Characterizing Dynamic Properties of the SopCast Overlay Network

Kênia Carolina Gonçalves[*]

Alex Borges[**], Jussara Almeida[*], Ana Paula Silva[**], Humberto Marques-Neto[***], Sérgio Campos[*]

*UFMG   **UFJF   ***PUC-Minas

Belo Horizonte - Brazil

Live Video Streaming Systems

- On-demand or **live** streaming

- Client-Server or **Peer-to-Peer (P2P)**
  - **SopCast**, PPLive, PPStream, ...

- 83 million users in 2013 (predicted)
  - Sentinelli et. al. - Will IPTV Ride The Peer-to-Peer Stream? - 2007

SopCast

Overlay Network

- Logical network to data transmission
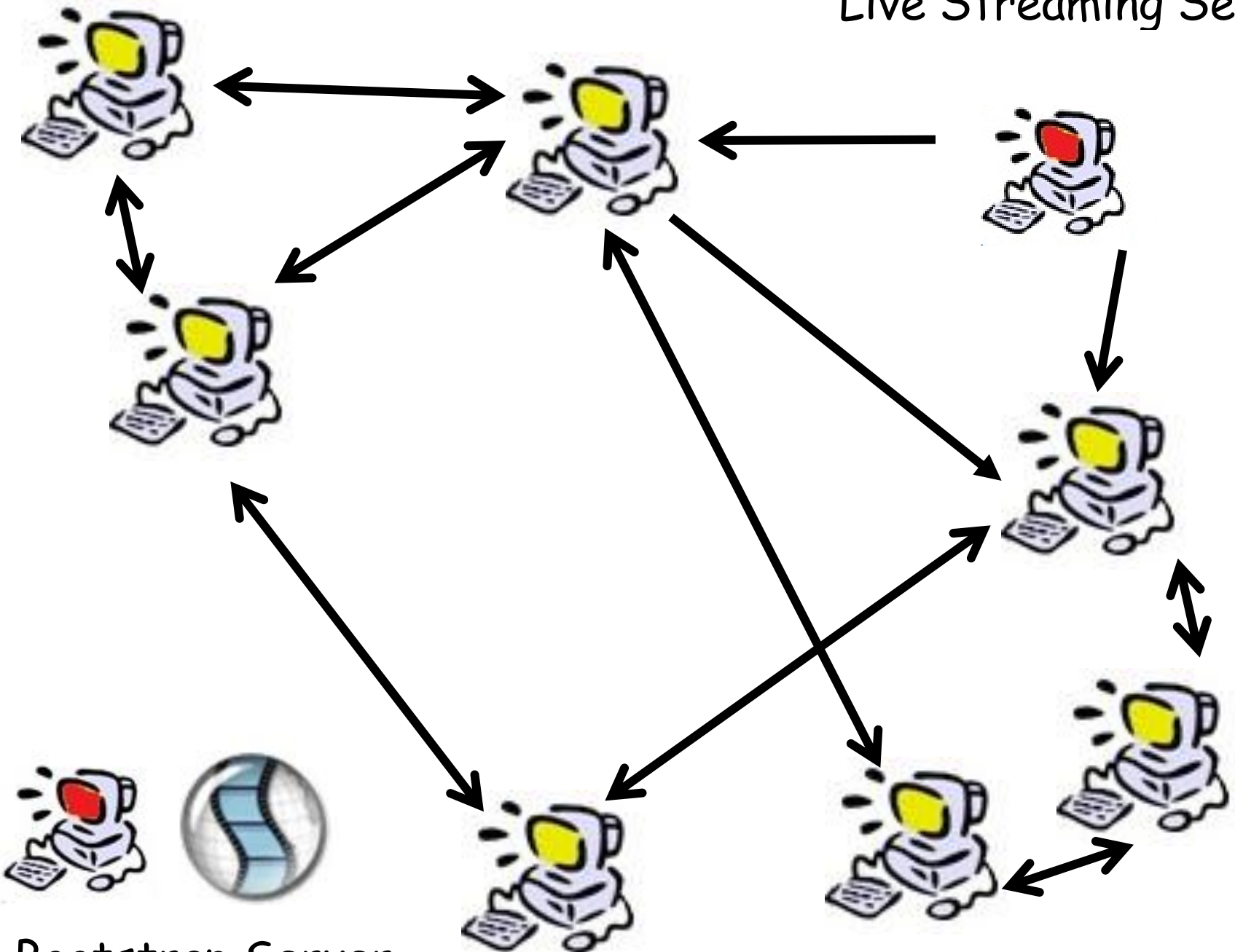
- Bootstrap Server

