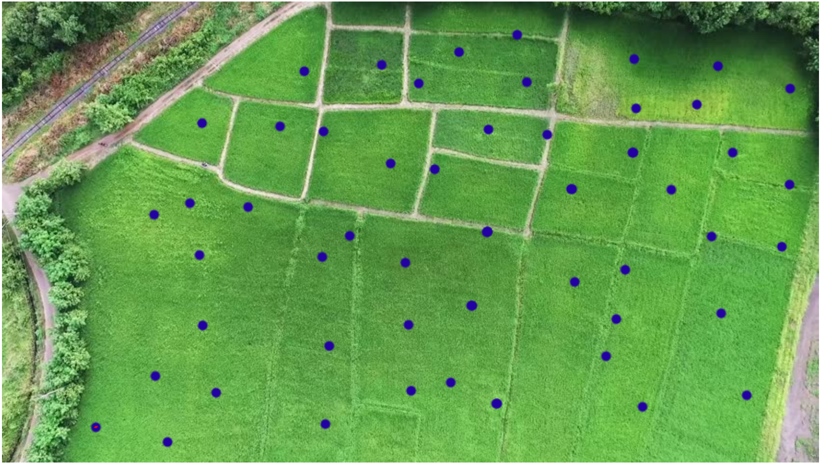
Probabilistic forwarding of coded packets on networks

Vinay Kumar B R, Navin Kashyap
Indian Institute of Science

IEEE International Symposium on Information Theory (ISIT) - 2019
11 Jul, 2019
Paris, France

Motivating example

Motivating example



Sensors in a field

Motivating example

Sensors in a field

Motivating example

Sensors in a field

Update sensing parameters among all nodes

Network of IoT nodes

Over-the-air programming of the nodes

Motivating example

Sensors in a field

Update sensing parameters among all nodes

Network of IoT nodes

Over-the-air programming of the nodes

Broadcast information with the following constraints

Nodes:

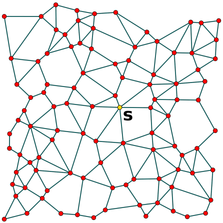
- Energy constrained
- Limited computational ability
- Limited knowledge of the network

Algorithm:

- Completely distributed
- Run in finite time

Abstraction

- **Broadcast over an ad-hoc network**



A network of nodes $\mathcal{G} = (V, E)$.
$|V| = N$.
A source node s has k data packets to be broadcast to all the other nodes in an ad-hoc network.

Flooding

- Each node forwards every received packet to all its one-hop neighbours.
- Subsequent receptions of the same packet are neglected.

- Total number of transmissions $= kN$.
- Wasteful [1]: nodes will receive same packet multiple times

[1] Y.-C. Tseng, S.-Y. Ni, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," Wireless Networks, vol. 8, no. 2/3, pp. 153–167, 2002.

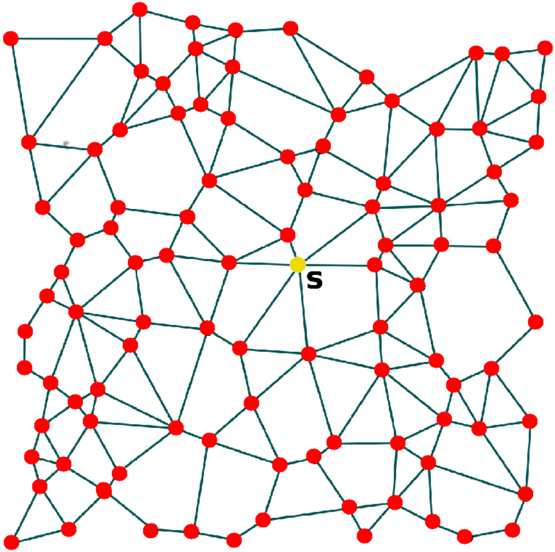
Probabilistic forwarding

- Retransmission probability p.
- Source transmits all k packets to all its neighbors.
- Each node, upon receiving packet $\#j$ for the first time, forwards it to all its neighbours with probability p; does nothing with probability $1 - p$.
- Subsequent receptions of the same packet are neglected.

