

Distributed Scheduling Algorithms for Optimizing Information Freshness in Wireless Networks

Rajat Talak, Sertac Karaman, and Eytan Modiano

Abstract—Age of Information (AoI), measures the time elapsed since the last received information packet was generated at the source. We consider the problem of AoI minimization for single-hop flows in a wireless network, under pairwise interference constraints and time varying channel. We consider simple, yet broad, class of distributed scheduling policies, in which a transmission is attempted over each link with a certain attempt probability. We obtain an interesting relation between the optimal attempt probability and the optimal AoI of the link, and its neighboring links. We then show that the optimal attempt probabilities can be computed by solving a convex optimization problem, which can be done distributively.

I. INTRODUCTION

Timely delivery of information updates is gaining increasing relevance with the advent of technologies such as autonomous flying vehicles, internet of things, and other cyber physical systems. In autonomous flying vehicles, for example, timely exchange of position, velocity, and other control information can improve network safety [1], [2].

Packet delay, the traditionally used measure, accounts for the average time taken for a packet to reach the destination node, but does not measure the ‘information lag’ at the destination node, especially when the information update generation can be controlled [3]–[5]. For example, if two packets carry the same information, and one arrives early at the destination, then there is no need in accounting for the delay in reception of the second packet.

Age of information (AoI), a newly proposed metric [3], [6], is defined to be the time elapsed since the last received information update was generated at the source node. AoI, upon reception of a new update packet drops to the time elapsed since generation of the packet, and grows linearly otherwise. AoI, therefore, more accurately captures the ‘information lag’ at the destination node.

AoI has been analyzed for various queueing models [3], [7]–[11], but very little attention has been paid to AoI minimization in wireless networks. A problem of scheduling finitely many update packets under physical interference constraints was shown to be NP-hard in [12]. Age for a broadcast network, where only a single link can be activated at any time, was studied in [13]–[15]. Some preliminary analysis of age for a slotted ALOHA like random access was done in [16],

while multi-hop networks have been considered in [5], [17]. More recent work includes [4], [18], [19], which propose centralized scheduling policies for AoI minimization, under general interference constraints and time varying channels.

In a wireless network, it is not always possible to implement centralized schemes. In this paper, we consider a simple, yet broad, class of distributed scheduling policies for wireless networks, in which each link attempts transmission with a certain probability. We consider time varying channel and a pairwise interference model, in which each link interferes with a specified set of links in the network.

We first derive an interesting relation between the optimal attempt probability, and the optimal age of the given link and the interfering links. We then characterize the optimal distributed policy π_D , as a solution to a convex optimization problem. We further show that this convex problem can be solved distributively.

II. SYSTEM MODEL

We model a wireless communication network as a graph $G = (V, E)$, where V is the set of nodes and E is the set of directed communication links between the nodes. Each link $e \in E$ is associated with a source-destination pair. Time is slotted, with the duration of each slot set equal to the time it takes to transmit a single update packet. For simplicity, we normalize the slot duration to unity.

Transmission of update packets cannot occur simultaneously over all links due to wireless interference constraints [20]. For each link $e \in E$, there is a subset of links $N_e \subset E$, that interfere with link e , i.e., if link e and a link $l \in N_e$ attempt simultaneous transmission, then the transmission over link e fails due to interference. We call this the *pairwise interference model*. Popular models such as 1-hop and 2-hop interference models [20] are special cases of this model.

Channel uncertainty can also lead to failure in reception of update packets. We consider a two state channel process $S_e(t) \in \{0, 1\}$ for all time t and links e . Channel state $S_e(t) = 1$ implies that link e ’s channel is ON at time t , and a non-interfering transmission over e is successfully received. When $S_e(t) = 0$, even a non-interfering transmission over link e fails due to bad channel conditions. We assume the channel process $\{S_e(t)\}_{t,e}$ to be independent and identically distributed (i.i.d.) across time t , and only independent across links e . We let $\gamma_e = \Pr(S_e(t) = 1)$ denote the channel success probability for link e . We consider the case when the channel statistics, namely γ_e for all $e \in E$, are known but the current channel state $S_e(t)$ cannot be observed.

The authors are with the Laboratory for Information and Decision Systems (LIDS) at the Massachusetts Institute of Technology (MIT), Cambridge, MA. {talak, sertac, modiano}@mit.edu

This work was supported by NSF Grants AST-1547331, CNS-1713725, and CNS-1701964, and by Army Research Office (ARO) grant number W911NF-17-1-0508.

