On Periodic Gossiping with an Optimal Weight

Oren Mangoubi Shaoshuai Mou Ji Liu A. Stephen Morse

Abstract—By the distributed averaging problem is meant the problem of computing the average value \( \bar{y} \) of a set of numbers possessed by the agents in a distributed network using only communication between neighboring agents. Gossiping is a well-known approach to the problem which seeks to iteratively arrive at a solution by allowing each agent to interchange information with at most one neighbor at each iterative step. A pair of neighboring agents \( i \) and \( j \) gossip at time \( t \) by first interchanging their current estimates of \( y \) and then by updating these estimates using the rules \( x_i(t+1) = (1-w)x_i(t) + wx_j(t) \) and \( x_j(t+1) = (1-w)x_j(t) + wx_i(t) \) respectively where \( w \) is a real constant weight between 0 and 1. This paper studies the effect of different values of \( w \) on convergence rate. This is done for “-periodic gossiping” protocols where by \( T \)-periodic gossiping is meant a centrally scheduled, deterministic protocol which stipulates that each agent must gossip with the same neighbor exactly once every \( T \) time units. It is shown both by analysis of specific examples and by computer studies that the optimal value of \( w \) may result in significantly faster convergence than that provided by the straight average weight \( w = 0.5 \).

I. INTRODUCTION

In recent years there has been a great deal of interest in understanding the process of distributed averaging in depth [1]–[5]. The distributed averaging problem arises in many applications, including achieving consensus in distributed sensor networks [6] as well as in the coordination of the movements of groups of robots or unmanned aerial vehicles [7].

Distributed averaging deals with a network of \( n > 1 \) agents under the constraint that each agent \( i \) is able to communicate only with certain other agents called agent \( i \)’s neighbors. Neighbor relations are described by a simple, connected graph \( \mathbb{N} \), called a neighbor graph, where vertices correspond to agents and edges indicate neighbor relations. Thus the neighbors of an agent \( i \) have the same labels as the vertices in \( \mathbb{N} \) that are adjacent to vertex \( i \). Initially, each agent holds or acquires a real number \( y_i \), for instance a quantity measured by the agent such as the local temperature near the agent. The distributed averaging problem is to devise a protocol that will enable each agent to compute the average \( \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \) using only information acquired from its neighbors. This paper assumes that the \( y_i \) are real numbers and that \( \mathbb{N} \) does not depend on time.

Gossiping is a well-known approach to the distributed averaging problem that seeks to iteratively arrive at a solution by allowing each agent \( i \) to exchange information with at most one neighbor \( j \) at each iterative time step. In this paper, agents \( i \) and \( j \) gossip at time \( t \) by interchanging their current estimates of \( y \) with each other and then updating their estimates according to the convex combination rules

\[
\begin{align*}
x_i(t+1) &= (1-w)x_i(t) + wx_j(t) \\
x_j(t+1) &= (1-w)x_j(t) + wx_i(t)
\end{align*}
\]

respectively, where \( w \) is a real number between 0 and 1. Therefore at any given time \( t \), each agent \( i \) only needs to store a single number \( x_i(t) \), minimizing the memory capacity required of each agent. In most work \( w = 0.5 \) in which case the preceding update rules are simple averages [1], [2], [8]. This paper studies the effect of different values of \( w \) on convergence rate. This will be done for “-periodic gossiping” protocols where by \( T \)-periodic gossiping is meant a centrally scheduled, deterministic protocol which stipulates that each agent must gossip with the same neighbor exactly once every \( T \) time units. Thus, \( T \)-periodic gossiping avoids a situation in which an agent receives requests to gossip from more than one neighbor at the same time, reducing the communication capacity required of each agent.

The convergence rate question for more general distributed averaging has been studied in [5], [9], [10]. In [8] the convergence rate question is addressed for probabilistic gossiping algorithms. A modified gossiping algorithm intended to speed up convergence is proposed in [11] without proof of correctness, but with convincing experimental results. The algorithm has recently been analyzed in [12]. Recent results concerning convergence rates appear in [2], [13]–[15] for periodic gossiping and in [1], [16] for deterministic aperiodic gossiping.