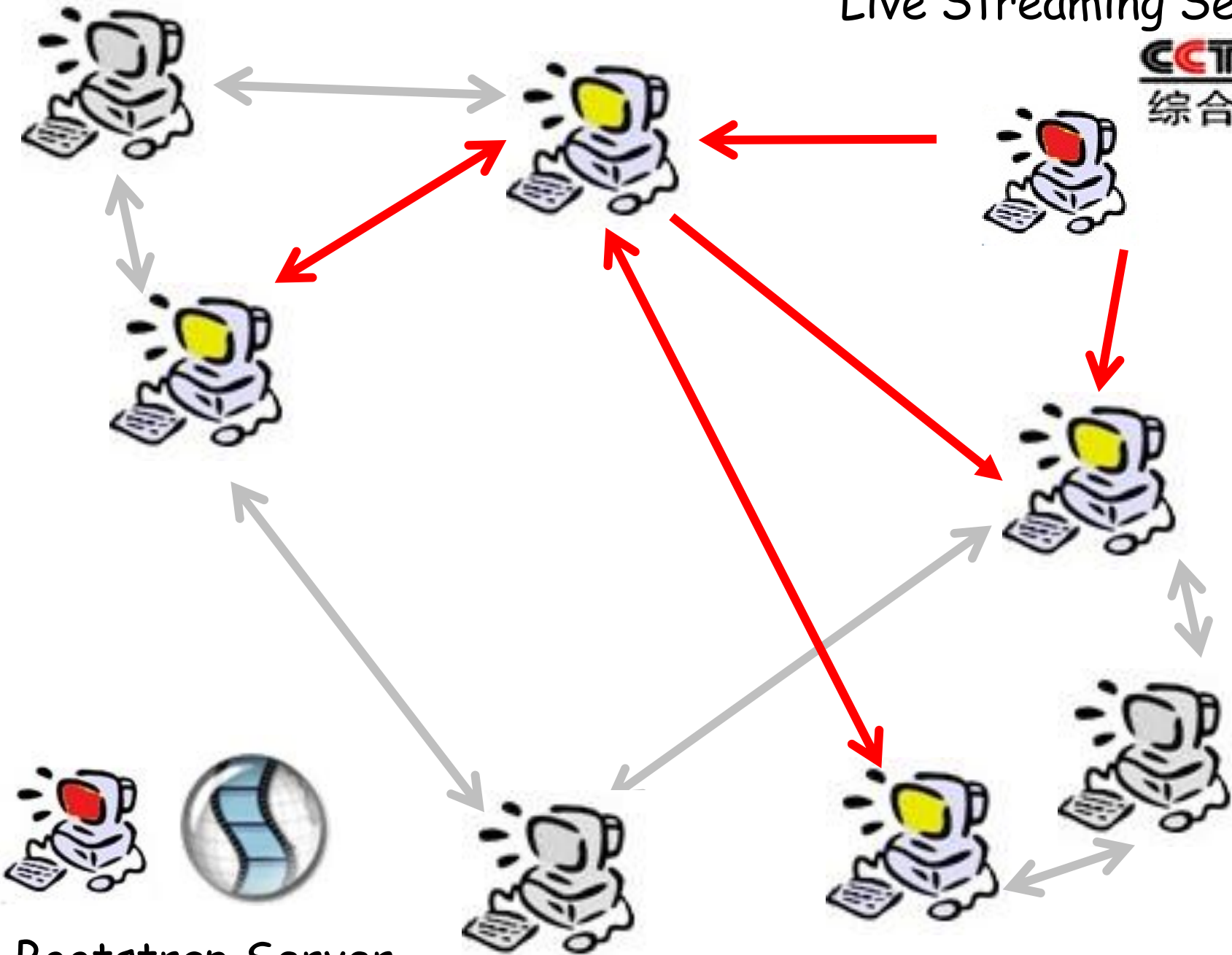- Live Streaming Server

- Clients (peers)
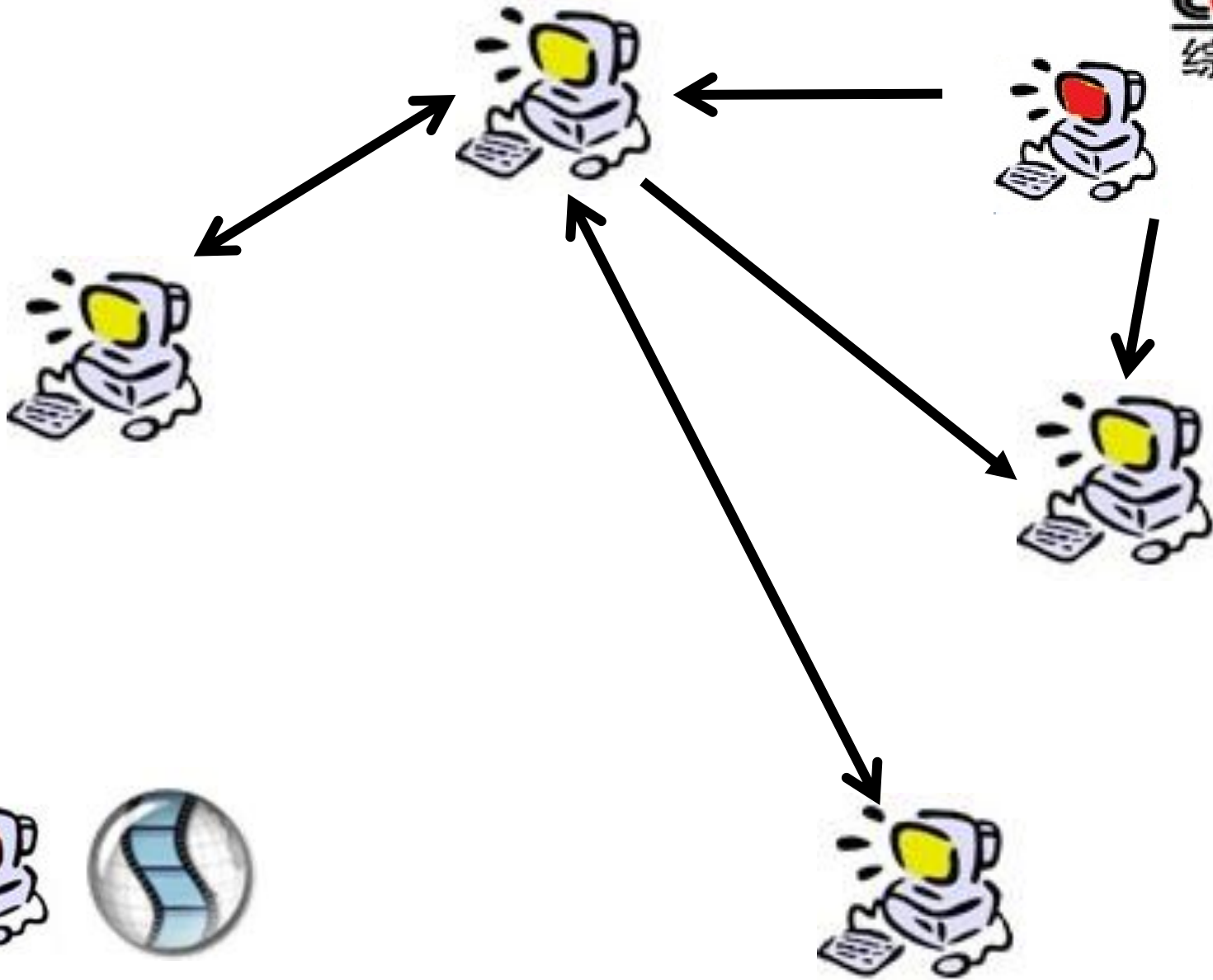
Live Streaming Server

Bootstrap Server

Live Streaming Server

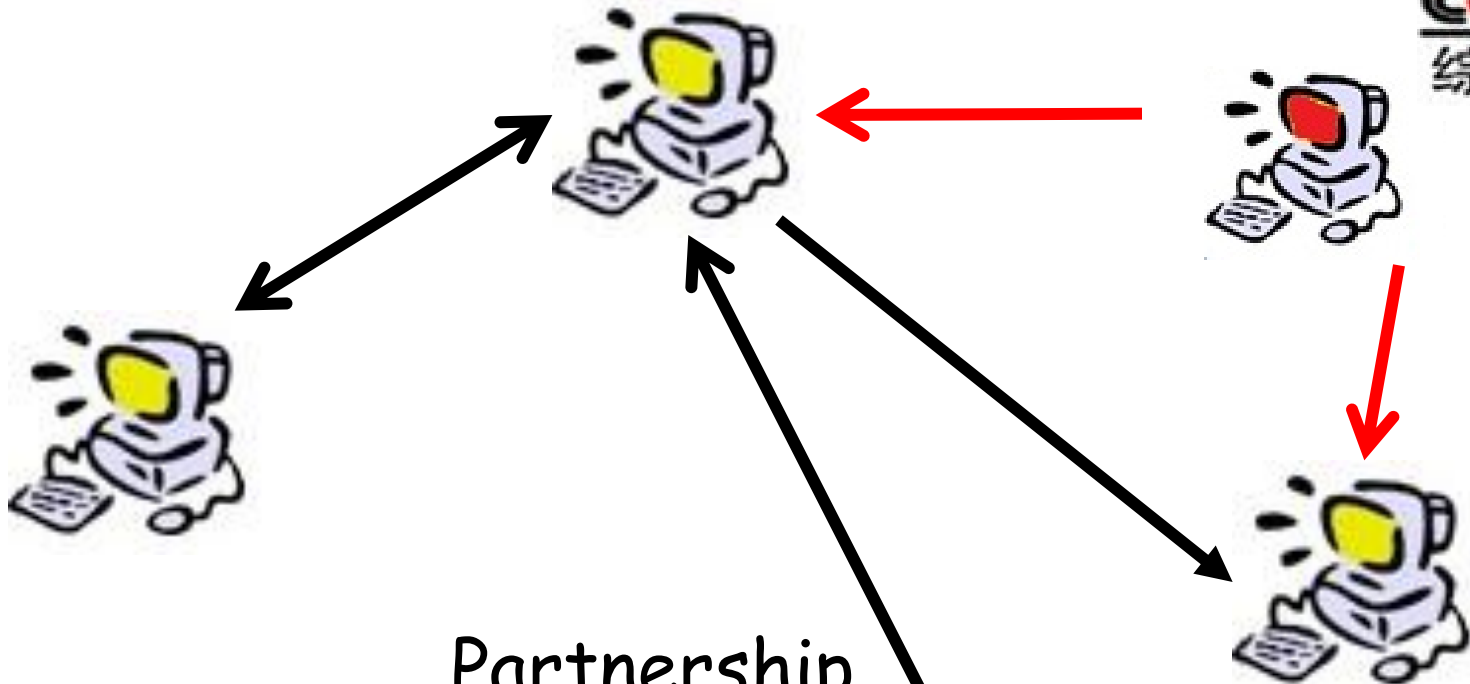Bootstrap Server

Live Streaming Server

Bootstrap Server

Live Streaming Server

Partnership

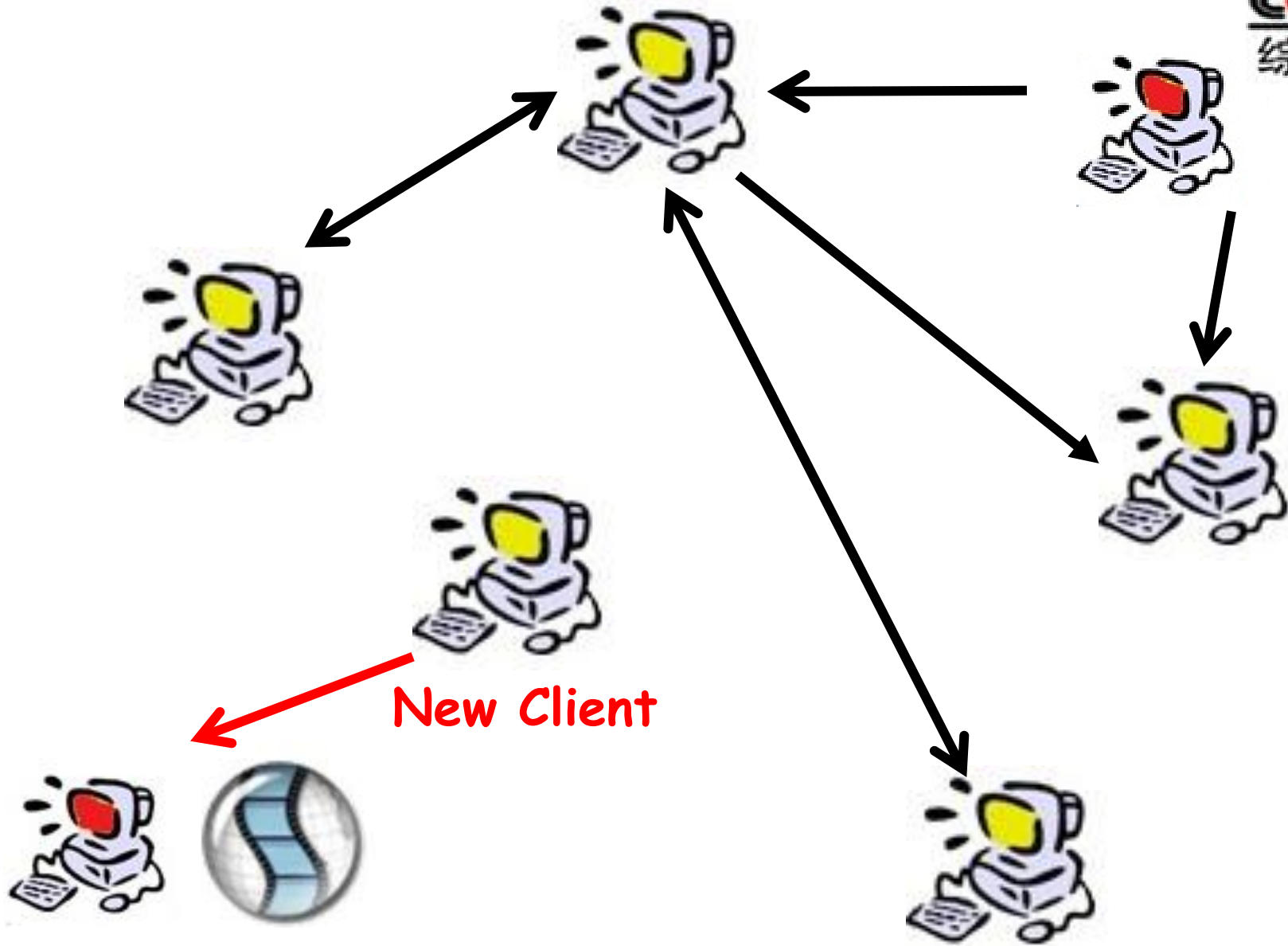Bootstrap Server

Live Streaming Server

CCTV1 综合频道

New Client

Bootstrap Server

Live Streaming Server

New Client

Bootstrap Server

Live Streaming Server

Bootstrap Server

Live Streaming Server
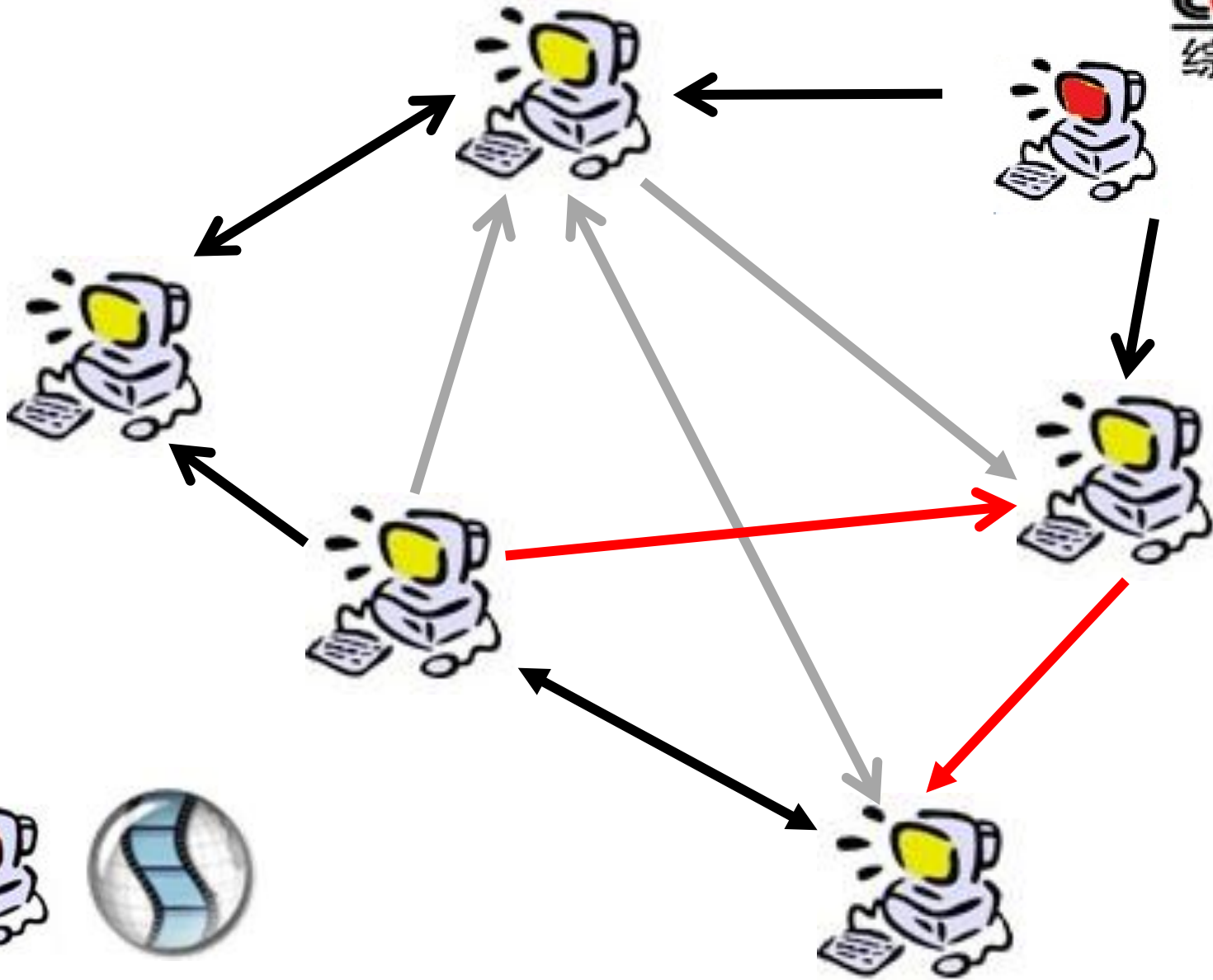
Bootstrap Server

HOW DOES THE STRUCTURE OF THE OVERLAY NETWORK EVOLVE OVER TIME DURING A LIVE TRANSMISSION?

Goals

- Characterize the temporal evolution of the P2P overlay network over a transmission