We use $U_e(t)$ to denote scheduling decisions at time t , for each link e . $U_e(t) = 1$ when link e attempts transmissions in slot t , and $U_e(t) = 0$ otherwise. Thus, a successful transmission occurs over link e if and only if a transmission is attempted on link e , the link e 's channel is ON, and no transmission is attempted over links $e' \in N_e$. This is equivalent to

$$\hat{U}_e(t) \triangleq S_e(t)U_e(t) \prod_{e' \in N_e} (1 - U_{e'}(t)) = 1. \quad (1)$$

A. Age of Information

AoI of a network is a function of AoI of each source-destination pair. Thus, we define age $A_e(t)$ for each link, to be the time since the last received update was generated. We assume source nodes to be *active sources*, i.e., it can generate a fresh update packet before every transmission.

When all sources are active, the age $A_e(t)$, is equal to the time elapsed since the last successful transmission over e . Thus, the evolution of age $A_e(t)$ can be written as

$$A_e(t+1) = \begin{cases} A_e(t) + 1 & \text{if } \hat{U}_e(t) = 0 \\ 1 & \text{if } \hat{U}_e(t) = 1 \end{cases}, \quad (2)$$

for all $e \in E$. We define average age for link e to be

$$A_e^{\text{ave}} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_e(t), \quad (3)$$

whenever the limit exists. We define the peak age to be the average of all the age peaks:

$$A_e^{\text{p}} = \limsup_{T \rightarrow \infty} \frac{1}{\mathcal{N}_e(T)} \sum_{i=1}^{\mathcal{N}_e(T)} A_e(T_e(i)), \quad (4)$$

where $T_e(i)$ denote the i th instance of successful transmission over link e , i.e., the i th instance when $\hat{U}_e(t) = 1$, and $\mathcal{N}_e(T)$ is the number of successful transmissions over link e until time T . We define average and peak AoI of the network to be a weighted sum of link AoI:

$$A^{\text{ave}} = \sum_{e \in E} w_e A_e^{\text{ave}}, \quad \text{and} \quad A^{\text{p}} = \sum_{e \in E} w_e A_e^{\text{p}}. \quad (5)$$

Our objective is to design distributed scheduling policies that minimize the network average and peak age.

B. Distributed Stationary Policies

We focus our attention on policies in which scheduling decisions are made by the source of each link e , and not by a centralized scheduler. In particular, we consider policies in which each link e attempts transmission with probability $p_e > 0$, independent across links and slots. We refer to these policies as the *distributed stationary policies*.

The probability that a non-interfering transmission is attempted over link e is given by $f_e = p_e \prod_{e' \in N_e} (1 - p_{e'})$. We will refer to f_e as the *link activation frequency* of link e , and use $\mathbf{f} = (f_e)_{e \in E}$ to denote the link activation frequency vector.

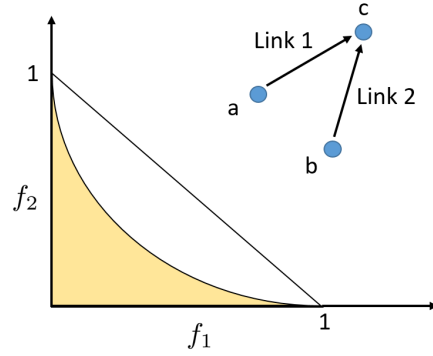


Fig. 1. Example of two interfering links, 1 and 2. The yellow region is \mathcal{F}_D . Also shown is the set of feasible link activation frequencies for centralized schedule: $\mathcal{F} = \{(f_1, f_2) \mid f_1 + f_2 \leq 1 \text{ and } f_i \geq 0\}$.

The space of all link activation frequencies \mathbf{f} , attainable using distributed stationary policies, is given by

$$\mathcal{F}_D = \left\{ \mathbf{f} \in \mathbb{R}^{|E|} \mid f_e = p_e \prod_{e' \in N_e} (1 - p_{e'}) \text{ and } 0 \leq p_e \leq 1 \forall e \in E \right\}. \quad (6)$$

Figure 1 shows the set of feasible link activation frequencies for distributed policies for the two interfering link example. Also, shown is the set of all link activation frequencies \mathcal{F} attainable using a centralized scheduler. It can be seen that the set \mathcal{F}_D is non-convex, and strictly smaller than \mathcal{F} . The gap between these two sets indicates the price one has to pay in resorting to distributed scheduling.

III. OPTIMAL POLICY

In this section, we characterize an age optimal policy, and propose an implementation that requires only local information. We first show that for any distributed stationary policy, the peak and average age are equal, and can be written as a simple convex function in link activation frequencies f_e .

