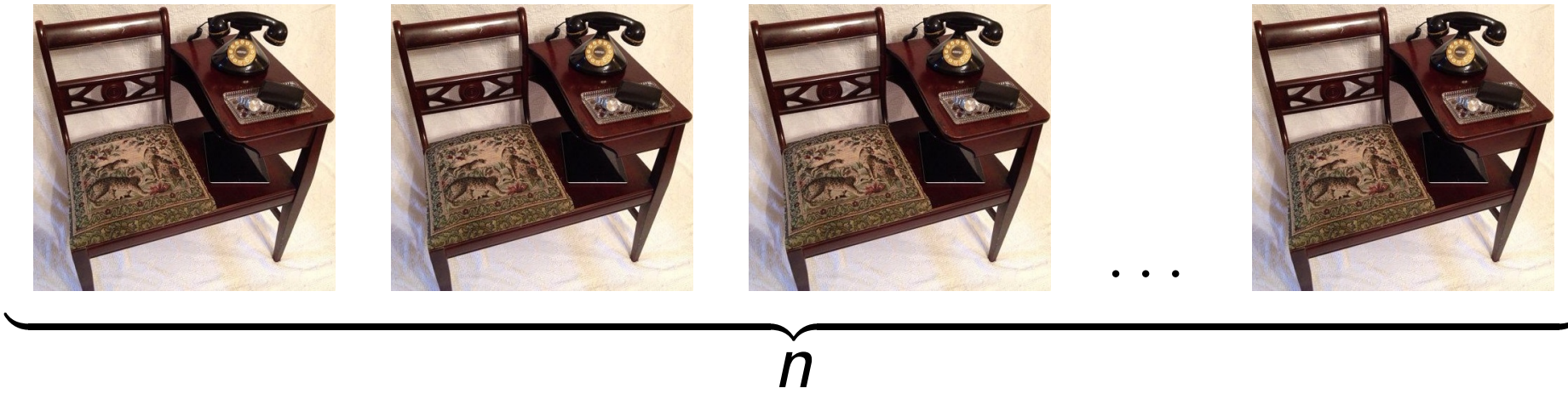
The Gossip Monoid

Peter Fenner
Joint work with Marianne Johnson and Mark Kambites

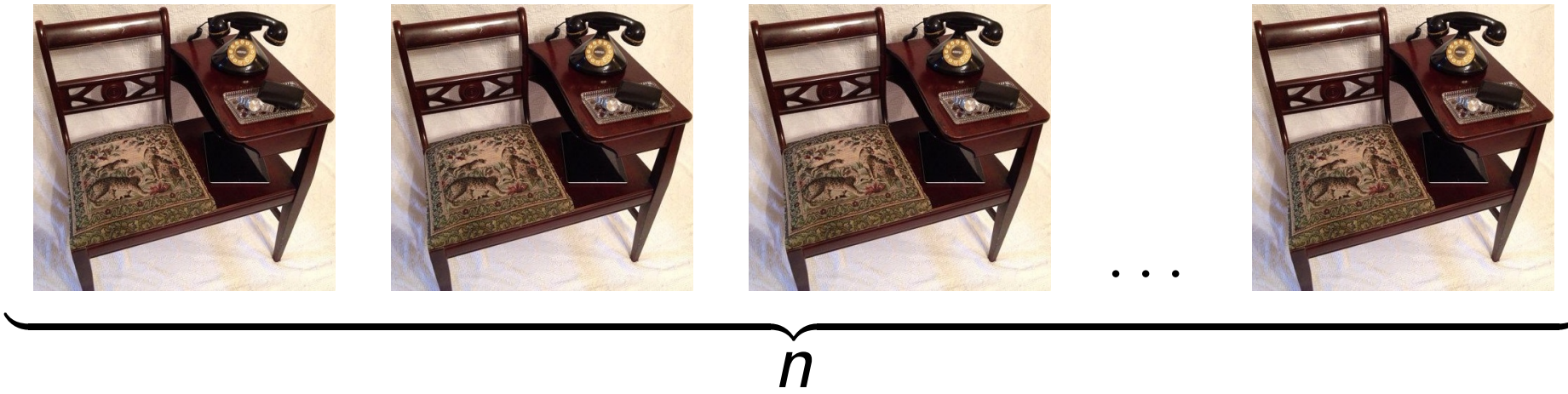
University of Manchester

17/06/2017

The Gossip Problem



The Gossip Problem



The Gossip Problem

Consider n people, each knowing a story unknown to the others. The people can communicate by phone, and in each phone call the participants share every story they know. What is the minimum number of phone calls required before everybody knows every story?

The Gossip Problem

The Gossip Problem

Consider n people, each knowing a story unknown to the others. The people can communicate by phone, and in each phone call the participants share every story they know. What is the minimum number of phone calls required before everybody knows every story?

This problem was studied in the 70s and the solution was found independently by Tijdeman, Baker and Shorstak, Hajnal, Milner and Szemeredi, and many others.

The Gossip Problem

The Gossip Problem

Consider n people, each knowing a story unknown to the others. The people can communicate by phone, and in each phone call the participants share every story they know. What is the minimum number of phone calls required before everybody knows every story?

This problem was studied in the 70s and the solution was found independently by Tijdeman, Baker and Shorstak, Hajnal, Milner and Szemeredi, and many others.

Solution

The solution is:

$$0 \text{ if } n = 1, \qquad 1 \text{ if } n = 2,$$
