




Article

The Role of Gossiping in Information Dissemination over a Network of Agents

Melih Bastopcu , Seyed Rasoul Etesami  and Tamer Başar* 

Coordinated Science Laboratory, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA; bastopcu@illinois.edu (M.B.); etesami1@illinois.edu (S.R.E.)

* Correspondence: basar1@illinois.edu

Abstract: We consider information dissemination over a network of gossiping agents. 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. Then, 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 the 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 and then analyze the system behavior under both high and low gossip rates. We develop an adaptive policy that the source can use to determine its current transmission capacity m based on its past transmission rates and the accuracy of the information at the nodes. Finally, we implement a clustered gossiping network model, to further improve the information dissemination.

Keywords: information dissemination; gossip networks; gossiping effect; social networks; Markov chains



Citation: Bastopcu, M.; Etesami, S.R.; Başar, T. The Role of Gossiping in Information Dissemination over a Network of Agents. *Entropy* **2024**, *26*, 9. <https://doi.org/10.3390/e26010009>

Academic Editors: Adam Lipowski, Photios A. Stavrou, Nikolaos Pappas, Giulia Cervia and Serkan Sartaş

Received: 27 September 2023

Revised: 9 December 2023

Accepted: 15 December 2023

Published: 21 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. 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, needs to be disseminated as quickly and as 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 binary 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, local information is shared between each pair of receiver nodes, 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.

1.1. Related Work

In the gossip-network literature, a model where only one node tries to spread its information to the entire network was considered in [1] and named *single-piece dissemination*. Multi-piece spreading, where all nodes try to spread their individual information to the remaining nodes, was studied in [2]. Moreover, the problem of finding the average of all nodes' initial information in a gossip network was studied under the framework of *distributed averaging* in [3,4]. The main goal of these works was to analytically characterize either the information spreading time [1,2] or the averaging time [3,4] in the entire network. In all these settings, the information was considered to be *static*, i.e., it did not change over time.

Reference [5] considered the problem of gossiping *dynamic information*. As the gossip network may consist of asynchronous agents where there is no central clock, in order to maintain the information flow in the gossip network, *timestamping* is a commonly used technique, where the agents keep the generation time of their local status updates [6]. During the gossiping phase, information updates among the agents are determined based on whoever has the largest timestamp of particular information, which indicates the *information freshness* of the local agents.

In another line of research, to measure information freshness, the age of information was defined as the difference between the current time and the timestamp of the last status update received by the agents. For a more detailed review of the age of information, we refer to [7,8]. Recently, scaling of the age of information was considered in gossip networks [9–11]. In [9], the stochastic hybrid system (SHS) method was used to characterize the version age in arbitrarily connected gossip networks, and scaling of the version age was studied in the symmetric ring and fully connected networks. By using the idea of clustering, scaling of the age of information was further improved in [10]. Then, scaling of the binary freshness metric [12–14], which takes either the value 1 when the information is fresh, or 0 otherwise, was studied in the gossip networks in [11].

In all these aforementioned works, the timestamp of the information plays a critical role in determining information dissemination in gossip networks. As the timestamp of the information increases, as new versions of the information are generated, either the size grows without bound—in which case, the agents spend most of their capacity in exchanging large numbers and comparing the values of these large timestamps [5], to determine the freshest information—or, in the case of a bounded timestamp, when overflow happens, the order for information freshness can be lost [6]. In certain applications, an external adversary may interrupt the information flow and alter the timestamp of the information, such that the older versions of the information may be re-branded as fresh information [15]. Recently, the effect of timestamping on age scaling was explored in [16].

Unlike the earlier works on gossip networks, as in [1,2], we considered in this paper a time-varying information source and, instead of tracking the information-spreading time, we studied the average percentage of the nodes that had access to the most recent information at the source before it was updated. Compared to the dynamic information dissemination in [6], in our work we did not use the timestamps of the information. Instead, to maintain the information flow, we used instantaneous signaling from the source to the nodes, to synchronize the nodes. We implemented an information updating mechanism consisting of two phases: in the first phase, only the source could send updates to m nodes; in the second phase, i.e., in the gossiping phase, only the nodes could share their local information. Thus, in the gossiping phase, incorrect information in the network could also spread. Works [9–11] considered the age of information in gossip networks, where each information update at the source was treated as a new update. In our work, on the other hand, we considered a binary dynamic information source. Thus, the content of the

information affected our error metric. Furthermore, in [9–11], the nodes updated their information only if they received fresher information. By contrast, in our work, the nodes that did not receive any update directly from the source made decisions based on the majority of the updates that they received from the other nodes. As a result, the error metric and the information updating model that we considered differed from the earlier works in [9–11].

The binary information structure appears in real-world applications such as robotic networks, where a group of robots working on a horizontal line wants to decide whether a neighbor robot is in front of the group or behind it. Here, the binary information structure is sufficient to represent the relative position of the robots in the network. Inspired by this example, by using only the sign of the relative states, reference [17] considered a decentralized-online-convex-optimization problem with time-varying loss functions, and reference [18] solved a distributed discrete-time optimization over multi-agent networks. As another notable example, reference [19] considered a model where the actual opinion of the public evolved as a continuous variable in $[0, 1]$ but the expressed opinions took only discrete binary values $\{0, 1\}$ that resembled the opinion polls. Motivated by all these examples, in our work, we focused our attention on binary information dissemination as an initial step in analyzing the role of gossiping in information dissemination.

Finally, dissemination of misinformation on social networks has attracted significant interest in recent years. The *network immunization* problem has been considered, to prevent the diffusion of harmful information that can *infect* the network [20–22]. More specifically, reference [20] proposed an algorithm that utilizes the community structure for network immunization. Reference [21] proposed a comprehensive solution for the immediate detection and containment of harmful content, aiming to curb its propagation across the network. Reference [22] applied deep neural networks to develop context-aware algorithms that can detect fake news.

1.2. Contributions

In this work, we first characterize the equations necessary to obtain the steady-state distribution of the average error (which was also appeared in our preliminary work in [23]). Then, we provide analytical results for the high and low gossip rates. When the gossip rate is high, we show that the probability of obtaining correct information converges to a step function where if the majority of the nodes have the correct information then all the nodes are able to estimate the information correctly with probability 1. In other words, as the gossip rate increases, the information at all nodes becomes mutually available to them, and all the nodes in the network behave like a single node. However, when the gossip rate is low, the gossiping phase can be approximated by either not receiving any updates, in which case the nodes hold on to their prior information, or receiving a single update. Based on this approximation, we characterize analytically the gain obtained through gossiping, and we find an adaptive selection policy for the source, which suggests that the source should send updates to more nodes when the nodes have mostly incorrect information. Then, to further reduce the average error, we implement the idea of clustering, where, instead of sending information to all nodes, the source sends its information only to a smaller number of cluster heads. Then, the cluster heads forward the information to nodes within their clusters. For this network model, we characterize the equations to find the long-term average error at the cluster heads and at the nodes in the clusters. Finally, we provide extensive simulations, to illustrate the effect of gossiping and clustering on information dissemination.

2. System Model and Problem Formulation

We considered an information updating system consisting of a source and n receiver nodes, as shown in Figure 1. The source kept the most up-to-date information about a state of the world that took binary values of 0 or 1. The information at the source was updated following a Poisson process with rate λ_e . We defined the time interval between

the j th and $j + 1$ th information update at the source as the j th update cycle and denoted it by I_j . We assumed that the source was able to send instantaneous signals to the nodes. After receiving these signals, the nodes knew that information at the source had been updated, but they did not know what information had been realized at the source. Such instantaneous signalings exist in many practical systems. For example, consider a news provider making news either to support or oppose a topic of interest. After the news is published, the news provider can send an instantaneous notification to its subscribers about the occurrence of the news through push notifications on smart devices or headlines in TV broadcasts or their websites. However, after receiving these notifications, individuals still do not know the actual update until they enter the news provider's website or watch the TV broadcast. As another application, consider a city emergency-warning system, or anomaly detection in security applications where warning signals can occur over time. As warning signals can also happen due to false alarms, upon receiving such warning signals, individuals do not know whether there is an actual anomaly or not, until further test results can confirm the actual status. Thus, motivated by the aforementioned examples, we utilized the synchronization signal, which can indicate the information update at the source but does not provide any information about the source's state realization.

We denoted 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 was equal to $x_s(j + 1) = x_s(j)$ with probability $1 - p$ and to $x_s(j + 1) = 1 - x_s(j)$ with probability p , i.e.,

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

for all j , where $0 < p < \frac{1}{2}$. As $0 < p < \frac{1}{2}$, the nodes kept their state estimation unchanged whenever a new update cycle started. (Our results are extendable to the setting where $0.5 < p < 1$. In this case, the optimal decision taken by each node should be to revert their belief at the beginning of each update cycle.)

The source updated 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 considered a fully connected network where each node was connected to every other node with equal update rates. The total update rate of a node was λ . Thus, in this network, each node updated its neighbor nodes following a Poisson process with rate $\frac{\lambda}{n-1}$. We denoted the information at node i at update cycle I_j as $x_i(j)$. The nodes wanted to follow the most up-to-date information prevailing at the source as accurately as possible, based on the updates received from the source as well as from the neighbor nodes during an update cycle.