  - Local View: individual nodes
  - Global View: whole network

- Build knowledge to:
  - Create more realistic P2P synthetic workloads
  - Drive future protocol designs and evaluations

Previous Characterizations of P2P Live Systems

- ## Structural properties
  - Static view of the network
  - Few crawlers (≤ 70)
    - Possibly less representative view of the network

- ## Temporal evolution
  - Focused on peer degree only

DATA COLLECTION METHODOLOGY

Data Crawling Setup
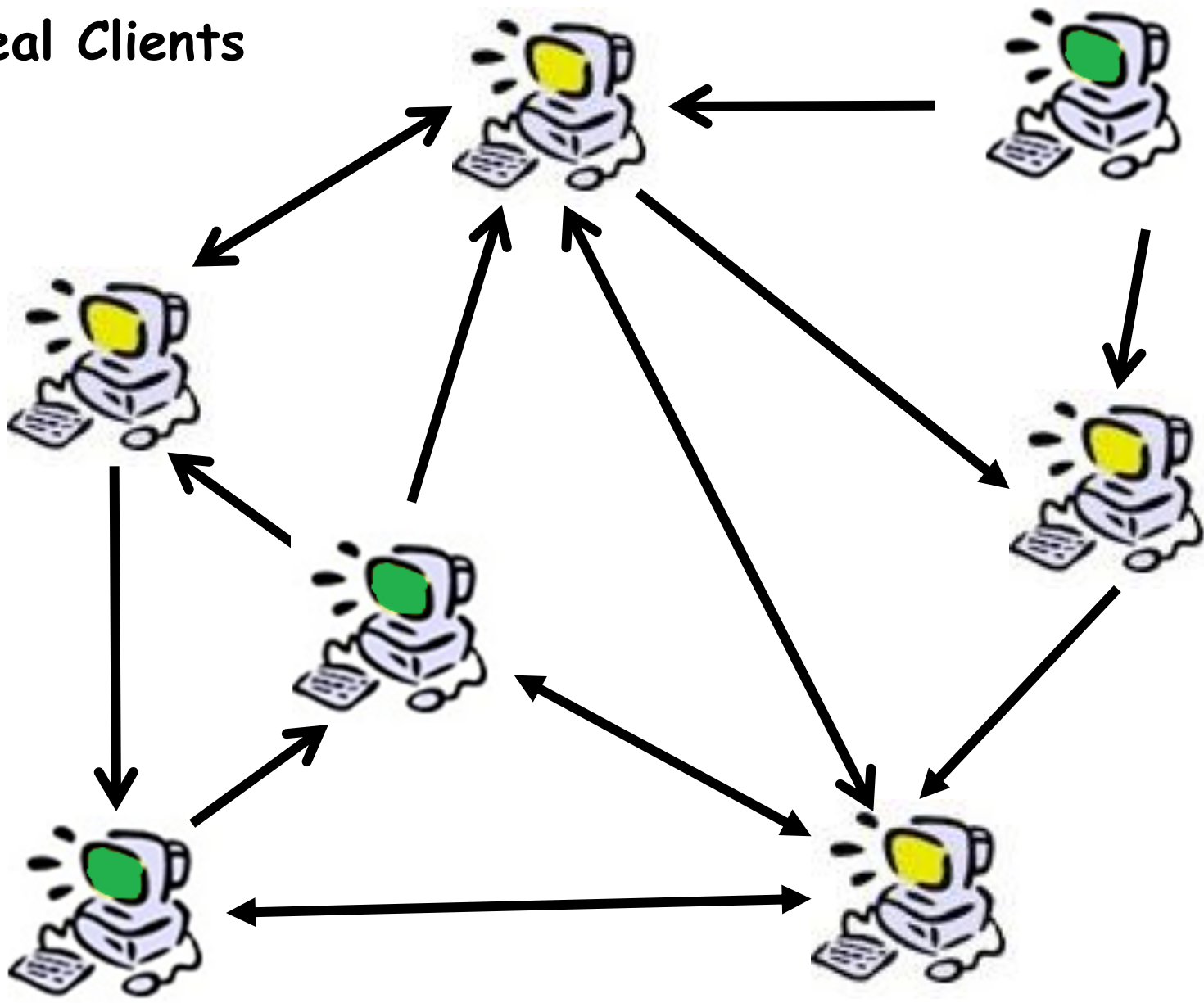
- SopCast Clients running on PlanetLab nodes
  - CCTV-1 channel
  - 8pm (China local time)

- Wireshark – UDP/TCP port

- Time synchronization (NTP)

- Unconstrained peer bandwidth (upload and download)
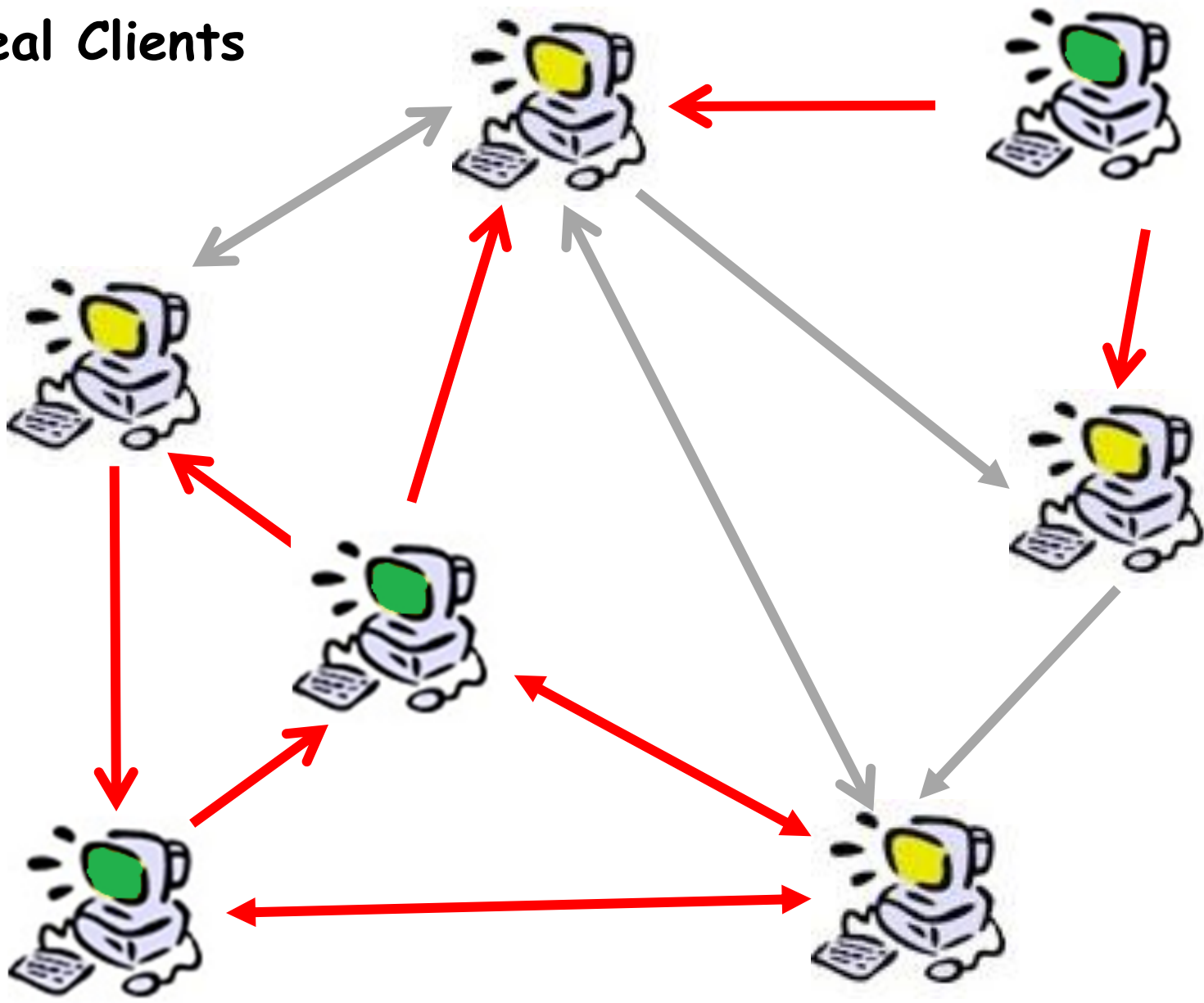
Crawlers on PlanetLab
Real Clients