$$3 \text{ if } n = 3, \qquad 2n - 4 \text{ if } n \geq 4.$$

The Gossip Problem

The gossip monoid is an algebraic structure related to the gossip problem. When studying the gossip monoid, we are not concerned with the solution. Instead we are interested in the set up of the problem.

The Gossip Problem

The gossip monoid is an algebraic structure related to the gossip problem. When studying the gossip monoid, we are not concerned with the solution. Instead we are interested in the set up of the problem.

The problem has the following rules:

- n gossiping people (gossips) $g_1, \ldots, g_n$
- n scandalous stories (scandals) $s_1, \ldots, s_n$
- At first, g_i knows s_i for each i (and that's it)
- Information is only learned in phone calls
- In each phone call, every known scandal is shared.

The Gossip Problem

The gossip monoid is an algebraic structure related to the gossip problem. When studying the gossip monoid, we are not concerned with the solution. Instead we are interested in the set up of the problem.

The problem has the following rules:

- n gossiping people (gossips) $g_1, \ldots, g_n$
- n scandalous stories (scandals) $s_1, \ldots, s_n$
- At first, g_i knows s_i for each i (and that's it)
- Information is only learned in phone calls
- In each phone call, every known scandal is shared.

If you prefer, imagine networked computers sharing data rather than people spreading rumours.

Boolean Matrices

Definition (Boolean Semiring)

Define the boolean semiring as

$$\mathbb{B} = (\{0, 1\}, +, \times),$$

with $1 + 1 = 1$.