In this paper, we considered an information updating mechanism where at the beginning of each update cycle I_j the source sent its current information to m nodes where $1 \leq m \leq n$, as shown in Figure 1a. Here, we assumed that the source knew (or was able to sense/monitor) the information prevailing at the nodes and, thus, it sent updates to the nodes that carried different information compared to the source. (This approach was motivated especially by online advertisements, whereby companies such as Amazon and Google are able to monitor whether a potential customer is interested in their target products by the customer's search, view, and click history and, thus, present their advertisements accordingly. They can sense the final opinion of their potential customer by observing the potential customer's behavior, such as buying an advertised product). During this phase, if the information at the source was updated, then another update cycle started and, thus, the j th update cycle could be terminated before sending updates to m nodes. If the source sent updates to m nodes, it sent another instantaneous signal to start the gossiping among the nodes. Then, we entered the gossiping phase in the update cycle I_j . During the gossiping phase, illustrated in Figure 1b, the nodes shared their local information with one another. When the information at the source was updated, the gossiping phase ended. At the end of

the gossiping period, the nodes that did not get an update directly from the source updated their information based on the majority of the updates they received during the gossiping period. If a node did not get any updates from the source or the other nodes, it kept its local information unchanged. We denoted 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 defined the error metric for node i at the update cycle j as

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

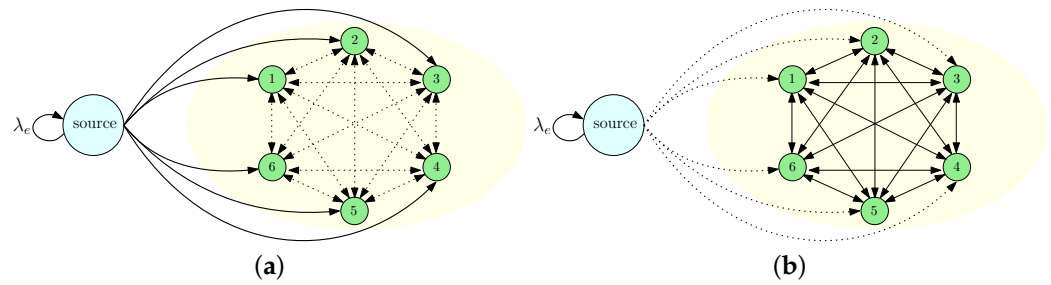


Figure 1. A communication system that consists of a source and fully connected n nodes where (a) only the source sends updates to the nodes, and (b) the nodes share their local information, called the gossiping phase.

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

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

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

3. 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 an 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):

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

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

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

For an update cycle with $N < n - m$, the network enters the gossiping phase with probability $\mathbb{P}(K_s = m | N < n - m) = \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^m$, 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 . As in [24], in the gossiping phase, node i receives K_i updates with the following 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 geometric distribution with parameter $\frac{\lambda_e}{\lambda + \lambda_e}$, i.e., $K_i \sim \text{Geo}\left(\frac{\lambda_e}{\lambda + \lambda_e}\right)$.

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., $R_i \sim \text{Bin}(k_i, \frac{N+m-1}{n-1})$. Similarly, for the nodes 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

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

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 an 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

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

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 that $\mathbb{P}(K_i = 0)$ is excluded. In the next theorem, we state the long-term average error.

Theorem 1. Under the proposed gossiping network, the long-term average error Δ is given by

$$\Delta = \sum_{j=0}^n \sum_{n''=0}^n (\pi_{0,j} + \pi_{1,j}) \mathbb{P}(N'' = n'' | N = j) \frac{n - n''}{n}, \tag{11}$$

where $\mathbb{P}(N'' = n'' | N = j)$ is provided in (13) and $\boldsymbol{\pi} = [\pi_{0,0}, \dots, \pi_{0,n}, \pi_{1,0}, \dots, \pi_{1,n}]$ is the row vector of steady-state probabilities of the Markov chain over the state space (x_s, N) . The unique stationary distribution is given by the solution of $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$ for a stochastic matrix $\mathbf{P} \in \mathbb{R}^{2(n+1) \times 2(n+1)}$, where the transition probabilities of \mathbf{P} are given later in (14).

Proof. We note that at the end of an update cycle with a gossiping phase, m nodes that obtain information directly from the source will have $x'_i = x_s$ (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). 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:

$$\mathbb{P}(N' = n') = \sum_{\ell_1=\ell_{lower}}^{\ell_{upper}} \mathbb{P}(N'_1 = \ell_1) \mathbb{P}(N'_2 = n' - \ell_1), \tag{12}$$

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 the gossiping phase starts and, as a result, $N''(j) = m + n'$ nodes will have $x_s(j)$ with probabilities

$\mathbb{P}(K_i = m)\mathbb{P}(N' = n')$, where $n' = 0, \dots, n - m$. Thus, the probability distribution of N'' for a given N is given by

$$\mathbb{P}(N'' = n''|N) = \begin{cases} \mathbb{P}(K_s = k_s|N < n - m), & \text{if } n'' = k_s + N < m, \\ \mathbb{P}(K_s = m|N < n - m)\mathbb{P}(N' = n'), & \text{if } m \leq n'' = m + n' < N, \\ \mathbb{P}(K_s = n'' - N|N < n - m) \\ + \mathbb{P}(K_s = m|N < n - m)\mathbb{P}(N' = n'' - m), & \text{if } m \leq N \leq n'' < N + m, \\ \mathbb{P}(K_s = m|N < n - m)\mathbb{P}(N' = n'' - m), & \text{if } N + m \leq n'' \leq n, \\ \mathbb{P}(K_s = n'' - N|N \geq n - m), & \text{if } n - m \leq N \leq n'' \leq n. \end{cases} \quad (13)$$

With the pmf of N'' as provided in (13), 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 \mathbf{P} , where $P_{a,b}$ denotes the probability of moving from state a to state b and is given by

$$P_{a,b} = \begin{cases} (1 - p)\mathbb{P}(N'' = b - 1|N = a - 1), & \text{if } 1 \leq a \leq n + 1, 1 \leq b \leq n + 1, \\ \frac{p}{1 - p}P_{a,2n+3-b}, & \text{if } 1 \leq a \leq n + 1, n + 1 \leq b \leq 2(n + 1), \\ \frac{p}{1 - p}P_{a,2n+3-b}, & \text{if } n + 1 \leq a \leq 2(n + 1), 1 \leq b \leq n + 1, \\ (1 - p)\mathbb{P}(N'' = b - n - 2|N = a - n - 2), & \text{if } n + 1 \leq a \leq 2(n + 1), n + 1 \leq b \leq 2(n + 1). \end{cases} \quad (14)$$

We note that the stochastic matrix \mathbf{P} in (14) is irreducible, as every state b is accessible from any state a in a finite update cycle duration. As $P_{a,a} > 0$ for all a in (14), the Markov chain induced by \mathbf{P} is also aperiodic. Thus, the above Markov chain admits a unique stationary distribution given by the solution of $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$, such that $\sum_{i=0}^1 \sum_{j=0}^n \pi_{ij} = 1$, $\pi_{ij} \geq 0, \forall i, j$. Finally, we characterize the long-term average error among all the nodes by (11). \square

In the following section, we proceed to approximate the probabilities $\mathbb{P}_{T,1}(N)$ and $\mathbb{P}_{T,2}(N)$ provided in this section, to understand the effect of gossiping better when the gossip rate λ is low and high compared to the information change rate at the source λ_e .

4. Analysis for High and Low Gossip Rates

In this section, we develop approximations for $\mathbb{P}_{T,1}(N)$ and $\mathbb{P}_{T,2}(N)$, which are the probabilities of choosing x_s at the end of a gossiping period when the nodes have the same prior information with the source and when they do not, respectively. First, by assuming sufficiently large n and N , we can approximate the conditional pmfs for R_i given in (7) and (8) by the binomial distribution $\mathbb{P}(R_i|K_i = k_i) \sim \text{Bin}(k_i, \frac{N+m}{n})$. Let us denote the corresponding $\mathbb{P}_{T,1}(N)$ and $\mathbb{P}_{T,2}(N)$ obtained by substituting this binomial approximation into (9) and (10) by $\hat{\mathbb{P}}_{T,1}(N)$ and $\hat{\mathbb{P}}_{T,2}(N)$, respectively. As $\hat{\mathbb{P}}_{T,1}(N) = \hat{\mathbb{P}}_{T,2}(N) + \mathbb{P}(K_i = 0)$, for the rest of this section we will only approximate $\hat{\mathbb{P}}_{T,2}(N)$, and can find the probability $\hat{\mathbb{P}}_{T,1}(N)$ accordingly. Next, for sufficiently large values of k_i , we can approximate $\mathbb{P}(R_i \geq \frac{k_i}{2}|K_i = k_i)$ as

$$\mathbb{P}\left(R_i \geq \frac{k_i}{2} | K_i = k_i\right) \approx Q\left(\sqrt{k_i}A(N)\right), \quad (15)$$

where $A(N) = \frac{\frac{1}{2} - \frac{N+m}{n}}{\sqrt{\frac{N+m}{n}\left(1 - \frac{N+m}{n}\right)}}$ and $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du$. We note that (15) is due to the normal approximation of binomial distribution by using the central limit theorem (CLT). In the following proposition, we show that $\hat{\mathbb{P}}_{T,2}(N)$ can be approximated by a summation of Q -functions.

Proposition 1. When λ is sufficiently large compared to λ_e , $\hat{\mathbb{P}}_{T,2}(N)$ can be approximated by

$$\mathbb{P}_{T,app}(N) = \sum_{k_i=1}^{\infty} Q\left(\sqrt{k_i}A(N)\right)\mathbb{P}(K_i = k_i). \tag{16}$$

Proof. Using the CLT, there exists a sufficiently large K , such that the difference between the probabilities $\mathbb{P}(R_i \geq \frac{k_i}{2} | K_i = k_i)$ and $Q(\sqrt{k_i}A(N))$ is smaller than $\epsilon > 0$. Then, we have

$$|\hat{\mathbb{P}}_{T,2}(N) - \mathbb{P}_{T,app}(N)| \leq \sum_{k_i=1}^K \left| \mathbb{P}(R_i \geq \frac{k_i}{2} | K_i = k_i) - Q(\sqrt{k_i}A(N)) \right| \mathbb{P}(K_i = k_i) + \epsilon \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{K+1},$$

where $\mathbb{P}(K_i = k_i) = \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{k_i} \frac{\lambda_e}{\lambda + \lambda_e}$, for $k_i = 0, \dots, \infty$. The above expression can be further upper-bounded by

$$|\hat{\mathbb{P}}_{T,2}(N) - \mathbb{P}_{T,app}(N)| \leq 1 - \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{K+1} + \epsilon \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{K+1}.$$

As the term $1 - \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{K+1}$ can be made smaller than ϵ by choosing $\lambda > \frac{\lambda_e(1-\epsilon)^{1/(K+1)}}{1-(1-\epsilon)^{1/(K+1)}}$, the difference between $\hat{\mathbb{P}}_{T,2}(N)$ and $\mathbb{P}_{T,app}(N)$ can be smaller than 2ϵ for every $\epsilon > 0$ by choosing sufficiently large λ . \square

Next, we show that $\hat{\mathbb{P}}_{T,2}(N)$ can be approximated by the summation of Q -functions when λ is sufficiently large. In the following proposition, we show that as $\lambda \rightarrow \infty$ the probability $\hat{\mathbb{P}}_{T,2}(N)$ converges to a step function.

