On the Age of Information in Multi-Source Multi-Hop Wireless Status Update Networks

Shahab Farazi†, Andrew G. Klein‡, John A. McNeill† and D. Richard Brown III†

†Worcester Polytechnic Institute, 100 Institute Rd., Worcester, MA 01609, Email: {sfarazi,mcneill,drb}@wpi.edu
‡Western Washington University, 516 High St., Bellingham, WA 98225, Email: andy.klein@wwu.edu

Abstract—This paper studies a multi-source “age of information” problem in multi-hop wireless networks with packetized status updates and explicit channel contention. Specifically, the scenario considered in this paper assumes that each node in the network is both a source and a monitor of information. Nodes take turns broadcasting their information to other nodes in the network while also maintaining tables of status updates for the information received from all other nodes in the network. Lower bounds on the peak and average age of information are derived and are found to be a function of fundamental graph properties including the connected domination number of the graph and the average shortest path length. In addition to these converse results, achievability results are developed through the presentation of an explicit algorithm for constructing near-optimal status update schedules along with an analytical upper bound for the average and peak age of these schedules. Finally, numerical results are presented that compute the bounds, construct schedules, and compute the achieved average and peak ages of these schedules exhaustively over every connected network topology with nine or fewer nodes. The results show that the the developed schedules achieve a peak age exactly matching the lower bounds and an average age within a multiplicative factor of 1.035 of the lower bound in all tested cases.

Index Terms—Age of Information, peak and average age, multi-source, multi-hop, explicit contention, graph theory.

I. INTRODUCTION

Information freshness is of critical importance in a variety of networked monitoring and control systems such as intelligent vehicular systems, channel state feedback, and environmental monitoring as well as applications such as financial trading and online learning. In these types of applications, stale information can lead to incorrect decisions, unstable control loops, and even compromises in safety and security. A recent line of research has considered information freshness from a fundamental perspective under an Age of Information (AoI) metric first proposed in [1] and further studied in [2]–[26]. The central idea is that there are one or more sources of information along with one or more monitors. A source generates timestamped status updates which are received by a monitor after some delay. The “age of information” is defined as the difference between the current time and the timestamp of the most recent status update at the monitor. A common theme of the AoI literature is to study and/or optimize the statistics of AoI, i.e., average age and/or peak age, as a function of the system parameters and update strategies.

Much of the work in this area has focused on studying AoI in the single-source, single-monitor setting, e.g., [2]–[8]. The delay in delivering the updates from the source to the monitor in these studies is typically modeled as a random delay through a queue. As such, these papers consider AoI in an “implicit contention” setting in the sense that the other nodes contending for the channel resources are not explicit in the system model.

This paper studies AoI in a general multi-source, multi-hop, time-slotted network setting with explicit contention in the sense that all delays between sources and monitors are due to explicit channel uses by other nodes in the network. Each node in the system is both a source and monitor of information. Since the only assumption on the network is that it is connected, some nodes in the network also serve as relays to facilitate multi-hop dissemination of information between nodes that are not directly connected. While [1], [13]–[26] also consider this fully-explicit contention setting, only [23]–[26] consider multi-hop networks. The analysis in [23]–[25] is restricted to specific network structures, e.g., line or ring networks, however, and the schedules developed for these networks are not easily extended to general network structures. The recent work in [26] considers more general network structures, but assumes a pre-defined source/monitor pairs, and analyzes achievable AoI under certain simplifying assumptions, e.g., stationary scheduling policies.

The main contributions of this paper are (i) the development of fundamental bounds on the peak and average age in a general multi-source, multi-hop network setting and (ii) the presentation of an explicit algorithm for constructing near-optimal update schedules. The development of fundamental bounds on peak and average age has received relatively little attention in the AoI literature. The bounds derived in this paper hold for all connected network topologies and minimum length periodic schedules in which one node transmits per time slot. The explicit update schedules developed in this paper are analytically shown to achieve a peak age exactly matching the lower bound and numerically shown to achieve an average age within a multiplicative factor of 1.035 of the lower bound for every connected network topology with nine or fewer nodes.

