The Dissemination of Time-Varying Information over Networked Agents with Gossiping

Melih Bastopcu, S. Rasoul Etesami, and Tamer Başar
Coordinated Science Laboratory
University of Illinois Urbana-Champaign, Urbana, IL, 61801
(bastopcu,etesami1,basar1)@illinois.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834729

***Abstract*—We consider information dissemination over a network of gossiping agents (nodes). In this model, a source keeps the most up-to-date information about a time-varying binary state of the world, and n receiver nodes want to follow the information at the source as accurately as possible. When the information at the source changes, the source first sends updates to a subset of $m \leq n$ nodes. After that, the nodes share their local information during the *gossiping period* to disseminate the information further. The nodes then estimate the information at the source using the majority rule at the end of the gossiping period. To analyze information dissemination, we introduce a new error metric to find the average percentage of nodes that can accurately obtain the most up-to-date information at the source. We characterize the equations necessary to obtain the steady-state distribution for the average error. Through numerical results, we first show that when the source's transmission capacity m is limited, gossiping can be harmful as it causes incorrect information to disseminate. We then find the optimal gossip rates to minimize the average error for a fixed m.**

I. Introduction

Motivated by many applications such as autonomous vehicular systems, content advertising on social media, and city emergency warning systems, information dissemination over the networks has gained significant attention. For instance, in the case of autonomous vehicular systems or city emergency warning systems, timely-critical information such as accident alerts or tornado warnings need to be disseminated as quickly and accurately as possible. As another example, companies often want to let their potential customers know about their latest products through advertisements over social media. In both of these examples, there is a single information source where the most up-to-date information is disseminated to multiple receivers over time.

In this paper, we consider a communication system with a source and n receiver nodes. The source keeps the most recent information about the state of the world, which takes values 0 or 1, and changes according to an exponential distribution. Upon each information update, the source wants to let the receiver nodes know about the most recent information. As the source has limited transmission capacity, it cannot send information to more than $m \leq n$ nodes, and each information transmission at the source takes an exponentially distributed length of time. After sending updates to m nodes, in order to further disseminate information, each pair of receiver nodes share their local information between each other, a process we shall refer to as *gossiping*. The gossiping period continues until the information at the source is updated again. At the end of each gossiping period, each receiver node that did not get the most recent information directly from the source comes up with an estimate based on the majority of the information it received from the other nodes. In order to measure the accuracy of the information dissemination at the end of each update cycle, we consider an error metric that takes value 1 for a receiver node that has a different estimate compared to the information at the source.

In gossip network literature, the model where only one node tries to spread its information to the entire network has been considered in [1] and named *single-piece dissemination*. The multi-piece spreading where all nodes try to spread their individual information to the remaining nodes has been studied in [2]. Moreover, the problem of finding the average of all nodes' initial information on a gossip network has been studied under the framework of *distributed averaging* in [3], [4]. The main goal of these works has been to analytically characterize either the information spreading time [1], [2] or the averaging time [3], [4] in the entire network. In another line of research, to measure the timeliness of information, age of information has been proposed in [5] and it has been extensively studied in multi-hop multicast networks [6]–[13], content freshness in the web [14]–[17], and timely remote estimation of random processes [18]–[24]. For a more detailed review of the age of information, we refer to [25], [26]. Recently, scaling of the version age has been considered in gossip networks [27]–[30].

Different from the earlier works on gossip networks as in [1], [2], we consider here a time-varying information source. Moreover, instead of tracking the information spreading time, we study the average percentage of the nodes that have access to the most recent information at the source before it is updated. Compared to [1], [2], our information updating is different and consists of two phases, where in the first phase, only the source can send updates to m nodes, and in the second phase, i.e., in the gossiping phase, only the nodes can share their local information. Thus, in the gossiping phase, incorrect information in the network can also spread. The works [27]–[30] have considered the age of information in gossip networks where each information update at the source is treated as a new update and content of the information has not been considered. In this work, we consider a binary information source that

*This work was supported in part by the NSF CAREER Award under Grant EPCN-1944403, and in part by the ARO MURI Grant AG285.

changes its state based on Poisson updates. Furthermore, in [27]–[30], the nodes update their information only if they receive fresher information. In contrast, in our work, the nodes that do not receive any update directly from the source make decisions based on the majority of the updates that they receive from the other nodes. As a result, the error metric and the information updating model that we consider differs from the earlier works in [27]–[30].

