Maximal Independent Sets in Radio Networks

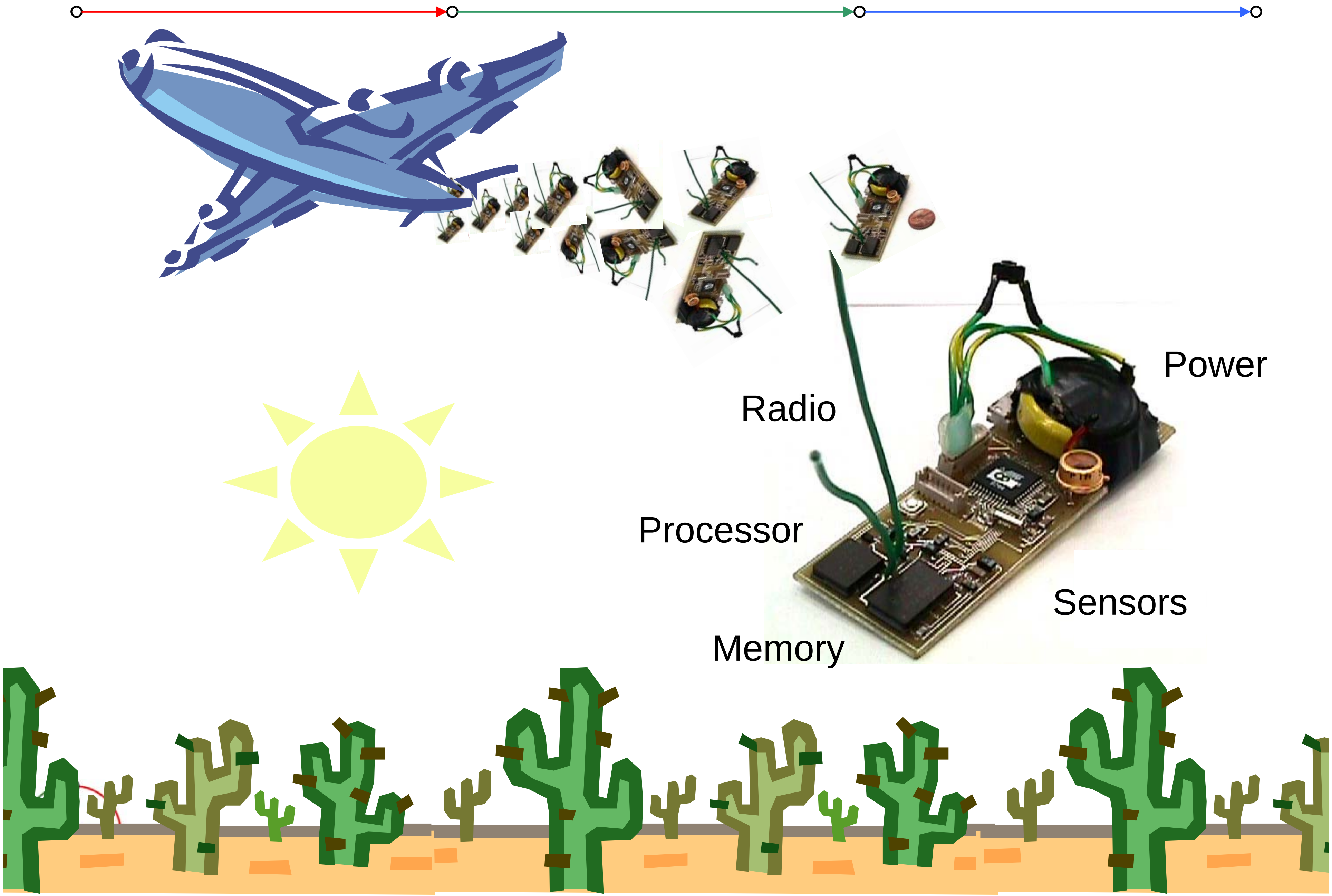
Thomas Moscibroda
Roger Wattenhofer

Distributed Computing Group