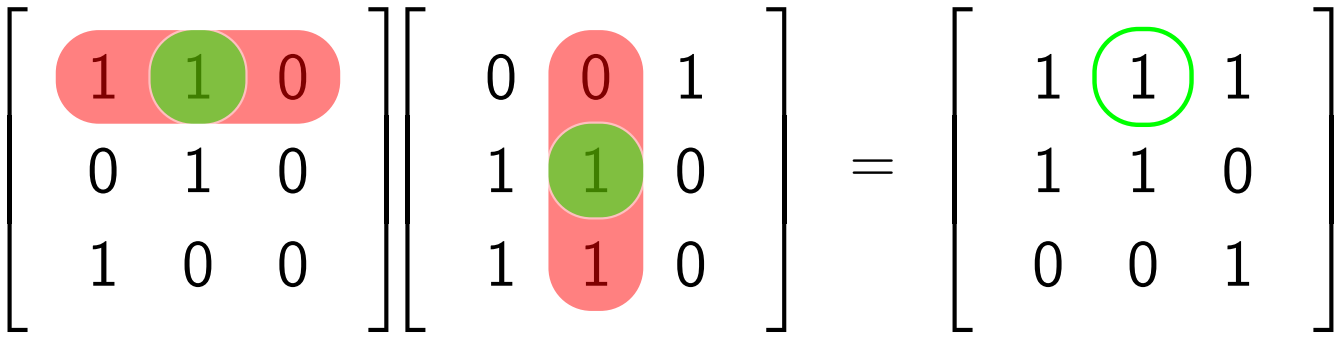
We will consider $n \times n$ matrices over $\mathbb{B}$. We write B_n for the set of all $n \times n$ boolean matrices. This set forms a monoid under multiplication.

Boolean Matrices

This multiplication simplifies as follows:
Let $C = AB$. Then

$$c_{i,j} = 1 \iff \exists k, a_{i,k} = b_{k,j} = 1.$$

Example:



Boolean Matrices and Binary Relations

There is an obvious connection between boolean matrices and binary relations.
Given a binary relation R on $\{1, \ldots, n\}$ we can define an $n \times n$ boolean matrix A by

$$a_{i,j} = 1 \iff R(i, j).$$

Boolean Matrices and Binary Relations

There is an obvious connection between boolean matrices and binary relations.
Given a binary relation R on $\{1, \ldots, n\}$ we can define an $n \times n$ boolean matrix A by

$$a_{i,j} = 1 \iff R(i, j).$$

By this map the boolean monoid is isomorphic to the monoid of binary relations.

Boolean Matrices and Binary Relations

There is an obvious connection between boolean matrices and binary relations.
Given a binary relation R on $\{1, \ldots, n\}$ we can define an $n \times n$ boolean matrix A by

$$a_{i,j} = 1 \iff R(i,j).$$

By this map the boolean monoid is isomorphic to the monoid of binary relations.
We will use the notation

$$\left[\, R(i,j) \, \right]_n$$

for the $n \times n$ matrix defined by the relation R.

The Gossip Monoid

- n gossips, $g_1, \ldots, g_n$
- n scandals, $s_1, \ldots, s_n$

We can represent a state of knowledge between n gossips and n scandals by the matrix

$$\left[\text{Gossip } g_j \text{ knows scandal } s_i \right]_n.$$

The Gossip Monoid

- n gossips, $g_1, \ldots, g_n$
- n scandals, $s_1, \ldots, s_n$

We can represent a state of knowledge between n gossips and n scandals by the matrix

$$\left[\text{Gossip } g_j \text{ knows scandal } s_i \right]_n.$$

- At first, g_i knows s_i for each i (and that's it)

This state of knowledge is represented by the identity matrix.

The Gossip Monoid

- Information is only learned in phone calls
- In each phone call, every known scandal is shared.

Definition (Phone call matrix)

Given $k, l \leq n$, let $C[k, l]$ be the matrix

$$C[k,l] = \left[\, i = j \text{ or } \{i,j\} = \{k,l\} \, \right]_n .$$

We call this a *phone call matrix*.

$$C[k,l] = \begin{array}{c} \\ \\ k \\ \\ l \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} & & k & & l & & \end{array} \\ \left[\begin{array}{ccccccc} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & 1 & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & 1 & & 1 & & \\ & & & & & 1 & \end{array} \right] \end{array}$$

The Gossip Monoid

If matrices A and B represent states of knowledge before and after a phone call between gossips g_k and g_l, then we have

$$B = AC[k, l].$$

In this way, right multiplication by a phone call matrix represents a phone call.