Crawlers on PlanetLab

Real Clients

Data Crawlling

- Crawlers remain connected throughout transmission

- For each packet received/sent by each crawler:
    - Date and time of the transmission
    - Source IP Address
    - Destination IP Address
    - Packet size

- Merge data collected by all crawlers

- Snapshots of the network every 60 seconds

Overview

Number of experiments	7
Period of experiments	10/28/10 to 11/17/10
Number of crawlers	200 until 465
Channel	CCTV-1
Local time	8 pm
Transmission duration	40 minutes
Duration of each snapshot	60 seconds

CHARACTERIZATION

Main Steps

- Properties of individual nodes
  - Centrality profiles
  - Changes in the profile of a node over time
  - Changes in the list of partners over time

- Properties of network as a whole

Main Steps

➢ Properties of individual nodes

  ➢ Centrality profiles

  – Changes in the profile of a node over time

  – Changes in the list of partners over time

• Properties of network as a whole

Centrality Profiles

What is the importance of a node in terms of its centrality in the network?

- Centrality Metrics:
  - Degree: number of partnerships
  - Betweenness:
    - Probability of a node to be in a shortest path between two other nodes
  - Closeness:
    - Average distance between a node and all other nodes in a network (reachable from it)
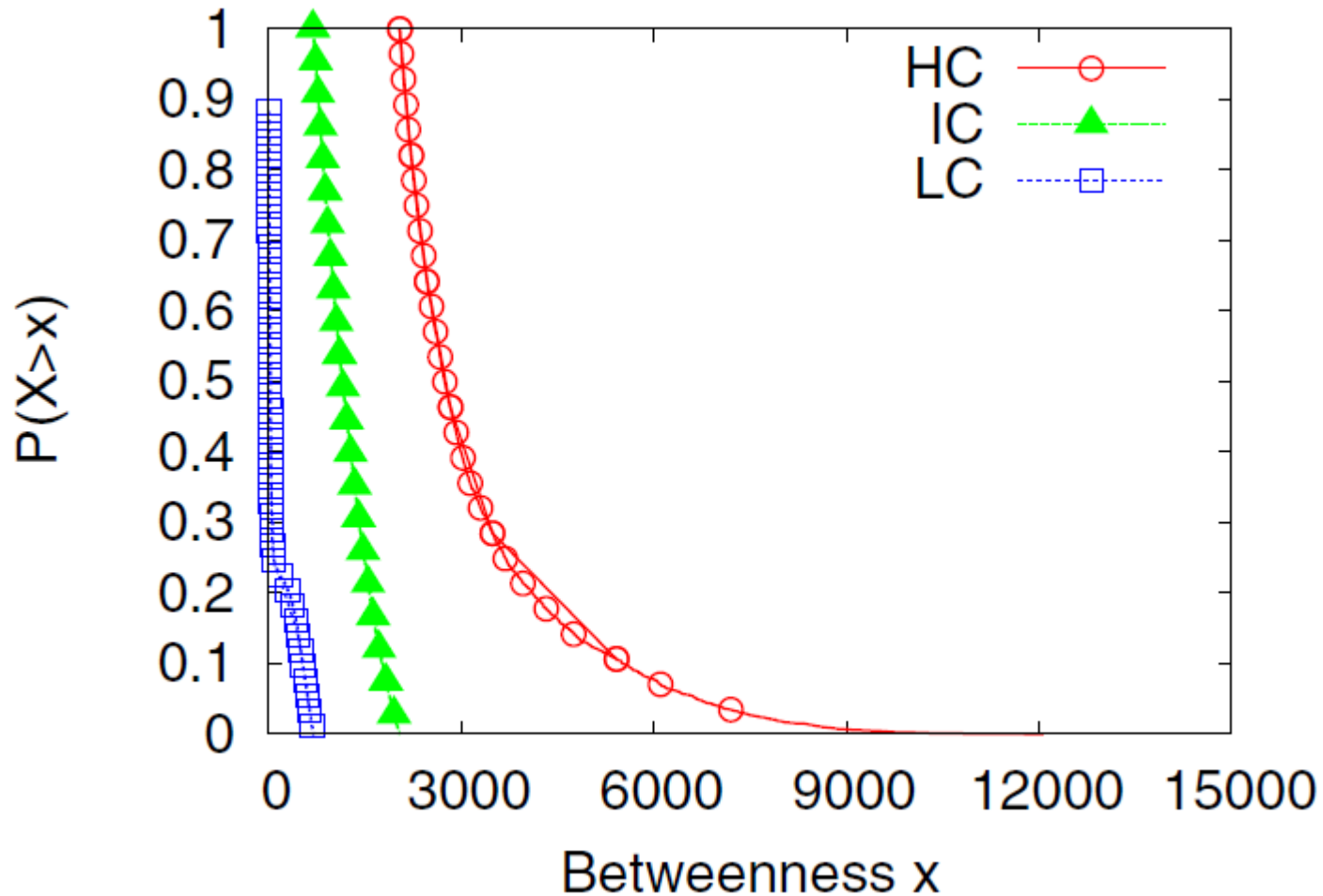
Centrality Profiles

- During each transmission, each node is represented by 3 features:
  - Average degree
  - Average betweenness
  - Average closeness

- k-means clustering algorithm

- Identification of three profiles (in all experiments)
  - High Centrality (HC)
  - Intermediate Centrality (IC)
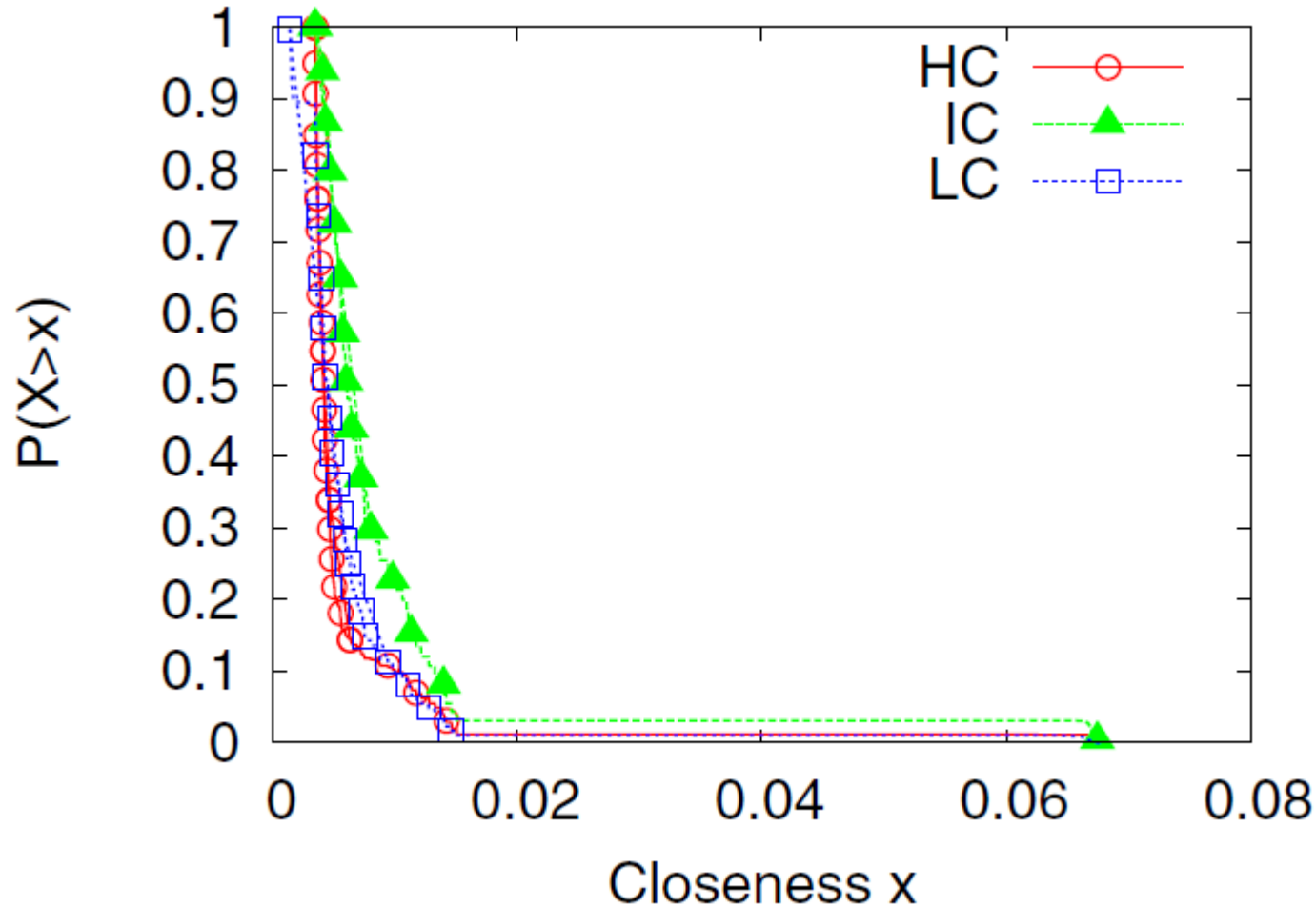 - Low Centrality (LC)