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

PODC 2005

Wireless Ad Hoc and Sensor Networks

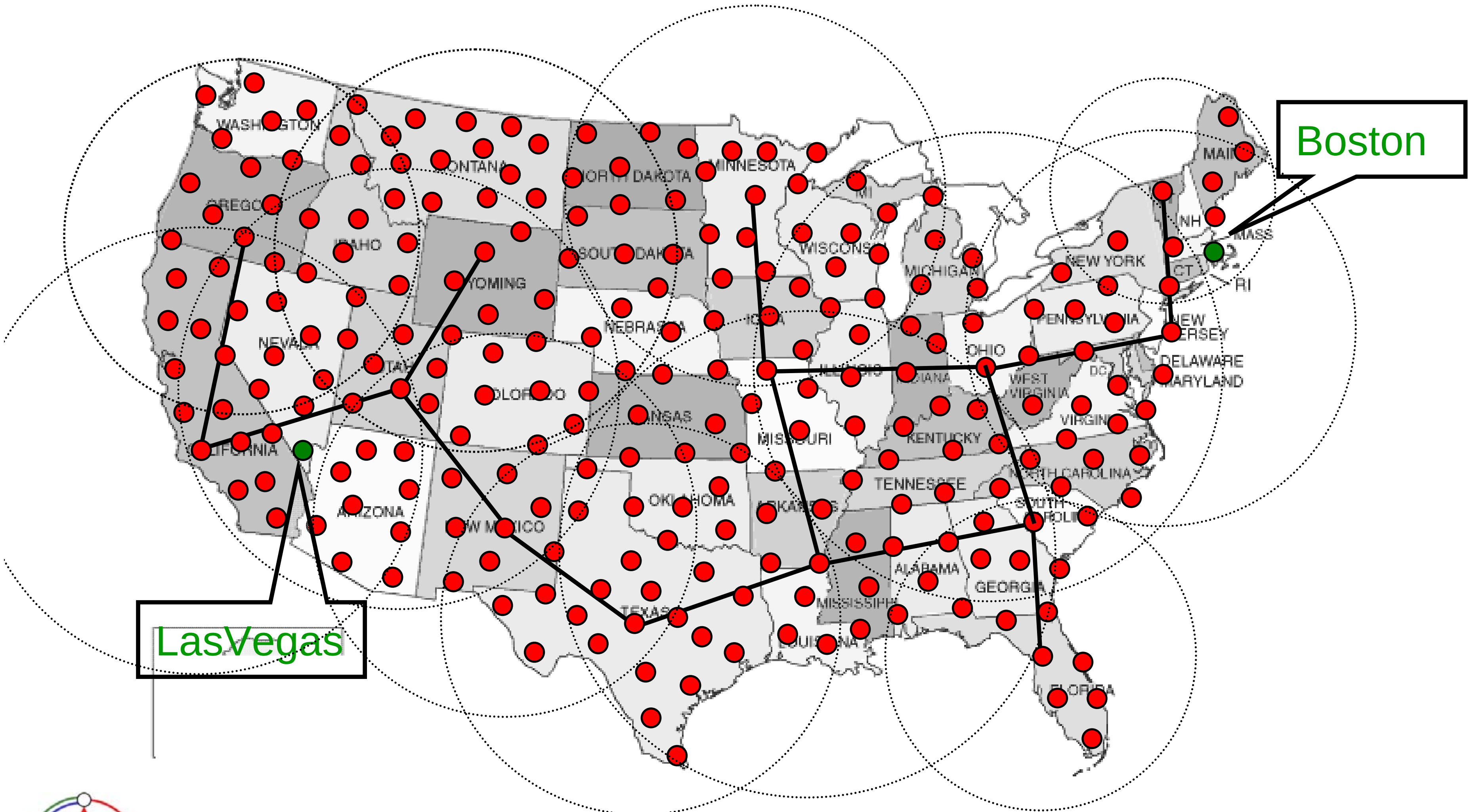


The Importance of Being Clustered... (Part A)

