Distributed Averaging via Gossiping

A. S. Morse

Yale University
Consensus and averaging
Linear iterations
Gossiping
Periodic gossiping
Multi-gossip sequences
Request-based gossiping
Double linear iterations
Consider a group of $n$ agents labeled 1 to $n$

The groups’ **neighbor graph** $N$ is an undirected, connected graph with vertices labeled $1,2,...,n$.

The **neighbors** of agent $i$, correspond to those vertices which are adjacent to vertex $i$.

Each agent $i$ controls a real, scalar-valued, time-dependent, quantity $x_i$ called an **agreement variable**.

The goal of a consensus process is for all $n$ agents to ultimately reach a consensus by adjusting their individual agreement variables to a common value.

This is to be accomplished over time by sharing information among neighbors in a distributed manner.
A consensus process is a recursive process which evolves with respect to a discrete time scale.

In a standard consensus process, agent $i$ sets the value of its agreement variable at time $t+1$ equal to the average of the current value of its own agreement variable and the current values of its neighbors’ agreement variables.

$$x_i(t + 1) = \frac{1}{(1 + d_i)} \left( x_i(t) + \sum_{j \in N_i} x_j(t) \right)$$

$N_i =$ set of indices of agent $i$'s neighbors.

$d_i =$ number of indices in $N_i$

Average at time $t$ of values of agreement variables of agent $i$ and the neighbors of agent $i$. 
An averaging process is a consensus process in which the common value to which each agreement variable is supposed to converge, is the average of the initial values of all agreement variables:

\[ x_{avg} = \frac{1}{n} \sum_{i=1}^{n} x_i(0) \]

Application: distributed temperature calculation

Generalizations:
- Time-varying case - \( N \) depends on time
- Integer-valued case - \( x_i(t) \) is to be integer-value
- Asynchronous case - each agent has its own clock

Implementation Issues:
- How much network information does each agent need?
- To what extent is the protocol robust?

Performance metrics:
- Convergence rate
- Number of transmissions needed

General Approach:
- Probabilistic
- Deterministic

Standing Assumption:
- \( N \) is a connected graph
ROADMAP

Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

Multi-gossip sequences

Request-based gossiping

Double linear iterations
Linear Iteration

\[ x_i(t + 1) = w_{ii}x_i(t) + \sum_{j \in N_i} w_{ij}x_j(t) \]

\( N_i \) = set of indices of agent \( i \)'s neighbors.

\( w_{ij} \) = suitably defined weights

\[ x(t + 1) = Wx(t) \]

Want \( x(t) \rightarrow x_{avg} \)

If \( A \) is a real \( n \times n \) matrix, then \( A^t \) converges as \( t \rightarrow 1 \), to a rank one matrix of the form \( qp^0 \) if and only if \( A \) has exactly one eigenvalue at value 1 and all remaining \( n-1 \) eigenvalues are strictly smaller than 1 in magnitude.

If \( A \) so converges, then \( Aq = q \), \( Ap = p \) and \( p^q = 1 \).

Thus if \( W \) is such a matrix and \( q = 1; \ p = \frac{1}{n} \) then \( W^t \rightarrow \frac{1}{n}1^n \)

\[ x(t) \rightarrow \frac{1}{n}1^n x(0) = \frac{1}{n}x_{avg} \]
Linear Iteration with Nonnegative Weights

A square matrix $S$ is **stochastic** if it has only nonnegative entries and if its row sums all equal 1.

$$S^1 = 1$$

$$\|S\|_1 = 1$$

Spectrum $S$ contained in the closed unit circle

All eigenvalue of value 1 have multiplicity 1

For the nonnegative weight case, $x(t)$ converges to $x_{avg}1$ if and only if $W$ is doubly stochastic and its single eigenvalue at 1 has multiplicity 1.

