2 = \frac{T_{h,0}}{2} .$$

REFERENCES

AKYILDIZ, I. F., SU, W., SANKARASUBRAMANIAM, Y., AND CAYIRCI, E. 2002. A survey on sensor networks. IEEE Communications Magazine 40, 8 (August), 102–114.

BORGHINI, M., CUOMO, F., MELODIA, T., MONACO, U., AND RICCIATO, F. 2005. Optimal data delivery in wireless sensor networks in the energy and latency domains. In Proc. of the First International Conference on Wireless Internet (WICON'05). ACM, Budapest, Hungary, 138– 145.

CAGALJ, M., HUBAUX, J.-P., AND ENZ, C. 2002. Minimum-energy broadcast in all-wireless networks:: NP-completeness and distribution issues. In Proc. of the 8th Annual ACM/IEEE Int. Conf. on Mobile Computing and Networking (MobiCom '02). IEEE, Atlanta, GA, USA, 172– 182.

CHANG, J.-H. AND TASSIULAS, L. 2000. Energy conserving routing in wireless ad-hoc networks. In Proc. of the 19th Conference of the IEEE Communications Society (Infocom '00). Vol. 1. IEEE, Tel Aviv, Israel, 22–31.

CHANG, N. AND LIU, M. 2004. Revisiting the TTL-based controlled flooding search: Optimality and randomization. In Proc. of the 10th Annual ACM/IEEE Int. Conf. on Mobile Computing and Networking (MobiCom '04). IEEE, Philadelphia, PA, USA, 85–99.

CHEN, P., O'DEA, B., AND CALLAWAY, E. 2002. Energy efficient system design with optimum transmission range for wireless ad hoc networks. In Proc. of IEEE ICC. Vol. 2. IEEE, New York, NY, USA, 945–952.

FEENEY, L. M. AND NILSSON, M. 2001. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In Proc. of the 20th Conference of the IEEE Communications Society (Infocom ’01). Vol. 3. IEEE, Anchorage, AK, US, 1548–1557.
GAO, J. L. 2002. Analysis of energy consumption for ad hoc wireless sensor networks using a bit-meter-per-joule metric. IPN Progress Report 42-150, JPL. August.
HEINZELMAN, W. B., CHANDRAKASAN, A. P., AND BALAKRISHNAN, H. 2002. An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. on Wireless Communications 1, 4 (October), 660–670.
HOU, T.-C. AND LI, V. O. K. 1986. Transmission range control in multihop packet radio networks. IEEE Trans. on Commun. COM-34, 1 (January), 38–44.
INTANAGONWIWAT, C., GOVINDAN, R., ESTRIN, D., HEIDEMANN, J., AND SILVA, F. 2003. Directed diffusion for wireless sensor networking. IEEE/ACM Trans. on Networking 11, 1 (February), 2–16.
JURDAK, R., LOPES, C. V., AND BALDI, P. 2006. Battery lifetime estimation and optimization for underwater sensor networks. 397–420.
KULIK, J., HEINZELMAN, W., AND BALAKRISHNAN, H. 2002. Negotiation-based protocols for disseminating information in wireless sensor networks. Wirel. Netw. 8, 2/3, 169–185.
PERILLO, M., CHENG, Z., AND HEINZELMAN, W. 2004. On the problem of unbalanced load distribution in wireless sensor networks. In Proc. of IEEE GLOBECOM, Workshop on Wireless Ad Hoc and Sensor Networks. IEEE, Dallas, TX, USA, 74–79.
PERILLO, M., CHENG, Z., AND HEINZELMAN, W. 2005. An analysis of strategies for mitigating the sensor network hot spot problem. In Proc. of the 2nd Annual Int. Conf. on Mobile and Ubiquitous Systems: Networking and Services. IEEE, San Diego, CA, USA, 474–478.
RODOPLU, V. AND MENG, T. H. 1999. Minimum energy mobile wireless networks. IEEE Journals on Selected Areas of Communications 17, 1333–1344.
SANKAR, A. AND LIU, Z. 2004. Maximum lifetime routing in wireless ad-hoc networks. In Proc. of the 23rd Conference of the IEEE Communications Society (Infocom ’04). IEEE, Hong Kong, P. R. China, 1089–1098.
TAKAGI, H. AND KLEINROCK, L. 1984. Optimal transmission ranges for randomly distributed packet radio terminals. IEEE Trans. on Commun. COM-32, 3 (March), 246–257.
YU, Y., GOVINDAN, R., AND ESTRIN, D. 2001. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical Report TR-01-0023, University of California, Los Angeles, Computer Science Department.
ZHAO, F. AND GUIHAS, L. 2004. Wireless Sensor Networks: An Information Processing Approach. Morgan Kaufmann.