- Communication between distant nodes
  - Because of mobility, maintaining routing-tables is expensive
  - Hence, flooding is often a key component of routing protocols
  - Flooding is inefficient in dense networks

→ Idea: Routing on Virtual Backbones

Virtual Backbone Routing



The Importance of Being Clustered... (Part A)

- Communication between distant nodes
  - Because of mobility, maintaining routing-tables is expensive
  - Hence, flooding is often a key component of routing protocols
  - Flooding is inefficient in dense networks

→ Idea: Routing on Virtual Backbones

1) Route to your clusterhead

2) Clusterhead routes message to clusterhead of destination

3) Clusterhead sends message to destination

compute clusters & connect clusterheads!

In this talk!

The Importance of Being Clustered...

- In wireless multi-hop networks,...
- ... clustering is important for structuring the network.

- Particularly, clustering helps in...

A) ...facilitating communication between distant nodes
  - Backbone routing

B) ...organizing communication between adjacent nodes
  - MAC layer, spatial multiplexing
  - Cluster-Leader can assign collision-free time-slots

C) ...improving energy efficiency
  - Synchronized Sleep/Awake schedules within a cluster
  - Only cluster-leader needs to remain awake

Maximal Independent Set (MIS)

- Clustering: Choose clusterhead such that:

  a) Each node has a clusterhead in its communication range.

  b) No two clusterheads interfere

- When modeling the network as a graph G=(V,E), this leads to the well-known Maximal Independent Set (MIS) problem.
- A Maximal Independent Set (MIS) is a subset $S \subseteq V$ of the nodes s.t.
  a) No two nodes $u,v \in S$ are neighbors.
  b) Every node $w \in V \setminus S$ has at least one neighbor in S.

In this talk, we study the distributed complexity of computing a MIS in a radio network model.

Overview

- Clustering in Ad Hoc and Sensor Networks
- Related Work
- Model
- Algorithm & Analysis
- Conclusions & Open Problems

MIS - Related Work (MIS)

- Distributed complexity of MIS is of fundemantal interest
  - Prototypically captures the notion of symmetry breaking
  - Numerous important applications

- Upper Bounds in general graphs
  - Randomized algorithm time O(log n) [Luby 86]
  - There are deterministic algorithms running in $o(n^c)$ for any c>0. [Awerbuch, Goldberg, Luby, Plotkin, FOCS 89], [Srinivasan, Panconesi, STOC 92]
  - Whether there is a deterministic O(polylog n) MIS algorithm is an outstanding open question in distributed complexity. [Linial 92]

- Lower Bound in general graphs
  - Every MIS algorithm requires at least time $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ [Kuhn, Moscibroda, Wattenhofer, PODC 2004]

MIS - Related Work (MIS)

- General graphs may be a too pessimistic model for describing wireless networks
- MIS also studied in restricted graph models

  (Unit Disk Graphs, Unit Ball Graphs,...)

- When nodes know the distances, MIS in time $O(\log^* n)$

  [Kuhn, Moscibroda, Wattenhofer, PODC 2005]
- Deterministic without distance-information in time $O(\log \Delta \cdot \log^* n)$.

  [Kuhn, Moscibroda, Nieberg, Wattenhofer, DISC 2005]
- Lower Bound $\Omega(\log^* n)$ [Linial 92]

All of these algorithms work in

message passing models

MIS - Related Work (Radio Networks)