How does one choose the $w_{ij} \geq 0$ so that $W$ has these properties?
\[ x(t + 1) = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \frac{1}{g} \begin{bmatrix} L \end{bmatrix} x(t) \]

Adjacency matrix of N: matrix of ones and zeros with \( a_{ij} = 1 \) if N has an edge between vertices i and j.

\[
D = \begin{bmatrix}
-\, d_1 & 0 & \cdots & 0 \\
0 & d_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & d_n
\end{bmatrix}
\]

\[ L = D - A \]

\[
I - \frac{1}{g} D \geq 0
\]

\[
1'(I - \frac{1}{g} L) = 1'
\]

\[
(I - \frac{1}{g} L) 1 = 1
\]

The eigenvalue of \( L \) at 0 has multiplicity 1 because N is connected

\[
x_i(t + 1) = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \frac{d_i}{g} x_i(t) + \frac{1}{g} \sum_{j \in N_i} x_j(t)
\]

Each agent needs to know \( \max \{d_1, d_2, \ldots, d_n\} \) to implement this.
Each agent needs to know the number of neighbors of each of its neighbors.

\[ L = QQ' \]

\( Q \) is a \(-1,1,0\) matrix with rows indexed by vertex labels in \( N \) and columns indexed by edge labels such that \( q_{ij} = 1 \) if edge \( j \) is incident on vertex \( i \), and \(-1\) if edge \( i \) is incident on vertex \( j \) and \( 0 \) otherwise.

\[ i = \frac{1}{(1 + \max f d_i; d_j g)} \]

\[ I - Q^\alpha Q^0 \]

\( \alpha = \text{diagonal} \{ 1; 2; \ldots \} \)

\[ x(t + 1) = (I - Q^\Lambda Q')x(t) \]

**Why is** \( I - Q^\alpha Q^0 \) **doubly stochastic?**

**Metropolis Algorithm**

\[ x_i(t+1) = \sum_{j \in \mathcal{N}_i} \frac{1}{(1 + \max f d_i; d_j g)} A x_i(t) + \sum_{j \in \mathcal{N}_i} \frac{1}{(1 + \max f d_i; d_j g)} x_j(t) \]

Each agent needs to know the number of neighbors of each of its neighbors.

Total number of transmissions/iteration:

\[ d_i = nd_{\text{avg}} \quad d_{\text{avg}} = \frac{1}{n} \sum_{i=1}^{n} d_i \]
Agent $i$'s queue is a list $q_i(t)$ of agent $i$'s neighbor labels.

Agent $i$'s preferred neighbor at time $t$, is that agent whose label is in the front of $q_i(t)$.

$M_i(t) = \{\text{set of labels of agent } i\}'s preferred neighbor together with the labels of all neighbors who send agent $i$ their agreement variables at time $t$.

$m_i(t) = \text{the number of labels in } M_i(t)$

Between times $t$ and $t+1$ the following steps are carried out in order.

Agent $i$ transmits its label $i$ and $x_i(t)$ to its current preferred neighbor. \hfill $n$ transmissions

At the same time agent $i$ receives the labels and agreement variable values of those agents for whom agent $i$ is their current preferred neighbor.

Agent $i$ transmits $m_i(t)$ and its current agreement variable value to each neighbor with a label in $M_i(t)$. \hfill at most $2n$ transmissions

Agent $i$ then moves the labels in $M_i(t)$ to the end of its queue maintaining their relative order and updates as follows:

\begin{align*}
    x_i(t+1) &= \sum_{j \in M_i(t)} x_j(t) \\
    x_i(t+1) &= \sum_{j \in N_i} x_j(t)
\end{align*}

\begin{align*}
    x_i(t+1) &= \sum_{j \in M_i(t)} x_j(t) \\
    x_i(t+1) &= \sum_{j \in N_i} x_j(t)
\end{align*}

Modification

$n d_{avg} > 3$

$3n$

$nd_{avg}$
Randomly chosen graphs with 200 vertices. 10 random graphs for each average degree.

$x(0) = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ 200 \end{bmatrix}$
Consider $A$ real symmetric with row and column sums $= 1$; i.e., $A\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^T A = \mathbf{1}$.

Original system:

$$x(t + 1) = A x(t)$$

$x(t)$ converges to $x_{avg}$ iff $\frac{3}{2} < 1$ where $\frac{3}{2}$ is the second largest singular value of $A$.