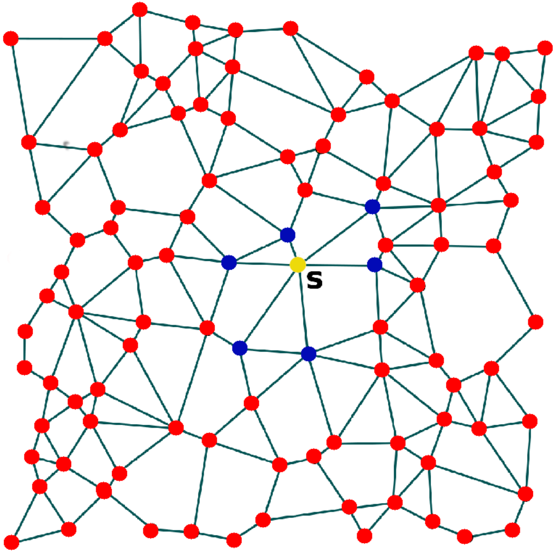
[1] Y. Sasson, D. Cavin, and A. Schiper, "Probabilistic broadcast for flooding in wireless mobile ad hoc networks," in Proc. IEEE Wireless Communications and Networking Conf. (WCNC) 2003, vol. 2, March 16–20, 2003, pp. 1124–1130.

[1] Z. J. Haas, J. Y. Halpern, and L. Li, "Gossip-based ad hoc routing," IEEE/ACM Trans. Networking, vol. 14, no. 3, pp. 479–491, 2006.

