Energy-Delay Region of Low Duty Cycle Wireless Sensor Networks for Critical Data Collection

Xinyu Mao*, Shiyu Liu*, Lin Zhang*, and Victor O.K. Li+

*Department of Electronic Engineering, Tsinghua University, Beijing, P.R.China , Email: {maox09, liu-sy11}@mails.tsinghua.edu.cn, linzhang@mail.tsinghua.edu.cn
+Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong, China

Abstract—We investigate the trade-off between energy consumption and delay for critical data collection in low duty cycle wireless sensor networks, where a causality constraint exists for routing and link scheduling. We characterize the energy-delay region (E-D region) and formulate a combinatorial optimization problem to determine the link scheduling with the causality constraint. A new multiple-degree ordered (MDO) coloring method is proposed to solve this problem with near-optimal delay performance. The impacts of many system parameters on the E-D region are evaluated by extensive simulation, providing an insightful frame of reference for design of critical data collection wireless sensor networks.

Index Terms—Wireless Sensor Networks, Energy-delay Trade-off, Cross-layer Optimization

I. INTRODUCTION

Data collection is one of the basic services for wireless sensor networks (WSNs). A typical scenario for data collection includes a large number of battery-powered wireless sensor nodes that are geographically distributed in the field, collecting and transmitting sensory data back to a base station for processing [1]. Continuous physical field reconstruction is a typical and well-studied application that requires data collection service. In this case, the sensory data are highly temporally and spatially correlated. Therefore, the data collection service can be lossy, and in-network compression schemes can be adopted to reduce the communication and energy cost.

However, lossy continuous data collection service is not suitable for critical data collection. For instance, in an intrusion detection sensor network, every event needs to be transmitted back to the base station in a timely fashion, and the user cannot afford loss of data due to compression or transmission errors. Other typical applications include radio radiation monitoring for nuclear facilities, consumer behavior monitoring network in super markets, etc. The critical data collection service has two distinguishing characteristics: 1) every sensory data must be collected with reasonably small delay, 2) energy efficiency is an important consideration as in most WSN problems, so low duty cycle scheduling is widely adopted.

We investigate the energy-delay (E-D) region in WSNs for critical data collection. Due to the criticality and bursty nature of sensory data, data compression is not feasible in the system. Every sensor packet generated at the source node is routed back to the data sink node along a spanning tree during the active phase of the network when all the nodes are awake. Fig. 1 illustrates the scheduling scheme, in which the duty cycle is normally very low to achieve low energy consumption and long network lifetime. Unlike in the continuous data collection scenarios, the link scheduling needs to consider the packet arrival sequence so that packets from all the nodes can be delivered to the data sink by the end of the active phase. We call this the causality constraint. Violation to this constraint will cause the packets be stuck in intermediate nodes during the very long sleep phase, and not be delivered until the next active phase. This will incur undesirable long delay.

While there is prior work [2] [3] on continuous aggregated data collection and [4] on low duty WSNs, this paper is the first to explore the E-D region for critical data collection in low duty cycle WSNs, where the causality constraint influences the delay and energy performance. The contributions of this paper are:

- We characterize the E-D region of WSNs with causality constraint.
- We propose an algorithm MDO which combines the spanning-tree enumeration and conflict graph coloring link scheduling to find the E-D region.
- We analyze the impact of different factors on the E-D region, including the channel number, power control scheme, network connectivity, and interference model. This gives a frame of reference and provides insights for the real-world engineered system design.

The remainder of the paper is organized as follows. In Section II, we review related work on network scheduling and argue the novelty of our work. Section III presents the system model. In Section IV, we define and characterize the E-D region and propose the algorithm to calculate the E-D region. The simulation results are evaluated and discussed in Section V. Finally, Section VI concludes the paper.
II. RELATED WORK

There is much work on energy-efficient scheduling in WSNs. This field ranges widely including energy-consumption modeling [5], predictive wake-up scheme [6] [7] [8], and joint optimization among energy, delay and throughput [9] [10] [11] [12]. However, to the best of our knowledge, none of them studies the joint energy and delay optimization and the E-D region for critical data collection in low duty cycle WSNs.