In our analysis, we represent the gossiping process as a finite dimensional discrete-time linear system. The worst case convergence rate is determined by \( T \) and by the second largest eigenvalue (in magnitude) of the stochastic matrix which the gossiping define over a period [1], [2]. In order to compare convergence rates between protocols with different \( w \) in each of the examples, we compute the magnitude of the second largest eigenvalue (in magnitude) for this matrix.

In Section II we review the \( T \)-periodic gossiping protocol and its convergence properties. In Section III we provide examples of some classes of neighbor graphs and determine the magnitude of the second largest eigenvalue (in magnitude) of stochastic matrix which the gossiping define over a
period for arbitrary fixed weights, allowing us to compare
the convergence rates of the straight average weight and
the optimal weight for each of the cases. In Section IV we
discuss the convergence rate of the protocol on a class of
random graphs, through statistical simulations.

II. T-PERIODIC Gossiping

Let \( x(t) \) be the \( n \)-vector whose \( i \)th entry is agent \( i \)'s
estimate of \( y \) at time \( t \). Then each single gossip between
neighbors \( i \) and \( j \) can be represented by an \( n \times n \) primitive
gossip matrix \( P_{ij} \) for which \( p_{ii} = p_{jj} = 1 - w, p_{ij} =
w, p_{ki} = 1 \) for each \( k \not\in \{i,j\} \), and all remaining entries
equal zero. Each sequence of gossips over a period can be
represented by the gossip matrix

\[
M = P_1 P_2 \cdots P_T
\]

(1)

where \( P_r \) is the primitive gossip matrix of the \( r \)th gossip
in the sequence and \( T \) is the number of gossips over a period
[1].

As in [1] we say that a gossip sequence \( \Sigma \) and its
associated matrix \( M_{\Sigma} \) are complete if the graph the gossips
in the sequence induce\(^1\) is a connected spanning subgraph
of \( \mathbb{N} \). A gossip sequence or corresponding gossip matrix is
minimally complete, if it is complete and if there is no other
complete gossip sequence of shorter length. In the special but
important case when \( \mathbb{N} \) is itself a tree \( T \), more can be said.
In this case, a minimally complete gossip sequence is one in
which, for each edge in \( T \), there is exactly one corresponding
individual gossip in the sequence.

We also note that if consecutive gossips in (1) involve
noninteracting pairs of agents, they can be performed in the
same time step as a single “multi-gossip”, reducing the
time the protocol takes to complete each period [2]. Since
the ability of multi-gossips reduces the convergence time by
the same factor regardless of \( w \), we will not investigate this
aspect of convergence time reduction, as it has already been
investigated in [2] for \( w = 0.5 \).

One of the advantages of \( T \)-periodic gossiping is that the
protocol is guaranteed to converge exponentially fast [1], [2].
For clarity, we will restate generalizations of some of the
convergence results in [1], [2], which were previously proved
for the special case of \( w = 0.5 \). All of the convergence results
in [1] and [2] easily generalize to gossiping with any fixed
weight \( w \in (0,1) \). The first result, proved in [1] for the
special case of \( w = 0.5 \), guarantees the convergence of \( T-
persistent gossiping. It states that \( k \) is used. Hence

\[
\tau = \frac{\log \lambda(0.5)}{\log \lambda(w^*)}
\]

Depending on the application, we may either be interested in
comparing \( \rho \) or \( \tau \) achieved by using \( w^* \) in place of \( 0.5 \).
As we shall see later, \( \rho \) tends to be larger and \( \tau \) tends to be
smaller when \( \lambda(0.5) \) and \( \lambda(w^*) \) are both closer to 1, since
in this case the stopping time is very long. The opposite is
true when \( \lambda(0.5) \) and \( \lambda(w^*) \) are both much smaller than 1.