Lemma 1: For any distributed stationary policy, the average and peak age is given by

$$A^{\text{ave}} = A^{\text{p}} = \sum_{e \in E} \frac{w_e}{\gamma_e f_e}, \quad (7)$$

where $f_e = p_e \prod_{e' \in N_e} (1 - p_{e'})$ is the link activation frequency of link e .

Proof: A similar result was proved in [4] which considered a class of centralized policies, that are not necessarily stationary. Here, we provide a proof for distributed stationary policies.

Consider a link e . Then, under any distributed stationary policy, the event that link e is successfully activated in slot t is independent and identically distributed across slots, with success probability of $\gamma_e f_e$. Thus, the time since last activation

We use $U_e(t)$ to denote scheduling decisions at time t , for each link e . $U_e(t) = 1$ when link e attempts transmissions in slot t , and $U_e(t) = 0$ otherwise. Thus, a successful transmission occurs over link e if and only if a transmission is attempted on link e , the link e 's channel is ON, and no transmission is attempted over links $e' \in N_e$. This is equivalent to

$$\hat{U}_e(t) \triangleq S_e(t)U_e(t) \prod_{e' \in N_e} (1 - U_{e'}(t)) = 1. \quad (1)$$

A. Age of Information

AoI of a network is a function of AoI of each source-destination pair. Thus, we define age $A_e(t)$ for each link, to be the time since the last received update was generated. We assume source nodes to be *active sources*, i.e., it can generate a fresh update packet before every transmission.

When all sources are active, the age $A_e(t)$, is equal to the time elapsed since the last successful transmission over e . Thus, the evolution of age $A_e(t)$ can be written as

$$A_e(t+1) = \begin{cases} A_e(t) + 1 & \text{if } \hat{U}_e(t) = 0 \\ 1 & \text{if } \hat{U}_e(t) = 1 \end{cases}, \quad (2)$$

for all $e \in E$. We define average age for link e to be

$$A_e^{\text{ave}} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_e(t), \quad (3)$$

whenever the limit exists. We define the peak age to be the average of all the age peaks:

$$A_e^{\text{p}} = \limsup_{T \rightarrow \infty} \frac{1}{\mathcal{N}_e(T)} \sum_{i=1}^{\mathcal{N}_e(T)} A_e(T_e(i)), \quad (4)$$

where $T_e(i)$ denote the i th instance of successful transmission over link e , i.e., the i th instance when $\hat{U}_e(t) = 1$, and $\mathcal{N}_e(T)$ is the number of successful transmissions over link e until time T . We define average and peak AoI of the network to be a weighted sum of link AoI:

$$A^{\text{ave}} = \sum_{e \in E} w_e A_e^{\text{ave}}, \quad \text{and} \quad A^{\text{p}} = \sum_{e \in E} w_e A_e^{\text{p}}. \quad (5)$$

Our objective is to design distributed scheduling policies that minimize the network average and peak age.

B. Distributed Stationary Policies

We focus our attention on policies in which scheduling decisions are made by the source of each link e , and not by a centralized scheduler. In particular, we consider policies in which each link e attempts transmission with probability $p_e > 0$, independent across links and slots. We refer to these policies as the *distributed stationary policies*.

The probability that a non-interfering transmission is attempted over link e is given by $f_e = p_e \prod_{e' \in N_e} (1 - p_{e'})$. We will refer to f_e as the *link activation frequency* of link e , and use $\mathbf{f} = (f_e)_{e \in E}$ to denote the link activation frequency vector.

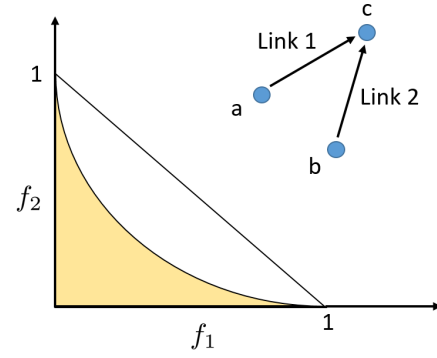


Fig. 1. Example of two interfering links, 1 and 2. The yellow region is \mathcal{F}_D . Also shown is the set of feasible link activation frequencies for centralized schedule: $\mathcal{F} = \{(f_1, f_2) \mid f_1 + f_2 \leq 1 \text{ and } f_i \geq 0\}$.