Proposition 2. As $\lambda \rightarrow \infty$ the probability $\hat{\mathbb{P}}_{T,2}(N)$ converges to a step function given by

$$\lim_{\lambda \rightarrow \infty} \hat{\mathbb{P}}_{T,2}(N) \approx \begin{cases} 0, & \text{when } \frac{N+m}{n} < \frac{1}{2}, \\ \frac{1}{2}, & \text{when } \frac{N+m}{n} = \frac{1}{2}, \\ 1, & \text{when } \frac{N+m}{n} > \frac{1}{2}. \end{cases} \tag{17}$$

Proof. First, we consider the case when $\frac{N+m}{n} < \frac{1}{2}$. In this case, we note that $Q(\sqrt{k_i}A(N))$ is a decreasing function of k_i . Thus, for any arbitrary $\epsilon_1 > 0$, there exists an L , such that $Q(\sqrt{k_i}A(N)) < \epsilon_1, \forall k_i > L$. Therefore, we have

$$\hat{\mathbb{P}}_{T,2}(N) < \sum_{k_i=1}^L Q\left(\sqrt{k_i}A(N)\right)\mathbb{P}(K_i = k_i) + \epsilon_1 \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{L+1}.$$

As $Q(\sqrt{k_i}A(N)) < \frac{1}{2}$ for $k_i \geq 1$, as in the proof of Proposition 1, by choosing sufficiently large λ , one can show that $\hat{\mathbb{P}}_{T,2}(N) < 2\epsilon_1$. Thus, if $\frac{N+m}{n} < \frac{1}{2}$, we have $\lim_{\lambda \rightarrow \infty} \hat{\mathbb{P}}_{T,2}(N) = 0$.

Next, we consider the case when $\frac{N+m}{n} > \frac{1}{2}$. As $Q(x) = 1 - Q(-x)$ for all x , we have

$$\hat{\mathbb{P}}_{T,2}(N) = \sum_{k_i=1}^{\infty} \left(1 - Q\left(-\sqrt{k_i}A(N)\right) \right) \mathbb{P}(K_i = k_i).$$

Note that $Q(-\sqrt{k_i}A(N))$ is a decreasing function of k_i . Thus, for any $\epsilon_2 > 0$, there exists a large \hat{L} , such that $Q(-\sqrt{k_i}A(N)) < \epsilon_2$. Therefore, we can write

$$\hat{\mathbb{P}}_{T,2}(N) > \frac{\lambda}{\lambda + \lambda_e} - \sum_{k_i=1}^{\hat{L}} Q\left(-\sqrt{k_i}A(N)\right)\mathbb{P}(K_i = k_i) - \epsilon_2 \left(\frac{\lambda}{\lambda + \lambda_e} \right)^{\hat{L}+1}.$$

Now, similar to the first part of the proof, we can show that $\hat{\mathbb{P}}_{T,2}(N) > \frac{\lambda}{\lambda + \lambda_e} - 2\epsilon_2$ by selecting a sufficiently large λ . Thus, when $\frac{N+m}{n} > \frac{1}{2}$, we have $\lim_{\lambda \rightarrow \infty} \hat{\mathbb{P}}_{T,2}(N) = 1$. Finally, when $\frac{N+m}{n} = \frac{1}{2}$, we note that the $A(N)$ terms in (16) become 0, which implies $\hat{\mathbb{P}}_{T,2}(N) \approx \mathbb{P}_{T,app}(N) = \frac{1}{2}$. \square

In Proposition 2, we showed that when the gossip rate λ is sufficiently large, the nodes start to have access to information from all other nodes. As a result, all the nodes in the network collectively start to behave like a single node, where at the end of a gossiping period the information is updated based on the majority of the information at all nodes. In other words, if the majority of the nodes have the same information as the source, which happens if $\frac{N+m}{n} > \frac{1}{2}$, all the nodes update their information to x_s and, thus, they will have the same information as the source at the end of the gossiping period. On the other hand, when the majority of the nodes have the incorrect information $1 - x_s$, which happens if $\frac{N+m}{n} < \frac{1}{2}$, then all the nodes will have the incorrect information at the end of the gossiping period. Therefore, when the information at the source changes frequently (i.e., λ_e is large) and the source has limited total update rate capacity (i.e., λ_s is small), a high gossip rate λ can cause incorrect information to disseminate in the network. As a result, gossiping can be harmful in these scenarios. On the other hand, when the source has high transmission rates, at each update cycle, it is enough for the source to send its information to the number of nodes that achieves the majority, i.e., $\frac{N+m}{n} > \frac{1}{2}$. After that, the remaining nodes can obtain the correct information during the gossiping phase. Thus, when the source has enough transmission rate, high gossip rates among the nodes can be utilized by sending the updates to at most half the network.

Next, we consider the case in which the gossip rate λ is relatively low compared to the rate of information change at the source, λ_e . When the gossip rate is low, the nodes either do not get any updates, in which case they hold on to their prior information, or they mostly get only one update from the other nodes and, hence, update their information based on the only received update. In the following proposition, we approximate the probability $\hat{\mathbb{P}}_{T,2}(N)$ when λ is low.

Proposition 3. *When λ is sufficiently small, the probability $\hat{\mathbb{P}}_{T,2}(N)$ can be approximated by*

$$\mathbb{P}_{T,app}^{low}(N) = \frac{\lambda}{\lambda + \lambda_e} \frac{N + m}{n}. \tag{18}$$

Proof. When λ is sufficiently low, the nodes may not receive any updates or receive a single update packet from the other nodes in the gossiping phase. Thus, the nodes that have the incorrect information $1 - x_s$ as prior information obtain x_s with probability $(1 - \mathbb{P}(K_i = 0))^{\frac{N+m}{n}}$, which is equal to

$$\mathbb{P}_{T,app}^{low}(N) = \frac{\lambda}{\lambda + \lambda_e} \frac{N + m}{n}.$$

Next, we consider the difference between $\hat{\mathbb{P}}_{T,2}(N)$ and $\mathbb{P}_{T,app}^{low}(N)$, which is given by

$$|\hat{\mathbb{P}}_{T,2}(N) - \mathbb{P}_{T,app}^{low}(N)| \leq \sum_{k_i=2}^{\infty} \mathbb{P}\left(R_i \geq \frac{k_i}{2} | K_i = k_i\right) \mathbb{P}(K_i = k_i) + \left(\frac{N+m}{n}\right) \left(\frac{\lambda}{\lambda + \lambda_e}\right)^2.$$

As $\mathbb{P}(R_i \geq \frac{k_i}{2} | K_i = k_i) \leq 1$, we have

$$|\hat{\mathbb{P}}_{T,2}(N) - \mathbb{P}_{T,app}^{low}(N)| \leq \left(1 + \frac{N+m}{n}\right) \left(\frac{\lambda}{\lambda + \lambda_e}\right)^2. \tag{19}$$

Thus, when the gossip rate λ is sufficiently low compared to λ_e , the upper bound on (19) can be made arbitrarily small, making the approximation $\hat{\mathbb{P}}_{T,2}(N) \approx \mathbb{P}_{T,app}^{low}(N)$ tight. \square

Gossip Gain and an Adaptive Policy for Selecting Transmission Capacity

As a result of gossiping, when λ is low, the nodes that have the correct information x_s as prior information keep their information as x_s with probability $\mathbb{P}_{T,app}^{low}(N) + \mathbb{P}(K_i = 0)$, which is given by $\hat{\mathbb{P}}_{T,1}(N) \approx \frac{\lambda}{\lambda + \lambda_e} \frac{N+m}{n} + \frac{\lambda_e}{\lambda + \lambda_e}$. Thus, when λ is small, the probability $\hat{\mathbb{P}}_{T,1}(N)$ can be equivalently approximated by

$$\hat{\mathbb{P}}_{T,1}(N) \approx 1 - \frac{n - N - m}{n} \frac{\lambda}{\lambda + \lambda_e}. \tag{20}$$

Therefore, when the gossip rate is low, we have

$$\begin{aligned} \mathbb{E}[N'_1|N] &= N \hat{\mathbb{P}}_{T,1}(N) = N - \frac{\lambda}{\lambda + \lambda_e} \frac{N}{n} (n - N - m), \\ \mathbb{E}[N'_2|N] &= (n - N - m) \hat{\mathbb{P}}_{T,2}(N) = \frac{\lambda}{\lambda + \lambda_e} \frac{m + N}{n} (n - N - m). \end{aligned}$$

Thus, at the end of the gossiping period, there are $\mathbb{E}[N'_1 + N'_2|N] + m = N + m + \frac{\lambda}{\lambda + \lambda_e} \frac{m}{n} (n - N - m)$ nodes that have the same information as the source x_s . If we consider the system with no gossiping, where only the source can send updates to m nodes, at the end of an update cycle most $N + m$ nodes have the same information as the source. Thus, compared to the system with no gossiping, the gain (error reduction) obtained as a result of gossiping can be computed as

$$G(N) = \frac{m}{n^2} (n - N - m) \left(\frac{\lambda}{\lambda + \lambda_e} \right) \left(\frac{\lambda_s}{\lambda_s + \lambda_e} \right)^m, \tag{21}$$

which is obtained by subtracting $N + m$ from $\mathbb{E}[N'_1 + N'_2|N] + m$ and dividing the result by n due to the definition of Δ . Note that the last term $\left(\frac{\lambda_s}{\lambda_s + \lambda_e} \right)^m$ in (21) is equal to the probability of entering the gossiping phase.