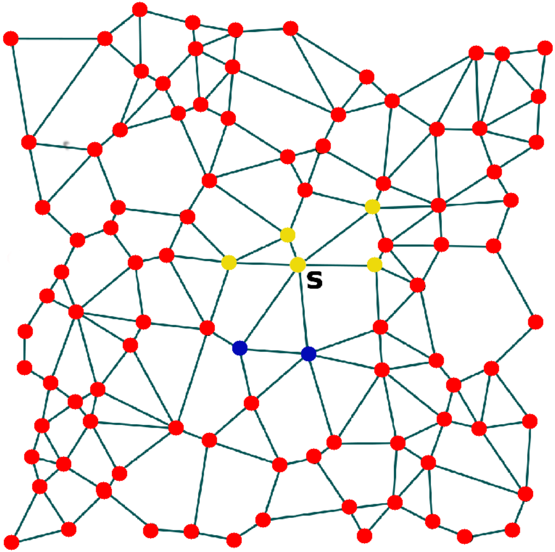
Probablistic forwarding



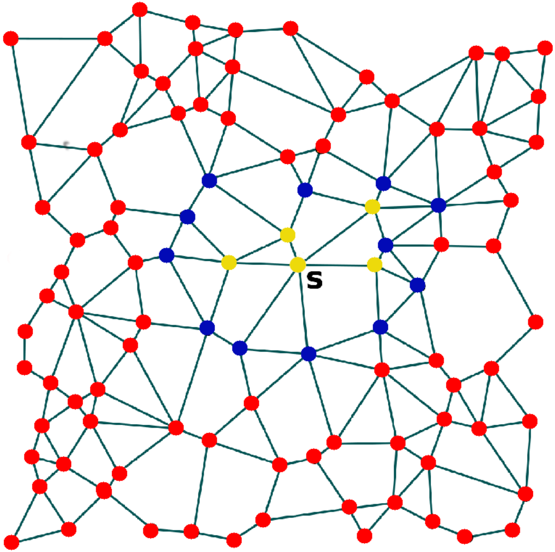
Probablistic forwarding



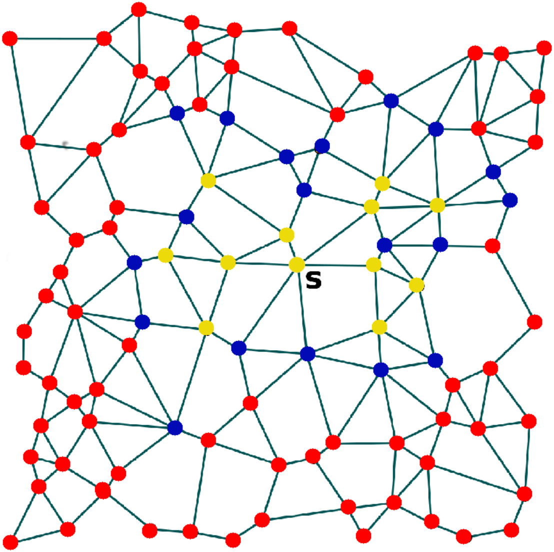
Probablistic forwarding



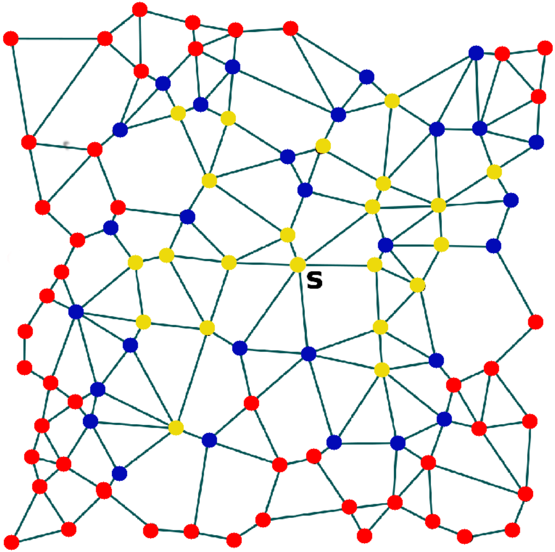
Probablistic forwarding



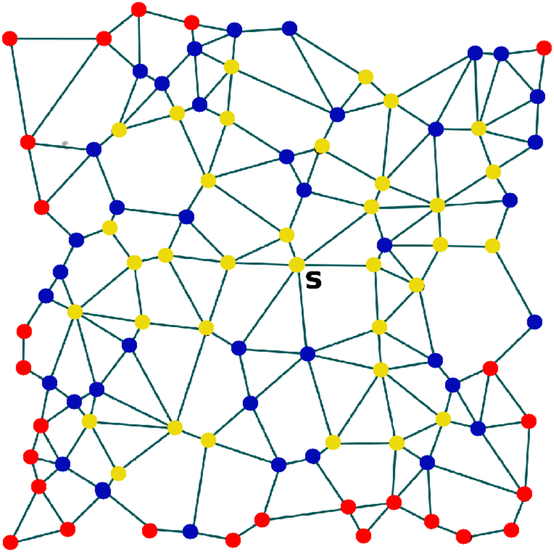
Probablistic forwarding



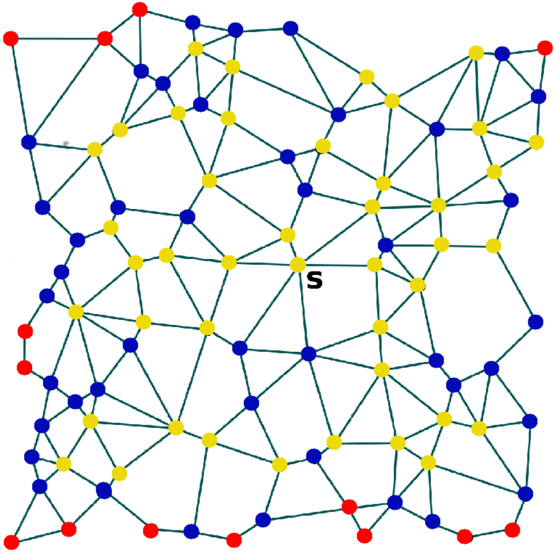
Probablistic forwarding



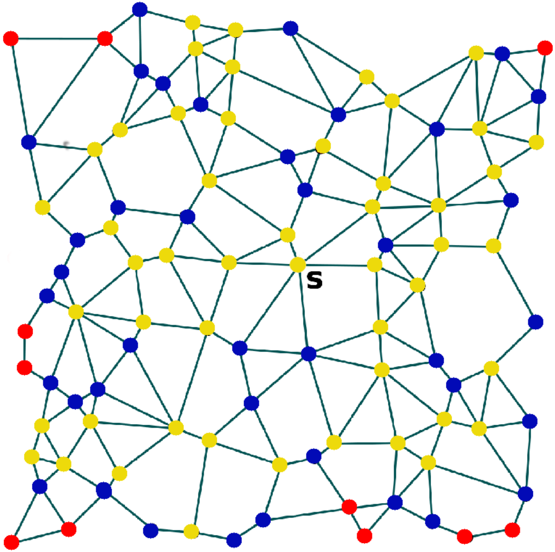
Probablistic forwarding



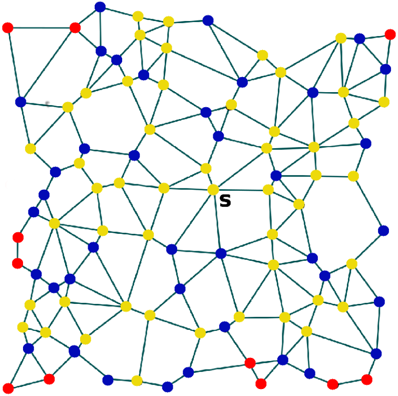
Probablistic forwarding



Probablistic forwarding



Probabilistic forwarding



Lesser number of transmissions (kNp) compared to flooding (kN).