[13] studied the problem of scheduling packet transmission for data collection in WSNs and abstracted the packet flow as a many-to-one communication flow over a data aggregation tree structure. The optimization problem was to minimize the overall energy consumption of all sensor nodes in the aggregation tree subject to the delay constraint. Though the non-monotonic energy model was explored, there was no consideration of the MAC layer interference. Furthermore, the results of the energy-delay trade-off curve were dependent on the routing tree and network topology. This algorithm guarantees minimal energy satisfying certain delay. [14] formulated the objective function as a weighted sum of the delay and the energy consumption and Pareto-optimal energy-delay trade-off curve was computed in a particular string topology and under the QPSK scheme. However, the exact E-D region is not obtained, and this is the focus in this paper.

[15] proposed a distributed time-slot assignment scheme to minimize the Single-channel-TDMA (SC-TDMA) schedule length for critical data gathering. In [1], the lower bound on the TDMA-based schedule length was computed for critical data collection in tree-based WSNs and the lower bound can be achieved by exploring transmission power control and multi-channel-TDMA (MC-TDMA) schemes to eliminate interferences. Also three channel assignment methods were discussed including joint frequency-time-slot scheduling (JFTSS) [1], receiver-based channel assignment (RBCA) [2] and tree-based multichannel protocol (TMCP) [16]. The results demonstrated that JFTSS performed the best for critical data collection in low duty cycle network with two or sixteen alternative channels and therefore JFTSS is chosen for comparison with our proposed MDO scheme which obtains the near-optimal E-D region. However, there was no consideration of the energy factor in [1] and the E-D region was not available.

In [3], the factors which can affect energy and delay were analyzed, including source-sink placements, the density of the network and the communication radius, considering three different routing schemes, namely, center at nearest source (CNS), SPT, and greedy incremental tree (GIT). In [2], transmission power control and MC-TDMA techniques were explored to minimize the number of time slots needed for specific data aggregation trees. Taking into account the limited discrete power levels for power control, they found that the number of time slots can be decreased by 15-20 percent compared to no power control. Also the impacts of interference models and multiple channels were analyzed. Thus, the impact on the E-D region can be observed and evaluated to guide the protocol design for data collection in WSNs with energy and delay constraints.

III. SYSTEM MODEL

A. Network Model

Our network model follows the dominant IEEE 802.15.4e standard [17].

A multi-hop many-to-one data collection WSN can be represented by a graph $G(V,E)$. $V = \{V_i : i = 0, 1, \ldots, N\}$ is the set of all network nodes, where $V_0$ denotes the sink node and the rest represent source nodes. $E$ is the set of all available wireless links. In a square area of length $L$, there are $N$ identical half-duplex source nodes randomly deployed. Their maximum transmission ranges are the same and equal $R_{c}^{MAX}$. If node $v \in V$ is within $R_{c}^{MAX}$, the radius of node $u \in V$, the edge $(u,v) \in E$. We assume symmetry in communication, i.e., $(u,v) \in E$ iff $(v,u) \in E$.

We assume an MC-TDMA network, in which time is divided into equal-sized slots, labelled $t_i$, $i = 1, 2, \ldots$. All network nodes can achieve time synchronization and transmit or receive over one of the $m$ orthogonal channels.

We adopt the protocol interference model, and define $R_i = \lambda R_c$ as the interference radius, where $\lambda = 2 \sim 3$ [18], and $R_c$ is the communication range. A collision occurs when a receiving node is within radius $R_i$ of some transmitting nodes other than its sender, inducing a packet transmission failure.

We assume three different schemes for power control. In the ideal power control scheme, a node can fine-tune its transmission power to achieve any transmission range $R_c < R_{c}^{MAX}$. In the partial power control scheme, a node can tune its transmission power to several pre-defined levels, resulting in several discrete transmission ranges. While in the no-power-control scheme, all nodes transmit at the maximum power, with transmission range $R_{c}^{MAX}$. Most real-world radio transceivers have partial power control capability, and the ideal power control scheme and the no-power control scheme are extremes when the number of power levels goes to infinity or degrades to one, respectively. We study the impact of these three power control schemes on the E-D region.