II. SYSTEM MODEL

Consider a wireless network modeled by an undirected graph $G = (\mathcal{V}, \mathcal{E})$. The vertex set $\mathcal{V}$ represents the nodes and the edge set $\mathcal{E}$ represents the channels between the nodes in the network. Two vertices $i, j \in \mathcal{V}$ are adjacent if edge $e_{i,j}$ is

This work was supported by the National Science Foundation award CCF-1319458.
in set $\mathcal{L}$. Equivalently, there exists a channel between nodes $i$ and $j$; as such, any wireless transmission broadcast from node $i$ is received at all adjacent nodes. Here, we only consider networks with a connected topology, i.e., there exists a path between any two distinct vertices $i, j \in \mathcal{V}$. Each node $i \in \mathcal{V}$ can generate samples of a local random process $H_i(t)$ at any time $t$. In addition to information on the status of its own process, every node in the network is also interested in updates of the status of the remaining $N - 1$ processes in the network. We denote the status of process $H_i(t)$ from the perspective of node $j$ at time $t$ by $H_i^{(j)}(t)$. So, at any time each of the $N$ nodes keeps a table of its most recently obtained status updates of each of the $N$ processes, giving a total of $N^2$ parameters throughout the network. Out of the $N$ parameters at each node, one is obtained locally by direct observation, and $N - 1$ are obtained by indirect observation from the status updates disseminated by other nodes. Overall, there exist $N$ directly and $N^2 - N$ indirectly obtained parameters throughout the network. We assume transmissions of status update packets with a fixed length of one unit of time. Each packet includes information about the one process that is being transmitted and a time stamp indicating the time that the information was generated. We denote the time interval $((n - 1), n]$ by time slot $n$ for an integer $n$.

The following definitions formalize the notation and age metrics considered in this paper. First, we review some common graph theoretic concepts. We use the notation $d(i, j)$ to denote the shortest path length between vertices $i$ and $j$, and $\ell_G \triangleq \frac{1}{N - N^2} \sum_{i, j \in \mathcal{V}, i \neq j} d(i, j)$ denotes the average shortest path length of the graph $G$. Recall that a set $S \subset \mathcal{V}$ of vertices in a graph is called a dominating set if every vertex not in $S$ is adjacent to a vertex in $S$ [27]. A minimum connected dominating set (MCDS) $S \subset \mathcal{V}$ is a dominating set with the properties that (i) the subgraph induced by $S$, $G[S]$ is connected and (ii) $S$ has the smallest cardinality among all connected dominating sets of $G$. The cardinality of any MCDS is called the connected domination number of $G$ and denoted by $\gamma_c$. In general, the MCDS is not unique [28], [29].

**Definition 1** (Pseudo-leaf vertex). We refer to a vertex as a pseudo-leaf if it is not a member of any MCDS. That is $i \in \mathcal{V}$ is a pseudo-leaf if $i \notin U$ where the $S_m \subset \mathcal{V}$ for $m = \{1, 2, \ldots, M\}$ represent all $M$ possible MCDS’s of $G$ and

$$U \triangleq S_1 \cup S_2 \cup \ldots \cup S_M.$$  

Further, we refer to the set of all pseudo-leaf vertices of $G$ by $\mathcal{L} \triangleq \mathcal{V} - U$.

Under this definition, every true leaf (i.e., every vertex with degree one) is also a pseudo-leaf. An example illustrating the notion of pseudo-leaf vertices and MCDS’s is shown in Fig. 1.

**Definition 2** (Age). Assume the most recent status update of the $H_i$ process received at node $j$ was timestamped at time $t'$. The age of status update $H_i^{(j)}$ at time $t \geq t'$ is defined as $\Delta_i^{(j)}(t) \triangleq t - t'$ for $j \neq i$.

Since each node is assumed to have zero-delay access to the status of its local process, we have $\Delta_i^{(j)}(t) = 0$ for any $i \in \mathcal{V}$ and $t$. To capture the timeliness of all of the $N^2 - N$ indirectly-obtained status update parameters throughout the network, we define the peak and average age metrics in the following.

**Definition 3** (Peak age). The peak age is defined as

$$\Delta_{\text{peak}} \triangleq \sup_{i,j \in \mathcal{V}, i \neq j} \Delta_i^{(j)}(t)$$

for $T$ sufficiently large such that all nodes have complete status update tables.

**Definition 4** (Average age). The average age is defined as

$$\Delta_{\text{avg}} \triangleq \lim_{T \to \infty} \frac{1}{N^2 - N} \sum_{i,j \in \mathcal{V}, i \neq j} \frac{1}{T - t} \int_t^T \Delta_i^{(j)}(t) \, dt$$

for $T$ sufficiently large such that all nodes have complete status update tables.