BUT

Information lost even if a single packet out of the k packets is not received.

Introducing Coded Packets

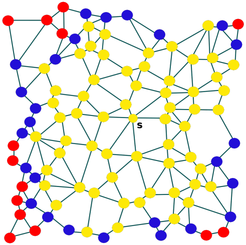
Coding:

- The source node encodes the k data packets into n coded packets using a Maximum Distance Separable (MDS) code.
- MDS code ensures that reception of any k of the n coded packets by any node, suffices to recover the original k data packets.

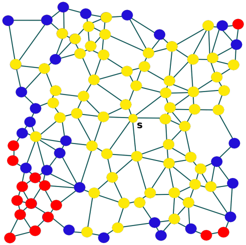
Probabilistic forwarding of coded packets:

- Source node transmits all n coded packets to its one-hop neighbours.
- Thereafter, the probabilistic forwarding protocol takes over. Each packet is forwarded independently of other packets and other nodes.
- Nodes which receive at least k out of n packets are termed successful receivers
- A near-broadcast is when the expected fraction of successful receivers is $\geq 1 - \delta$.

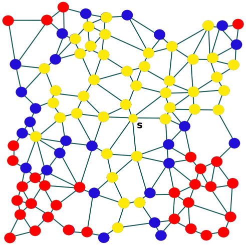
Illustration



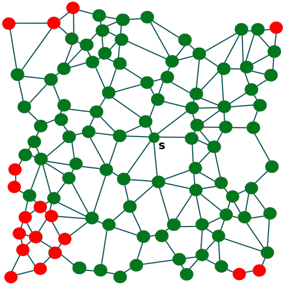
Packet 1



Packet 2



Packet 3



Successful receivers

- $k = 2, n = 3$
- (yellow) - Received and transmitted the packet.
- (blue) - Received but did not transmit the packet.
- (red) - Did not receive the packet.
- (green) - Successful receivers

Formal Problem Statement

Given:

- a connected graph $\mathcal{G}$ with N nodes
- number of data packets, k
- number of coded packets, n
- δ close to 0.
- retransmission probability p

Define

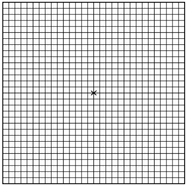
- $R_{k,n}$ = # nodes that receive at least k out of n coded packets

Want to find

- $p_{k,n,\delta}$ = minimum p such that $\mathbb{E}_p[\frac{1}{N} R_{k,n}] \geq 1 - \delta$
- $\tau_{k,n,\delta} = \mathbb{E}_{p_{k,n,\delta}}[\text{total \# transmissions over all } N \text{ nodes of } \mathcal{G}]$

On what graphs is coding along with probabilistic forwarding beneficial?

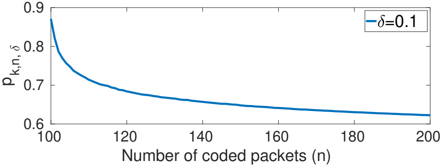
On well-connected graphs



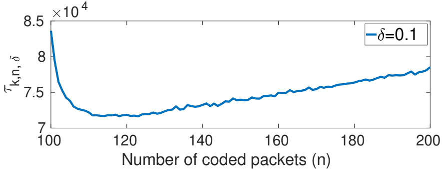
31 × 31 grid

1. $p_{k,n,\delta}$ decreases (to 0) as n increases; for fixed k and δ.
2. $\tau_{k,n,\delta} = \sum_{i=1}^{n} T_i$, where T_i is the expected number of transmissions of packet i.
    - Each term T_i decreases as n increases since $p_{k,n,\delta}$ decreases.
    - The number of terms in the above equation increases with n.

$k = 100$ packets and $\delta = 0.1$



Minimum forwarding probability

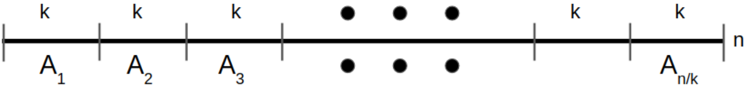


Expected total number of transmissions

Why $p_{k,n,\delta} \searrow 0$?