Let us denote the average error for a system with no gossiping (that is, $\lambda = 0$) by Δ_{ng} . If the gossip rate is low, the overall gain obtained from gossiping, $|\Delta - \Delta_{ng}|$, can be approximated by

$$|\Delta - \Delta_{ng}| \approx B(p) \sum_{N=0}^{n-m} (\pi_{0N} + \pi_{1N}) G(N), \tag{22}$$

where $B(p)$ is a scaling function, in terms of p , to represent the effect of gossiping on the steady-state distribution π .

When the gossip rate among the nodes is low, the gossip gain $G(N)$ in (21) depends on the selection of m . Therefore, if the source is allowed to dynamically choose its transmission capacity m in terms of N , a natural choice is to adaptively select an m which maximizes the gossiping gain by solving $\frac{\partial G(N)}{\partial m} = 0$. Solving this equation in terms of m gives us

$$m^*(N) = \frac{n - N}{2} - \frac{1}{\log \rho_s} - \sqrt{\left(\frac{n - N}{2} \right)^2 + \left(\frac{1}{\log \rho_s} \right)^2}, \tag{23}$$

where $\rho_s = \frac{\lambda_s}{\lambda_s + \lambda_e}$. (We note that $\frac{\partial G(N)}{\partial m} = 0$ has two solutions. The other solution is equal to $m^*(N)$ in (23) except that the square-root term has a positive sign. One can show that this root is always larger than $n - N$ and, thus, cannot be a feasible selection for m). In fact, it is easy to see from (23) that the optimal solution $m^*(N)$ always lies in the range $0 \leq m^*(N) \leq \frac{n - N}{2}$.

When the source has infinite transmission capacity, we have $\lim_{\lambda_s \rightarrow \infty} m^*(N) = \frac{n - N}{2}$, which suggests that the source should send its information to at most half of the nodes that carry incorrect information. In the other extreme case, when the source's transmission

capacity is equal to 0, we have $\lim_{\lambda_s \rightarrow 0} m^*(N) = 0$, in which case the source should not send its information to any other nodes. In general, for a given λ_s , $m^*(N)$ in (23) is a decreasing function of N , which means that when N is small, i.e., when most of the nodes have incorrect information, the source should send updates to a higher number of nodes. As N increases, the source should send updates to a smaller number of nodes, as most nodes carry the same information as the source. In the following section, in order to reduce the average error, we implement clustering in the gossip network.

5. Average Error in Clustered Networks

In this section, we explore the idea of clustering in the gossip network, in order to further reduce the average error. As illustrated in Figure 2, the gossip network is partitioned into m_s clusters with equal cluster size n_c , i.e., without loss of generality, we assume that n is divisible by m_s and, thus, we have $n_c = \frac{n}{m_s}$. Each cluster has a designated cluster head that is connected to the source directly. In the clustered network, when the information at the source is updated, instead of sending updates to individual nodes directly, the source only sends its information to cluster heads that carry different information compared to the source. Thus, the source can send its information to a smaller number of cluster heads with higher update rates. Furthermore, as the cluster heads behave like an information source of each cluster, by using cluster heads, we can increase the total update rates going from the source to the individual nodes while decreasing the total number of connections at the gossip network. The downside of clustering is that if the information is updated during transmission from the source to the cluster heads, information may not be disseminated to individual nodes in the gossip network. Thus, we need to choose the number of cluster heads m_s optimally, to minimize the average error.

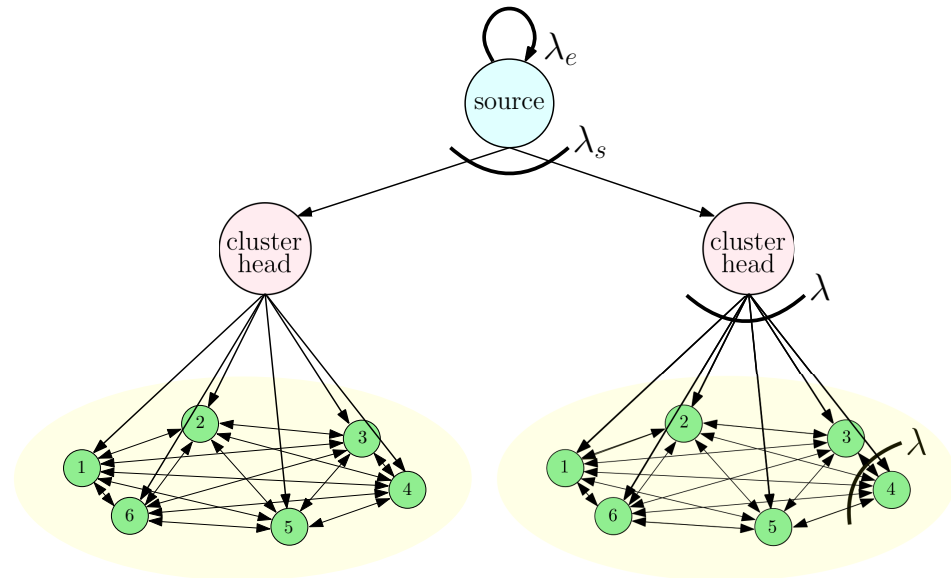


Figure 2. A clustered gossip network that consists of a source and $m_s = 2$ cluster heads and fully connected $n_c = 6$ nodes.

At the update cycle $I(j)$, let the number of cluster heads that have the same information as the source be $N_s(j)$, where $0 \leq N_s(j) \leq m_s$. Before the information at the source is updated again, the source sends updates to $K_{sc}(j)$ cluster heads with the following probability distribution:

$$\mathbb{P}(K_{sc} = k_{sc} | 0 \leq N_s \leq m_s) = \begin{cases} \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^{k_{sc}} \frac{\lambda_e}{\lambda_s + \lambda_e}, & \text{if } k_{sc} = 0, \dots, m_s - N_s - 1, \\ \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^{m_s - N_s}, & \text{if } k_{sc} = m_s - N_s. \end{cases} \quad (24)$$

If $K_{sc} < m_s - N_s$, it means that the information at the source is updated before all the cluster heads obtain the information at the source. In this case, a new update cycle begins and the source starts sending information to the cluster heads again. If all the cluster heads obtain the most current information at the source, then the cluster heads start sending their information to m_c nodes within their corresponding cluster that carries different information compared to the source with a total update rate of λ . (In this section, we introduce the cluster heads as special nodes of the network, as in [10]. However, using these special nodes may not always be possible, as they may result in additional costs for the system. Considering these factors, we take the total update rate of the cluster heads as λ , which is the same as that of the regular nodes at the clusters, which means that in the absence of these special nodes, some of the nodes in the gossip network can be used as cluster heads. As all the clusters in the network are identical, we focus on a typical cluster and obtain the average error of a node within that cluster). In a typical update cycle I , we define N_c as the number of nodes that carry the same information at the source. When $N_c < n_c - m_c$, a cluster head sends updates to K_c nodes with the following pmf:

$$\mathbb{P}(K_c = k_c | N_c < n_c - m_c) = \begin{cases} \left(\frac{\lambda}{\lambda + \lambda_e}\right)^{k_c} \frac{\lambda_e}{\lambda + \lambda_e}, & \text{if } k_c = 0, \dots, m_c - 1, \\ \left(\frac{\lambda}{\lambda + \lambda_e}\right)^{m_c}, & \text{if } k_c = m_c. \end{cases} \quad (25)$$

If $N_c \geq n_c - m_c$, then we have

$$\mathbb{P}(K_c = k_c | N_c \geq n_c - m_c) = \begin{cases} \left(\frac{\lambda}{\lambda + \lambda_e}\right)^{k_c} \frac{\lambda_e}{\lambda + \lambda_e}, & \text{if } k_c = 0, \dots, n_c - N_c - 1, \\ \left(\frac{\lambda}{\lambda + \lambda_e}\right)^{n_c - N_c}, & \text{if } k_c = n_c - N_c. \end{cases} \quad (26)$$

When $N_c < n_c - m_c$, if the cluster head sends updates to $K_c = m_c$ nodes, then the gossiping phase starts and all the nodes within the same cluster share their local information with one another. When the information at the source is updated, the gossiping phase ends, and the nodes that do not get information directly from the cluster head update their information based on the majority of the updates that they received during the gossiping phase.

In the next lemma, for a given state N_s , we provide the expression for the long-term average error at the nodes within the clusters.

Lemma 1. Under the proposed clustered network structure, for a given N_s , the long-term average error at the nodes within the clusters, denoted by $\Delta_{c|N_s}$, is given by

$$\Delta_{c|N_s} = \sum_{j=0}^{n_c} \sum_{n''=0}^{n_c} (\pi_{0,j|N_s} + \pi_{1,j|N_s}) \mathbb{P}(N_c'' = n_c'' | N_c = j, N_s) \frac{n_c - n_c''}{n_c}, \quad (27)$$

where $\mathbb{P}(N_c'' = n_c'' | N_c = j, N_s)$ is later provided in (30) and $\boldsymbol{\pi}_{N_s} = [\pi_{0,0|N_s}, \dots, \pi_{0,n_c|N_s}, \pi_{1,0|N_s}, \dots, \pi_{1,n_c|N_s}]$ is the row vector of the steady-state distribution of the Markov chain formed over the state space (x_s, N_c) . The unique stationary distribution is given by the solution of $\boldsymbol{\pi}_{N_s} = \boldsymbol{\pi}_{N_s} \mathbf{P}_{N_s}$ for a stochastic matrix $\mathbf{P}_{N_s} \in \mathbb{R}^{2(n_c+1) \times 2(n_c+1)}$, where the transition probabilities of \mathbf{P}_{N_s} can be derived by replacing N'' , N , and n in (14) by N_c'' , N_c , and n_c , respectively.