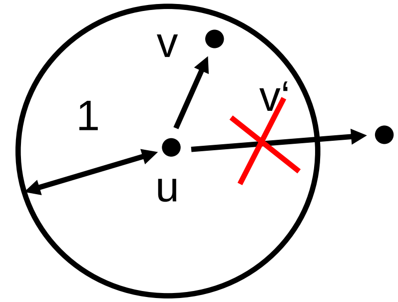
- In radio networks: Much work on communication primitives (broadcast, wake-up, gossiping,…)
- Much less is known about computation of local network structures.
- In single-hop radio networks → wake-up problem!
  - Lower Bound of $\Omega(\log^2 n / \log\log n)$ [Jurdzinski, Stachowiak, ISAAC 02]
  - This lower bound also holds for the MIS case!
- In multi-hop case
  - Many linear time algorithms [Wan, Alzoubi, Frieder, INFOCOM 02], …
  - $O(\log^2 n)$ algorithm when nodes wake up simultaneously and know 2-hop neighborhood [Gandhi, Parthasarathy, FSTTCS 04]
  - $O(\log^3 n)$ algorithm using three communication channels. [Moscibroda, Wattenhofer, MASS 04]
  - $O(\Delta \log n)$ algorithm for $O(\Delta)$ coloring. [Moscibroda, Wattenhofer, SPAA 05]

Overview

- Clustering in Ad Hoc and Sensor Networks
- Related Work
- Model
- Algorithm & Analysis
- Conclusions & Open Problems

Unstructured Radio Network Model

- Multi-Hop
- No collision detection
  - Not even at the sender!
- No knowledge about (the number of) neighbors
- Asynchronous Wake-Up
  - Nodes are not woken up by messages !
- Unit Disk Graph (UDG) to model wireless multi-hop network
  - Two nodes can communicate iff Euclidean distance is at most 1
- Upper bound n for number of nodes in network is known
  - This is necessary due to Ω(n / log n) lower bound
    [Jurdzinski, Stachowiak, ISAAC 2002]



Our algorithm computes a MIS in time $O(\log^2 n)$ with high probability.

Optimal up to O(loglog n) factor

Time-Complexity: Every node decides $O(\log^2 n)$ time-slots after its wake-up whether or not it joins the MIS!

Overview

- Clustering in Ad Hoc and Sensor Networks
- Related Work
- Model
- Algorithm & Analysis
- Conclusions & Open Problems

MIS Algorithm - Overview

<u>Goal of first phase</u>: Determine a limited number of *candidates*

- Upon waking up do:

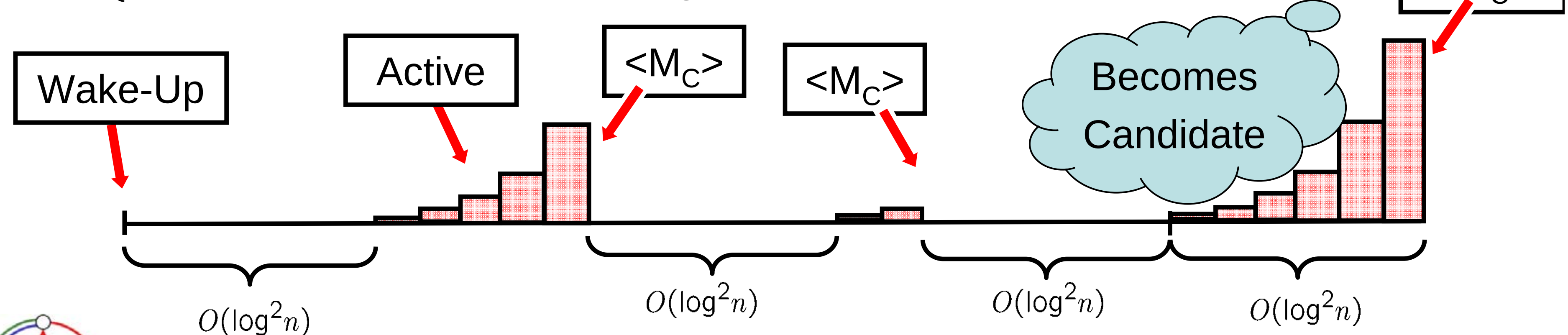
1) For $O(\log^2 n)$ time-slots do
   - If receive message from MIS node $\rightarrow$ stop
   - If receive message $<M_C>$ from active node $\rightarrow$ restart phase 1

Initial waiting phase

2) For i=0 to log n do
   - Set sending probability $p_i := C*2^i/n$
   - For $\lambda \log n$ time-slots do
     - Send message $<M_C>$ with probability p_i $\rightarrow$ become Candidate
     - If receive message from MIS node $\rightarrow$ stop
     - If receive message $<M_C>$ from active node $\rightarrow$ restart phase 1