The Gossip Monoid

If matrices A and B represent states of knowledge before and after a phone call between gossips g_k and g_l, then we have

$$B = AC[k, l].$$

In this way, right multiplication by a phone call matrix represents a phone call.
Example:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

The Gossip Monoid

Every state of knowledge which can be obtained in the gossip problem is the result of applying phone calls to the initial state of knowledge (which is represented by the identity matrix.)

The Gossip Monoid

Every state of knowledge which can be obtained in the gossip problem is the result of applying phone calls to the initial state of knowledge (which is represented by the identity matrix.)

Therefore every obtainable state of knowledge is represented by a product of phone call matrices.

The Gossip Monoid

Every state of knowledge which can be obtained in the gossip problem is the result of applying phone calls to the initial state of knowledge (which is represented by the identity matrix.)

Therefore every obtainable state of knowledge is represented by a product of phone call matrices.

So the monoid generated by the phone call matrices is precisely the set of matrices which represent obtainable states of knowledge.

The Gossip Monoid

Definition (Gossip Monoid)

Given $n \in \mathbb{N}$, the gossip monoid G_n is the submonoid of B_n generated by the phone call matrices.

$$G_n = \langle C[a, b] : a, b \leq n \rangle .$$

The Gossip Monoid

Definition (Gossip Monoid)

Given $n \in \mathbb{N}$, the gossip monoid G_n is the submonoid of B_n generated by the phone call matrices.

$$G_n = \langle C[a, b] : a, b \leq n \rangle .$$

This is a $\mathcal{J}$-trivial, idempotent generated monoid.

It's idempotent generated because phone calls are idempotent: if you phone somebody and share all information with them, then immediately phone them again, there's no new information to share.

It's J-trivial because multiplying A on either side by a phone call matrix can only add 1s and not remove them.

Questions

How many elements are there in G_n?

Andries Brouwer, Jan Draisma, and Bart Frenk have computed this up to $n = 9$.

n	$\lvert G_n \rvert$
1	1
2	2
3	11
4	189
5	9,152
6	1,092,473
7	293,656,554
8	166,244,338,221
9	188,620,758,836,916

Finding the size of G_n for general n is an open problem.

Questions

What is the structure of G_n like?

Since G_n is $\mathcal{J}$-trivial, this comes down to understanding the $\mathcal{J}$-order.

Results

Theorem (Brouwer, Draisma, Frenk 2014)

Any sequence of phone calls among n gossiping parties such that in each phone call both participants exchange all they know, and at least one of the parties learns something new, has length at most $\binom{n}{2}$*, and this bound is attained.*

Results

Theorem (Brouwer, Draisma, Frenk 2014)

Any sequence of phone calls among n gossiping parties such that in each phone call both participants exchange all they know, and at least one of the parties learns something new, has length at most $\binom{n}{2}$*, and this bound is attained.*

As a result of this, any element of G_n can be written as a product of at most $\binom{n}{2}$ phone call matrices.

Results

Theorem

There is a one to one correspondence between the idempotents of G_n and the equivalence relations on $\{1, \ldots, n\}$ via the map

$$\sim \mapsto \left[i \sim j \right]_n.$$

Results

Theorem

There is a one to one correspondence between the idempotents of G_n and the equivalence relations on $\{1, \ldots, n\}$ via the map

$$\sim \mapsto \left[i \sim j \right]_n.$$

Corollary

The number of idempotents of G_n is equal to the nth Bell number.

Results

Any state of knowledge represented by an element of the gossip monoid can be achieved through a sequence of phone calls such that in each call somebody learns something new.

Results

Any state of knowledge represented by an element of the gossip monoid can be achieved through a sequence of phone calls such that in each call somebody learns something new.

Theorem

Any state of knowledge represented by an element of the gossip monoid can be achieved through a sequence of **conference calls** *such that in each call* **every participant** *learns something new.*

Results

Any state of knowledge represented by an element of the gossip monoid can be achieved through a sequence of phone calls such that in each call somebody learns something new.