Proof. For a given N_s , the average error analysis with a clustered gossip network similarly follows from Section 3. During the gossiping phase, node i receives K_i updates with the pmf in (6). Similar to (7) and (8), we can rewrite $\mathbb{P}(R_i = r | K_i = k_i, x_i = x_s)$ and $\mathbb{P}(R_i = r | K_i = k_i, x_i \neq x_s)$ by replacing N and m with N_c and m_c , respectively. Then, we can define $\mathbb{P}_{T,1}(N_c)$ and $\mathbb{P}_{T,2}(N_c)$ as in (9) and (10), respectively. Before starting the gossiping period, N_c nodes have x_s and $n_c - N_c - m_c$ nodes have $1 - x_s$ as their prior information. At the end of the gossiping period, we have $N_c' = N_{c,1}' + N_{c,2}'$ nodes that have the information x_s , where $N_{c,1}' \sim \text{Bin}(N_c, P_{T,1}(N_c))$ and $N_{c,2}' \sim \text{Bin}(n_c - N_c - m_c, P_{T,2}(N_c))$. The pmf of N_c' follows from (12), with $\ell_{lower} = \max\{0, n_c' + N_c + m_c - n_c\}$ and $\ell_{upper} = \min\{N_c, n_c'\}$.

Next, we define N_c'' to be the number of nodes in a cluster that have the same information as the source at the end of an update cycle. For a given K_{sc} , N_s , and N_c , similar to (13), the probability distribution of N_c'' is given by

$$\mathbb{P}(N_c'' = n_c'' | N_c, N_s, K_{sc} = m_s - N_s) = \begin{cases} \mathbb{P}(K_c = k_c | N_c < n_c - m_c), \\ \quad \text{if } n_c'' = k_c + N_c < m_c, \\ \mathbb{P}(K_c = m_c | N_c < n_c - m_c) \mathbb{P}(N_c' = n_c'), \\ \quad \text{if } m_c \leq n_c'' = m_c + n_c' < N_c, \\ \mathbb{P}(K_c = n_c'' - N_c | N_c < n_c - m_c) \\ \quad + \mathbb{P}(K_c = m_c | N_c < n_c - m_c) \mathbb{P}(N_c' = n_c'' - m_c), \\ \quad \text{if } m_c \leq N_c \leq n_c'' < N_c + m_c, \\ \mathbb{P}(K_c = m_c | N_c < n_c - m_c) \mathbb{P}(N_c' = n_c'' - m_c), \\ \quad \text{if } N_c + m_c \leq n_c'' \leq n_c, \\ \mathbb{P}(K_c = n_c'' - N_c | N_c \geq n_c - m_c), \\ \quad \text{if } n_c - m_c \leq N_c \leq n_c'' \leq n_c. \end{cases} \quad (28)$$

When $K_{sc} < m_s - N_s$, we have

$$\mathbb{P}(N_c'' = n_c'' | N_c, N_s, K_{sc} < m_s - N_s) = \begin{cases} 1, & \text{if } n_c'' = N_c, \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

From (28) and (29), we obtain

$$\mathbb{P}(N_c'' = n_c'' | N_c, N_s) = \mathbb{P}(N_c'' = n_c'' | N_c, N_s, K_{sc} = m_s - N_s) \mathbb{P}(K_{sc} = m_s - N_s | N_s) + \mathbb{P}(N_c'' = n_c'' | N_c, N_s, K_{sc} < m_s - N_s) \mathbb{P}(K_{sc} < m_s - N_s | N_s). \quad (30)$$

Finally, for a given N_s , the states (x_s, N_c) form a Markov chain. Similar to the average error analysis in Section 3, we label the first $n_c + 1$ states $(0, 0), (0, 1), \dots, (0, n_c)$ by $1, 2, \dots, n_c + 1$, and the last $n_c + 1$ states $(1, 0), (1, 1), \dots, (1, n_c)$ by $n_c + 2, n_c + 3, \dots, 2n_c + 2$. The stochastic matrix P_{N_s} consists of $P_{a,b}^c(N_s)$, which denotes the probability of moving from state a to state b , and can be derived by replacing N'' , N , and n in (14) with N_c'' , N_c , and n_c , respectively. Then, we arrive at a unique stationary distribution $\pi_{N_s} = \pi_{N_s} P_{N_s}$ that satisfies $\sum_{i=0}^1 \sum_{j=0}^{n_c} \pi_{i,j|N_s} = 1$ and $\pi_{i,j|N_s} \geq 0 \forall i, j$. Thus, for a given N_s , we can characterize the long-term average error among all the nodes within the same cluster by (27). \square

In the following theorem, we state the long-term average error of the nodes at the clusters and at the cluster heads.

Theorem 2. Under the proposed clustered network structure, the long-term average error of the nodes at the clusters, denoted by $\Delta_c = \mathbb{E}[\Delta_{c|N_s}]$, is given by

$$\Delta_c = \sum_{N_s=0}^{m_s} (\pi_{0,N_s}^s + \pi_{1,N_s}^s) \Delta_{c|N_s}, \quad (31)$$

where $\pi_s = [\pi_{0,0}^s, \dots, \pi_{0,m_s}^s, \pi_{1,0}^s, \dots, \pi_{1,m_s}^s]$ is the row vector of the steady-state distribution of the Markov chain with the state space (x_s, N_s) . The unique stationary distribution is given by the solution of $\pi_s = \pi_s P_s$ for a stochastic matrix $P_s \in \mathbb{R}^{2(m_s+1) \times 2(m_s+1)}$, where the transition probabilities of P_s are given later in (34). Furthermore, the average error of the nodes at the cluster heads, denoted by Δ_s , is given by

$$\Delta_s = \sum_{j=0}^{m_s} \sum_{n_s''=0}^{m_s} (\pi_{0,j}^s + \pi_{1,j}^s) \mathbb{P}(N_s'' = n_s'' | N_s = j) \frac{m_s - n_s''}{m_s}, \quad (32)$$

where $\mathbb{P}(N_s'' = n_s'' | N_s = j)$ is provided in (33).

Proof. In order to obtain the long-term average error $\Delta_c = \mathbb{E}[\Delta_c | N_s]$, we need to find the probability distribution for N_s . For that, we note that in the clustered network, the information at the source and the number of cluster heads that have the same information at the source, i.e., (x_s, N_s) , also form a Markov chain. During the source’s update transmission to the cluster heads, by using (24) we write the probability distribution for transition to state N_s'' from state N_s , as follows:

$$\mathbb{P}(N_s'' = n_s'' | N_s) = \begin{cases} \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^{n_s'' - N_s} \left(\frac{\lambda_e}{\lambda_s + \lambda_e}\right), & \text{if } N_s \leq n_s'' < m_s, \\ \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^{m_s - N_s}, & \text{if } n_s'' = m_s. \end{cases} \quad (33)$$

The Markov chain formed by (x_s, N_s) has the states $(0, 0), \dots, (0, m_s), (1, 0), \dots, (1, m_s)$ where we label these states from 1 to $2(m_s + 1)$, correspondingly. Then, the stochastic matrix P_s , consisting of $P_{a,b}^s$, which denotes the probability of moving from state a to state b and is given by (34). Then, we can arrive at the unique stationary distribution $\pi_s = \pi_s P_s$ that satisfies $\sum_{i=0}^1 \sum_{j=0}^{m_s} \pi_{i,j}^s = 1$, and $\pi_{i,j}^s \geq 0, \forall i, j$. Finally, by using $\Delta_c = \mathbb{E}[\Delta_c | N_s]$, we obtain the average error of a node in a cluster in (31).

Similarly, the average error at the cluster heads Δ_s can be obtained by using (32) with the stationary distribution π_s and $\mathbb{P}(N_s'' = n_s'' | N_s)$ in (33). □

$$P_{a,b}^s = \begin{cases} (1-p)\mathbb{P}(N_s'' = b-1 | N_s = a-1), & \text{if } 1 \leq a \leq b \leq m_s + 1, \\ \frac{p}{1-p} P_{a,2m_s+3-b}, & \text{if } 1 \leq a, b - m_s - 1 \leq m_s + 1, \\ \frac{p}{1-p} P_{a,2m_s+3-b}, & \text{if } 1 \leq a - m_s - 1, b \leq m_s + 1, \\ (1-p)\mathbb{P}(N_s'' = b - m_s - 2 | N_s = a - m_s - 2), & \text{if } m_s + 2 \leq a, b \leq 2(m_s + 1). \end{cases} \quad (34)$$

In general, the clustered networks can model a system where not all the nodes have access to the source directly. In a way, cluster heads constitute a small group of nodes that have the privilege of accessing the information source directly. These nodes can be considered as paid subscribers to the source, while regular nodes can have free access to the information through these paid subscribers and gossiping. Thus, looking at the average difference between the errors at the cluster heads, Δ_s , and at the regular nodes, Δ_c , tells us how much a regular node can increase its quality of information through subscription. We can also imagine the clustered gossip networks in a way such that if every node is connected (subscribed) to the source directly, the information quality at the individual nodes may decrease due to the limited update capacity of the source. Instead, these nodes may choose some nodes as subscribers and share the cost of the subscription. As a result, through clustering, the nodes can decrease the cost of accessing the information while increasing the overall quality of their information.

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

6. Numerical Results

This section has three subsections: in the first one, we discuss the numerical results of 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 second one, we provide simulation results to corroborate the analytical results in Section 4; in the third one, simulations illustrate the results of Section 5—that is, the effects of clustering on information dissemination.