In this work, we first characterize the equations necessary to obtain the steady-state distribution of the average error. Then, in the numerical results, we show that when the source's transmission capacity m is limited, gossiping can be harmful as it increases incorrect information at the network. For the given source's total update capacity, there is an optimal gossip rate that minimizes the average error. When the network size increases, we should also increase both the source's transmission capacity m and the total update rate λ_s proportional to n to keep the average error at a constant level.

II. System Model and Problem Formulation

We consider an information updating system consisting of a source and n receiver nodes as shown in Fig. 1. The source keeps the most up-to-date information about a state of the world that takes binary values of 0 or 1. The information at the source is updated following a Poisson process with rate λ_e. We define the time interval between the jth and $j+1$th information update at the source as the jth update cycle and denote it by I_j. We assume that the source is able to send instantaneous signals to the nodes. After receiving these signals, the nodes know that information at the source is updated, but they do not know which information is realized at the source. We denote the information at the source at update cycle j as $x_s(j)$. For a given $x_s(j)$, the information at the source at the $j+1$th update cycle is equal to $x_s(j+1) = x_s(j)$ with probability $1-p$ and $x_s(j+1) = 1 - x_s(j)$ with probability p, i.e.,

$$\mathbb{P}(x_s(j+1)|x_s(j)) = \begin{cases} 1-p, & \text{if } x_s(j+1) = x_s(j)\ , \\ p, & \text{if } x_s(j+1) = 1 - x_s(j), \end{cases} \tag{1}$$

for all j, where $0 < p < 1$.

The source updates each receiver node according to a Poisson process with rate $\frac{\lambda_s}{n}$. In this system, in addition to the update arrivals from the source, each node can share its local information with the other nodes, a process called *gossiping*. Specifically, in this work, we consider a fully connected network where each node is connected to every other node with equal update rates. The total update rate of a node is λ. Thus, in this network, each node updates other neighbor nodes following a Poisson process with rate $\frac{\lambda}{n-1}$. We denote the information at node i at update cycle I_j as $x_i(j)$. The nodes want to follow the most up-to-date information prevailing at the source as accurately as possible based on the updates that they receive from the source as well as from the neighbor nodes during an update cycle.

In this paper, we consider an information updating mechanism where at the beginning of each update cycle I_j,

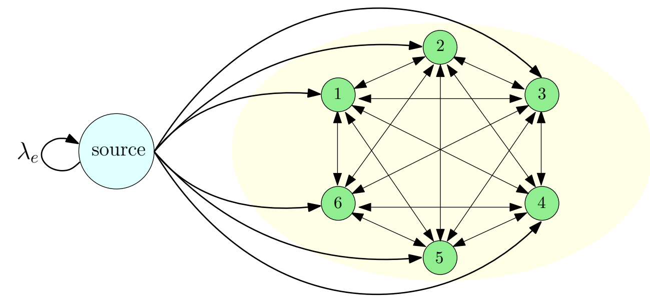

Fig. 1. A communication system that consists of a source and fully connected n nodes.

the source sends its current information to m nodes where $1 \leq m \leq n$. Here, we assume that the source knows (or is able to sense/monitor) the information prevailing at the nodes, and thus, it sends updates to the nodes that carry different information compared to the source. During this phase, if the information at the source is updated, then another update cycle starts, and thus the jth update cycle can be terminated before sending updates to m nodes. If the source sends updates to m nodes, we enter the gossiping phase in the update cycle I_j. During the gossiping phase, the nodes share their local information with each other. When the information at the source is updated, the gossiping phase ends. At the end of the gossiping period, the nodes that did not get an update directly from the source update their information based on the majority of the updates they receive during the gossiping period. If a node does not get any updates from the source or the other nodes, it keeps its local information unchanged. We denote the information at node i at the end of the gossiping period by $x_i'(j)$. In order to measure the performance of the information dissemination process, we define the error metric for node i at the update cycle j as

$$\Delta_i(j) = |x_s(j) - x_i'(j)|. \tag{2}$$

Then, the average estimation error over all nodes equals $\Delta(j) = \frac{1}{n}\sum_{i=1}^{n} \Delta_i(j)$, and the long-term average estimation error over all nodes is

$$\Delta = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} \Delta(j). \tag{3}$$

In the next section, we provide detailed analyses to characterize the long-term average error Δ.