Augmented system:

$$x(t + 1) = (g + 1) A x(t) - g y(t)$$ $$y(t + 1) = x(t)$$

Augment system is fastest if $g = \frac{1}{1 + \frac{1}{1} \frac{\sqrt{2}}{1}} = \frac{1}{1 + \frac{1}{1} \frac{\sqrt{2}}{1}}$.
ROADMAP

Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

Multi-gossip sequences

Request-based gossiping

Double linear iterations
Gossip Process

A gossip process is a consensus process in which at each clock time, each agent is allowed to average its agreement variable with the agreement variable of at most one of its neighbors.

\[ x_i(t + 1) = \frac{1}{2}(x_i(t) + x_j(t)) \]
\[ x_j(t + 1) = \frac{1}{2}(x_i(t) + x_j(t)) \]

The index of the neighbor of agent \( i \) which agent \( i \) gossips with at time \( t \).

If agent \( i \) gossips with neighbor \( j \) at time \( t \), then agent \( j \) must gossip with agent \( i \) at time \( t \).

This is called a gossip and is denoted by \((i, j)\).

In the most commonly studied version of gossiping, the specific sequence of gossips which occurs during a gossiping process is determined probabilistically.

In a deterministic gossiping process, the sequence of gossips which occurs is determined by a pre-specified protocol.
A gossip process is a consensus process in which at each clock time, each agent is allowed to average its agreement variable with the agreement variable of at most one of its neighbors.

\[
x_i(t + 1) = \frac{1}{2}(x_i(t) + x_j(t))
\]

\[
x_j(t + 1) = \frac{1}{2}(x_i(t) + x_j(t))
\]

\[
x_i(t + 1) + x_j(t + 1) = x_i(t) + x_j(t)
\]

1. The sum total of all agreement variables remains constant at all clock steps.

2. Thus if a consensus is reached in that all agreement variables reach the same value, then this value must be the average of the initial values of all gossip variables.

3. This is not the case for a standard consensus process.
A finite sequence of multi-gossips induces a spanning sub-graph $M$ of $N$ where $(i, j)$ is an edge in $M$ iff $(i, j)$ is a gossip in the sequence.

If more than one pair of agents gossip at a given time, the event is called a multi-gossip.

For purposes of analysis, a multi-gossip can be thought of as a sequence of single gossips occurring in zero time. The arrangement of gossips in the sequence is of no importance because individual gossip pairs do not interact.

A gossip $(i, j)$ is allowable if $(i, j)$ is the label of an edge on $N$.

A finite sequence of multi-gossips induces a spanning sub-graph $M$ of $N$ where $(i, j)$ is an edge in $M$ iff $(i, j)$ is a gossip in the sequence.
State Space Model

\[
\begin{align*}
x(t + 1) &= M(t)x(t) \\
x(t) &= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \\
x_k(t + 1) &= x_k(t), \quad k \neq i, j
\end{align*}
\]

For \( n = 7, i = 2, j = 5 \)

A doubly stochastic matrix

For each \( t, \ M(t) \) is a primitive gossip matrix.

Each primitive gossip matrix is a doubly stochastic matrix with positive diagonals.

Doubly stochastic matrices with positive diagonals are closed under multiplication.
State Space Model

\[ x(t + 1) = M(t)x(t) \]

Doubly stochastic matrices with positive diagonals are closed under multiplication.

For each \( t \), \( M(t) \) is a primitive gossip matrix.

Each gossip matrix is a product of primitive gossip matrices.

Each gossip matrix is a doubly stochastic matrix with positive diagonals.

Doubly stochastic matrices with positive diagonals are closed under multiplication.
Claim: If $A$ is a complete gossip matrix then

$$A^i \leq \frac{1}{n} 11^0$$

as fast as $\lambda_i$ where $\lambda$ is the second largest eigenvalue {in magnitude} of $A$. 
Facts about nonnegative matrices

A square nonnegative matrix $A$ is **irreducible** if there does not exist a permutation matrix $P$ such

$$PAP^{-1} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

where $B$ and $D$ are square matrices.

A square nonnegative matrix $M$ is **primitive** if $M^q$ is a positive matrix for $q$ sufficiently large.

**Fact 1:** $A$ is irreducible iff its graph is strongly connected.

**Fact 2:** $A$ is irreducible iff $I + A$ is primitive

Consequences if $A$ has all diagonal elements positive:

$A$ is irreducible iff $A$ is primitive.