The space of all link activation frequencies \mathbf{f} , attainable using distributed stationary policies, is given by

$$\mathcal{F}_D = \left\{ \mathbf{f} \in \mathbb{R}^{|E|} \mid f_e = p_e \prod_{e' \in N_e} (1 - p_{e'}) \text{ and } 0 \leq p_e \leq 1 \forall e \in E \right\}. \quad (6)$$

Figure 1 shows the set of feasible link activation frequencies for distributed policies for the two interfering link example. Also, shown is the set of all link activation frequencies \mathcal{F} attainable using a centralized scheduler. It can be seen that the set \mathcal{F}_D is non-convex, and strictly smaller than \mathcal{F} . The gap between these two sets indicates the price one has to pay in resorting to distributed scheduling.

III. OPTIMAL POLICY

In this section, we characterize an age optimal policy, and propose an implementation that requires only local information. We first show that for any distributed stationary policy, the peak and average age are equal, and can be written as a simple convex function in link activation frequencies f_e .

Lemma 1: For any distributed stationary policy, the average and peak age is given by

$$A^{\text{ave}} = A^{\text{p}} = \sum_{e \in E} \frac{w_e}{\gamma_e f_e}, \quad (7)$$

where $f_e = p_e \prod_{e' \in N_e} (1 - p_{e'})$ is the link activation frequency of link e .

Proof: A similar result was proved in [4] which considered a class of centralized policies, that are not necessarily stationary. Here, we provide a proof for distributed stationary policies.

Consider a link e . Then, under any distributed stationary policy, the event that link e is successfully activated in slot t is independent and identically distributed across slots, with success probability of $\gamma_e f_e$. Thus, the time since last activation

is geometrically distributed with mean $\frac{1}{\gamma_e f_e}$. It turns out that the peak and average age for link e are indeed equal to this quantity, i.e.,

$$A_e^{\text{ave}} = A_e^{\text{p}} = \frac{1}{\gamma_e f_e},$$

and therefore, the network age is $\sum_{e \in E} \frac{w_e}{\gamma_e f_e}$. The detailed proof is given in Appendix A. ■

An immediate consequence of Lemma 1 is that both peak and average age minimization problems, over the space of all distributed stationary policies, can be written as

$$\text{Minimize}_{\mathbf{f} \in \mathcal{F}_D} \sum_{e \in E} \frac{w_e}{\gamma_e f_e} \quad (8)$$

Notice that, although the objective function is convex, the constraint set \mathcal{F}_D is non-convex, see Figure 1. Furthermore, obtaining optimal link activation frequencies \mathbf{f}^* does not suffice as the distributed policy is characterized by the attempt probabilities p_e . The optimal link attempt probabilities \mathbf{p}^* , that solve (10), yield a distributed stationary policy, that is both peak and average age optimal. We call this policy π_D .

The following result characterizes the optimal link attempt probabilities \mathbf{p}^* in terms of the optimal link age A_e^* .

Theorem 1: The attempt probabilities $\mathbf{p}^* = (p_e^* | e \in E)$, that solve (8), are given by

$$p_e^* = \frac{w_e A_e^*}{w_e A_e^* + \sum_{e' : e \in N_{e'}} w_{e'} A_{e'}^*}, \quad (9)$$

for all $e \in E$, where $A_e^* = [\gamma_e p_e^* \prod_{e' \in N_e} (1 - p_{e'}^*)]^{-1}$ is the optimal peak/average age of link e .

Proof: Substituting \mathcal{F}_D , from (6), in (8) we get

$$\begin{aligned} & \text{Minimize}_{\mathbf{p} \in [0,1]^{|E|}, \mathbf{f} \in \mathbb{R}^{|E|}} \sum_{e \in E} \frac{w_e}{\gamma_e f_e} \\ & \text{subject to} \quad f_e = p_e \prod_{e' \in N_e} (1 - p_{e'}) \quad \forall e \in E \quad (10) \\ & \quad \quad \quad f_e \geq 0 \text{ for all } e \in E \end{aligned}$$

Now, substituting $q_e = 1 - p_e$, the optimization problem (10) reduces to

$$\begin{aligned} & \text{Minimize}_{\mathbf{p} \geq 0, \mathbf{q} \geq 0} \sum_{e \in E} \frac{w_e}{\gamma_e p_e \prod_{e' \in N_e} q_{e'}} \quad (11) \\ & \text{subject to} \quad p_e + q_e \leq 1 \quad \forall e \in E \end{aligned}$$

This is a convex program in standard form, and can be solved using KKT conditions to obtain Theorem 1. The details are given in Appendix B. ■

Preliminary results on age optimization for a ALOHA type network model were presented in [16] for a network in which only one link can be activated at any given time. As a heuristic, it was argued in [16] that the attempt rates should be

$$p_e = \frac{1/\sqrt{\gamma_e}}{\sum_{e' \in E} 1/\sqrt{\gamma_{e'}}}. \quad (12)$$

However, Theorem 1 shows that the attempt probability

$$p_e^* = \frac{A_e^*}{\sum_{e' \in E} A_{e'}^*}, \quad (13)$$

is both peak and average age optimal. This gives us a more precise answer to the question posed in [16]. Our result, however, holds under the more general pairwise interference constraints, and non-equal weights $w_e > 0$.