III. Characterizing the Long-Term Average Error

In this section, we characterize the long-term average error Δ. Let us consider a generic update cycle I_j, and for simplicity of presentation, let us drop the index j from the variables in the rest of the analysis. At the beginning of the update cycle, we denote the number of nodes that have the same information as the source by $N \in \{0, \dots, n\}$. In this phase, either the source sends an update to a node after an exponential time with the rate λ_s or the information at the source is updated after an exponential time with the rate λ_e. Thus, the source sends update to a node with probability $\frac{\lambda_s}{\lambda_s + \lambda_e}$ or the information at the source is updated and the next update cycle starts with probability $\frac{\lambda_e}{\lambda_s+\lambda_e}$. Therefore, during a typical update cycle I with $N < n-m$, the source sends K_s updates with the following probability mass function (pmf)

$$\mathbb{P}(K_s = k_s|N<n-m)=\begin{cases}\left(\frac{\lambda_s}{\lambda_s+\lambda_e}\right)^{k_s}\frac{\lambda_e}{\lambda_s+\lambda_e}, \\ \quad \text{if } k_s=0,\dots,m-1, \\ \left(\frac{\lambda_s}{\lambda_s+\lambda_e}\right)^{m}, \quad \text{if } k_s=m.\end{cases} \tag{4}$$

Similarly, if $N \geq n-m$, we have

$$\mathbb{P}(K_s=k_s|N\geq n-m)=\begin{cases}\left(\frac{\lambda_s}{\lambda_s+\lambda_e}\right)^{k_s}\frac{\lambda_e}{\lambda_s+\lambda_e}, \\ \quad \text{if } k_s=0,\dots,n-N-1, \\ \left(\frac{\lambda_s}{\lambda_s+\lambda_e}\right)^{n-N}, \text{if } k_s=n-N.\end{cases} \tag{5}$$

For an update cycle with $N < n-m$, the network enters the gossiping phase with probability $\mathbb{P}(K_s = m|N < n-m)$ in (4), which decreases with m. In other words, choosing a large m decreases the probability of entering the gossiping phase. On the other hand, choosing a small m results in sending updates to a small number of nodes, and thus, in the gossiping phase, incorrect information can be spread. Therefore, there is an optimal m that achieves the smallest average error Δ.

If the source sends updates to m nodes before the information at the source is updated, then the gossiping phase starts. During the gossiping phase, either each node receives an update from the other nodes after an exponential time with rate λ or the information at the source is updated after an exponential time with rate λ_e. Thus, similar to [31], during the gossiping phase, node i receives K_i updates with the pmf:

$$\mathbb{P}(K_i = k_i) = \left(\frac{\lambda}{\lambda+\lambda_e}\right)^{k_i}\frac{\lambda_e}{\lambda+\lambda_e}, \quad k_i = 0,1,\dots. \tag{6}$$

In other words, K_i has a geometric distribution with parameter $\frac{\lambda_e}{\lambda+\lambda_e}$, i.e., $K_i \sim \text{Geo}(\frac{\lambda_e}{\lambda+\lambda_e})$.

At the beginning of the gossiping phase, there are $N+m$ nodes with the same information as the source and $n-N-m$ nodes with incorrect information. For the nodes with $x_i = x_s$, conditioned on the total number of updates $K_i = k_i$ that they received during the gossiping phase, the distribution of the number of updates that are equal to x_s is given by

$$\mathbb{P}(R_i = r|K_i = k_i, x_i = x_s) = \binom{k_i}{r}\left(\frac{N+m-1}{n-1}\right)^{r}\left(\frac{n-N-m}{n-1}\right)^{k_i-r}, \quad r=0,\dots,k_i, \tag{7}$$

where R_i is a random variable denoting the number of updates that are equal to x_s. In other words, for a node i that has $x_i = x_s$, conditioned on $K_i = k_i$, the random variable R_i has a binomial distribution with parameters $(k_i, \frac{N+m-1}{n-1})$, i.e., $R_i \sim \text{Bin}(k_i, \frac{N+m-1}{n-1})$. For a node i with $x_i \neq x_s$, we have

$$\mathbb{P}(R_i = r|K_i = k_i, x_i \neq x_s) = \binom{k_i}{r}\left(\frac{N+m}{n-1}\right)^{r}\left(\frac{n-N-m-1}{n-1}\right)^{k_i-r}, \quad r=0,\dots,k_i. \tag{8}$$