$A$ is primitive iff its graph is strongly connected.
Perron Frobenius Theorem

Theorem: Suppose $A$ is primitive. Then $A$ has a single eigenvalue $\lambda$, of maximal magnitude and $\lambda$ is real. There is a corresponding eigenvector which is positive.

Consequence:

Suppose $A$ is a stochastic matrix with positive diagonals and a strongly connected graph. Then $A$ has exactly one eigenvalue at 1 and all remaining $n-1$ eigenvalues have magnitudes strictly less than 1.

So to show that completeness of $A$ implies that

$$A^i \rightarrow \frac{1}{n} 1^n$$

as fast as $\lambda^i \rightarrow 0$ where $\lambda$ is the second largest eigenvalue (in magnitude) of $A$, it is enough to show that the graph of $A$ is strongly connected.
If $A$ is a complete gossip matrix, then the graph of $A$ is strongly connected.

**Blackboard proof**

If $A$ is a complete gossip matrix, its second largest singular value is less than 1

A complete implies ° $(A)$ is strongly connected.

Thus ° $(A^0 A)$ is strongly connected because ° $(A^0 A) = ° (A^0 ± ° (A))$ and arcs are conserved under composition since $A$ and $A^0$ have positive diagonals.

But $A^0 A$ is stochastic because $A$ is doubly stochastic.

Thus the claim is true because of the Perron Frobenius Theorem

If $A$ is doubly stochastic, then $|A|_2 = $ second largest singular value of $A$

**Blackboard proof**

If $A$ is a complete gossip matrix, then $A$ is a semi-contraction w/r $|\phi|_2$
ROADMAP

Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

Multi-gossip sequences

Request-based gossiping

Double linear iterations
Periodic Gossiping

\[ A = P_{12} P_{34} P_{35} P_{32} P_{56} P_{57} \]

\[ x(iT + 1) = A^i x(1); \quad i \geq 1 \]

\[ A = P_{12} P_{34} P_{35} P_{32} P_{56} P_{57} \]

\[ x(iT + 1) = A^i x(1); \quad i \geq 1 \]

We have seen that because \( A \) is complete

\[ A^i \propto \frac{1}{n} 1^n \]

as fast as \( \lambda \) where \( \lambda \) is the second largest eigenvalue \{in magnitude\} of \( A \).

convergence rate = \[ \frac{\lambda}{\lambda_j - \lambda_1} \]
A = P_{12}P_{34}P_{35}P_{32}P_{56}P_{57}

B = P_{35}P_{56}P_{35}P_{12}P_{34}P_{57}

How are the second largest eigenvalues \{in magnitude\} related?

If the neighbor graph N is a tree, then the spectrums of all possible minimally complete gossip matrices determined by N are the same!
multi-gossips:
(3,4) and (7,5)
(3,5) and (1,2)
(2,3) and (5,6)

Period $T = \text{number of multi-gossips}$
$T = 3$

Minimal number of multi-gossips for a given neighbor graph $N$ is the same as the minimal number of colors needed to color the edges of $N$ so that no two edges incident on any vertex have the same color.

The minimal number of multi-gossips \{and thus the minimal value of $T$\} is the \textbf{chromatic index} of $N$ which, if $N$ is a tree, is the degree of $N$. 

convergence rate = $\gamma \sum_{j,j}$

Minimizing $T$
Modified Gossip Rule

Suppose agents $i$ and $j$ are to gossip at time $t$.

Standard update rule:

$$x_i(t + 1) = \frac{1}{2}x_i(t) + \frac{1}{2}x_j(t)$$

$$x_j(t + 1) = \frac{1}{2}x_i(t) + \frac{1}{2}x_j(t)$$

Modified gossip rule:

$$x_i(t + 1) = \oplus x_i(t) + (1 \oplus) x_j(t)$$

$$x_j(t + 1) = (1 \oplus) x_i(t) + \oplus x_j(t)$$

$$0 < \oplus < 1$$
$|\lambda_2| \ vs \ \text{®}$

100 Vertices, $P(\text{edge} = 0.05)$ i.i.d., 1 trial

$\min_{\alpha} \{ |\lambda_2(\alpha) / \lambda_2(0.5)| \} = 0.6935$
ROADMAP

Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

Multi-gossip sequences

Request-based gossiping

Double linear iterations
Suppose that the infinite sequence of multi-gossips under consideration is repetitively complete with period $T$ if the finite sequence of gossips which occur in any given period of length $T$ is complete.