Active Node

Exponentially increase sending probability

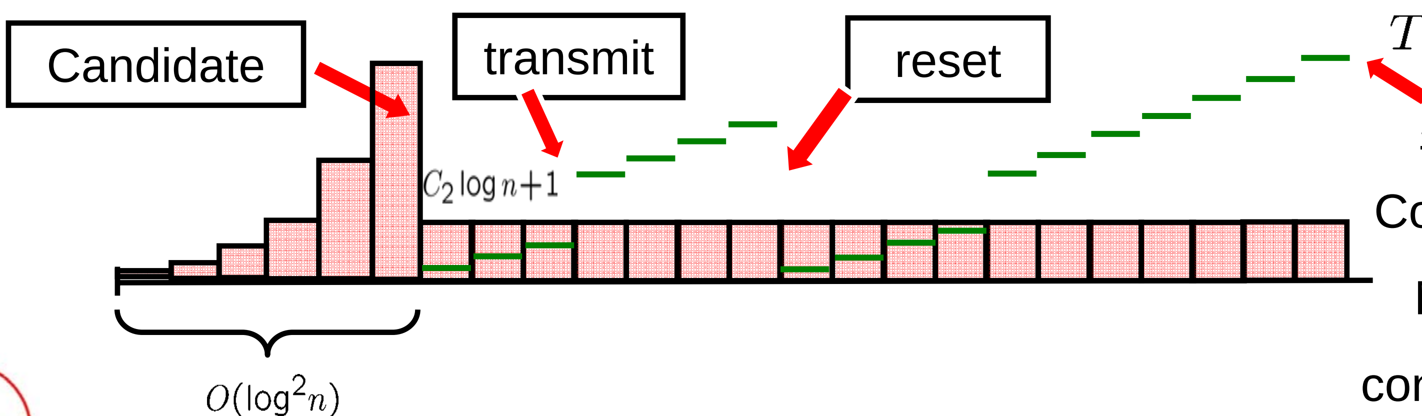


MIS Algorithm - Overview

<u>Goal of second phase</u>: Choose MIS Nodes from among candidates.

- Upon becoming candidate do:
- Each candidate has a local counter, initialized to 0.

3) In every time-slot
   - Transmit the local counter with probability $p \in \Theta(\log^{-1} n)$
     → set local counter to max{counter, $C_2 \cdot \log n+1$}
   - If receive message with neighboring counter
     AND if difference between counters is at most $C_2 \cdot \log n$
     THEN reset local counter to 0
   - If receive message from MIS node → stop
   - Increase counter by 1, if it reaches a threshold $T \in O(\log^2 n)$, join MIS



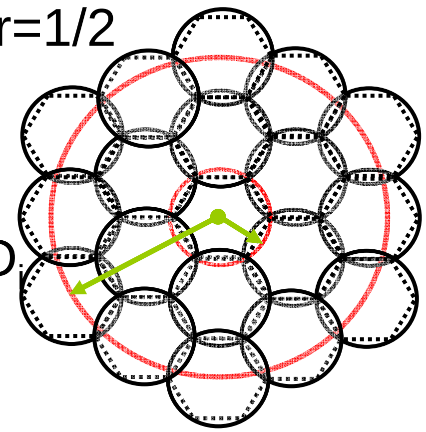
<u>Join the MIS:</u>

Continue sending Message with constant probability

MIS Algorithm – Basic Proof Idea

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

- Cover the plane with (imaginary) disks D_i of radius r=1/2
- Let E_i be the disk with radius R=3/2
- A node in D_i can hear all nodes in D_i
- Nodes outside of E_i cannot interfere with nodes in D_i



Throughout the execution, three properties hold w.h.p. (probabilistic invariants)

(1) The *sum of sending probabilities* $s(t) := \sum_{k \in C_i} p_k(t)$ in a disk D_i is bounded by a constant α.