At the end of the gossiping period, based on the majority of the updates, the nodes i that have x_s as their prior information estimate the information at the source as $x_i' = x_s$ with probability $\mathbb{P}_{T,1}(N)$, which is given by

$$\begin{aligned}\mathbb{P}_{T,1}(N) = &\sum_{k_i=1}^{\infty}\mathbb{P}(R_i \geq \lfloor\frac{k_i}{2}\rfloor+1|K_i = k_i, x_i=x_s)\mathbb{P}(K_i=k_i) \\ &+\frac{1}{2}\sum_{k_i=1}^{\infty}\mathbb{P}(R_i=k_i|K_i=2k_i, x_i=x_s)\mathbb{P}(K_i=2k_i) \\ &+\mathbb{P}(K_i=0).\end{aligned} \tag{9}$$

We note that the first summation term in (9) corresponds to the case where a node receives a strictly higher number of x_s during the gossiping period. The second summation term in (9) refers to the case where a node receives equal number of x_s and $1-x_s$. In this case, a node estimates the information as either x_s or $1-x_s$ with equal probabilities. If a node does not get any updates during the gossiping phase, it keeps its current information that is given by the last term in (9). Similarly, for a node i that has prior information $x_i \neq x_s$, we can derive an expression for the probability of updating its information to x_s, denoted by $\mathbb{P}_{T,2}(N)$, as

$$\begin{aligned}\mathbb{P}_{T,2}(N) = &\sum_{k_i=1}^{\infty}\mathbb{P}(R_i \geq \lfloor\frac{k_i}{2}\rfloor+1|K_i=k_i, x_i\neq x_s)\mathbb{P}(K_i = k_i) \\ &+\frac{1}{2}\sum_{k_i=1}^{\infty}\mathbb{P}(R_i = k_i|K_i= 2k_i, x_i \neq x_s)\mathbb{P}(K_i= 2k_i).\end{aligned} \tag{10}$$

Note that this expression is identical to that in (9), except that in the summations, we use the probabilities $\mathbb{P}(R_i = r|K_i = k_i, x_i \neq x_s)$ given in (8), and the term $\mathbb{P}(K_i = 0)$ is excluded.

At the end of an update cycle with gossiping phase, m nodes that obtain information directly from the source will have $x_i' = x_s$.[1] There are N nodes that have prior information x_s. These nodes will update their information to $x_i' = x_s$ with probability $P_{T,1}(N)$ and to $x_i' = 1-x_s$ with probability $1-P_{T,1}(N)$. Thus, the total number of nodes that update their information to x_s, denoted by N_1', has the binomial distribution $N_1' \sim \text{Bin}(N, P_{T,1}(N))$. On the other hand, there are $n-N-m$ nodes that have prior information $1-x_s$. At the end of the gossiping phase, these nodes will update their information to $x_i' = x_s$ with probability $P_{T,2}(N)$, and to $x_i' = 1-x_s$ with probability $1-P_{T,2}(N)$. Thus, the total number of nodes that change their information to x_s, denoted by N_2', obeys the binomial distribution $N_2' \sim \text{Bin}(n-N-m, P_{T,2}(N))$. Therefore, at the end of the gossiping period, the total number of the nodes that have x_s is equal to $m+N'$, where $N' = N_1' + N_2'$ has the following pmf

$$\mathbb{P}(N' = n') = \sum_{\ell_1=\ell_{lower}}^{\ell_{upper}}\mathbb{P}(N_1' = \ell_1)\mathbb{P}(N_2' = n'-\ell_1), \tag{11}$$

[1]In the gossiping phase, these nodes send information to other nodes with rate λ, but they do not update their information based on the updates received from the other nodes.

for $n' = 0, \dots, n-m$, where $\ell_{lower} = \max\{0, n'+N+m-n\}$ and $\ell_{upper} = \min\{N, n'\}$.

Next, let us define $N''(j)$ to be the number of nodes that have the same information with the source at the end of the update cycle I_j, i.e., $x'_i(j) = x_s(j)$. If the update cycle I_j ends before entering the gossiping phase, then either $N(j) < n-m, K_s < m$ or $N(j) \geq n-m$. In these cases, the source sends updates to k_s nodes with probability distributions given in (4) and (5), respectively. If the source is able to send updates to m nodes, then gossiping phase starts and as a result, $N''(j) = m + n'$ nodes will have $x_s(j)$ with probabilities $\mathbb{P}(K_s = m)\mathbb{P}(N' = n')$, where $n' = 0, \dots, n-m$. Thus, the probability distribution of N'' for a given N is given in (12).