For example, a periodic gossiping sequence with complete gossiping matrices over each period is repetitively complete.

Suppose that the infinite sequence of multi-gossips under consideration is repetitively complete with period $T$.

Then there exists a nonnegative constant $\lambda < 1$ such that for each initial value of $x(0)$, all $n$ gossip variables converge to the average value $x_{\text{avg}}$ as fast as $\lambda^t$ converges to zero as $t \to 1$. 
A = P_{12}P_{34}P_{35}P_{32}P_{56}P_{57}\\
(5,6) (5,7) (5,6) (3,2) (3,5) (3,4) (1,2) (5,6) (3,4) (1,2) (2,6) (5,7) (3,5) (5,7) (5,6)

x(T + 1) = Ax(1)

A = P_{12}P_{34}P_{35}P_{32}P_{56}P_{57}

B = P_{35}P_{57}P_{26}P_{12}P_{34}P_{56}

A and B are complete gossip matrices
Repetitively Complete Multi-Gossip Sequences

Thus we are interested in studying the convergence of products of the form

\[ \prod_{i=1}^{k} G_i = G_k G_{k-1} \ldots G_2 G_1 \]

as \( k \to 1 \) where each \( G_i \) is a complete gossip matrix.

\[ C = \text{set of all complete multi-gossip sequences of length at most } T \]

\( \lambda_{\frac{3}{4}} = \text{the second largest singular value of the complete gossip matrix determined by multi-gossip sequence } \frac{3}{4} \frac{2}{C} \)

The convergence rate of a repetitively complete gossip sequence of period \( T \) is no slower than

\[ \lambda_{\frac{3}{4}} \]

where:

\[ \lambda_{\frac{3}{4}} = \max_{\frac{3}{4} C} \lambda_{\frac{3}{4}} \]
Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

**Multi-gossip sequences**

Request-based gossiping

Double linear iterations
Request-Based Gossiping

A process in which a gossip takes place between two agents whenever one of the agents accepts a request placed by the other.

An event time of agent $i$ is a time at which agent $i$ places a request to gossip with a neighbor.

Assumption: The time between any two successive event times of agent $i$ is bounded above by a positive number $T_i$.

Deadlocks can arise if a requesting agent is asked to gossip by another agent at the same time. A challenging problem to deal with!

Distinct Neighbor Event Time Assumption: All of the event times of each agent differ from all of the event times of each of its neighbors.
One way to satisfy the distinct neighbor event time assumption is to assign each agent a set of event times which is disjoint with the sets of event times of all other agents. But it is possible to do a more efficient assignment

**Distinct Neighbor Event Time Assumption:** All of the event times of each agent differ from all of the event times of each of its neighbors.
Distinct Neighbor Event Time Assumption: All of the event times of each agent differ from all of the event times of each of its neighbors.

One way to satisfy the distinct neighbor event time assumption is to assign each agent a set of event times which is disjoint with the sets of event times of all other agents.

But it is possible to do a more efficient assignment.

The minimal number of event time sequences needed to satisfy the distinct neighbor event time assumption is the same as the minimal number of colors needed to color the vertices of N so that no two adjacent vertices have the same color.

Minimum number of different event times sequences needed equals the chromatic number of N which does not exceed 1 + the maximum vertex degree of N.
Request-Based Protocol

Agent $i$'s queue is a list $q_i(t)$ of agent $i$'s neighbor labels.

Protocol:

1. If $t$ is an event time of agent $i$, agent $i$ places a request to gossip with that neighbor whose label is in front of the queue $q_i(t)$.

2. If $t$ is not an event time of agent $i$, agent $i$ does not place a request to gossip.

3. If agent $i$ receives one or more requests to gossip, it gossips with that requesting neighbor whose label is closest to the front of the queue $q_i(t)$ and then moves the label of that neighbor to the end of $q_i(t)$.