A. Path Graphs

In this section we first provide analytical solutions for the
eigenvalues of the minimally complete gossip matrix of the
two-edge path graph, and compare the convergence rates for
different \( w \). It can be verified that the set of eigenvalues \( \Lambda
\)

\( \begin{align*}
\text{CONFIDENTIAL. Limited circulation. For review only.}
\end{align*} \)

\( \begin{align*}
\text{Preprint submitted to 51st IEEE Conference on Decision and Control .}
\end{align*} \)

\( \begin{align*}
\text{Received March 4, 2012.}
\end{align*} \)

\( ^1 \)The graph induced by a given sequence of single gossips (or correspond-

\( ^2 \)We define extended star graphs as graphs that are composed of \( k > 2 \)
paths connected at one end to a common vertex. An extended start graph
called uniform if its \( k \) paths have the same length.
of the minimally complete gossip matrix for the two-edge path graph is

$$\Lambda = \left\{ 1, \frac{1}{2} \left( 2 - 4w + w^2 \pm w \sqrt{4 - 8w + w^2} \right) \right\}$$ (2)

Using (2) we can get \( w^* = 4 - 2\sqrt{3} = 0.536 \) which minimizes the magnitude of the second largest eigenvalue (in magnitude). In particular, this optimal value of \( w \) is strictly greater than 0.5. This is different from the \( k = 1 \) case where it is easy to show that \( w^* = 0.5 \).

Next we provide numerical solutions for the eigenvalues of the minimally complete gossip matrix of the \( k \)-edge path graph for various values of \( k \), and compare the convergence rates for different \( w \). From the graphs in Figure 1 we see that the optimal value of \( w \) increases with the number of edges \( k \), and is always strictly greater than 0.5 for \( k > 1 \). The ratio \( \rho \) of the optimal and straight average convergence rates increases with \( k \), approaching 1 for large \( k \) (some values of \( (k, \rho) \) are \((2, 0.287)\), \((5, 0.445)\), and \((15, 0.700)\)). However, the ratio \( \tau \) of the total amount of time taken decreases with \( k \), approaching 0 for large \( k \) (some values of \( (k, \tau) \) are \((2, 0.526)\), \((5, 0.262)\), and \((15, 0.098)\)).

B. Star Graphs

In this section we provide numerical solutions for the eigenvalues of the minimally complete gossip matrix of star graphs for various values of the number of edges \( k \), and compare the convergence rates for different \( w \). From Figure 2 we see that the optimal value of \( w \) briefly increases and then decreases with the number of edges \( k \), being greater than 0.5 for small \( k \) but quickly becoming strictly less than 0.5 for somewhat larger \( k \) (e.g., for \( k \geq 6 \)).

The ratio \( \rho \) of the optimal and straight average convergence rates is close to 1 for \( k > 2 \), and approaches 1 as \( k \) gets very large (e.g., some values of \( (k, \rho) \) are \((2, 0.287)\), \((3, 0.934)\), \((10, 0.940)\), and \((100, 0.973)\)). However, the ratio \( \tau \) of the total amount of time taken at first increases with \( k \) and then decreases with \( k \) after \( k = 6 \), approaching 0 for large \( k \) (e.g., some values of \( (k, \tau) \) are \((2, 0.526)\), \((6, 0.961)\), \((10, 0.799)\), and \((100, 0.122)\)).

C. Uniform Extended Star Graphs

In this section we provide numerical solutions for the eigenvalues of the minimally complete gossip matrix of the uniform extended star graph (as defined in [2]) for various central vertex degrees \( k \) and fixed uniform “arm” length \( m = 6 \), and compare the convergence rates for different \( w \). From Figure 3 we see that the optimal weight \( w^* \) increases with the central vertex degree \( k \) for \( k \leq 3 \) and decreases with \( k \) for \( k \geq 6 \). Although the optimal weight is strictly greater than 0.5 for \( k \leq 10 \), for \( k \geq 20 \) it is strictly less than 0.5. Note that the extended star graph requires a larger central vertex degree \( k \) than the simple star graph for the optimal value of \( w \) to begin to decrease with \( k \) and become less than 0.5.

The fraction \( \tau \) of the total amount of time taken when using the weight \( w^* \) in place of 0.5 is smaller for values of \( k \) when the optimal value of \( w \) is far from 0.5, i.e., when \( k \ll m \) or \( k \gg m \) (e.g., some values of \( (k, \tau) \) are \((1, 0.355)\), \((2, 0.294)\), \((3, 0.522)\), \((8, 0.988)\), \((20, 0.761)\), and \((40, 0.439)\)); some values of \( (k, \rho) \) are \((1, 0.593)\), \((8, 0.999)\), and \((40, 0.992)\)).