With the pmf of N'' as provided in (12), we can fully characterize the transition probabilities of going from N nodes that have x_s at the beginning of an update cycle to N'' nodes that have x_s at the end of that update cycle. Now let us consider a Markov chain over the state space (x_s, N), where by abuse of notation, we label the first $n+1$ states $(0,0), (0,1), \dots, (0,n)$ by $1, 2, \dots, n+1$, and the last $n+1$ states $(1,0), (1,1), \dots, (1,n)$ by $n+2, n+3, \dots, 2n+2$. We can then represent the transition probabilities between different states $a, b \in \{1, 2, \dots, 2n+2\}$ using a stochastic matrix $\boldsymbol{P} \in \mathbb{R}^{2(n+1)\times 2(n+1)}$, where $P_{a,b}$ denotes the probability of moving from state a to state b, and is given by

$$P_{a,b} = \begin{cases} (1-p)\mathbb{P}(N'' = b-1 | N = a-1), \\ \quad \text{if } 1 \leq a \leq n+1,\ 1 \leq b \leq n+1, \\ \frac{p}{1-p} P_{a,2n+3-b}, \\ \quad \text{if } 1 \leq a \leq n+1,\ n+1 \leq b \leq 2(n+1), \\ \frac{p}{1-p} P_{a,2n+3-b}, \\ \quad \text{if } n+1 \leq a \leq 2(n+1),\ 1 \leq b \leq n+1, \\ (1-p)\mathbb{P}(N'' = b-n-2 | N = a-n-2), \\ \quad \text{if } n+1 \leq a \leq 2(n+1),\ n+1 \leq b \leq 2(n+1). \end{cases} \tag{13}$$

We note that the stochastic matrix $\boldsymbol{P}$ in (13) is irreducible as every state b is accessible from any state a in a finite update cycle duration. Since $P_{a,a} > 0$ for all a in (13), the Markov chain induced by $\boldsymbol{P}$ is also aperiodic. Thus, the above Markov chain admits a unique stationary distribution given by the solution of $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$, where $\boldsymbol{\pi} = [\pi_{0,0}, \dots, \pi_{0,n}, \pi_{1,0}, \dots, \pi_{1,n}]$ is the row vector of steady-state probabilities of being at different states such that $\sum_{i=0}^{1}\sum_{j=0}^{n} \pi_{ij} = 1$, $\pi_{ij} \geq 0, \forall i, j$. Finally, we find the long-term average error among all the nodes by

$$\Delta = \sum_{j=0}^{n} \sum_{n''=0}^{n} (\pi_{0j} + \pi_{1j}) \mathbb{P}(N'' = n'' | N = j) \frac{n - n''}{n}. \tag{14}$$

In the next section, we provide numerical results to shed light on the effects of gossiping on information dissemination.

IV. Numerical Results

In this section, we provide numerical results to evaluate the effects of various parameters such as transmission capacity m, rate of information change λ_e, information transmission rate at the source λ_s, gossip rate λ, and the number of nodes n on information dissemination in gossip networks.

In the first numerical result, we take $p = 0.4$, $\lambda_e = 1$, $\lambda_s = 10$, and $n = 60$. We find the average error Δ with respect to m when $\lambda = \{0, 10, 20\}$. Note that $\lambda = 0$ corresponds to the case of no gossiping among the nodes. We see in Fig. 2 that when m is small, i.e., when the source can send updates to a small number of nodes, the average error Δ increases with gossip rate λ. Since m is small and the information change rate $p = 0.4$ is high, incorrect information disseminates due to gossiping in the network. As a result, the system with no gossiping ($\lambda = 0$) achieves the lowest average error. When we increase m sufficiently, the nodes start to have access to the same information as the source, and gossiping helps to disseminate the correct information. That is why the systems with gossiping, i.e., $\lambda = 10, 20$, achieve lower average error compared to the system with no gossiping. The lowest average error Δ is achieved when $m = 25$ for $\lambda = 10, 20$ and $m = 55$ for $\lambda = 0$. Here, we also note that the average error Δ is lower when $\lambda = 10$ compared to $\lambda = 20$, which shows that for a given m, there is an optimal gossip rate that achieves the lowest average error. Finally, increasing m further decreases the probability of entering the gossiping phase, and that is why all the curves in Fig. 2 overlap when $m \geq 40$.