4. If $t$ is not an event time of agent $i$ and agent $i$ does not receive a request to gossip, then agent $i$ does not gossip.
Let $E$ denote the set of all edges $(i, j)$ in the neighbor graph $N$.

Suppose that the distinct neighbor event time assumption holds.

Then the infinite sequence of gossips generated by the request based protocol will be repetitively complete with period

$$T = \max_{(i,j) \in E} \min \left\{ \sum_{j \in N_i} \vartheta \left( T_i \times d_j ; T_k \times d_j \right) \right\},$$

where for all $i$,

- $T_i$ is an upper bound on the time between two successive event times of agent $i$,
- $d_j$ is the number of neighbors of agent $j$ and
- $N_i$ is the set of labels of agent $i$'s neighbors.
ROADMAP

Consensus and averaging

Linear iterations

Gossiping

Periodic gossiping

Multi-gossip sequences

Request-based gossiping

Double linear iterations
Double Linear Iteration

\[ y(t + 1) = S(t)y(t); \quad y(0) = x(0) \]
\[ z(t + 1) = S(t)z(t); \quad z(0) = 1 \]

\[ x_i(t) = \frac{y_i(t)}{z_i(t)} \]

\[ y_i = \text{unscaled agreement variable} \]
\[ z_i = \text{scaling variable} \]

Suppose

\[ \lim_{t \to 1} S(t)S(t-1) \cdots S(1) = q_1^n \]

\[ \lim_{t \to 1} y(t) = q_1^n x(0) = q_1^n x_{\text{avg}} \]
\[ \lim_{t \to 1} z(t) = 1'x(0) = \sum_{i=1}^n x_i(0) \]

Suppose each \( S(t) \) has positive \( \phi \)

\[ = n \left( \frac{1}{n} \sum_{i=1}^n x_i(0) \right) < 1 \]

\[ \lim_{t \to 1} x_i(t) = \lim_{t \to 1} \frac{y_i(t)}{z_i(t)} = \frac{q_i n x_{\text{avg}}}{q_i n} = x_{\text{avg}}; \quad i = 1; 2; \ldots; n \]
Broadcast-Based
Double Linear Iteration

Initialization: \[ y_i(0) = x_i(0) \quad z_i(0) = 1 \]

Transmission: Agent \( i \) broadcasts the pair \([y_i(t), z_i(t)]\) to each of its neighbors.

Update: \[
\begin{align*}
x_i(t) &= \frac{y_i(t)}{z_i(t)} \\
y_i(t + 1) &= \frac{1}{1 + d_i} y_i(t) + \sum_{j \in N_i} \frac{1}{1 + d_j} y_j(t) \\
z_i(t + 1) &= \frac{1}{1 + d_i} z_i(t) + \sum_{j \in N_i} \frac{1}{1 + d_j} z_j(t)
\end{align*}
\]

Agent’s require same network information as Metropolis
\( n \) transmissions/iteration
Works if \( N \) depends on \( t \)

Why does it work?
\[
\begin{align*}
y(t + 1) &= S y(t) \quad y(0) = x(0) \\
z(t + 1) &= S z(t) \quad z(0) = 1
\end{align*}
\]

\( A = \) adjacency matrix of \( N \)
\( S = (I + A)(I + D)^{-1} \)
\( D = \) degree matrix of \( N \)
\( S = \) left stochastic with positive diagonals
\[ S^t \rightarrow q1' \] because N is connected

Connectivity of N implies strong connectivity of the graph of \((I+A)\)

\[ \mathbb{N} \text{ connected implies } (I + A)^{n-1} > 0 \]

\[ (((I + A)(I + D)^{-1})^{n-1} > 0 \]

\[ (S')^{n-1} > 0 \]

S' is irreducible

Therefore \(S^0\) has a strongly connected graph

Therefore \(S^0\) has a rooted graph

\[ (S')^t \rightarrow 1q' \text{ as } t \rightarrow \infty \]
\( S^t \rightarrow q1' \) because N is connected

\[
q = \frac{1}{\sum_{i=1}^{n} (1 + d_i)} \begin{bmatrix}
1 + d_1 \\
1 + d_2 \\
\vdots \\
1 + d_n
\end{bmatrix}
\]