One way to view this behavior is that when \( k \) is small in comparison to \( m \), the uniform extended star graph behaves
more like a path graph. When \( k \) is large in comparison to \( m \), the uniform extended star graph behaves more like a simple star graph.

Fig. 3. The magnitude of the second largest eigenvalue (in magnitude) \( \lambda \) vs. the weight \( w \) for for any minimally complete gossip matrix of uniform extended star graph. The number closest to a given curve indicates the curve’s corresponding value of \( k \).

IV. CONNECTED ERDŐS-RÉNYI RANDOM GRAPHS

In this section we discuss the optimal weight \( w^* \) for a class of periodic gossiping protocols on connected Erdős-Rényi random graphs [17]. To construct a connected Erdős-Rényi random graph, we start with a fixed number of vertices \( n \) \( (n = 50 \) in the simulations of Figure 4). Then \( nd/2 \) edges are successively added to the graph at random with each remaining edge in the set of all possible edges \( \{i, j\}, \ i, j \in \{1, 2, \ldots, n\} \), having equal probability of being chosen (each edge can only be included once and no self-loops are allowed). Thus the average degree of the graph is \( d \). If the graph is not connected, it is discarded and the construction is repeated until a connected graph is constructed. These random graphs therefore have the same distribution as the Erdős-Rényi graph with \( n \) vertices and \( nd/2 \) edges, conditioned on the event that the graph is connected. The order of gossips over a period is chosen randomly from the uniform distribution on all permutations of the edges included in the constructed connected random graph. Thus the period \( T \) equals the number of edges \( nd/2 \). It is worth emphasizing that in this case the convergence rate depends on the order in which gossips over a period take place since the graph is not, in general, a tree.

Every time a connected random graph and the order of gossips over a period were chosen at random in the simulations, the complete gossip matrix determined by the gossips over a period was repeatedly constructed for twenty equally spaced values of \( w \) on the interval \((0, 1)\) using the same graph and gossip sequence over a period, and the second largest eigenvalue (in magnitude) was computed at each of these values of \( w \). This process was repeated in 100 separate trials at each value of \( d \) investigated. The number of vertices \( n \) was fixed at \( n = 50 \), and the average degree \( d \) was varied from \( d = 3 \) to \( d = 49 \) in increments of size 1.

From the simulation results in Figure 4, we see that the average optimal weight \( \bar{w}^* \) is in general not 0.5 and decreases sharply with \( d/n \). Moreover, \( \bar{\tau} \), the average value of \( \tau \), is small for all values of \( d/n \) except for the values of \( d/n \) around 0.2 where \( w^* \) is close to 0.5, implying that the optimal weight is substantially faster than the straight average weight for most values of \( d/n \).

V. CONCLUDING REMARKS

Periodic gossiping protocols are investigated for some classes of tree graphs (path, star, uniform extended star) that benefit from using the optimal weight \( w = w^* \) in place of the straight average \( (w = 0.5) \). In some cases the fraction \( \tau \) of the total amount of time taken when using the weight \( w^* \) in place of 0.5 is very small (down to \( \frac{1}{10} \) for some of the graphs investigated, most likely arbitrarily small for certain arbitrarily large graphs), while in other cases the improvement in the total amount of time taken is more modest. The improvement in convergence time tends to be more significant for graphs where \( w^* \) is far from 0.5.

For periodic gossiping in connected Erdős-Rényi random graphs, we find that the optimal weight \( w^* \) is in general far from 0.5, and that the fraction \( \tau \) of the total amount of time taken when using \( w^* \) in place of 0.5 was small for most values of \( d/n \). Hence, we expect that using the optimal weight will substantially speed up the convergence of periodic gossiping for more general classes of both tree and non-tree graphs arising from random distributions.
The average value of optimal weight $w^*$ (top) and the average value of $\tau$ (bottom) vs. $d/n$ for periodic gossiping in connected Erdős-Rényi random graphs with fixed number of vertices $n = 50$. Both averages are taken over 100 independent trials at each value of $d/n$.

Fig. 4. The average value of optimal weight $w^*$ (top) and the average value of $\tau$ (bottom) vs. $d/n$ for periodic gossiping in connected Erdős-Rényi random graphs with fixed number of vertices $n = 50$. Both averages are taken over 100 independent trials at each value of $d/n$.

REFERENCES