In the second numerical result, we consider the same variable selections as in the previous example except that we take $m = \{5, 10, 15\}$ and change λ from 0 to 40. We see in Fig. 3 that increasing the gossip rate λ initially helps to reduce the average error Δ. Then, increasing λ further increases Δ as incorrect information among the nodes becomes more available. We see in Fig. 3 that the minimum average error is obtained when $\lambda = 1$ for $m = 5$, $\lambda = 3$ for $m = 10$, and $\lambda = 6$ or $\lambda = 7$ for $m = 15$. We note that as the source sends updates to more nodes, the optimal gossip rate increases.

In the third numerical result, we consider $p = 0.2$, $\lambda_e = 1$, $\lambda = 5$, and $n = 60$. We increase λ_s from 1 to 400 for $m = \{5, 10, 15\}$. We see in Fig. 4 that increasing λ_s initially decreases the average error Δ faster. However, as Δ depends also on the other parameters such as m and the gossip rate λ, increasing λ_s further does not improve the average error Δ and it converges to 0.348 for $m = 5$, 0.21 for $m = 10$, and 0.144 for $m = 15$.

In the fourth numerical result, we consider the effect of the network size n on the information dissemination. For that, first, we take $p = 0.2$, $\lambda_e = 1$, $\lambda = 10$, $m = 8$, $n = \{10, 20, \dots, 150\}$ and increase $\lambda_s = \{0.1n, 0.2n, 0.5n\}$ with the network size n. In this case, as the network size increases, the source's transmission rate also increases. However, we keep the total number of nodes that the source can send updates to the same, i.e., $m = 8$ for all n. In Fig. 5(a), when $\lambda_s = \{0.1n, 0.2n\}$, we see that the average error Δ initially decreases with n as λ_s is initially a primary limiting factor. Increasing n further increases Δ as m becomes more important. That is why all these three curves overlap between each other when λ_s is sufficiently large. Then, we consider a scenario where we keep $\lambda_s = 4$ and only increase

$$\mathbb{P}(N''=n''|N)=\begin{cases}\mathbb{P}(K_s=k_s|N<n-m), & \text{if } n''=k_s+N<m,\\ \mathbb{P}(K_s=m|N<n-m)\mathbb{P}(N'=n'), & \text{if } m\leq n''=m+n'<N,\\ \mathbb{P}(K_s=n''-N|N<n-m)+\mathbb{P}(K_s=m|N<n-m)\mathbb{P}(N'=n''-m), & \text{if } m\leq N\leq n''<N+m,\\ \mathbb{P}(K_s=m|N<n-m)\mathbb{P}(N'=n''-m), & \text{if } N+m\leq n''\leq n,\\ \mathbb{P}(K_s=n''-N|N\geq n-m), & \text{if } n-m\leq N\leq n''\leq n.\end{cases} \quad (12)$$

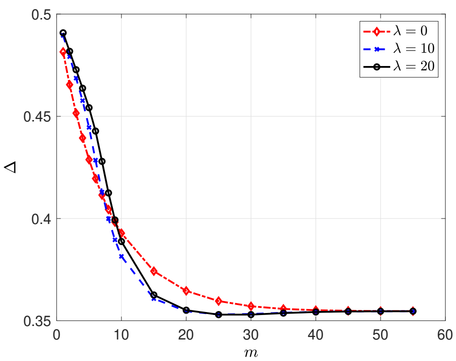

Fig. 2. The average error Δ with respect to m when $\lambda \in \{0, 10, 20\}$.

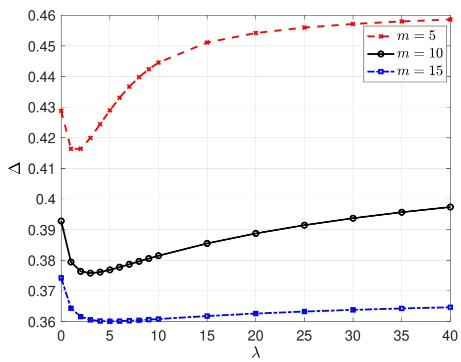

Fig. 3. The average error Δ with respect to the gossip rate λ for $m \in \{5, 10, 15\}$.