(2) The *number of candidates* in any disk D_i is at most O(log n).

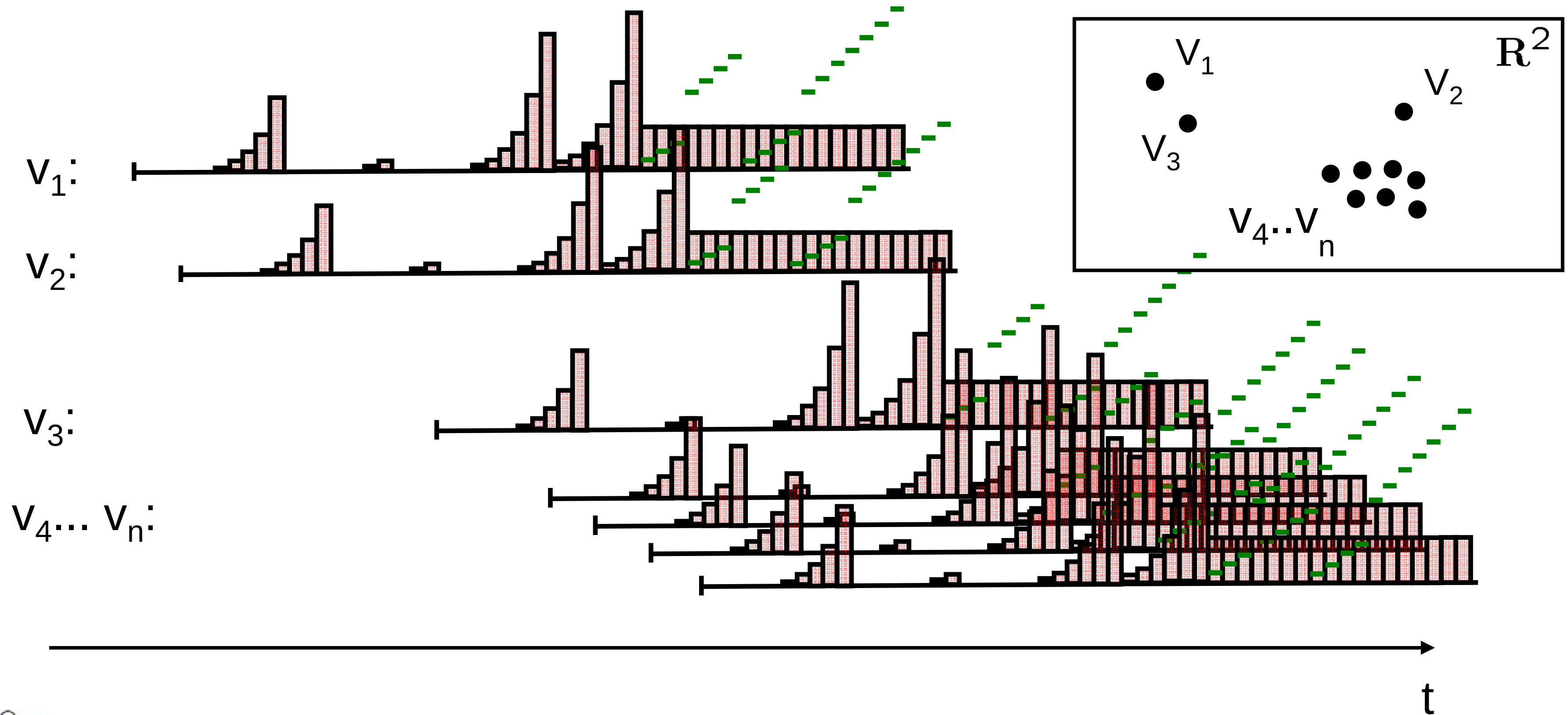
(3) The set of MIS nodes is *independent*

- At the beginning, all properties hold.
- Show that w.h.p., none of the properties is violated first.