$R_{k,n}$ = # nodes that receive at least k out of n coded packets.

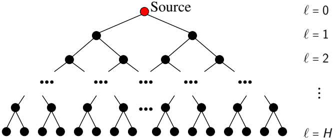
$p_{k,n,\delta} = \inf\{p \mid \mathbb{E}_p[\frac{1}{N} R_{k,n}] \geq 1 - \delta\}$

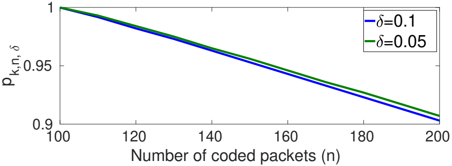


- For $j = 1, 2, \cdots, \lfloor \frac{n}{k} \rfloor$, let A_j be the event that the jth set of k coded packets is received by at least $1 - \delta/2$ fraction of the nodes.
- The events A_j are mutually independent and have the same probability of occurrence.
- For any $p > 0$, we have $P(A_j)$ being small but strictly positive.
- Hence, $P(\text{at least one } A_j \text{ occurs}) \geq 1 - \delta/2$ for all sufficiently large n, so that $\frac{1}{N} R_{k,n} \geq 1 - \delta/2$ with probability at least $1 - \delta/2$.
- Thus, for any $p > 0$, we have $p_{k,n,\delta} \leq p$ for sufficiently large n.

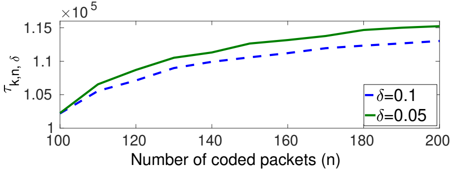
On trees

Rooted binary tree of height $H = 10$ with $k = 100$ and $\delta = 0.1$





Minimum forwarding probability



Expected total number of transmissions

- Large fraction of nodes on the leaves
- Unique path from the source to any node on the tree
- It can be shown that

$$p_{k,n,\delta} \approx c\left(\frac{k}{n}\right)^{\frac{1}{H-1}}$$

for some constant c dependent on H and δ.

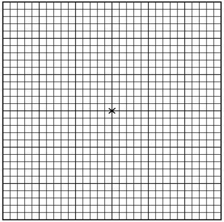
- This gives

$$\tau_{k,n,\delta} \approx n\,\frac{(2c)^H\left(\frac{k}{n}\right)^{\frac{H}{H-1}}-1}{2c\left(\frac{k}{n}\right)^{\frac{1}{H-1}}-1}$$

which is increasing in n.

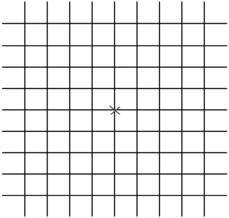
On the grid

Probabilistic forwarding on the $m \times m$ grid, Γ_m



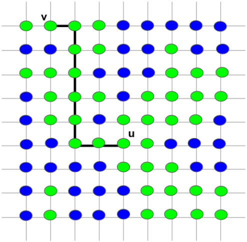
$\xrightarrow{m \to \infty}$

Probabilistic forwarding on the $\mathbb{Z}^2$ lattice



$\downarrow$

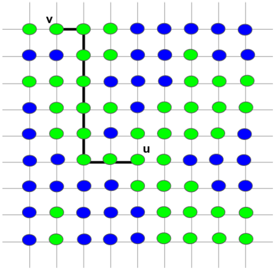
Site percolation on the $\mathbb{Z}^2$ lattice



We will see that the site percolation process on $\mathbb{Z}^2$ is a faithful model for the probabilistic forwarding mechanism on $\mathbb{Z}^2$.

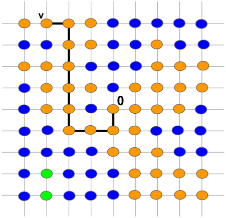
Site percolation on $\mathbb{Z}^2$ - Definitions

- Associate each vertex (site) u of $\mathbb{Z}^2$ with a Bernoulli(p) random variable X_u. The vertex is open if $X_u = 1$; else closed.
- Two open sites u and v are said to be connected by an open path ($u \longleftrightarrow v$), if there is a path of open sites from u to v.