Degree Distribution
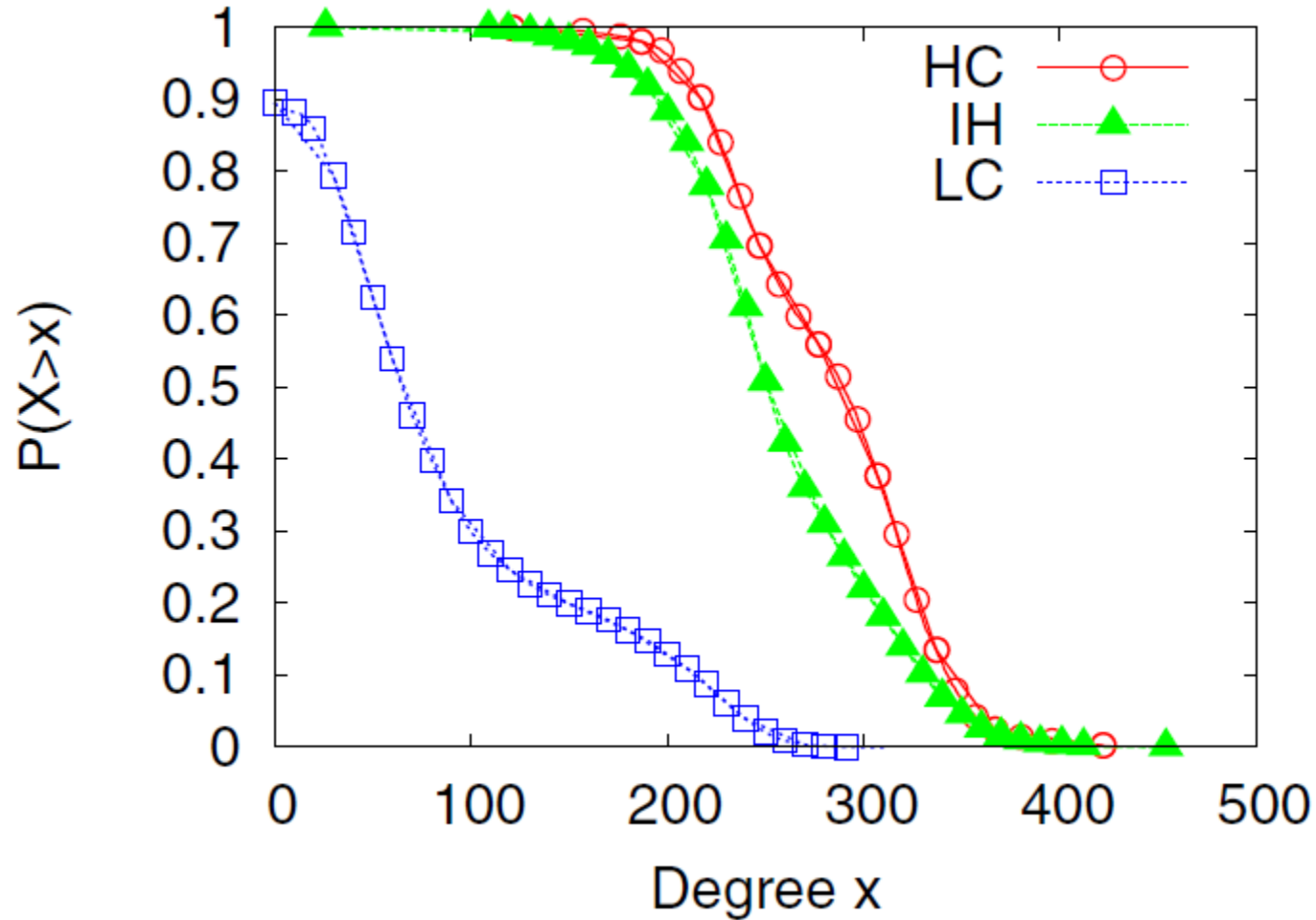
Nodes with HC profile have much more partnerships

Betweenness Distribution

Nodes with HC profile are located in the path of many more nodes than the other profiles
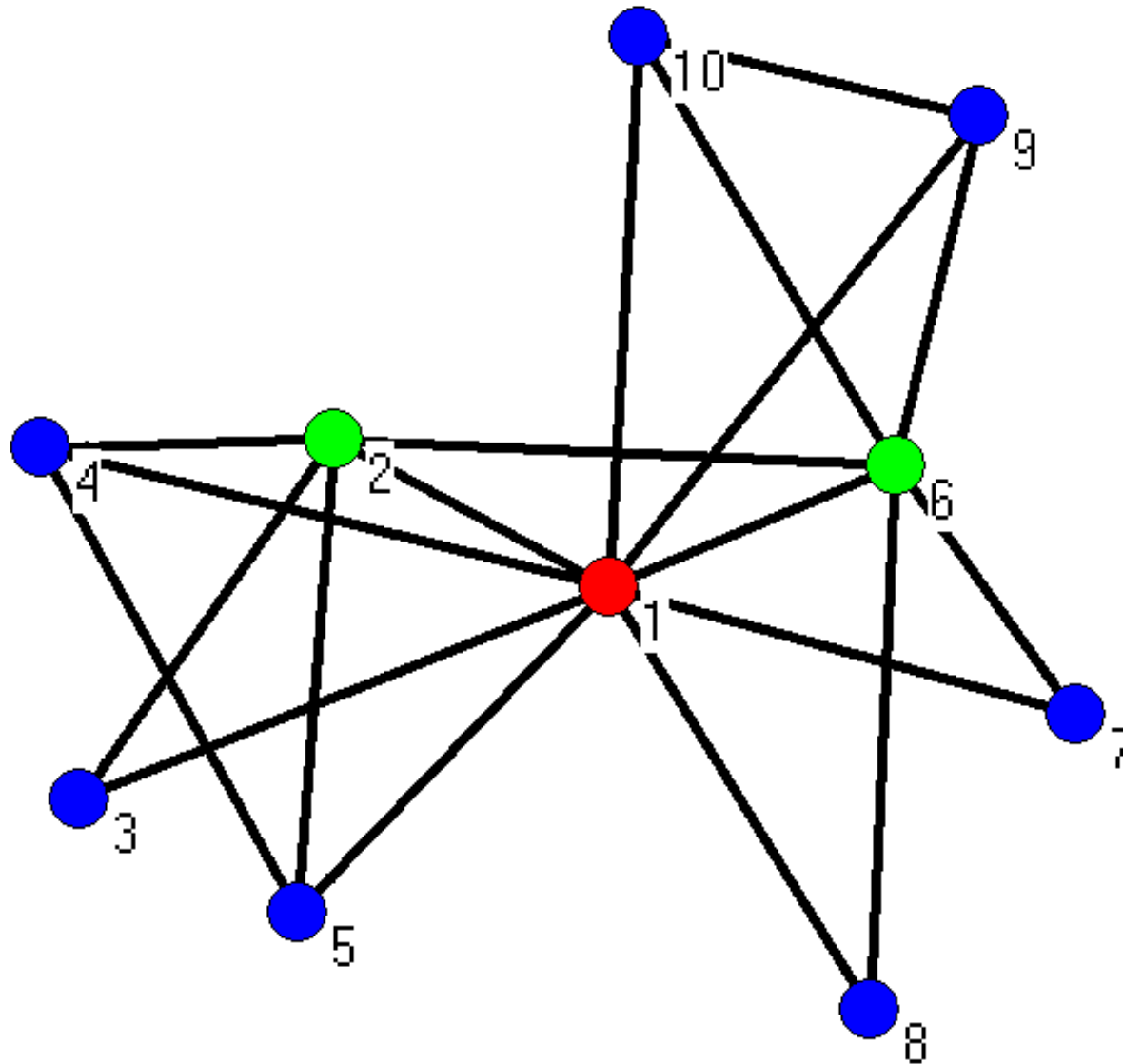
Closeness Distribution

Closeness is not able to clearly distinguish the profiles

Main Steps

➢ Properties of individual nodes

✓ Centrality profiles

➢ Changes in the profile of a node over time

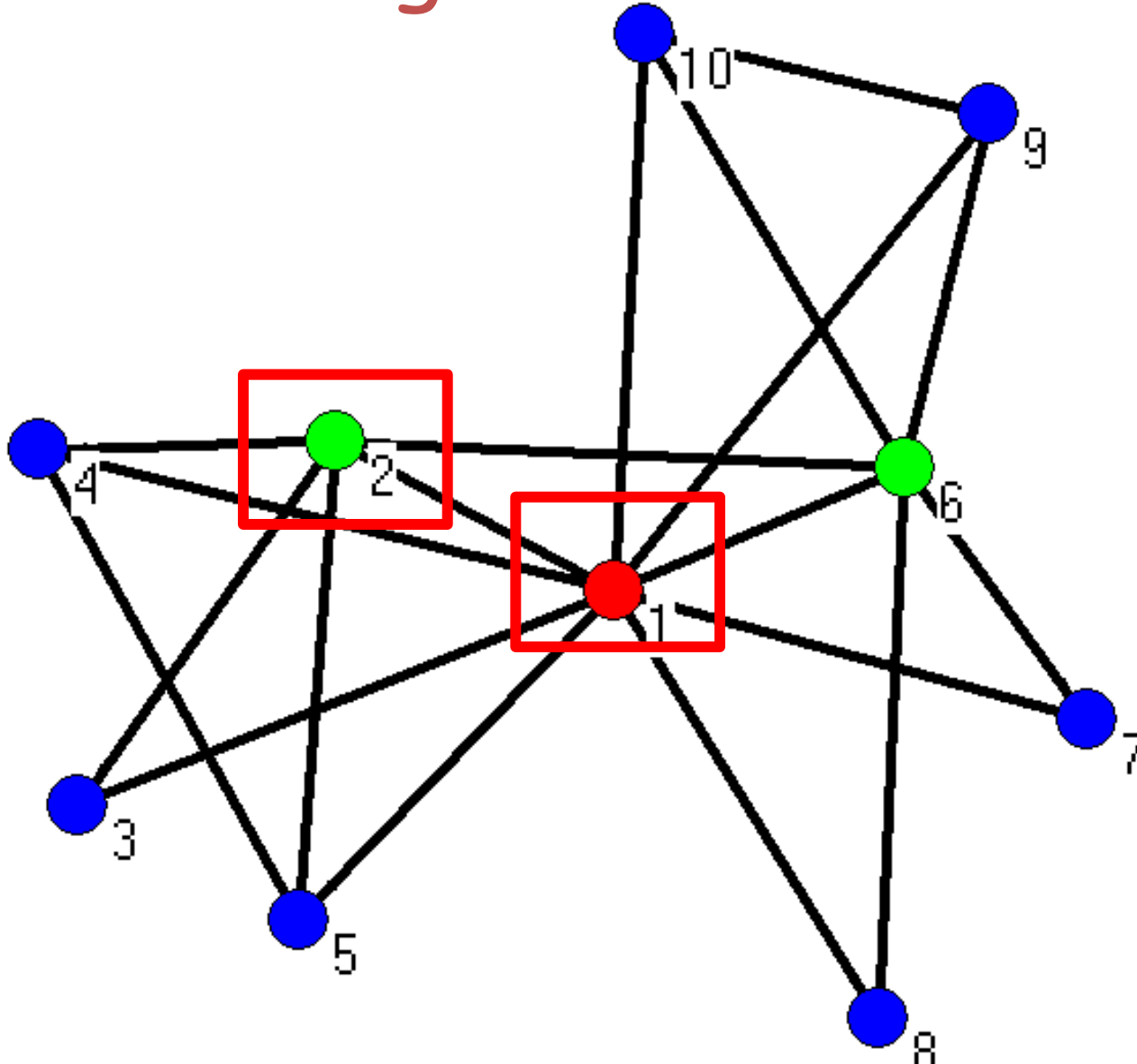– Changes in the list of partners over time

• Properties of network as a whole

How does the profile of a node change over time?