Theorem

Any state of knowledge represented by an element of the gossip monoid can be achieved through a sequence of **conference calls** *such that in each call* **every participant** *learns something new.*

Definition (Conference call matrix)

Given $S \subseteq \{1, \ldots, n\}$, let $C[S]$ be the matrix

$$C[S] = \left[\, i = j \text{ or } \{i, j\} \subseteq S \, \right]_n .$$

Computational Complexity

An obvious question to ask is:

Gossip Membership Problem

Given $A \in B_n$, is $A \in G_n$?

This problem is NP-complete.

Computational Complexity

An obvious question to ask is:

Gossip Membership Problem

Given $A \in B_n$, is $A \in G_n$?

This problem is NP-complete.

Each element of G_n can be written as a product of at most $\binom{n}{2}$ phone call matrices.
A NDTM can compute a product of $\binom{n}{2}$ or fewer phone call matrices in polynomial time.
The Gossip Membership Problem can be solved by checking if this product is equal to A.

Computational Complexity

NP-hardness is shown with a polynomial time reduction from the Dominating Set Problem.

Definition (Dominating Set)

A dominating set of a graph $H = (V, E)$ is a subset of vertices $D \subset V$ such that

$$v \in V \setminus D \implies \exists d \in D,\ \{d, v\} \in E.$$

Definition (Dominating Set Problem)

Given a graph $H = (V, E)$ and a natural number $0 < k < |V|$, is there a dominating set for H with size k?

Computational Complexity

- Given a graph $H = (V, E)$ and a natural number $0 < k < |V|$, we assume that $V = \{1, \ldots, n\}$ and define

$$M = \Big[(i = j) \text{ or } (i \text{ and } j \text{ are adjacent in } H) \Big]_n .$$

H has a dominating set of size k if and only if there is a set of k columns of M which between them have a 1 in each row.

Computational Complexity

- Given a graph $H = (V, E)$ and a natural number $0 < k < |V|$, we assume that $V = \{1, \ldots, n\}$ and define

$$M = \Big[(i = j) \text{ or } (i \text{ and } j \text{ are adjacent in } H) \Big]_n .$$

  H has a dominating set of size k if and only if there is a set of k columns of M which between them have a 1 in each row.

- Consider the following partial order on the column vectors of M:

$$x \preceq y \iff (x_i = 1 \Rightarrow y_i = 1) \text{ for all } i.$$

Computational Complexity

- Given a graph $H = (V, E)$ and a natural number $0 < k < |V|$, we assume that $V = \{1, \ldots, n\}$ and define

$$M = \Big[(i = j) \text{ or } (i \text{ and } j \text{ are adjacent in } H) \Big]_n .$$

  H has a dominating set of size k if and only if there is a set of k columns of M which between them have a 1 in each row.

- Consider the following partial order on the column vectors of M:

$$x \preceq y \iff (x_i = 1 \Rightarrow y_i = 1) \text{ for all } i.$$

- M is non-zero so it has at least one non-zero maximal column under this partial order. Choose one such column and call it m.

Computational Complexity

- Given a graph $H = (V, E)$ and a natural number $0 < k < |V|$, we assume that $V = \{1, \ldots, n\}$ and define

$$M = \Big[(i = j) \text{ or } (i \text{ and } j \text{ are adjacent in } H) \Big]_n .$$

  H has a dominating set of size k if and only if there is a set of k columns of M which between them have a 1 in each row.

- Consider the following partial order on the column vectors of M:

$$x \preceq y \iff (x_i = 1 \Rightarrow y_i = 1) \text{ for all } i.$$

- M is non-zero so it has at least one non-zero maximal column under this partial order. Choose one such column and call it m.
- Replace every non-maximal column of M with m and call the resulting matrix M' (Note that this can be done in polynomial time.)
  H has a dominating set of size k if and only if there is a set of k columns of M' which between them have a 1 in each row.