Proof : (2)+(3) → (1)

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

- Assume (2), (3) hold → what is the probability of (1) being violated?
- The proof is made difficult by asynchronous wake-up
- Executions of different nodes can be arbitrarily shifted in time

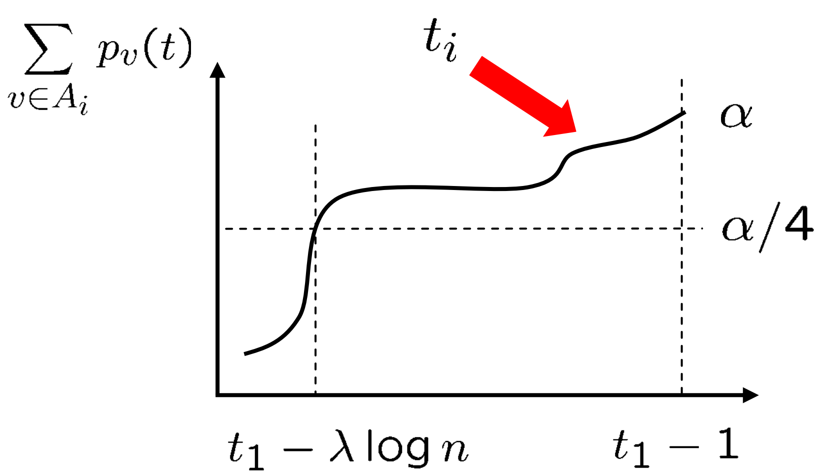


(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

Proof : (2)+(3) → (1)

- Assume (1) is *first* property to be violated at time t_1 !
- Consider λlog n time-slots betore t_1.
  - (2) and (3) hold in time-slots $t_1 - \lambda \log n, \ldots, t_1 - 1$
  - Every active node at time $t_1 - \lambda \log n$ will its probability exactly once
  - At most n new node may join with the initial sending probability

→ In interval $\mathcal{J} = t_1 - \lambda \log n, \ldots, t_1 - 1$, it holds that $\frac{1}{4}\alpha \leq \sum_{v \in A_i} p_v(t) \leq \alpha$



Consider an arbitrary time-slot $t_i \in \mathcal{J}$

Proof : (2)+(3) → (1)

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

- If exactly one active node $v_a \in D_a$ transmits and no other node in E_a
  → *Clearance* in D_a occurs
  → v_a becomes candidate, all other active nodes restart

Sending probability reduces to 0 in D_a

→ Probability of a clearance of D_a in $t_i \in \mathcal{J}$:

$$
\begin{aligned}
P_c(t_i) &\geq P[\text{one active node sends in } D_a] \cdot P[\text{no other node sends in } E_a] \\
&\geq \sum_{v \in \mathcal{A}_a} \left(p_v \prod_{w \in D_a \setminus \{v\}} (1 - p_w) \right) \cdot \prod_{q \in E_a} (1 - p_q) \\
&\geq \sum_{v \in \mathcal{A}_a} p_v \cdot \left(\frac{1}{4}\right)^{\sum_{w \in D_a} p_w} \cdot \left(\left(\frac{1}{4}\right)^{\sum_{w \in D_a} p_w} \right)^{O(1)} \geq \Omega(1).
\end{aligned}
$$

bounded by constant

O(1) if (2) and (3) hold

Wait! There can be infinitely many t_i → high probability?

→ There is a clearance in $\mathcal{J}$ w.h.p.

We show that unless (1) is violated, there are at most O(1) clearances in D_a.

Hence, (1) is not violated unless (2) or (3) is violated first w.h.p.

Proof : (1)+(3) → (2)

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

- Assume (1), (3) hold → what is the probability of (2) being violated?
- We show that the number of new candidates emerging in D_a is at most O(log n) between two clearances!

Definition: A time-slot is a *failure* in D_a if a new candidate emerges, but there is no clearance of D_a.

- We prove: The number of failures between two clearances is bounded by O(log n) w.h.p.
- Consider a time-slot in which at least one active node sends:

$$P_t[\text{clearance}|\text{at least one node sends}] \geq \ldots \geq \Omega(1)$$

Geometric Series:
At most O(log n) failures

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

Proof : (1)+(3) → (2)

- How many candidates can emerge during these O(log n) failures?