Site percolation on $\mathbb{Z}^2$- Transmitters

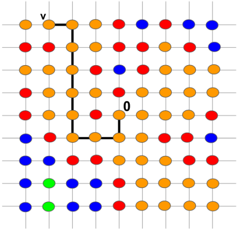
- Associate each vertex (site) u of $\mathbb{Z}^2$ with a Bernoulli(p) random variable X_u. The vertex is open if $X_u = 1$; else closed.
- Two open sites u and v are said to be connected by an open path ($u \longleftrightarrow v$), if there is a path of open sites from u to v.



- Probabilistic forwarding of a single packet over $\mathbb{Z}^2$ is modelled by site percolation on $\mathbb{Z}^2$ conditioned on the origin $\mathbf{0}$ being open.
- Nodes transmitting the j**th packet** (for fixed $j \in [n]$) may be viewed as open sites of a site percolation process which are connected to the origin by an open path. Call this cluster of nodes as C_j^o.
- The total number of transmissions is simply $\sum_{j=1}^{n} |C_j^o|$.

Site percolation on $\mathbb{Z}^2$ – *Receivers*

- The boundary, ∂C_j^o, is the set of all closed sites which are adjacent to a site in C_j^o.
- The set $C_j^{o+} := C_j^o \cup \partial C_j^o$ is called the extended cluster of the origin.



For site percolation on $\mathbb{Z}^2$, there exists $p_c \in (0,1)$ s.t. for $p > p_c$,

- There exists a unique infinite open cluster (IOC), C, almost surely. $p_c \approx 0.59$ for site percolation.
- Hence, there also exists a unique infinite extended cluster (IEC), $C^+ = C \cup \partial C$, a.s..
- $\theta(p) :=$ percolation probability, i.e., $\mathbb{P}(\mathbf{0} \in C)$
  $\theta^+(p) := \mathbb{P}(\mathbf{0} \in C^+)$

Lemma: $\theta^+(p) = \frac{\theta(p)}{p}$.
Proof: $\{\mathbf{0} \in C\} = \{\mathbf{0} \in C^+ \text{ and } \mathbf{0} \text{ is open}\}$

Ergodic theorems

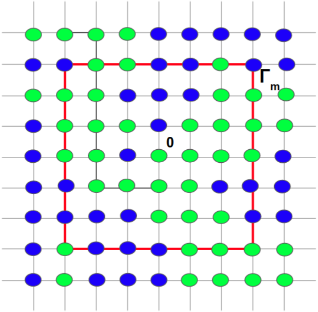
For $p > p_c$, if C is the unique IOC and C^+ is the unique IEC, then,

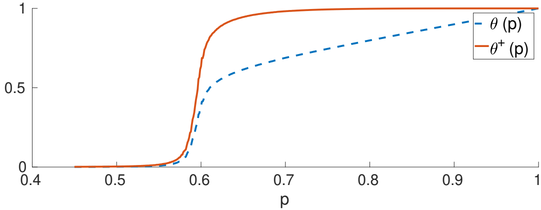
$$\lim_{m\to\infty} \frac{|C \cap \Gamma_m|}{m^2} = \theta(p) \qquad a.s.$$

Similarly,

$$\lim_{m\to\infty} \frac{|C^+ \cap \Gamma_m|}{m^2} = \theta^+(p) \qquad a.s.$$

Using DCT, expected values also converge.





Site percolation and probabilistic forwarding

- Prob. forwarding of a single packet over $\mathbb{Z}^2$ is modelled by site percolation on $\mathbb{Z}^2$ conditioned on the origin $\mathbf{0}$ being open.
- n pkts $\leftrightarrow$ n independent site percolations with $\mathbf{0}$ open in all.
- $\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m) :=$ $\{$sites in Γ_m that receive at least k out of n pkts$\}$
- We are interested in finding

$$p_{k,n,\delta} = \min\left\{p \;\middle|\; \mathbb{E}_p\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m)\right|\right] \geq 1-\delta\right\}$$

Site percolation and probabilistic forwarding

- Prob. forwarding of a single packet over $\mathbb{Z}^2$ is modelled by site percolation on $\mathbb{Z}^2$ conditioned on the origin $\mathbf{0}$ being open.
- n pkts $\leftrightarrow$ n independent site percolations with $\mathbf{0}$ open in all.
- $\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m) :=$ $\{$sites in Γ_m that receive at least k out of n pkts$\}$
- We are interested in finding

$$p_{k,n,\delta} = \min\left\{p \;\middle|\; \mathbb{E}_p\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m)\right|\right] \geq 1-\delta\right\}$$

Theorem

For $p > p_c$, we have

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

Total transmissions

Consider the transmissions of a single packet on $\mathbb{Z}^2$

- C^o := open cluster containing $\mathbf{0}$ (set of transmitters)
- The expected number of transmissions on a large grid, Γ_m, is given by $\mathbb{E}\big[|C^o \cap \Gamma_m| \;\big|\; \mathbf{0} \text{ is open}\big]$.