$m = \{0.1n, 0.2n, 0.5n\}$. In Fig. 5(b), increasing the maximum number of nodes that the source can send updates to in an update cycle alone does not reduce Δ as n increases. As we increase n, λ_s becomes the presiding factor and all the curves in Fig. 5(b) overlap. Finally, we increase both the source's transmission rate λ_s and capacity m with n, i.e., $\lambda_s = \{0.1n, 0.2n, 0.5n\}$ and $m = \{0.1n, 0.2n, 0.5n\}$. As a result, in Fig. 5(c), we observe that we can achieve a constant Δ by increasing λ_s and m proportional to n.

V. Conclusion and Future Directions

In this work, we considered information dissemination over gossip networks consisting of a source that keeps the most up-to-date information about a binary state of the world and n nodes whose goal is to follow the binary state of the world as accurately as possible. We first characterized the equations necessary to obtain the average error Δ over all nodes. In numerical results, we observed that gossiping could be harmful when the source's transmission capacity m is limited. Moreover, for a given m, we numerically determined the

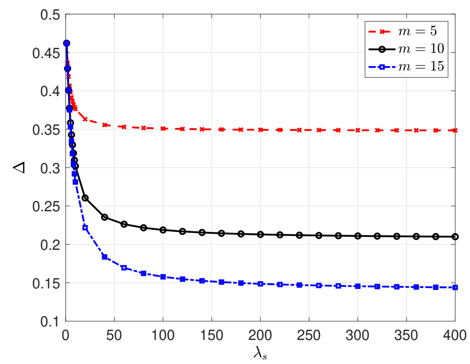

Fig. 4. The average error Δ with respect to the source's update rate λ_s for $m = \{5, 10, 15\}$.

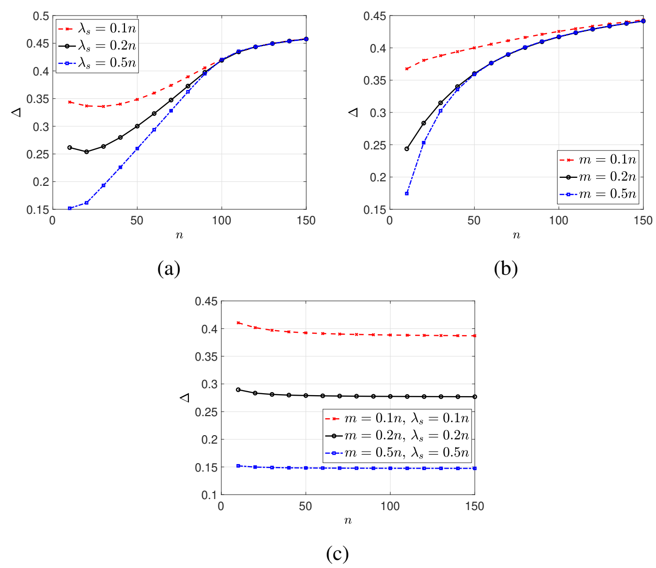

Fig. 5. Average error Δ with respect to n (a) when $\lambda_s \in \{0.1n, 0.2n, 0.5n\}$, (b) when $m \in \{0.1n, 0.2n, 0.5n\}$, and (c) when $m \in \{0.1n, 0.2n, 0.5n\}$, and $\lambda_s \in \{0.1n, 0.2n, 0.5n\}$.

optimal gossip rate that minimizes the average error Δ. When the network size n increases, in order to keep Δ the same, both the source's transmission capacity m and transmission rate λ_s need to be increased proportionally to n.

As a future direction, one can consider the problem where the information at the source can take $k > 2$ different values based on a known pmf. In this work, we only considered fully connected networks. Extending this work to arbitrarily connected networks could be an interesting direction. To further improve the average error Δ, the idea of clustering networks can be investigated. Finally, one can extend our work to the setting in which the source does not know the prior information of the nodes and, has to select m nodes at random.

References

[1] D. Shah. *Gossip algorithms*. Now Publishers Inc, 2009.