6.1. Simulations for the Effects of Various System Parameters on Information Dissemination

In the numerical results provided in this subsection, we provide real-time simulations over 200,000 update cycles, and we provide the sample average errors with the markers in Figures 3 and 4. In the first numerical study, we took $p = 0.4$, $\lambda_e = 1$, $\lambda_s = 10$, and $n = 60$. We found the average error Δ with respect to m when $\lambda = \{0, 10, 20\}$. Note that $\lambda = 0$ corresponded to the case of no gossiping among the nodes. We see in Figure 3a that when m was small, i.e., when the source could send updates to a small number of nodes, the average error Δ increased with gossip rate λ . As m was small and the information change rate $p = 0.4$ was high, incorrect information disseminated, due to gossiping in the network. As a result, the system with no gossiping ($\lambda = 0$) achieved the lowest average error. When we increased m sufficiently, the nodes started to have access to the same information as the source, and gossiping helped to disseminate the correct information. That is why the systems with gossiping—i.e., $\lambda = 10, 20$ —achieved lower average error compared to the system with no gossiping. The lowest average error Δ was achieved when $m = 25$ for $\lambda = 10, 20$ and $m = 55$ for $\lambda = 0$. Here, we also note that the average error Δ was 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 decreased the probability of entering the gossiping phase, and that is why all the curves in Figure 3a overlap when $m \geq 40$.

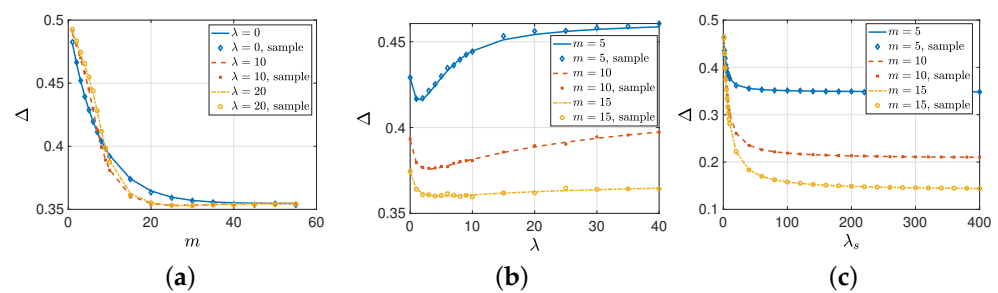


Figure 3. The average error Δ with respect to (a) m when $\lambda \in \{0, 10, 20\}$, (b) the gossip rate λ for $m \in \{5, 10, 15\}$, and (c) the source’s update rate λ_s for $m = \{5, 10, 15\}$.

In the second numerical study, we considered the same variable selections as in the previous example except that we took $m = \{5, 10, 15\}$ and changed λ from 0 to 40. We see in Figure 3b that increasing the gossip rate λ initially helped to reduce the average error Δ . Then, increasing λ further increased Δ as the incorrect information among the nodes became more available. We see in Figure 3b that the minimum average error was 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 sent updates to more nodes, the optimal gossip rate increased.

In the third numerical study, we considered $p = 0.2$, $\lambda_e = 1$, $\lambda = 5$, and $n = 60$. We increased λ_s from 1 to 400 for $m = \{5, 10, 15\}$. We see in Figure 3c that increasing λ_s initially decreased the average error Δ faster. However, as Δ depended also on the other parameters, such as m and the gossip rate λ , increasing λ_s further did not improve the average error Δ and it converged to 0.348 for $m = 5$, 0.21 for $m = 10$, and 0.144 for $m = 15$.

In the fourth numerical study, we considered the effect of the network size n on the information dissemination. For that, first, we took $p = 0.2$, $\lambda_e = 1$, $\lambda = 10$, $m = 8$, and $n = \{10, 20, \dots, 150\}$, and we increased $\lambda_s = \{0.1n, 0.2n, 0.5n\}$ with the network size n . In this case, as the network size increased, the source’s transmission rate also increased. However, we kept the total number of nodes that the source could send updated to the same, i.e., $m = 8$ for all n . In Figure 4a, when $\lambda_s = \{0.1n, 0.2n\}$, we see that the average error Δ initially decreased with n , as λ_s was initially a primary limiting factor. Increasing n further increased Δ as m became more important. That is why all these three curves overlap each other when λ_s is sufficiently large. Then, we considered a scenario where we kept $\lambda_s = 4$ and only increased $m = \{0.1n, 0.2n, 0.5n\}$. In Figure 4b, increasing the

maximum number of nodes that the source could send updates to in an update cycle alone did not reduce Δ as n increased. As we increased n , λ_s became the presiding factor, and all the curves in Figure 4b overlap. Finally, we increased 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 Figure 4c, we observe that we could achieve a constant Δ by increasing λ_s and m proportional to n .

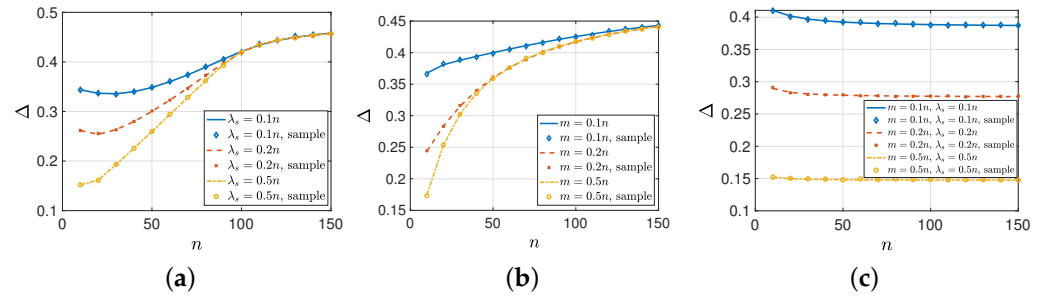


Figure 4. Average error Δ with respect to n (a) when $\lambda_s \in \{0.1n, 0.2n, 0.5n\}$, (b) when $m \in \{0.1n, 0.2n, 0.5n\}$, and (c) when $m \in \{0.1n, 0.2n, 0.5n\}$ and $\lambda_s \in \{0.1n, 0.2n, 0.5n\}$.

6.2. Simulations for High and Low Gossiping Rates

In this subsection, we provide numerical results for the analysis developed for high and low gossip rates in Section 4. Here, we also ran real-time simulations over 10,000,000 update cycles. As $m = 20$, $\lambda_s = 2$, and $\lambda_e = 1$, out of 10,000,000 update cycles, approximately in $10,000,000 \times \left(\frac{\lambda_s}{\lambda_s + \lambda_e}\right)^m \sim 3000$ update cycles, the system entered the gossiping phase. As $P_{T,1}(N)$ and $P_{T,2}(N)$ were the probabilities of individuals that were able to obtain the source’s information as a result of gossiping, the sample averages of $P_{T,1}(N)$ and $P_{T,2}(N)$ were obtained approximately over 3000 update cycles, where the system entered the gossiping phase. In the first numerical study, we verified the analytical results in Propositions 1 and 2. For this simulation, we numerically evaluated $\mathbb{P}_{T,2}(N)$ when $n = 200$, $m = 20$, $\lambda_s = 2$, $\lambda_e = 1$, $p = 0.2$ for $\lambda = \{20, 200, 400\}$. Then, we compared $\mathbb{P}_{T,2}(N)$ to $\mathbb{P}_{T,app}(N)$. In Figure 5, we observe that when λ was high compared to λ_e , $\mathbb{P}_{T,2}(N)$ could be approximated well by $\mathbb{P}_{T,app}(N)$, which was given by the summation of Q -functions in (16). Furthermore, due to Proposition 2, as we increased λ from 20 to 400, $\mathbb{P}_{T,app}(N)$ and, thus, $\mathbb{P}_{T,2}(N)$ converged to a step function, i.e., when $N < \frac{n}{2} - m = 80$, we observed that $\mathbb{P}_{T,2}(N)$ converged to 0, and when $N > \frac{n}{2} - m = 80$, $\mathbb{P}_{T,2}(N)$ converged to 1 while we had $\mathbb{P}_{T,2}(80) = 0.5$.

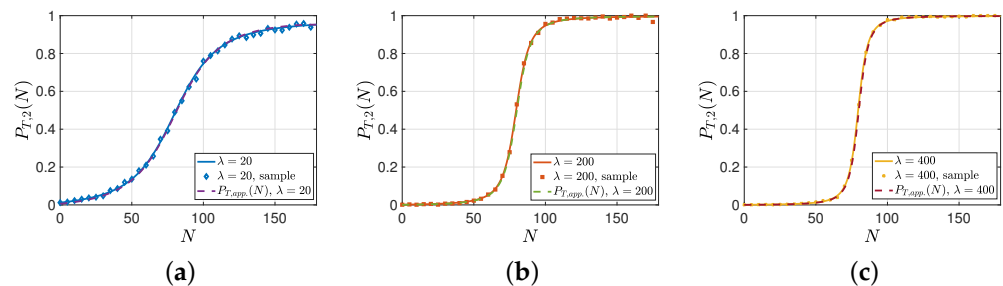


Figure 5. A sample evolution of $\mathbb{P}_{T,2}(N)$, which is approximated by $\mathbb{P}_{T,app}(N)$ in (16) when λ is high compared to λ_e for (a) $\lambda = 20$, (b) $\lambda = 200$, and (c) $\lambda = 400$.

In the remaining numerical studies, we considered the case when the gossip rate λ was low compared to λ_e . In the second simulation, we evaluated $\mathbb{P}_{T,1}(N)$ and $\mathbb{P}_{T,2}(N)$ with the same parameters except for $\lambda = \{0.1, 0.5, 1\}$. We have shown in Proposition 3 that when λ is low compared to λ_e , $\mathbb{P}_{T,2}(N)$ can be approximated by $\mathbb{P}_{T,app}^{low}(N)$ in (18). We see in Figure 6b that when $\lambda = 0.1$ and $\lambda = 0.5$, $\mathbb{P}_{T,2}(N)$ matched closely to $\mathbb{P}_{T,app}^{low}(N)$ in

(18). When $\lambda = \lambda_e = 1$, $\mathbb{P}_{T,2}(N)$ could still be approximated well by $\mathbb{P}_{T,app}^{low}(N)$, but their differences started to be noticeable. Similarly, for the low gossiping rate, we see in Figure 6a that the approximation for $\mathbb{P}_{T,1}(N)$ given in (20) was close when $\lambda = \{0.1, 0.5\}$. When the gossip rate λ was low, during the gossiping phase, the nodes either did not receive any updates, in which case they held on to their previous beliefs, or only got one update. That is why in Figure 6b, when N was low, $\mathbb{P}_{T,2}(N)$, which was the probability of having the correct information as a result of gossiping for a node that had incorrect prior information, was close to 0, and then it increased with N .