\( S^01' = 1 \)

\[
S = (I + A)(I + D)^{-1}
\]

\( z(t) > 0 \), \( 8 t < 1 \) because \( S \) has positive diagonals

\( z(t) > 0 \), \( 8 t \cdot 1 \) because \( z(1) = nq \) and \( q > 0 \)

So \( x_i(t) = \frac{y_i(t)}{z_i(t)} \) is well-defined

\[
y(t + 1) = Sy(t) \quad y(0) = x(0)
\]

\[
z(t + 1) = Sz(t) \quad z(0) = 1
\]

\( S = \text{left stochastic with positive diagonals} \)
Metropolis Iteration vs Double Linear Iteration

30 vertex random graphs

$|\lambda_2|$ vs $d_{avg}$
Round Robin - Based Double Linear Iteration

Initialization: \[ y_i(0) = x_i(0) \quad z_i(0) = 1 \]

Transmission: Agent \( i \) transmits the pair \( \{y_i(t), z_i(t)\} \) its preferred neighbor.

At the same time agent \( i \) receives the values \[ y_{j_1}(t), z_{j_1}(t), y_{j_2}(t), z_{j_2}(t), \ldots, y_{j_k}(t), z_{j_k}(t) \]

from the agents \( j_1, j_2, \ldots, j_k \) who have chosen agent \( i \) as their current preferred neighbor.

Update: Agent \( i \) then moves the label of its current preferred neighbor to the end of its queue and sets

\[
\begin{align*}
    y_i(t + 1) &= \frac{1}{2} (y_i(t) + y_{j_1}(t) + y_{j_2}(t) + \cdots + y_{j_k}(t)) \\
z_i(t + 1) &= \frac{1}{2} (z_i(t) + z_{j_1}(t) + z_{j_2}(t) + \cdots + z_{j_k}(t))
\end{align*}
\]

No required network information \( n \) transmissions/iteration

Why does it work?
\( S(t) = \) left stochastic, positive diagonals
\[
\begin{align*}
y(t + 1) &= S(t) y(t), & y(0) = x(0) \\
z(t + 1) &= S(t) z(t), & z(0) = 1
\end{align*}
\]

\( S(0), S(1), \ldots \) is periodic with period \( T = \text{lcm} \{d_1, d_2, \ldots, d_n\} \).
\[
P_1 = S(T - 1) S(T - 2) \cdots S(0) \\
P_2 = S(T) S(T - 1) \cdots S(1) \\
\vdots \\
P_\tau = S(T - 2 + \tau) S(T - 3 + \tau) \cdots S(\tau - 1), \quad \tau \in \{1, 2, \ldots, T\}
\]

\[
y(iT + \tau - 1) = P^i_\tau y(0), \quad i \geq 0 \\
z(iT + \tau - 1) = P^i_\tau z(0), \quad i \geq 0
\]

\( P_\tau \) is primitive \( P_\tau^k > 0 \) for some \( k > 0 \)

Perron-Frobenius: \( P_\tau \) has single eigenvalue at 1 and it has multiplicity 1. \( q_\tau > 0 \)

\[
\lim_{i \to \infty} P^i_\tau = q_\tau 1', \quad \tau \in \{1, 2, \ldots, T\}
\]

\[
\begin{bmatrix} y(t) \\ z(t) \end{bmatrix} \rightarrow \begin{bmatrix} n x_{\text{avg}} I & 0 \\ 0 & n I \end{bmatrix} \begin{bmatrix} q_\tau \\ q_\tau \end{bmatrix} : \tau \in \{1, 2, \ldots, T\}
\]

\( z(t) > 0, \quad 8 \ t < 1 \) because each \( S(t) \) has positive diagonals

\( z(t) > 0, \quad 8 \ t \cdot 1 \) because \( z(t) \) ! \{nq_1, nq_2, \ldots, nq_{\tau}\} and \( q_\tau > 0 \)

\[
\lim_{t! 1} x_i(t) = \lim_{t! 1} \frac{y_i(t)}{z_i(t)} = x_{\text{avg}}, \quad i \in \{1, 2; \ldots; n\}
\]