Proposition

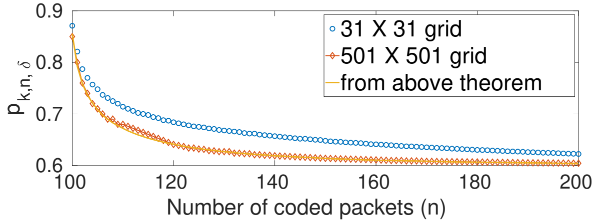
For $p > p_c$, we have

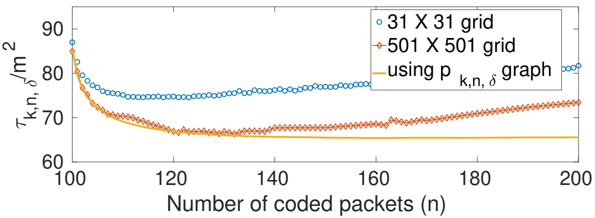
$$\lim_{m\to\infty} \frac{1}{m^2}\mathbb{E}\big[|C^o \cap \Gamma_m| \;\big|\; \mathbf{0} \textit{ is open}\big] = \frac{\theta(p)^2}{p}.$$

- Hence, for n coded packets, with each packet being transmitted with probability $p_{k,n,\delta} > p_c$, we obtain

$$\tau_{k,n,\delta} \approx nm^2 \frac{\theta^2(p_{k,n,\delta})}{p_{k,n,\delta}}.$$

Comparison with simulation results





Proof sketch

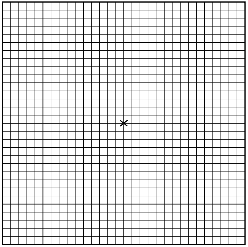
We want to show,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

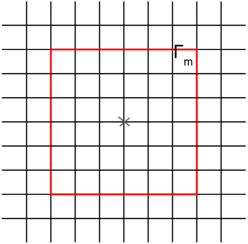
Step 1: $\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2) := \{\text{sites that receive at least } k \text{ out of } n \text{ pkts on } \mathbb{Z}^2\}$

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|\right]$$

Probabilistic forwarding on the $m \times m$ grid, Γ_m



Probabilistic forwarding on the $\mathbb{Z}^2$ lattice



Proof sketch

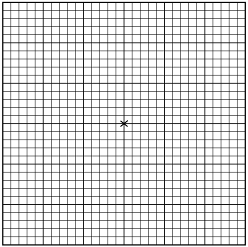
We want to show,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

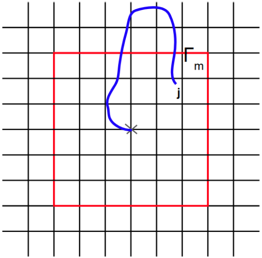
Step 1: $\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2) :=$ {sites that receive at least k out of n pkts on $\mathbb{Z}^2$}

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|\right]$$

Probabilistic forwarding on the $m \times m$ grid, Γ_m



Probabilistic forwarding on the $\mathbb{Z}^2$ lattice



Proof sketch

We want to show,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

Step 1: $\mathcal{R}_{k,n}^{pr.fwd}(\mathbb{Z}^2) := \{$sites that receive at least k out of n pkts on $\mathbb{Z}^2\}$

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\Gamma_m)\right|\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}_{k,n}^{pr.fwd}(\mathbb{Z}^2)\cap\Gamma_m\right|\right]$$

Step 2: Let $\mathcal{R}_{k,n}^{perc}(\mathbb{Z}^2) := \{$sites that are in at least k out of n C^+s$\}$
Using ergodicity, we show that for $p > p_c$,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{\left|\mathcal{R}_{k,n}^{perc}(\mathbb{Z}^2)\cap\Gamma_m\right|}{m^2}\right] = \underbrace{\sum_{j=k}^{n}\binom{n}{j}(\theta^+(p))^{j}(1-\theta^+(p))^{n-j}}_{\equiv\ \theta_{k,n}^+(p)}.$$

Proof sketch

We want to show,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

Step 1: $\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2) := \{$sites that receive at least k out of n pkts on $\mathbb{Z}^2\}$