Computational Complexity

M' is the result of replacing each non-maximal column M with m. Now let

$$A = \begin{array}{c} \overbrace{}^{n}\overbrace{}^{n}\overbrace{}^{n} \\ \left[\begin{array}{c|c|c} M' & 0 & 0 \\ \hline 0 & 1 & 1 \\ \hline 0 & 0 & 0 \end{array}\right] \end{array} \begin{array}{l} \} n \\ \} n \\ \} n \end{array} \qquad B = \begin{array}{c} \overbrace{}^{n}\overbrace{}^{n}\overbrace{}^{k}\overbrace{}^{n-k} \\ \left[\begin{array}{c|c|c|c} M' & M' & 1 & 0 \\ \hline 1 & 1 & 1 & \\ \hline 0 & 0 & 0 & \end{array}\right] \end{array} \begin{array}{l} \} n \\ \} n \\ \} n \end{array}.$$

Now there is a set of k columns of M' which between them have a 1 in each row if and only if there exists $G \in G_{3n}$ such that $AG = B$.

Computational Complexity

With A and B defined as above, there exists $G \in G_{3n}$ such that $AG = B$ if and only if

$$\begin{array}{c} \overbrace{}^{9n^2} \quad \overbrace{}^{3n} \quad \overbrace{}^{3n} \quad \overbrace{}^{3n} \quad \overbrace{}^{3n} \\ \left[\begin{array}{c|c|c|c|c} 1 & A & A & 0 & B \\ \hline 1 & 1 & 1 & 0 & 1 \\ \hline 0 & 0 & 1 & I_{3n} & 1 \\ \hline 0 & I_{3n} & 1 & I_{3n} & 1 \\ \hline 0 & I_{3n} & 1 & I_{3n} & 1 \end{array}\right] \begin{array}{l} \} 3n \\ \} 9n^2 \\ \} 3n \\ \} 3n \\ \} 3n \end{array} \end{array} \in G_{9n^2+12n}.$$

This completes the polynomial time reduction from the Dominating Set Problem to the Gossip Membership Problem.

Computational Complexity

Other decision problems:

Definition (Gossip Transformation Problem)

Given $A, B \in B_n$, is there a $G \in G_n$ such that $AG = B$?

Computational Complexity

Other decision problems:

Definition (Gossip Transformation Problem)

Given $A, B \in B_n$, is there a $G \in G_n$ such that $AG = B$?

Definition (Gossip $\mathcal{L}$-Order Problem)

Given $X, Y \in G_n$, is there a $U \in G_n$ such that $UY = X$?

Definition (Gossip $\mathcal{R}$-Order Problem)

Given $X, Y \in G_n$, is there a $V \in G_n$ such that $YV = X$?

Definition (Gossip $\mathcal{J}$-Order Problem)

Given $X, Y \in G_n$, are there $U, V \in G_n$ such that $UYV = X$?

Computational Complexity

Other decision problems:

Definition (Gossip Transformation Problem)

Given $A, B \in B_n$, is there a $G \in G_n$ such that $AG = B$?

Definition (Gossip $\mathcal{L}$-Order Problem)

Given $X, Y \in G_n$, is there a $U \in G_n$ such that $UY = X$?

Definition (Gossip $\mathcal{R}$-Order Problem)

Given $X, Y \in G_n$, is there a $V \in G_n$ such that $YV = X$?

Definition (Gossip $\mathcal{J}$-Order Problem)

Given $X, Y \in G_n$, are there $U, V \in G_n$ such that $UYV = X$?

All of these problems are NP-complete.

References

NP-completeness in the gossip monoid - P. Fenner, M. Johnson, M. Kambites
https://arxiv.org/abs/1606.01026

Lossy gossip and composition of metrics - A. Brouwer, J. Draisma, B. Frenk
https://arxiv.org/abs/1405.5979

Gossips and telephones - B. Baker, R. Shostak
http://www.sciencedirect.com/science/article/pii/0012365X72900015