How does the profile of a node change over time?

HC (red)
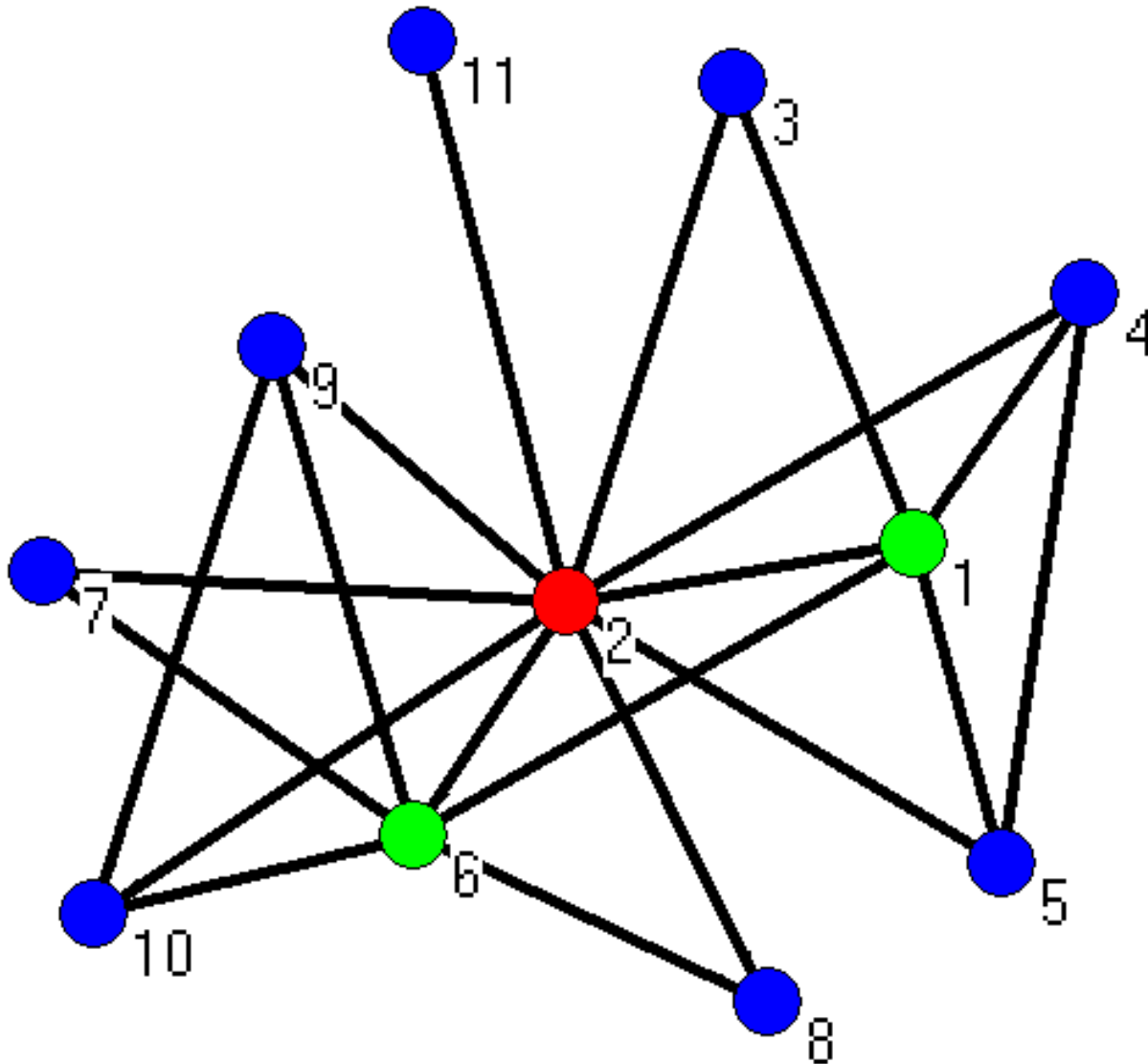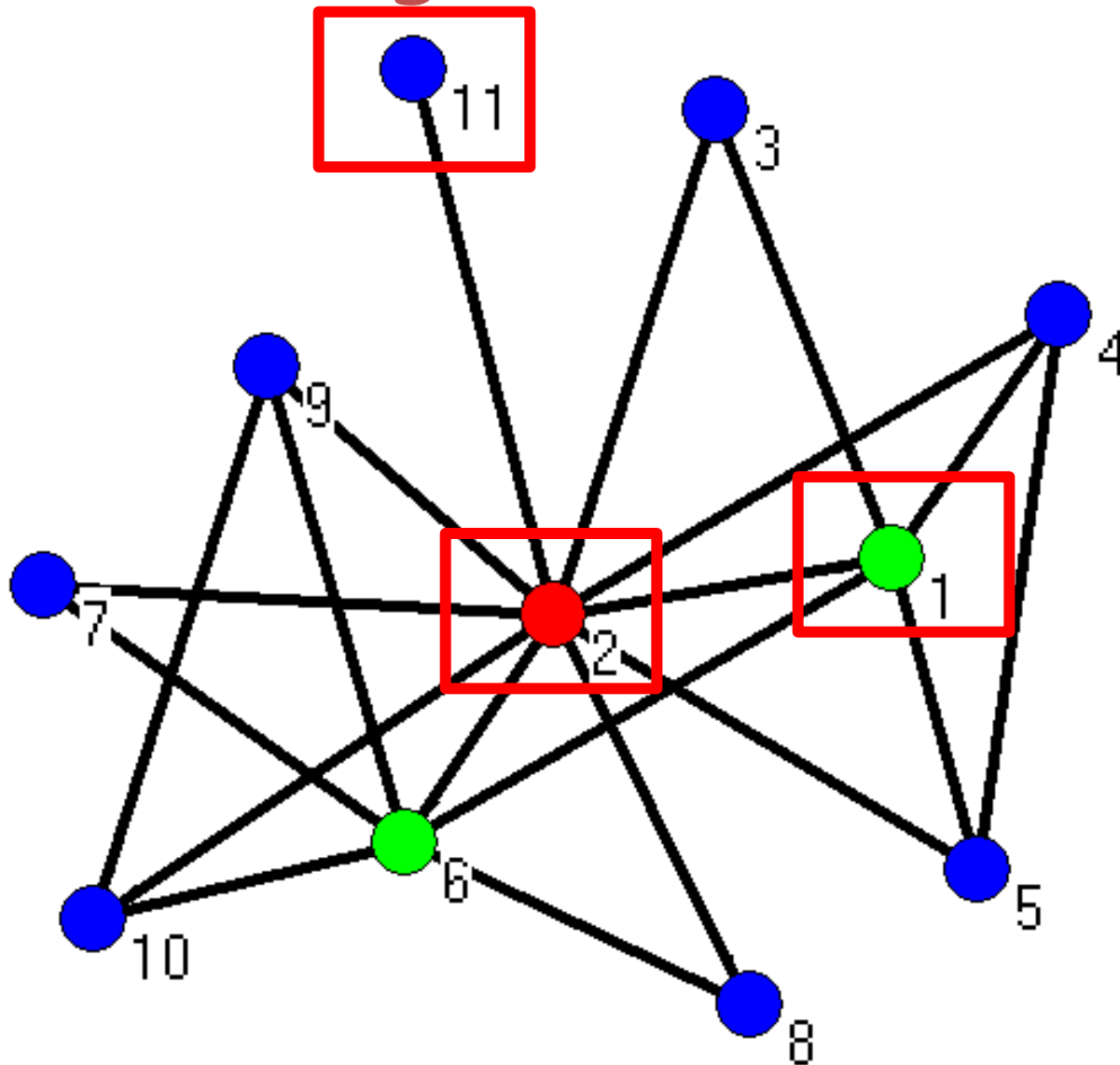IC (green)
LC (blue)

How does the profile of a node change over time?

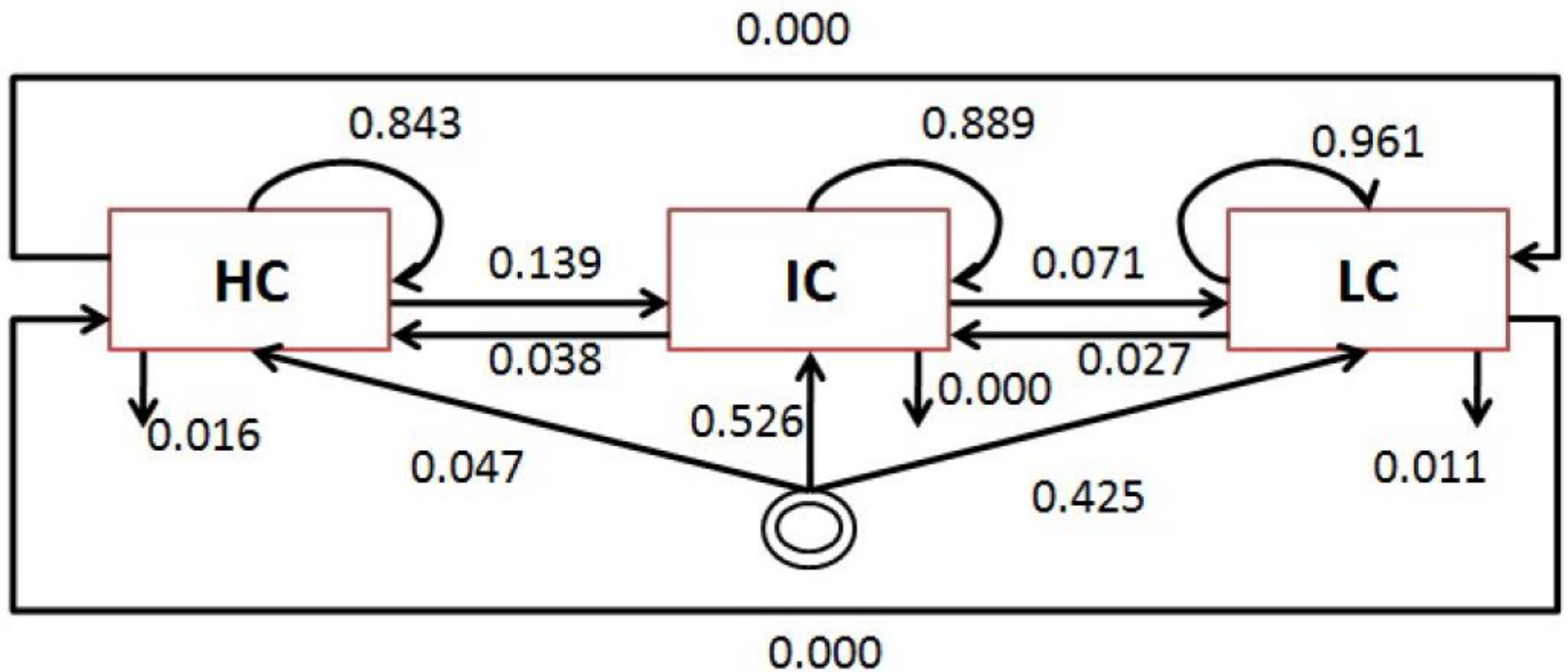# How does the profile of a node change over time?