A. Distributed Computation of π_D

Although, Theorem 1 gives a characterization of the optimal attempt probabilities, in terms of the optimal link age A_e^* , it does not provide for a simple method to computation.

We now provide for an alternate characterization of the network age minimization problem (8). In particular, we show that \mathbf{p}^* can be obtained by a simpler convex optimization problem, than the one in (11). We will also see that this simpler problem is akin to distributed computation of the policy π_D .

Theorem 2: The age minimization problem (8) is equivalent to the convex optimization problem:

$$\begin{aligned} & \text{Maximize}_{\lambda_e \geq 0} \sum_{e \in E} \left(\lambda_e + \sum_{e' : e \in N_{e'}} \lambda_{e'} \right) H \left(\frac{\lambda_e}{\lambda_e + \sum_{e' : e \in N_{e'}} \lambda_{e'}} \right) \\ & \quad + \sum_{e \in E} \lambda_e \left[1 + \log \left(\frac{w_e}{\lambda_e} \right) \right] \triangleq G(\boldsymbol{\lambda}), \quad (14) \end{aligned}$$

where $H(p) = p \log \left(\frac{1}{p} \right) + (1-p) \log \left(\frac{1}{1-p} \right)$ is the entropy function. If $\boldsymbol{\lambda}^*$ is the optimal solution to this problem, then the optimal attempt probabilities \mathbf{p}^* are given by

$$p_e^* = \frac{\lambda_e^*}{\lambda_e^* + \sum_{e' : e \in N_{e'}} \lambda_{e'}^*}, \quad (15)$$

for all $e \in E$.

Proof: See Appendix C. ■

We first note that the variable λ_e is a proxy for the weighted age $w_e A_e$ of link e , and the solution λ_e^* corresponds to optimal weighted age $w_e A_e^*$. This can also be intuited from (15), which is same as (9).

The objective function $G(\boldsymbol{\lambda})$ in (14) is a concave function; since the problem is convex. Therefore, a simple projected gradient ascent is guaranteed to converge to the optimal $\boldsymbol{\lambda}^*$ [21]. The gradient of the function $G(\boldsymbol{\lambda})$ is given by

$$\frac{\partial G(\boldsymbol{\lambda})}{\partial \lambda_e} = \log \left(\frac{w_e}{\lambda_e} \right) + \log \left(1 + \frac{\theta_e}{\lambda_e} \right) + \sum_{e' : e \in N_{e'}} \log \left(1 + \frac{\lambda_{e'}}{\theta_{e'}} \right),$$

for all $e \in E$, where $\theta_e = \sum_{e' : e \in N_{e'}} \lambda_{e'}$. Note that the gradient $\frac{\partial G(\boldsymbol{\lambda})}{\partial \lambda_e}$ depends only on λ_e and $\lambda_{e'}$, for e' that are ‘extended neighbors’ of e . This space dependence of $\frac{\partial G(\boldsymbol{\lambda})}{\partial \lambda_e}$ on $\boldsymbol{\lambda}$ makes the projected gradient descent algorithm akin to distributed implementation.

We give the projected gradient descent algorithm for distributively computing policy π_D in Algorithm 1. In it, we divide time into frames, each frame of $F \geq 1$ slots. For simplicity, we assume symmetric interference, i.e., e interferes with e' if and only if e' interferes with e which is same as the condition $N_e = \{e' \in E \mid e \in N_{e'}\}$ for all links e .

Link e , sets its attempt probability to $p_e(m)$, in frame $m \geq 1$. The source of link e tracks and updates two variables, namely, $\lambda_e(m)$ and $\theta_e(m)$. We use Π_ϵ to denote the projection on the set $[\epsilon, +\infty)$, for a $\epsilon > 0$. As seen in Steps 5 and 7 in Algorithm 1, variables $\lambda_e(m)$ and $\theta_e(m)$ need to be exchanged only between the neighboring links.