We adopt the first order radio model [19] to compute the energy consumption of packet transmissions. The transmitting energy consumption for one packet is computed by (1), which is composed of transmitter electronics and transmit amplifier energy dissipations. The energy consumption for receiving one packet can be obtained from (2), which is composed of receiver electronics energy dissipation only.

$$P^{Tx} = E_{elec} \times k + \varepsilon_{amp} \times k \times d_{ij}^2$$  \hspace{1cm} (1)

$$P^{Rx} = E_{elec} \times k$$  \hspace{1cm} (2)

where $E_{elec}$ is the radio dissipation parameter to run the transmitter and receiver circuitry, $\varepsilon_{amp}$ is the amplification factor for the transmit amplifier to achieve an acceptable SINR, $k$ is the number of bits per packet and $d_{ij}$ is the distance between the nodes $i$ and $j$. 
**B. The scheduling for data collection and the causality constraint**

To deliver data from many sensor nodes to the data sink with good energy efficiency, the spanning tree is considered as the energy-conserving routing method [20]. When the minimal energy routing scheme is determined, the network schedules transmissions on multiple channels and in multiple slots to eliminate the collision, and minimize the number of timeslots used. This is equivalent to minimizing the latency of the data collection. There are many methodologies to schedule the transmissions, and we will leave the details to the subsequent sections.

It shall be noted that in low duty cycle WSNs, transmission scheduling shall not only be collision-free, but must also be properly sequenced. When a spanning tree is formulated, the transmission scheduling algorithm shall consider the order of transmissions along the tree. Specifically, a node in the tree cannot relay data packets from its child nodes before they arrive. Failure to conform to this rule will leave packets stuck in the network until the next round of transmissions. In low duty cycle WSNs, this will introduce tremendous delay. We also have the following theorem:

**Definition 1 (Achievable E-D pair):** We say an energy-delay pair is achievable if there exists a sequence of cycles \( \{E_i, D_i\} \) such that

\[
E = \frac{1}{M} \sum_i E_i \quad \text{and} \quad D = \frac{1}{M} \sum_i D_i.
\]

An achievable E-D pair is the average energy and delay over multiple scheduling cycles. The E-D region is defined as the set of all achievable E-D pairs. Obviously, the E-D region of a data collection WSN depends on many factors, including the topology, number of channels, interference range, and power control schemes. To characterize the E-D region, we have two straightforward observations:

- Given \((E, D)\) is achievable, \((E', D)\) is achievable, if \(E' > E\).
- Given \((E, D)\) is achievable, \((E, D')\) is achievable, if \(D' > D\).

We also have the following theorem:

**Theorem 1 (Convexity of E-D region):** If \((E_1, D_1)\) and \((E_2, D_2)\) are both achievable, then \((\lambda E_1 + \bar{\lambda} E_2, \lambda D_1 + \bar{\lambda} D_2)\) is achievable, where \(\lambda + \bar{\lambda} = 1\) and \(\lambda, \bar{\lambda} > 0\).

**Proof:** Concatenating the scheduling sequences that achieve \((E_1, D_1)\) and \((E_2, D_2)\), we obtain a new scheduling sequence, in which the ratio between the number of \((E_1, D_1)\) achieving sequences to that of \((E_2, D_2)\) achieving sequences is \(\frac{\lambda}{\bar{\lambda}}\). The concatenated sequence achieves \((\lambda E_1 + \bar{\lambda} E_2, \lambda D_1 + \bar{\lambda} D_2)\).

The two observations and Theorem 1 together characterize the E-D region as shown in Fig. 2. The circles in the figure represent the E-D pairs that are achieved by the scheduling scheme in a single cycle. Every point in the shaded area is achievable. Note that when the portion of the E-D region characterized by Theorem 1 is not active, the shape of the E-D region will have a right angle, and minimal energy and delay can be achieved simultaneously, and no trade-off exists.

**IV. THE ENERGY-DELAY REGION**

In this section, we define and characterize the E-D region. Then we propose an algorithm to calculate the E-D region, consisting of a spanning tree enumeration algorithm and a transmission scheduling algorithm that guarantees the causality constraint.

**A. Definition and Characterization of the E-D Region**

As discussed in the previous section, data collection in WSNs can be accomplished by a spanning tree and the corresponding transmission scheduling, resulting in an E-D pair for that specific round.

**Definition 1 (Achievable E-D pair):** We say an energy-delay pair \((E, D)\) is achievable, if there exists a sequence of cycles \(\{E_i, D_i\}, i = 1, 2, \ldots, M\), such that

\[
E = \frac{1}{M} \sum_i E_i \quad \text{and} \quad D = \frac{1}{M} \sum_i D_i.
\]

An achievable E-D pair is the average energy and delay over multiple scheduling cycles. The E-D region is defined as the set of all achievable E-D pairs. Obviously, the E-D region of a data collection WSN depends on many factors, including the topology, number of channels, interference range, and power control schemes. To characterize the E-D region, we have two straightforward observations:

- Given \((E, D)\) is achievable, \((E', D)\) is achievable, if \(E' > E\).
- Given \((E, D)\) is achievable, \((E, D')\) is achievable, if \(D' > D\).