HC
IC
LC

Changes in The Profile of a Node Over Time

Does a node tend to change its profile very often over a transmission?

- Customer Behavior Model Graph (CBMG)

  - State transition model

  - States: centrality profiles

  - Transition: labeled with probability of a node changing between two profiles

  – Represents the dynamics of the nodes

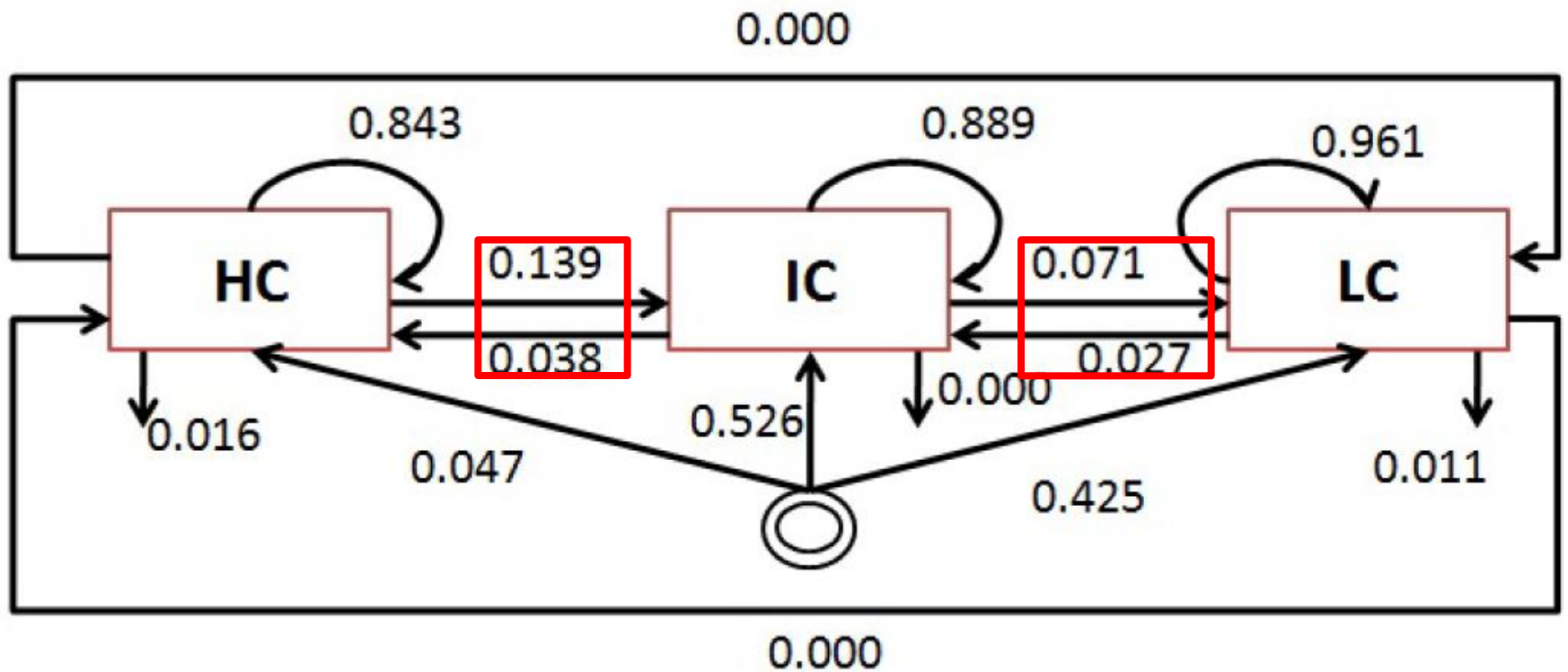# Changes in The Profile of a Node Over Time

Changes between centrality profiles for a given node

Changes in The Profile of a Node Over Time
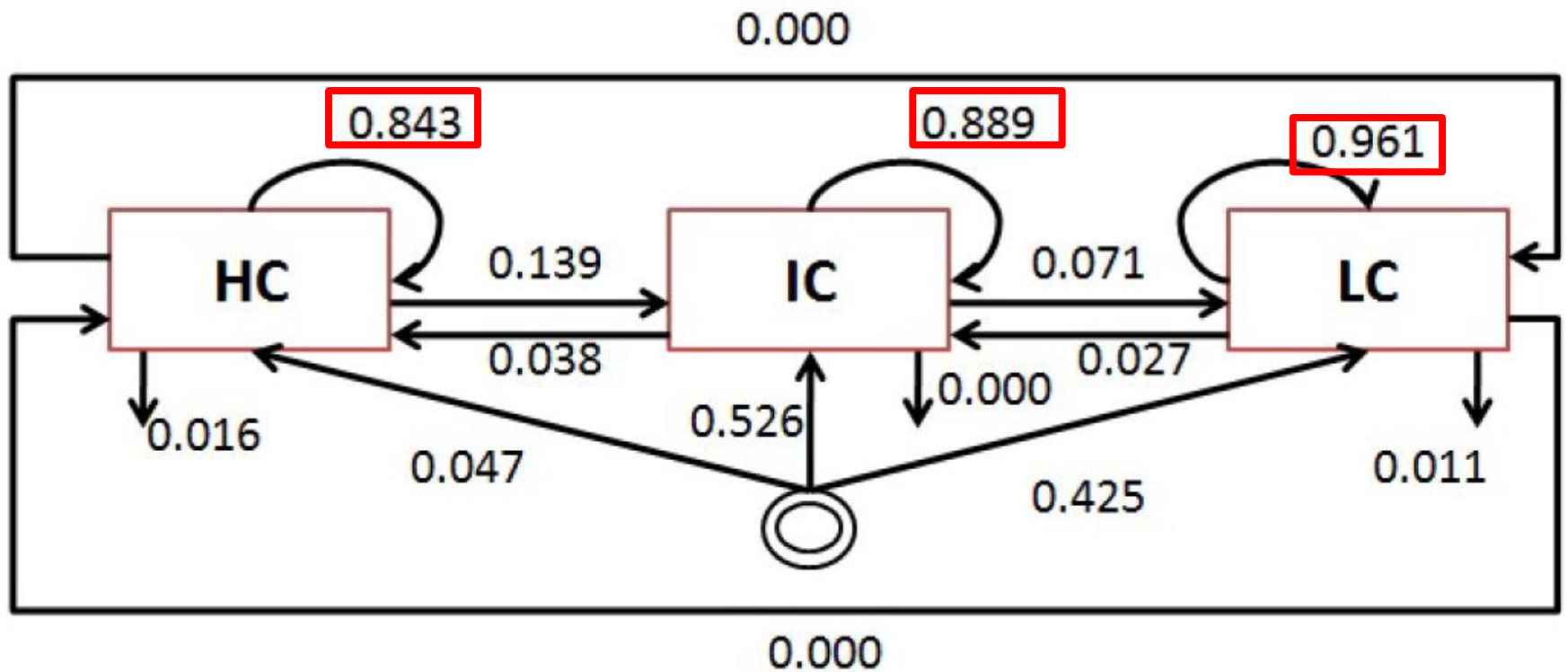
High probability of a node remaining with the same profile over a transmission

Changes in The Profile of a Node Over Time
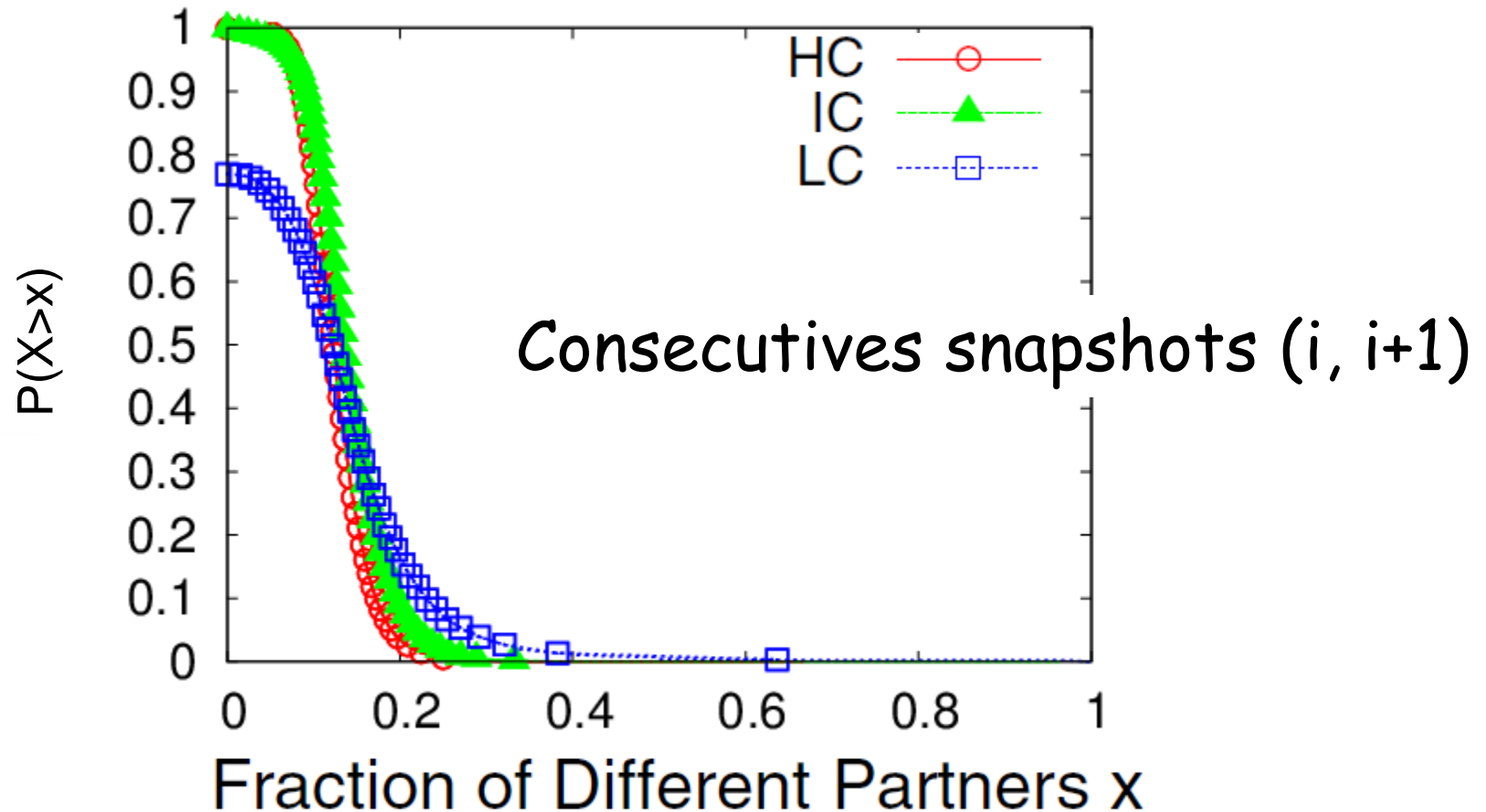