Algorithm 1 Distributed Computation of Policy π_D

- 1: $p_e(m)$: attempt probability of link e , in frame m
- 2: $\lambda_e(m), \theta_e(m)$: dual variables, in frame m
- 3: **Start:** Set $\lambda_e(0) = 1$ and $\theta_e(0) = |N_e|$ for all $e \in E$. Set $p_e(0) = 1/2$ and $m = 0$.
- 4: **for** each frame m **do**
- 5: **Send:** $\theta_e(m)$ to all $e' \in N_e$
- 6: Compute $\lambda_e(m+1)$:

$$\lambda_e(m+1) \leftarrow \Pi_\epsilon \left[\lambda_e(m) + \eta_m \left\{ \log \left(\frac{w_e}{\lambda_e(m)} \right) + \log \left(1 + \frac{\theta_e(m)}{\lambda_e(m)} \right) + \sum_{e' \in N_e} \log \left(1 + \frac{\lambda_{e'}(m)}{\theta_{e'}(m)} \right) \right\} \right]$$

- 7: **Send:** $\lambda_e(m+1)$ to all $e' \in N_e$
- 8: Update $\theta_e(m+1)$:

$$\theta_e(m+1) \leftarrow \sum_{e' \in N_e} \lambda_{e'}(m+1)$$

- 9: Update $p_e(m+1)$:

$$p_e(m+1) \leftarrow \frac{\lambda_e(m+1)}{\lambda_e(m+1) + \theta_e(m+1)}$$

- 10: $m \leftarrow m+1$
 - 11: **end for**
-

IV. CONCLUSION

We considered age minimization in a wireless network, with pairwise interference constraints, time varying channels, and single-hop flows. We considered a simple class of distributed scheduling policies, in which each link attempts transmission with probability p_e . We showed an interesting relationship between the optimal attempt probability for a link, and the optimal age of the link and it's neighboring links. We then showed that the optimal link attempt probabilities can be obtained by solving a convex optimization problem, which can be done distributively using the projected gradient ascent algorithm.

REFERENCES

- [1] O. K. Sahingoz, "Networking models in flying ad-hoc networks (FANETs): Concepts and challenges," *J. Intell. Robotics Syst.*, vol. 74, pp. 513–527, Apr. 2014.

- [2] R. Talak, S. Karaman, and E. Modiano, "Speed limits in autonomous vehicular networks due to communication constraints," in *Proc. CDC*, pp. 4998–5003, Dec. 2016.
- [3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. INFOCOM*, pp. 2731–2735, Mar. 2012.
- [4] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. Mobihoc (arXiv:1803.06467)*, Jun. 2018.
- [5] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *Proc. Allerton*, pp. 486–493, Oct. 2017.
- [6] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. SECON*, pp. 350–358, Jun. 2011.
- [7] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc. ISIT*, pp. 1681–1685, Jun. 2015.
- [8] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *Proc. ISIT*, pp. 66–70, Jul. 2013.
- [9] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. CISS*, pp. 1–6, Mar. 2012.
- [10] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *Proc. ISIT*, pp. 2574–2578, Jul. 2016.
- [11] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *Proc. ISIT*, pp. 1583–1587, Jun. 2014.
- [12] Q. He, D. Yuan, and A. Ephremides, "Optimizing freshness of information: On minimum age link scheduling in wireless systems," in *Proc. WiOpt*, pp. 1–8, May 2016.
- [13] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Proc. Allerton*, pp. 844–851, Sep. 2016.
- [14] Y.-P. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *Proc. ISIT*, pp. 1–5, Jun. 2017.
- [15] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. INFOCOM*, Apr. 2018.
- [16] S. K. Kaul and R. D. Yates, "Status updates over unreliable multiaccess channels," *Arxiv e-prints arXiv:1705.02521*, May 2017.
- [17] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal information updates in multihop networks," in *Proc. ISIT*, pp. 576–580, Jun. 2017.
- [18] R. Talak, I. Kadota, S. Karaman, and E. Modiano, "Scheduling policies for age minimization in wireless networks with unknown channel state," in *Proc. ISIT*, Jun. 2018.
- [19] R. Talak, S. Karaman, and E. Modiano, "Optimizing age of information in wireless networks with perfect channel state information," in *Proc. WiOpt*, May 2018.
- [20] A. Kumar, D. Manjunath, and J. Kuri, *Wireless Networking*. Morgan Kaufmann, 2008.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [22] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice Hall, 1 ed., 1989.