We refer to a schedule as an ordered sequence of transmitting nodes and the corresponding status update parameter that they disseminate in each time slot. To illustrate the concept of a schedule, consider the following example for a 3-node line network represented in Fig. 2.

$n = 0$: Assume the first status update packet is transmitted by node 1. Node 1 starts dissemination of $H_1^{(1)}(-1^+)$ which has age of 0 at the beginning of time slot 0. Node 2 receives this packet at time $t = 0$ at which point the information has aged by one time unit, giving $\Delta_1^{(2)}(0^+) = 1$.

$n = 1$: Since node 2 now has an update of the $H_1(t)$ process, it relays a packet containing $H_1^{(2)}(0^+)$ during time slot 1. Node 3 receives the packet at time $t = 1$ at which point the information has aged by one additional time unit, giving $\Delta_1^{(3)}(1^+) = \Delta_1^{(2)}(1^+)^2 = 2$. Node 1 also receives the packet, but discards it since node 1 always has fresher, local knowledge of the $H_1(t)$ process.

$n = 2$: Node 2 disseminates $H_2^{(3)}(1^+)$ during time slot 2. Both nodes 1 and 3 receive this packet at time $t = 2$ at which point the information has aged by one time unit, giving $\Delta_2^{(1)}(2^+) = \Delta_2^{(3)}(2^+) = 1$.

![Fig. 1. A 5-node pan network. The nodes are indexed by $\mathcal{V} = \{1, 2, 3, 4, 5\}$. Here there exist two MCDS’s, $S_1 = \{2, 3\}$ and $S_2 = \{2, 4\}$, so that $\mathcal{L} = \mathcal{V} - (S_1 \cup S_2) = \{1, 5\}$ is the set of pseudo-leaf nodes.](image-url)
n = 3: Node 3 disseminates \( H^{(3)}_3(2^+) \) during time slot 3. Node 2 receives this packet at time \( t = 3 \) at which point the information has aged by one time unit, giving \( \Delta^{(2)}_3(3^+) = 1 \).

n = 4: Since node 2 now has an update of the \( H_2(t) \) process, it relays a packet containing \( H^{(2)}_3(3^+) \) during time slot 4. Node 1 receives the packet at time \( t = 4 \) at which point the information has aged by one additional time unit, giving \( \Delta^{(1)}_3(4^+) = \Delta^{(2)}_3(4^+) = 2 \).

Next, the sequence repeats starting by node 1 transmitting \( H^{(1)}_1(4^+) \) during time slot 5. Figure 3 represents the age evolution of the 6 indirectly-obtained status update parameters throughout the network for 3 repetitions of the schedule described above. Table I summarizes this schedule.

### Table I

**Example schedule for the 3-node line network in Fig. 2**

<table>
<thead>
<tr>
<th>time slot</th>
<th>transmitting node</th>
<th>disseminated status update</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 0, 5, 10, \ldots )</td>
<td>1</td>
<td>( H^{(1)}_1((n - 1)^+) )</td>
</tr>
<tr>
<td>( n = 1, 6, 11, \ldots )</td>
<td>2</td>
<td>( H^{(2)}_2((n - 1)^+) )</td>
</tr>
<tr>
<td>( n = 2, 7, 12, \ldots )</td>
<td>2</td>
<td>( H^{(2)}_2((n - 1)^+) )</td>
</tr>
<tr>
<td>( n = 3, 8, 13, \ldots )</td>
<td>3</td>
<td>( H^{(3)}_3((n - 1)^+) )</td>
</tr>
<tr>
<td>( n = 4, 9, 14, \ldots )</td>
<td>2</td>
<td>( H^{(2)}_3((n - 1)^+) )</td>
</tr>
</tbody>
</table>

Fig. 2. A 3-node line network as an example for partially-connected networks. The nodes are indexed by \( \{1, 2, 3\} \).

Observe that there are two primary sources that cause the status update parameters to age when maintaining a table. First, some updates arrive relatively stale since they must be disseminated over several hops. We refer to this source as the multi-hop penalty. Second, status updates grow stale while awaiting a refresh. We refer to this source as the inter-arrival time. For example, consistent with the red curve in the bottom subplot of Fig. 3, any status update of the \( H_1 \) process obtained by node 3 must have an age greater than or equal to 2 since the update must be disseminated over two hops. Also, node 3 waits for 5 time slots to receive another update of the \( H_1 \) process, causing \( \Delta^{(3)}_3(t) \) to grow staler yet by 5 more time slots. Thus, the age of the status update of the \( H_1 \) process at node 3 grows as high as \( \Delta^{(3)}_3(6^-) = 7 \).