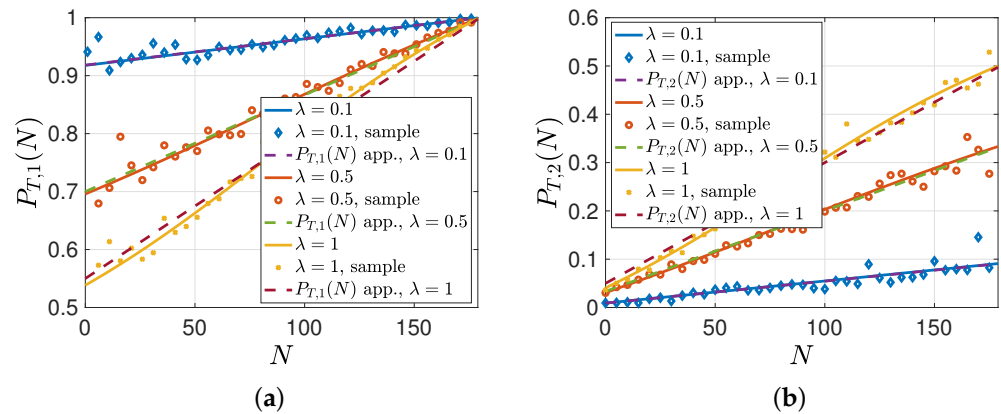


Figure 6. A sample evolution of (a) $\mathbb{P}_{T,1}(N)$ and (b) $\mathbb{P}_{T,2}(N)$ approximated by (20) and (18), respectively, when the gossiping rate is low.

In the third simulation study, when the gossip rate was low, we numerically found the gossip gain (22), which was the difference between the average error with no gossiping Δ_{ng} and the average error with gossiping Δ . For this example, we took $n = 80$, $\lambda = 0.4$, $\lambda_s = 10$, $\lambda_e = 1$, and $p = \{0.3, 0.5, 0.7\}$. We plotted $|\Delta - \Delta_{ng}|$ with respect to m in Figure 7a. We observed in Figure 7a that for all values of p , the gossip gain initially increased with m as the source sent correct information to a sufficient number of nodes. Then, increasing m further decreased the gossip gain as the probability of entering the gossiping phase decreased in an update cycle. We observe in Figure 7a that the optimum gain was obtained when $m = 8$ for all p values. We note that the scaling term $B(p)$ in (22) was equal to 1.7, 1.1, and 0.8 for $p = 0.2$, $p = 0.5$, and $p = 0.7$, respectively. We also note that $G(N)$ in (21) decreased N in the next update cycle with probability p and increased N with probability $1 - p$. Thus, the term $B(p)$ in (22), which was the amplitude of the gossip gain, decreased with p .

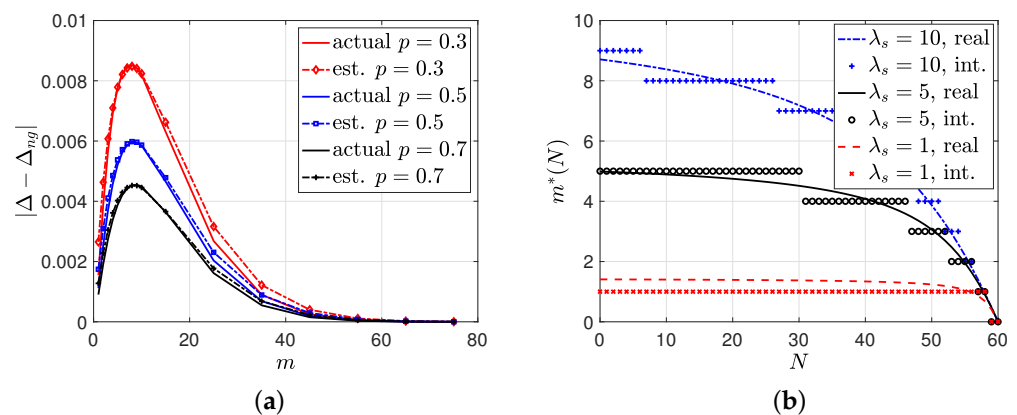


Figure 7. (a) The gossip gain $|\Delta - \Delta_{ng}|$ in (22) with regard to m for $p = \{0.3, 0.5, 0.7\}$. (b) A sample evolution of $m^*(N)$ in (23) and its rounding to the nearest integer for different values of λ_s .

Based on $G(N)$ in (21), we can find the optimal m that maximizes the gossip gain $G(N)$ for each N , which is provided as $m^*(N)$ in (23). So far, in this work, we have only considered the case where m is kept constant for all update cycles. However, $m^*(N)$ in (23) decreases with N , which suggests a policy that selects m adaptively, depending on N . In the next simulation result, we took $n = 60, p = 0.2, \lambda = 10, \lambda_e = 1$, and $\lambda_s = \{1, 5, 10\}$. In Figure 7b, we plotted $m^*(N)$ and their corresponding rounding to the nearest integer. We see in Figure 7b that the source sent updates to more nodes as λ_s increased.

In the last simulation study, we compared the performances of the proposed adaptive policy and the constant policy for selecting m . We considered $n = 60, p = 0.2, \lambda = \{0, 1, 5\}, \lambda_e = 1$, and varied λ_s from 1 to 200. We first implemented the adaptive- m transmission policy by using the nearest integer rounding of $m^*(N)$ in (23), which was denoted by $\bar{m}^*(N)$. We then found the stationary distribution π and calculated the average m , using $\mathbb{E}[\bar{m}^*] = \sum_{j=0}^n (\pi_{0j} + \pi_{1j}) \bar{m}^*(N)$, which is depicted in Figure 8b. In order to make a fair comparison, we took the nearest integer rounding of $\mathbb{E}[\bar{m}^*]$, which is shown with the dashed lines in Figure 8b, and implemented the constant m transmission policy. We see in Figure 8a that the adaptive m policy (even without gossiping) achieved significantly lower average error Δ compared to the constant m policy. In Figure 8a, we also observe that as the gossiping took place, especially when nodes had the correct information, the average error Δ decreased with the gossip rate λ . In the adaptive m selection policy, we see in Figure 8b that increasing gossip rate λ not only achieved lower Δ but also decreased the source's transmission capacity $\mathbb{E}[\bar{m}^*]$. Even though we found this policy for low gossip rates ($\lambda < \lambda_e$), we observed that it was an effective transmission policy even for the higher values of λ and could achieve lower Δ compared to the constant m policy.

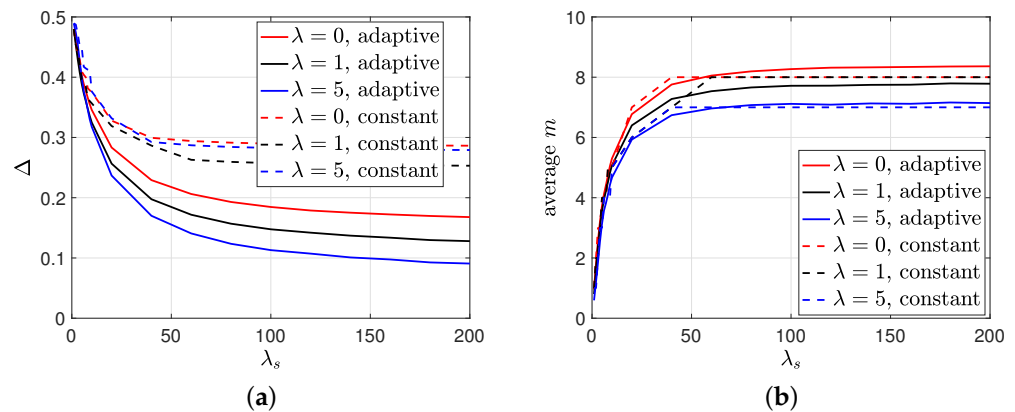


Figure 8. The comparison between (a) the average error Δ and (b) the average m for the adaptive m and constant m selection policies.

6.3. Simulations for the Clustered Networks

In this subsection, we provide the results of simulations that illustrate the effects of clustering on information dissemination. In the first numerical study, we chose $\lambda = 10, \lambda_s = 10, \lambda_e = 1, p = 0.4$, and $n = 120$. We took $m_c = 5$ and considered all m_s values that could divide n . In Figure 9, we plotted the long-term average error at the clusters, Δ_c , and at the cluster heads, Δ_s . We see that increasing the number of cluster heads initially helped to reduce Δ_c as the update rates from the cluster heads to the nodes increased. We see in Figure 9 that the minimum Δ_c was achieved when $m_s = 15$. Increasing m_s further increased Δ_c , as the average error at the cluster heads Δ_s became large.

In the second numerical study, we compared the performances of the gossiping networks with and without clustering when the source's transmission capacity m had an upper limit $m_{lim} = 12$. For this numerical study, we took the same set of variables as in the first numerical study, but we increased $n = 12, 24, \dots, 96$. For each n , we found the optimum m for the network model without clustering and the optimum m_s for the clustered

network that minimized the average error at the nodes. We plotted the minimum average error values in Figure 10a and the optimum m and m_s selections in Figure 10b. We see in Figure 10a that the average error with clustering, Δ_c , was smaller than the average error without clustering, Δ , for all values of n , although the source used its maximum capacity $m = m_{lim}$ for $n \geq 24$ in the network model without clustering, as shown in Figure 10b. For the clustered network model, the optimal number of cluster heads mostly increased with n and reached m_{lim} for $n \geq 84$.