APPENDIX

A. Proof of Lemma 1

Consider a distributed stationary policy with link activation frequencies $f_e > 0$ for each $e \in E$. For a link e , let $T_e(i)$ be the i th instance when link e was successfully activated, i.e., the i th instance when $\hat{U}_e(t) = 1$. Let $X_e(i) = T_e(i) - T_e(i-1)$ to be the i th inter-(successful) activation time for link e . Since all the processes involved, namely $S_e(t)$ and $U_e(t)$, are i.i.d. across time t , $X_e(i)$ is geometrically distributed with rate equal to $\mathbf{P}[\hat{U}_e(t) = 1] = \gamma_e f_e$, i.e., $\mathbf{P}[X_e(i) = k] = \gamma_e f_e (1 - \gamma_e f_e)^{k-1}$, for all $k \in \{1, 2, \dots\}$.

We now compute the average age A_e^{ave} for link e . Note that the age $A_e(t)$ over slots $t \in \{T_e(i)+1, \dots, T_e(i+1)\}$ increases from 1 to $X_e(i)$ in steps of 1. Therefore, $\sum_{t=T_e(i)+1}^{T_e(i+1)} A_e(t) =$

$\sum_{m=1}^{X_e(i)} m$ for all i . Then, using renewal theory [22], we can derive the average age to be

$$\begin{aligned} A_e^{\text{ave}} &= \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_e(t) = \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \sum_{m=1}^{X_e(i)} m}{\sum_{i=1}^N X_e(i)}, \\ &= \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \frac{1}{2} X_e(i) (X_e(i) + 1)}{\sum_{i=1}^N X_e(i)}, \\ &= \frac{1}{2} \frac{\mathbb{E}[X_e(i) (X_e(i) + 1)]}{\mathbb{E}[X_e(i)]} \quad \text{a.s.} \end{aligned} \quad (16)$$

Computing the moments of $X_e(i)$ we get $A_e^{\text{ave}} = \frac{1}{\gamma_e f_e}$ a.s..

For peak age, note that the i th peak is equal to the i th inter-(successful) activation time $X_e(i)$, i.e., $A_e(T_e(i)) = X_e(i)$. Therefore, we have

$$A_e^p = \limsup_{T \rightarrow \infty} \frac{1}{\mathcal{N}(T)} \sum_{i=1}^{\mathcal{N}(T)} A_e(T_e(i)) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_e(i),$$

which is equal to $\mathbb{E}[X_e(1)] = \frac{1}{\gamma_e f_e}$ a.s.. This proves the result.

B. Proof of Theorem 1

Consider the Lagrangian dual

$$L(\mathbf{p}, \mathbf{q}, \boldsymbol{\lambda}) = \sum_{e \in E} \frac{w_e}{\gamma_e p_e \prod_{e' \in N_e} q_{e'}} + \sum_{e \in E} \lambda_e (p_e + q_e - 1). \quad (17)$$

The problem (11) satisfies Slater's conditions, and thus, the KKT conditions are both necessary and sufficient. Setting partial derivatives of L with respect to p_e and q_e to zero, we obtain

$$p_e \lambda_e = w_e A_e, \quad \text{and} \quad q_e \lambda_e = \sum_{e': e \in N_{e'}} w_{e'} A_{e'}, \quad (18)$$

for all $e \in E$, where $A_e = \frac{1}{\gamma_e p_e \prod_{e' \in N_e} q_{e'}}$ is the age of link e . Equations (18) imply that λ_e cannot be zero. By complementary slackness criteria we must have $p_e + q_e = 1$. This, with (18), gives $p_e = \frac{w_e A_e}{w_e A_e + \sum_{e': e \in N_{e'}} w_{e'} A_{e'}}$ and $\lambda_e = w_e A_e + \sum_{e': e \in N_{e'}} w_{e'} A_{e'}$. This proves the result.

C. Proof of Theorem 2

The age minimization problem, given in (10), can be written as by substituting $q_e = 1 - p_e$:

$$\begin{aligned} &\text{Minimize}_{\mathbf{f} \geq 0, \mathbf{p} \geq 0, \mathbf{q} \geq 0} \sum_{e \in E} \frac{w_e}{\gamma_e f_e} \\ &\text{subject to} \quad f_e \leq p_e \prod_{e' \in N_e} q_{e'} \quad \forall e \in E \\ &\quad \quad \quad p_e + q_e \leq 1 \quad \forall e \in E \end{aligned} \quad (19)$$

Now, substituting $P_e = \log p_e$, $Q_e = \log q_e$, and $h_e = \log f_e$, the problem reduces to

$$\begin{aligned} &\text{Minimize}_{\mathbf{h}, \mathbf{P}, \mathbf{Q}} \sum_{e \in E} \frac{w_e}{\gamma_e} \exp\{-h_e\} \\ &\text{subject to} \quad h_e - P_e - \sum_{e' \in N_e} Q_{e'} \leq 0 \quad \forall e \in E \\ &\quad \quad \quad \log(\exp\{P_e\} + \exp\{Q_e\}) \leq 0 \quad \forall e \in E \end{aligned} \quad (20)$$

This follows by first making the the substitution $p_e = e^{P_e}$, $q_e = e^{Q_e}$, and $f_e = e^{h_e}$, and then taking log in each of the constraints. This is a standard technique in geometric programming [21]. However, unlike geometric programming, we don't need to apply log function to the objective as it is already convex, and also separable in variables h_e .