### III. Lower bounds on peak and average age

In this section we derive lower bounds on the peak and average age metrics for status update dissemination in general, partially-connected networks. Prior to developing the lower bound, we present the following useful Lemma.

**Lemma 1** (Lower bound on the schedule length to refresh all tables). *To update the status of all parameters throughout the network, at least \( T \geq T^* \triangleq \gamma_c N + |L| \) status update packets need to be disseminated.*

**Proof.** To provide an update of a single process \( H_i \) to all nodes in the network, an update of the process must be disseminated at least \( \gamma_c \) times, or at least \( \gamma_c + 1 \) times if node \( i \) is a pseudo-leaf node [30, Theorem 1]. From Definition 1, there are \( |L| \) pseudo-leaf vertices in the graph; thus, to update all \( N \) parameters throughout the network it follows that at least \( \gamma_c N + |L| \) packets must be transmitted.

The following lower bounds hold for minimum length periodic schedules such that all parameters throughout the network get updated exactly once during each period. We refer to this class of schedules as \( T^* \)-periodic schedules.

**Theorem 1** (Lower bound on peak age). *The peak age of information for any \( T^* \)-periodic schedule is lower-bounded by*

\[
\Delta_{\text{peak}} \geq \Delta_{\text{peak}}^* \triangleq \begin{cases} 
\gamma_c(N + 1) & |L| = 0 \\
\gamma_c(N + 1) + |L| + 1 & |L| \geq 1
\end{cases} \quad (3)
\]

Due to space constraints, we provide a sketch of the proof of each theorem. Refreshing the \( H^{(3)}_3 \) process throughout the network requires \( \gamma_c \) hops (plus one, if node \( i \) is a pseudo-leaf node), and in each hop the age increases by one time unit. Thus, the maximum multi-hop penalty over all parameters can...
be lower-bounded by $\gamma_c$, or by $\gamma_c + 1$ if at least one pseudo-leaf node is present, i.e. if $|\mathcal{L}| \geq 1$. Next, the maximum inter-arrival time of all of the parameters throughout the network is lower-bounded by $T^*$. This is because only one node transmits during each time slot. $T^*$ packets need to be transmitted to update all parameters, and there is therefore always a parameter that has not received a new status update within the last $T^*$ time units. Since the peak total age is the sum of the inter-arrival time and multihop penalty terms, the peak total age is lower-bounded by the sum of the individual lower bounds.

**Theorem 2** (Lower bound on average age). The average age of information for any $T^*$-periodic schedule is lower-bounded by

$$\Delta_{\text{avg}} \geq \Delta_{\text{avg}}^* \triangleq \frac{T^*}{2} + \ell_G^*.$$ 

Since it requires at least $d(i,j)$ hops for the update of information process $H_i$ to reach node $j$, the multi-hop penalty of the $H_i^{(j)}$ parameter can be lower-bounded by $d(i,j)$. Averaging over all $N^2 - N$ indirectly obtained status updates gives the lower bound of $\ell_G^*$ on the average multi-hop penalty. Next, the average inter-arrival times of all of the parameters throughout the network are lower-bounded by $T^*/2$. Since the total average age is the sum of these two terms, it is lower-bounded by the sum.

**IV. SCHEDULE DESIGN FOR STATUS UPDATE DISSEMINATION**

In this section we provide an algorithm that generates schedules for refreshing all of the status update parameters throughout the network with any arbitrary topology. Observe that in the following, “Depth-First Search($G[S], i$)” describes an ordered list of vertices generated by performing a depth-first search of the graph induced by $S$ where the search starts at root vertex $i$.

**Algorithm 1: Schedule design to disseminate status updates throughout the network**

**Step I:** initialize time, $t \leftarrow -1$.

**Step II:** for node $i = 1 : N$ do

- if $\exists \text{MCDS } \bar{S}$ s.t. $i \in \bar{S}$ then
  - $\bar{S} \leftarrow \bar{S}$.
- else
  - $\bar{S} \leftarrow \bar{S} \cup \{i\}$, for any MCDS $\bar{S} \subset \mathcal{V}$.