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|\right]$$

Step 2: Let $\mathcal{R}^{perc}_{k,n}(\mathbb{Z}^2) := \{$sites that are in at least k out of n C^+s$\}$
Using ergodicity, we show that for $p > p_c$,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{\left|\mathcal{R}^{perc}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|}{m^2}\right] = \underbrace{\sum_{j=k}^{n}\binom{n}{j}(\theta^+(p))^{j}(1-\theta^+(p))^{n-j}}_{\equiv\ \theta^+_{k,n}(p)}.$$

Recall:

- $\theta^+(p) = \mathbb{P}(\mathbf{0}\in C^+)$
- $\lim_{m\to\infty} \frac{|C^+\cap\Gamma_m|}{m^2} = \theta^+(p) \qquad a.s.$

Proof sketch

We want to show,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j}.$$

Step 1: $\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2) :=$ {sites that receive at least k out of n pkts on $\mathbb{Z}^2$}

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\Gamma_m)\right|\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|\right]$$

Step 2: Let $\mathcal{R}^{perc}_{k,n}(\mathbb{Z}^2) :=$ {sites that are in at least k out of n C^+s}
Using ergodicity, we show that for $p > p_c$,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{\left|\mathcal{R}^{perc}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|}{m^2}\right] = \underbrace{\sum_{j=k}^{n}\binom{n}{j}(\theta^+(p))^{j}(1-\theta^+(p))^{n-j}}_{\equiv\ \theta^+_{k,n}(p)}.$$

Step 3: Carefully relate $\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)$ and $\mathcal{R}^{perc}_{k,n}(\mathbb{Z}^2)$ to obtain,

$$\lim_{m\to\infty} \mathbb{E}\left[\frac{1}{m^2}\left|\mathcal{R}^{pr.fwd}_{k,n}(\mathbb{Z}^2)\cap\Gamma_m\right|\right] = \sum_{t=k}^{n}\sum_{\substack{T\subseteq[n]:\\|T|=t}}\theta^+_{k,t}(p)(\theta^+(p))^{t}(1-\theta^+(p))^{n-t}.$$

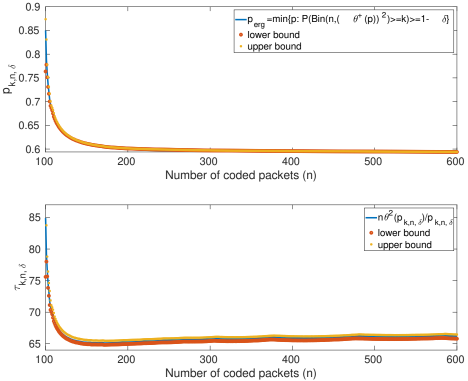
Some bounds

It can be shown that

$$\sum_{t=k}^{n}\sum_{j=k}^{t}\binom{n}{t}\binom{t}{j}(\theta^+(p))^{t+j}(1-\theta^+(p))^{n-j} = Pr(Y \geq k)$$

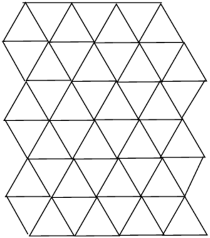
where $Y \sim Bin\left(n, (\theta^+(p))^2\right)$. So, for large grids, we have

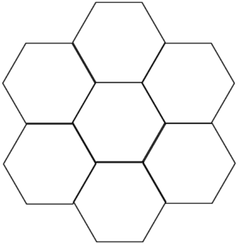
$$p_{k,n,\delta} = \inf\{p : Pr(Y \geq k) \geq 1-\delta\} \text{ and } \tau_{k,n,\delta} \approx nm^2\frac{\theta^2(p_{k,n,\delta})}{p_{k,n,\delta}}.$$



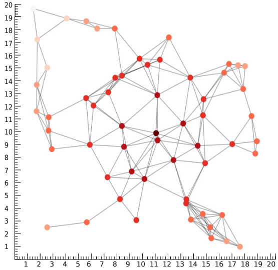
Other graphs

Lattices- triangular, hexagonal.

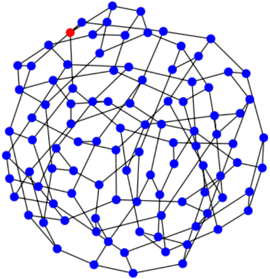




Random geometric graphs



Random regular graphs



Thank you

☺