The problem (20) is convex [21], and satisfies Slater's conditions. As a consequence, the duality gap is zero, and the problem (20) is equivalent to its Lagrangian dual problem.

Define the Lagrangian function of the optimization problem (20) to be

$$\begin{aligned} L(\mathbf{h}, \mathbf{P}, \mathbf{Q}, \boldsymbol{\lambda}, \boldsymbol{\nu}) &= \sum_{e \in E} \frac{w_e}{\gamma_e} e^{-h_e} + \sum_{e \in E} \nu_e \log(e^{P_e} + e^{Q_e}) \\ &\quad + \sum_{e \in E} \lambda_e (h_e - P_e - \sum_{e' \in N_e} Q_{e'}), \end{aligned} \quad (21)$$

for $\boldsymbol{\lambda} \geq 0$ and $\boldsymbol{\nu} \geq 0$. The dual problem for (20) is given by

$$\text{Maximize}_{\boldsymbol{\lambda} \geq 0, \boldsymbol{\nu} \geq 0} g(\boldsymbol{\lambda}, \boldsymbol{\nu}), \quad (22)$$

where $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is the dual objective function defines as

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \text{Minimize}_{\mathbf{h}, \mathbf{P}, \mathbf{Q}} L(\mathbf{h}, \mathbf{P}, \mathbf{Q}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \quad (23)$$

We will now show that the the optimization problem in Theorem 2 is in fact the dual problem (22).

First note that the Lagrangian function $L(\mathbf{h}, \mathbf{P}, \mathbf{Q}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ is strictly convex in \mathbf{h} , \mathbf{P} , and \mathbf{Q} [21]. Thus, first order conditions $\nabla_{\mathbf{h}, \mathbf{P}, \mathbf{Q}} L = 0$ yield the optimal \mathbf{h} , \mathbf{P} , and \mathbf{Q} . Setting these partial derivatives to 0, we get

$$\lambda_e = \frac{\nu_e e^{P_e}}{e^{P_e} + e^{Q_e}}, \quad \sum_{e': e \in N_{e'}} \lambda_{e'} = \frac{\nu_{e'} e^{Q_{e'}}}{e^{P_{e'}} + e^{Q_{e'}}}, \quad \text{and} \quad \lambda_e = \frac{w_e}{\gamma_e} e^{-h_e}, \quad (24)$$

for all $e \in E$. Adding and dividing the first two equations in (24) also yields

$$\nu_e = \lambda_e + \sum_{e': e \in N_{e'}} \lambda_{e'}, \quad \text{and} \quad P_e - Q_e = \log\left(\frac{\lambda_e}{\sum_{e': e \in N_{e'}} \lambda_{e'}}\right), \quad (25)$$

respectively, for all e . Substituting (24) and (25) in L , in (21), we recover $g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = G(\boldsymbol{\lambda})$, the objective function in (14), and therefore, the dual problem (22) is same as the optimization problem in Theorem 2. This proves the first part.

We now show that if $\boldsymbol{\lambda}^*$ is the solution to this problem, then the optimal attempt probability is given by (15). Let \mathbf{f}^* , \mathbf{p}^* , \mathbf{q}^* be the solution to (19), \mathbf{h}^* , \mathbf{P}^* , \mathbf{Q}^* be the solution to (20), and $\boldsymbol{\lambda}^*$ be the solution to the problem in Theorem 2.

Note that, for each link $e \in E$, we must have $f_e^* = p_e^* \prod_{e' \in N_e} q_{e'}^*$ and $p_e^* + q_e^* = 1$; as otherwise the objective function can be reduced by increasing f_e^* and/or p_e^* . Since $p_e^* = e^{P_e^*}$ and $q_e^* = e^{Q_e^*}$, we have $e^{P_e^*} + e^{Q_e^*} = 1$. Substituting this in (24), and using (25), we obtain

$$p_e^* = e^{P_e^*} = \frac{\lambda_e^*}{\lambda_e^* + \sum_{e': e \in N_{e'}} \lambda_{e'}^*}, \quad (26)$$

which proves the result.