[2] D. Mosk-Aoyama and D. Shah. Information dissemination via network coding. In *IEEE ISIT*, pages 1748–1752, July 2006.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Analysis and optimization of randomized gossip algorithms. In *IEEE CDC*, pages 5310–5315, December 2004.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[5] S. K. Kaul, R. D. Yates, and M. Gruteser. Real-time status: How often should one update? In *IEEE Infocom*, March 2012.

[6] R. Talak, S. Karaman, and E. Modiano. Minimizing age-of-information in multi-hop wireless networks. In *Allerton Conference*, October 2017.

[7] V. Tripathi and S. Moharir. Age of information in multi-source systems. In *IEEE Globecom*, December 2017.

[8] A. M. Bedewy, Y. Sun, and N. B. Shroff. The age of information in multihop networks. *IEEE/ACM Transactions on Networking*, 27(3):1248–1257, June 2019.

[9] J. Zhong, E. Soljanin, and R. D. Yates. Status updates through multicast networks. In *Allerton Conference*, October 2017.

[10] J. Zhong, R. D. Yates, and E. Soljanin. Multicast with prioritized delivery: How fresh is your data? In *IEEE SPAWC*, June 2018.

[11] B. Buyukates, A. Soysal, and S. Ulukus. Age of information in multihop multicast networks. *Journal of Communications and Networks*, 21(3):256–267, July 2019.

[12] K. S. A. Krishnan and V. Sharma. Minimizing age of information in a multihop wireless network. In *IEEE ICC*, June 2020.

[13] S. Farazi, A. G. Klein, and D. R. Brown III. Fundamental bounds on the age of information in multi-hop global status update networks. *Journal of Communications and Networks, special issue on Age of Information*, 21(3):268–279, July 2019.

[14] J. Cho and H. Garcia-Molina. Effective page refresh policies for web crawlers. *ACM Transactions on Database Systems*, 28(4):390–426, December 2003.

[15] Y. Azar, E. Horvitz, E. Lubetzky, Y. Peres, and D. Shahaf. Tractable near-optimal policies for crawling. *PNAS*, 115(32):8099–8103, August 2018.

[16] A. Kolobov, Y. Peres, E. Lubetzky, and E. Horvitz. Optimal freshness crawl under politeness constraints. In *ACM SIGIR Conference*, July 2019.

[17] B. E. Brewington and G. Cybenko. Keeping up with the changing web. *Computer*, 33(5):52–58, May 2000.

[18] M. Wang, W. Chen, and A. Ephremides. Reconstruction of counting process in real-time: The freshness of information through queues. In *IEEE ICC*, July 2019.

[19] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu. Remote estimation of the Wiener process over a channel with random delay. In *IEEE ISIT*, June 2017.

[20] Y. Sun and B. Cyr. Information aging through queues: A mutual information perspective. In *IEEE SPAWC*, June 2018.

[21] J. Chakravorty and A. Mahajan. Remote estimation over a packet-drop channel with Markovian state. *IEEE Transactions on Automatic Control*, 65(5):2016–2031, July 2020.

[22] C. Kam, S. Kompella, and A. Ephremides. Age of incorrect information for remote estimation of a binary Markov source. In *IEEE Infocom*, July 2020.

[23] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Sample, quantize and encode: Timely estimation over noisy channels. *IEEE Transactions on Communications*, 2021. To appear. Available on arXiv:2007.10200.

[24] M. Bastopcu and S. Ulukus. Who should Google Scholar update more often? In *IEEE Infocom*, July 2020.

[25] A. Kosta, N. Pappas, and V. Angelakis. Age of information: A new concept, metric, and tool. *Foundations and Trends in Networking*, 12(3):162–259, November 2017.

[26] R. D. Yates, Y. Sun, D. R. Brown III, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, May 2021.

[27] R. D. Yates. The age of gossip in networks. In *IEEE ISIT*, July 2021.

[28] M. Bastopcu, B. Buyukates, and S. Ulukus. Gossiping with binary freshness metric. July 2021. Available on arXiv:2107.14218.

[29] B Buyukates, M. Bastopcu, and S. Ulukus. Age of gossip in networks with community structure. In *IEEE SPAWC*, September 2021.

[30] B. Buyukates, M. Bastopcu, and S. Ulukus. Version age of information in clustered gossip networks. September 2021. Available on arXiv:2109.08669.

[31] M. Bastopcu and S. Ulukus. Information freshness in cache updating systems. *IEEE Transactions on Wireless Communications*, 20(3):1861–1874, March 2021.