We also have the following theorem:
A framework based on vertex degree ordering for the graph coloring problem is proposed in [24], in which the non-colored node with maximal degree is colored first and the minimal color \(\sigma\) is assigned. In this framework, the definition of vertex degree is the most important factor. In [22], the vertex degree is only determined by the conflict relationships in the conflict graph. In [25] and JFTSS proposed by [1], the packet load of each source node is also considered to improve scheduling performance. However, a node with more children in \(ST_i\) is more likely to be a potential bottleneck node, and should have priority to be scheduled. We propose an algorithm MDO to achieve this. First, we define the multiple degree of a node.

**Definition 2 (Multiple Degree):** The Multiple Degree (MD) of a node in \(G^C\) is denoted by a triplet \((PoT, Packet\_Load, D^C)\), where \(PoT\) potential packets to be transmitted is determined by its children number in specific \(ST_j\), \(Packet\_Load\) is the current queue length of the node, and \(D^C\) is the vertex degree of the node in \(G^C\). The priority in Multiple-Degree is ordered by \(PoT, Packet\_Load\), and \(D^C\), that is, \(MD_i > MD_j\) if \(PoT_i > PoT_j\) and if \(PoT_i = PoT_j\), the \(MD_i > MD_j\) if \(Packet\_Load_i > Packet\_Load_j\).

As Algorithm 1 shows, in MDO, the order of node coloring is determined by its \(MD\), considering whether the node to be colored satisfies the half-duplex constraint of the nodes colored before and whether it will collide with them.

**Algorithm 1: MDO Coloring Algorithm for MC-TDMA Scheduling**

\[
\text{input} : G = (V,E) \\
\text{output:} (Energy, Delay) = \{(P_{ST_i}, T_{ST_i})|i = 1, 2, ...,|ST|\} \\
A(G) = \text{GetAdjacencyMatrix}(G); \\
ST = \text{GetAllSpanningTree}(A(G)); \\
f = 1 \text{ to } |ST| \text{ do} \\
P_{ST_i} = \text{GetEnergy}(ST_i) \text{ according to (3);} \\
C_{ST_i}^C(ST_i), E_{ST_i}^C = \text{GetConflictGraph}(ST_i); \\
T_{ST_i} = 0; CN = \emptyset \text{ (Colored Node Set) while (causality constraint has not been satisfied) do} \\
V_{ST_i}^C = V_{ST_i}^C; \\
\text{for } j = 1 \text{ to } |V_{ST_i}^C| \text{ do} \\
v_j = \text{the vertex in } V_{ST_i}^C \text{ with maximal Multiple-Degree;} \\
\text{if } (v_j \neq V_0 \text{ or IsLoadEmpty}(v_j) \text{ or } \text{IsHalf-DuplexValid}(v_j, CN) \text{ then} \\
V_{ST_i}^C = V_{ST_i}^C - \{v_j\}; \\
\text{continue;} \\
\text{color } = 0; \\
\text{while } (v_k \in CN, \sigma(v_k) = \text{color and } \text{IsCollision}(v_j, v_k)) \text{ do} \\
\text{color } = \text{color } + 1; \\
\sigma(v_j) = \text{color}; CN = CN + \{v_j\}; \\
V_{ST_i}^C = V_{ST_i}^C - \{v_j\}; \\
\text{for } j = 1 \text{ to } m \text{ do} \\
\text{for } k = 1 \text{ to } |V_{ST_i}^C| \text{ do} \\
\text{if } \sigma(v_k) = j \text{ then} \\
v_k \leftarrow (t_{ST_i}, ch_j); \\
\text{Clear the coloring results; Update packet loads of all vertices in } V_{ST_i}^C; \\
T_{ST_i} = T_{ST_i} + 1; \\
\text{end while} \\
\text{end if} \\
\text{end do} \\
\text{end do} \\
\text{end while} \\
\text{end do} \\
\text{end do} \\
\text{end do} \\
\text{end while} \\
\text{end do} \\
\text{end do} \\
\text{end do} \\
\text{end do} \\
\text{end do} \\
\text{end do} \
\]

Fig. 3. Demonstration of a calculated E-D region.