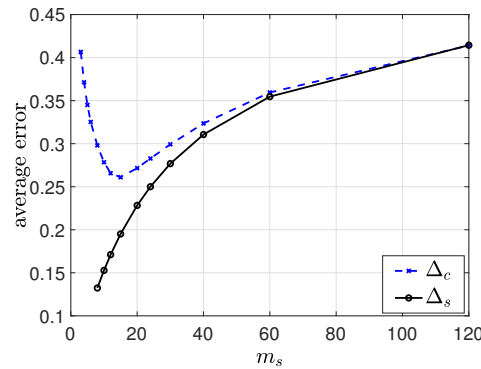


Figure 9. The long-term average error at the clusters, Δ_c , and at the cluster heads, Δ_s , as we increase the number of clusters m_s .

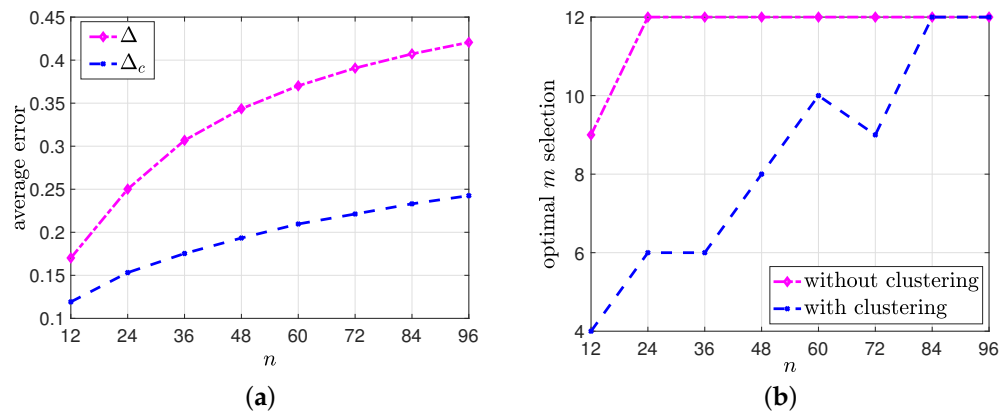


Figure 10. The comparison between (a) the average errors Δ and Δ_c and (b) the optimum m selections for the network models w and w/o clustering.

7. Conclusions 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 common 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 the nodes. Then, we provided analytical results for the high and low gossip rates. As information became available among the nodes in the high gossip rates, all the nodes behaved like a single node. In the low gossip case, we analyzed the gossip gain, which was the error reduction compared to the system with no gossiping, and we obtained $m^*(N)$, which maximized the gain. This suggests an adaptive m selection policy using $m^*(N)$, where the source sends updates to more nodes if most of them have incorrect prior information. Finally, we implemented a clustered gossiping network model and characterized the average errors at the cluster heads and at the nodes in the clusters.

We would like to note that, in this paper, information change probability p and update rate λ_e are taken as given exogenous parameters to the source. As time passes, the source generates updates based on Poisson ticking with rate λ_e among which information is reverted with rate $p\lambda_e$. Let us assume that $p\lambda_e$ is fixed (while $0 < p < 0.5$) and m is

constant. As $p\lambda_e$ is constant, information change rate over time does not change as we vary λ_e . When λ_e is low (and, thus, p is relatively large), then information at the source is flipped more frequently and the update cycle duration gets longer (as a result, the probability of entering the gossiping phase is higher). As the information is flipped more often, the majority of the nodes may have incorrect information. During the gossiping phase, this may increase the average error, as incorrect information may be disseminated further in the network. On the other hand, when λ_e is high while p is low, the information at the source is updated more frequently, but information does not get mutated much. In this case, the probability of entering the gossiping phase decreases and, thus, the system may not benefit from gossiping. Therefore, for a fixed $p\lambda_e$, there should be an optimal p and λ_e selection that minimizes the average error. We leave the optimization problem over λ_e as a future research direction.

As a future direction of research, one could consider the problem where the information at the source can take $k > 2$ different values based on a known pmf. Furthermore, here we have considered only fully connected networks, and extending these results to arbitrarily connected networks could be another interesting direction. One could consider a setting where the source does not have access to the prior information on the nodes, and has to select nodes randomly. In addition to the real-time simulation results, we would like to test our results with the real-world datasets provided in [25]. Finally, one can consider a setting where, although some nodes have the most accurate information, they maliciously send incorrect information to others during the gossiping phase, thus increasing average error.

Author Contributions: Conceptualization, M.B., S.R.E. and T.B.; methodology, M.B., S.R.E. and T.B.; software, M.B.; validation, M.B., S.R.E. and T.B.; formal analysis, M.B., S.R.E. and T.B.; investigation, M.B., S.R.E. and T.B.; resources, M.B., S.R.E. and T.B.; data curation, M.B.; writing—original draft preparation, M.B., S.R.E. and T.B.; writing—review and editing, M.B., S.R.E. and T.B.; visualization, M.B.; supervision, S.R.E. and T.B.; project administration, S.R.E. and T.B.; funding acquisition, S.R.E. and T.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Air Force Office of Scientific Research FA9550-23-1-0107, the NSF CAREER Award EPCN-1944403, and the ARO MURI Grant AG285.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shah, D. *Gossip Algorithms*; Now Publishers Inc.: Norwell, MA, USA, 2009.
2. Mosk-Aoyama, D.; Shah, D. Information dissemination via network coding. In Proceedings of the IEEE ISIT, Seattle, DC, USA, 9–14 July 2006; pp. 1748–1752.
3. Boyd, S.; Ghosh, A.; Prabhakar, B.; Shah, D. Analysis and optimization of randomized gossip algorithms. In Proceedings of the IEEE CDC, Nassau, Bahamas, 14–17 December 2004; pp. 5310–5315.
4. Boyd, S.; Ghosh, A.; Prabhakar, B.; Shah, D. Randomized gossip algorithms. *IEEE Trans. Inf. Theory* **2006**, *52*, 2508–2530. [[CrossRef](#)]
5. Mukund, M.; Sohoni, M. Keeping track of the latest gossip in a distributed system. *Distrib. Comput.* **1997**, *10*, 137–148. [[CrossRef](#)]
6. Israeli, A.; Li, M. Bounded time-stamps. In Proceedings of the 28th Annual Symposium on Foundations of Computer Science, Los Angeles, CA, USA, 12–14 October 1987; pp. 371–382.
7. Sun, Y.; Kadota, I.; Talak, R.; Modiano, E. Age of information: A new metric for information freshness. *Synth. Lect. Commun. Netw.* **2019**, *12*, 1–224.
8. Yates, R.D.; Sun, Y.; Brown, D.R.; Kaul, S.K.; Modiano, E.; Ulukus, S. Age of information: An introduction and survey. *IEEE JSAC* **2021**, *39*, 1183–1210. [[CrossRef](#)]
9. Yates, R.D. The age of gossip in networks. In Proceedings of the IEEE ISIT, Melbourne, Australia, 12–20 July 2021; pp. 2984–2989.
10. Buyukates, B.; Bastopcu, M.; Ulukus, S. Version age of information in clustered gossip networks. *IEEE JSAC* **2022**, *3*, 85–97. [[CrossRef](#)]

11. Bastopcu, M.; Buyukates, B.; Ulukus, S. Gossiping with binary freshness metric. In Proceedings of the IEEE Globecom Workshops, Madrid, Spain, 7–11 December 2021; pp. 1–6.
12. Cho, J.; Garcia-Molina, H. Effective page refresh policies for web crawlers. *Acem Trans. Database Syst.* **2003**, *28*, 390–426. [[CrossRef](#)]
13. Azar, Y.; Horvitz, E.; Lubetzky, E.; Peres, Y.; Shahaf, D. Tractable near-optimal policies for crawling. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 8099–8103. [[CrossRef](#)] [[PubMed](#)]
14. Kolobov, A.; Peres, Y.; Lubetzky, E.; Horvitz, E. Optimal freshness crawl under politeness constraints. In Proceedings of the ACM SIGIR Conference, Paris, France, 21–25 July 2019.
15. Minnaard, W.; de Laat, C.T.A.M.; van Loosen MSc, M. *Timestomping NTFS*; IMSc Final Research Project Report; University of Amsterdam: Amsterdam, The Netherlands, 2014.
16. Kaswan, P.; Ulukus, S. Susceptibility of age of gossip to timestomping. *arXiv* **2022**, arXiv:2205.08510.
17. Zhang, J.; You, K.; Başar, T. Distributed discrete-time optimization in multiagent networks using only sign of relative state. *IEEE Trans. Autom. Control.* **2019**, *64*, 2352–2367. [[CrossRef](#)]
18. Cao, X.; Başar, T. Decentralized online convex optimization based on signs of relative states. *Automatica* **2021**, *129*, 109676. [[CrossRef](#)]
19. Etesami, S.R.; Bolouki, S.; Nedić, A.; Başar, T.; Poor, H.V. Influence of conformist and manipulative behaviors on public opinion. *IEEE Trans. Control. Netw. Syst.* **2018**, *6*, 202–214. [[CrossRef](#)]
20. Coban, Ö.; Truică, C.O.; Apostol, E.S. CONTAIN: A Community-based Algorithm for Network Immunization. *arXiv* **2023**, arXiv:2303.01934.
21. Petrescu, A.; Truică, C.O.; Apostol, E.S.; Karras, P. Sparse shield: Social network immunization vs. harmful speech. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management, Virtual Event, 1–5 October 2021; pp. 1426–1436.
22. Truică, C.O.; Apostol, E.S.; Karras, P. DANES: Deep Neural Network Ensemble Architecture for Social and Textual Context-aware Fake News Detection. *arXiv* **2023**, arXiv:2302.01756.
23. Bastopcu, M.; Etesami, S.R.; Başar, T. The dissemination of time-varying information over networked agents with gossiping. In Proceedings of the IEEE ISIT, Espoo, Finland, 26 June–1 July 2022.
24. Bastopcu, M.; Ulukus, S. Information freshness in cache updating systems. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 1861–1874. [[CrossRef](#)]
25. Shu, K.; Mahudeswaran, D.; Wang, S.; Lee, D.; Liu, H. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data* **2020**, *8*, 171–188. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.