- In expectation, O(1) sending nodes (=new candidates) per failure.
- If there are at most O(log n) failures, then O(log n) new candidates in expectation.

- Using Chernoff Bound → At most O(log n) new candidates between two subsequent clearances w.h.p
  (*if* number of failures bounded by O(log n))

Unless property (2) is violated, there are at most O(1) clearances in D_a.

Hence, (2) is not violated unless (1) or (3) is violated first w.h.p.

Proof : (1)+(2) → (3) (i.e. Correctness)

(1) Sum of sending probabilities
(2) Number of candidates
(3) Independence

- Assume (1), (2) hold → Will the MIS be correct?
- We prove:

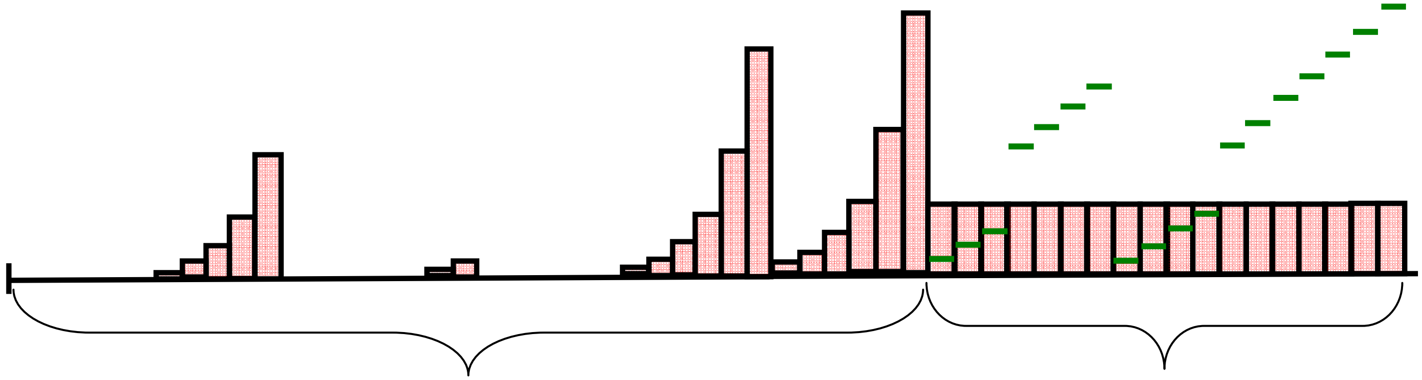
1) No two neighbors join MIS within O(log n) time w.h.p!
→ Proof by contradiction:
Consider $O(\log^2 n)$ time-slots before the first joined MIS.
We show:
second node receives message from first
→ second node resets its counter!

2) The first MIS node will successfully send a MIS message to all its neighbors within time O(log n) after joining MIS!

3) It follows: (3) holds → MIS is correct w.h.p. !

Proof : Running Time...

- How long does a node need to wait...?



How many times must a node restart?

How many times must a candidate reset?

→ <u>We prove:</u>

After at most $O(\log^2 n)$ slots, v is either covered or becomes a candidate!

→ <u>We prove:</u>

After at most $O(\log^2 n)$ slots, v is either covered or joins the MIS!

Every node decides within time $O(\log^2 n)$, w.h.p

Conclusions / Open Problems

We believe $O(\log^2 n)$ to be tight.

- MIS in harsh radio network model in time $O(\log^2 n)$ w.h.p.
  → Tight up to a factor of O(loglog n)
  → Compare to Luby's O(log n) in message passing models
  → Both algorithms are known to be close from optimal

- Future Directions / Open Problems:
  - Tighten upper or lower bound
  - Study the complexity of other local network structures in radio networks
  - Deterministic constructions?

Questions? Comments?

Questions?
Comments?

Distributed
Computing
Group

Thomas Moscibroda

Roger Wattenhofer