**C. Near Optimality of the MDO algorithm**

With a sufficiently large number of channels, the overall delay per cycle can be lower bounded by \(max(2n_{max}^{ST_i} - 1, N)\) [1], where \(n_{max}^{ST_i}\) is the maximum number of nodes in any subtree of the children of the root node in \(ST_i\). The interfering links can be eliminated if there are enough channels and ideal power control is utilized. With 25 sensing nodes, 25 available channels and ideal power control, Table I shows the simulation results of the MDO algorithm compared to JFTSS [1], where OptMDO and OptJFTSS represent, respectively, the cumulative number of schedules with delay performance equal to the lower bound utilizing MDO and JFTSS for each topology \(Topo\). NumSPT is the number of all spanning trees corresponding to a particular topology. Apparently, in all random scenarios given, MDO outperforms JFTSS and the scheduling result of MDO is equal to the lower bound, which demonstrates the near optimality of the MDO algorithm.
A. Impact of Channel Number

With 25 source nodes randomly deployed in a 100 × 100 square where $R_c^{MAX} = 25$ and $\lambda = 2.5$, we investigate the impact of channel number $m$ on the E-D Region. The ideal power control scheme is used. Fig. 4(a) shows the variation of the E-D region with $m$, and we can observe that the optimal strategy exists when three or four channels are available, that is, the minimal energy and delay may be achieved simultaneously by utilizing efficient routing and scheduling schemes. However, the trade-off between energy and delay should be considered when the number of channels is 1 or 2. Fig. 4(b) demonstrates that minimal energy consumption stays the same as channel number increases. This is because the set of all spanning trees is unchanged. However, as the channel number increases and spatial reuse is increased, the minimal delay can be reduced by more than half. It is also interesting to see that the lower bound of the delay may be achieved with three or more channels, which means that three available channels are sufficient in this scenario.

B. Impact of Power Control

With 25 source nodes randomly deployed in a 100 × 100 square where $R_c^{MAX} = 25$, $\lambda = 2.5$ and with two channels available, Fig. 5(a) demonstrates the variation of the E-D region with ideal power control, two-level graded power control and no power control. We observe that both minimal energy and delay decrease with more advanced power control methods.

As in the popular transceiver $CC2420$ [26], the transmission power is graded into $2^n$ levels, where $n = 0, 1, ..., 6$. Fig. 5(b) shows that the minimal energy and delay almost decrease linearly when the level changes from $2^0$ to $2^2$, which means that most of interferences among sensor nodes can be avoided. As the number of levels increases, the minimal energy and delay decline slowly and remain at the same values as in ideal power control when the level is equal to $2^6$, which suggests that higher level is not necessary in this scenario.

C. Impact of Network Connectivity

With 15 source nodes randomly deployed in a 100 × 100 square where $\lambda = 2.5$, two channels are available, and ideal power control is utilized, Fig. 6(a) and Fig. 6(b) demonstrate the variation of the E-D region, with minimal energy and delay under different network connectivity, noting that the network connectivity is only influenced by $R_c^{MAX}$ and with $R_c^{MAX}$ increasing from 30 to 42 it changes from 0.2 to 0.35. We observe that the minimal energy and delay both decrease with increasing connectivity. This happens because the set of all spanning trees with smaller connectivity is a subset of that with larger connectivity, and better strategies may be obtained. However, the communication ability of sensor nodes is limited in reality, that is, $R_c^{MAX}$ is bounded, so that $R_c^{MAX}$ should be chosen according to energy and delay demand in real application scenarios.

D. Impact of Protocol Interference Model

With 25 source nodes randomly deployed in a 100 × 100 square where $R_c^{MAX} = 25$, two channels are available and with ideal power control, we investigate the impact of the protocol interference model on the E-D region. Fig. 7(a) shows the variation of regions with $\lambda$. We observe that the minimal delay almost increases linearly as $\lambda$ changes from 2 to 3 in Fig. 7(b). This happens because the inter-interference between sensor nodes become more serious with increasing interference radius. However, the set of all spanning trees is unchanged so that the minimal energy stays the same.

VI. Conclusion

We investigated the E-D region in low duty cycle WSNs for critical data collection, where nodes communicate using MC-TDMA MAC and a routing tree to collect data. We defined
the causality constraint, critical in this scenario, and used the GM algorithm to generate all spanning trees. We developed a novel MDO coloring method for achieving near optimal scheduling. We found that the power control scheme and network connectivity influence the minimal energy and delay simultaneously. However, the channel number and interference protocol can only change the minimal latency. In addition, if minimal energy and delay can be achieved at the same time, the optimal strategy may exist, but the trade-off should be considered in other cases. The results help guide design for data collection in WSNs.

REFERENCES