* $S_{\text{sorted}} = \text{Depth-First Search($G[S], i$)}$

* for $k = 1 : |S_{\text{sorted}}|$ do

  - $j = S_{\text{sorted}}(k)$,
  - node $j$ transmits $H_i^{(j)}(t^+)$,
  - $t \leftarrow t + 1$.

**Step III:** repeat from Step II.

In the following we derive upper bounds on the achievable peak and average age of the schedules generated by Algorithm 1.

**Theorem 3** (Achievable peak age of Algorithm 1). The schedule generated by Algorithm 1 achieves $\Delta_{\text{peak}} = \Delta_{\text{peak}}^*$.

Algorithm 1 results in a schedule with the minimal length $T^*$, so the maximum inter-arrival time is exactly $T^*$. Meanwhile, the maximum multi-hop penalty is equal to $\gamma_c$, or $\gamma_c + 1$ if $|\mathcal{L}| \geq 1$. Thus, the lower bound on the peak age is achieved by this schedule.

**Theorem 4** (Upper bound on the achievable average age of Algorithm 1). The average age of the schedule generated by Algorithm 1 is at most

$$\Delta_{\text{avg}} \leq \frac{T^*}{2} + \gamma_c + \frac{|\mathcal{L}|}{N}.$$ 

For the schedule generated by Algorithm 1, the multi-hop penalty of all the parameters is upper-bounded by $\gamma_c$, or $\gamma_c + 1$ if $|\mathcal{L}| \geq 1$. Next, the average inter-arrival time of all of the parameters throughout the network is exactly $T^*/2$ since the algorithm constructs a schedule with the minimal length $T^*$. After some algebra the given upper bound is obtained.

**V. NUMERICAL RESULTS**

This section provides an exhaustive numerical example that illustrates the bounds on peak and average age, as well as the achievable average age of the schedule generated by Algorithm 1 for every connected network topology with $3 \leq N \leq 9$ nodes. We make use of a database [31] that exhaustively enumerates all connected network topologies with isomorphs removed. The achievable peak age is compared to the lower bound in Theorem 1 and the achievable average age is compared to the lower and upper bounds in Theorems 2 and 4, respectively. The results show that, indeed, $\Delta_{\text{peak,ach}} = \Delta_{\text{peak}}^*$ for all of the considered topologies. In addition, for the achievable average age values the numerical results show that

$$1 \leq \frac{\Delta_{\text{avg,ach}}(k)}{\Delta_{\text{avg}}^*(k)} \leq 1.035, \quad 1 \frac{\sum_{k=1}^{\mathcal{K}} \Delta_{\text{avg,ach}}(k)}{\sum_{k=1}^{\mathcal{K}} \Delta_{\text{avg}}^*(k)} \approx 1.008,$$

for $k = \{1, \ldots, \mathcal{K}\}$ where $\mathcal{K} = 273191$ represents the total number of networks with a connected topology and $N \leq 9$ nodes. In addition, for all $k = \{1, \ldots, \mathcal{K}\}$ we have $\Delta_{\text{avg,ach}}(k) \leq \Delta_{\text{avg,ub}}(k)$. Also, observe that the achievable average age is roughly half the the achievable peak age, i.e., $\Delta_{\text{avg,ach}}(k) \approx 0.5 \Delta_{\text{peak,ach}}(k)$ for all $k = \{1, \ldots, \mathcal{K}\}$.

**VI. CONCLUSION**

This work describes initial steps to quantify the age of information in multi-source multi-hop status update networks with any arbitrary partially-connected topology. Assuming that each source can generate status updates of its local process at
any time in a network with slotted transmissions, we derived fundamental lower bounds on the peak and average age of information over all of the status update parameters throughout the network. Next, an algorithm was proposed that constructs schedules for dissemination of the status updates in any given network topology. We derived upper bounds on the achievable peak and average age of the schedules constructed by the algorithm. The results showed that the schedule construction always achieves the lower bound on the peak age, and the gap between the achieved average age and the lower bound on the average age was not more than 3.5% of the achieved average age over all possible connected network topologies with \( N \leq 9 \) nodes.

Developing peak and average age bounds in a scenario where the transmission times are not fixed is an interesting extension. Characterizing the age improvement with multiple transmitting nodes during each time slot is another interesting future work direction to consider.

REFERENCES