Higher probability of a node centrality to decrease than to increase

Main Steps

➢ Properties of individual nodes

  ✓ Centrality profiles

  ✓ Changes in the profile of a node over time

➢ Changes in the list of partners over time

• Properties of network as a whole

Change The Nodes Partnerships

Consecutives snapshots (i, i+1)

- Nodes have up to ≈ 30% different partners

Change The Nodes Partnerships

Non-consecutives snapshots (i, i+3)

- Nodes have up to ≈ 50% different partners

Change The Nodes Partnerships

Non-consecutives Snapshots (i, i+5)
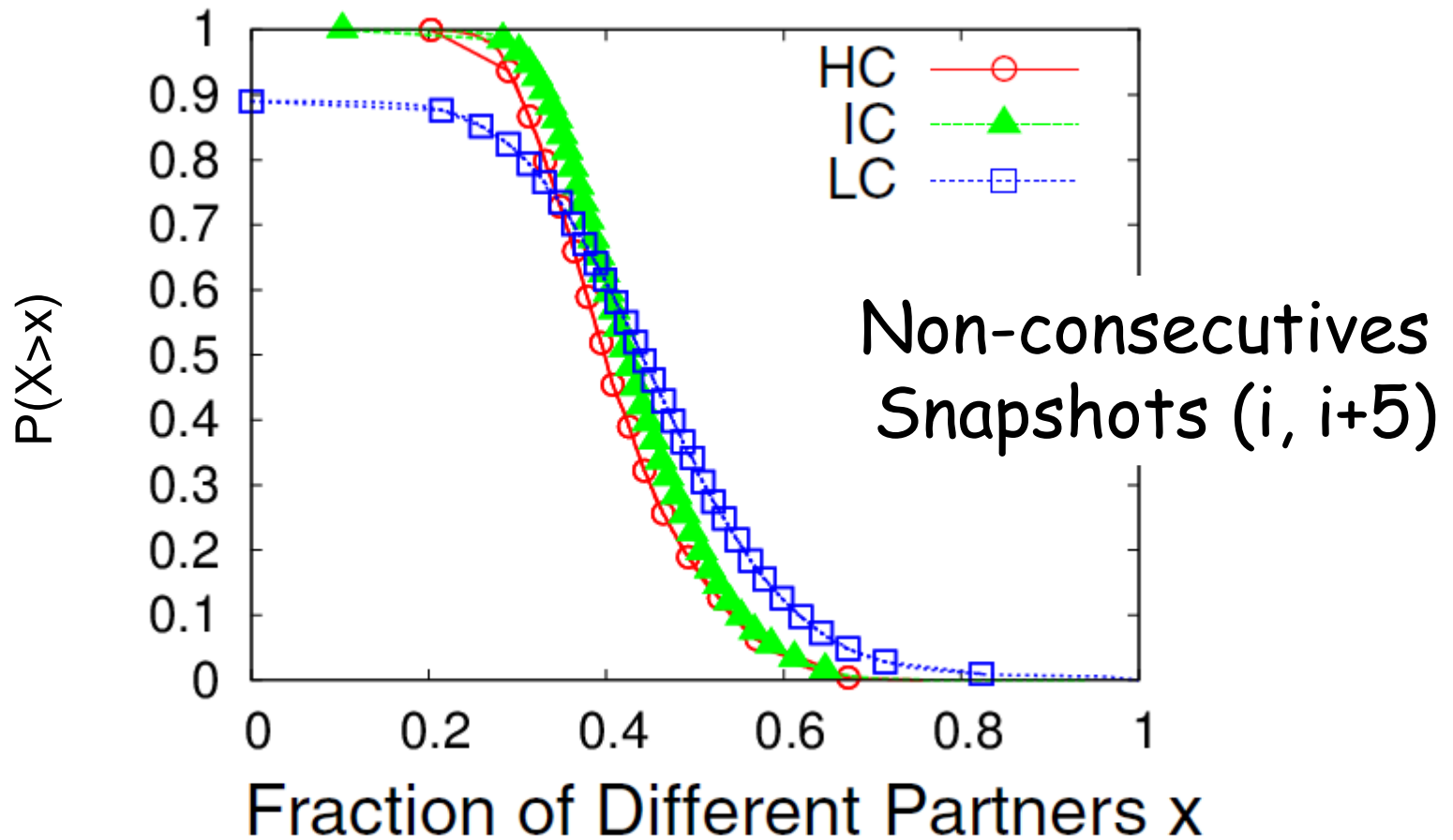
- Nodes have up to ≈ 70% different partners

Main Steps

✓ Properties of individual nodes

   ✓ Centrality profiles

   ✓ Changes in the profile of a node over time

   ✓ Changes in the list of partners over time

➢ Properties of network as a whole
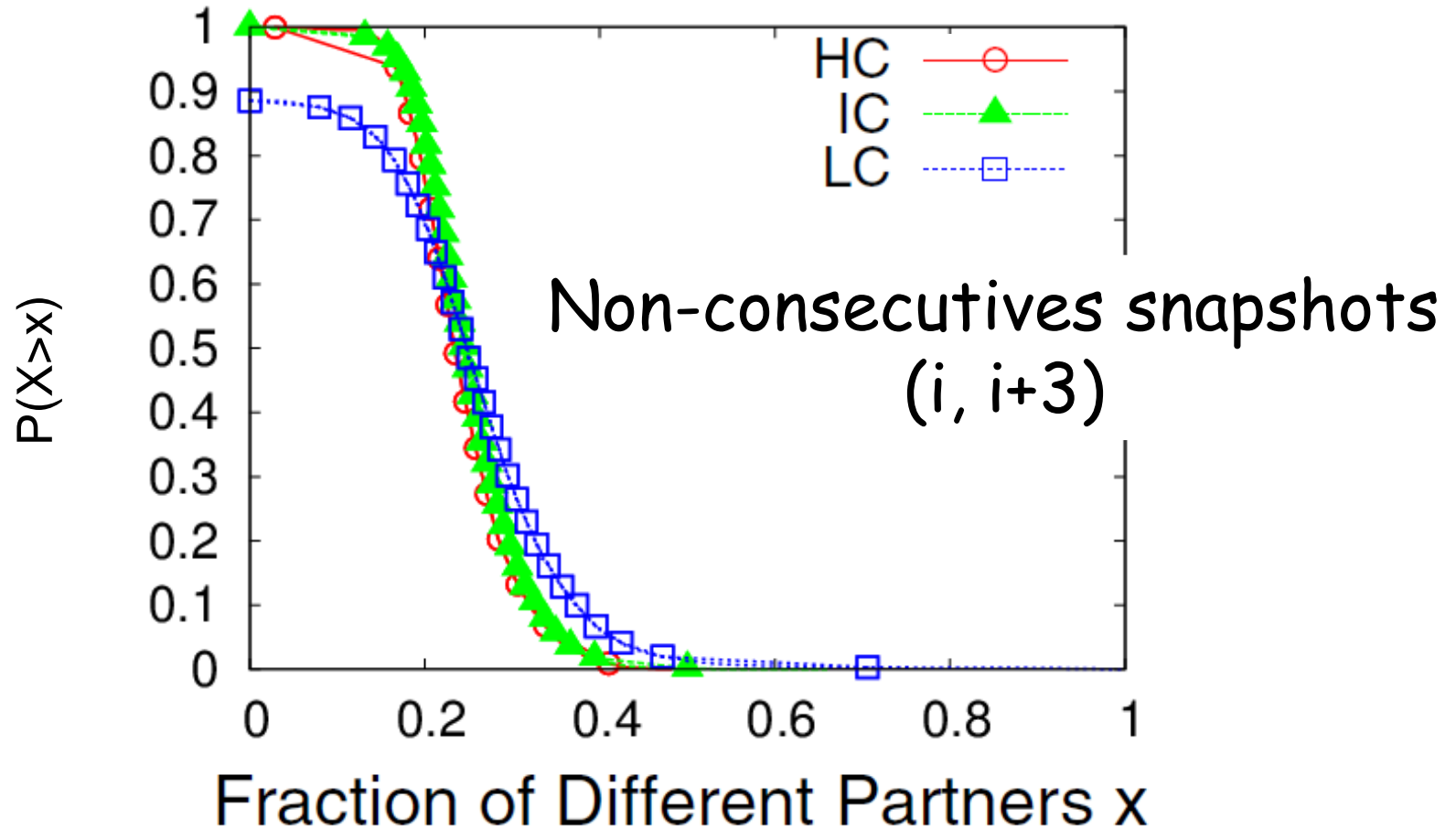
Network Properties Over a Transmission

- Diameter (network dispersion)
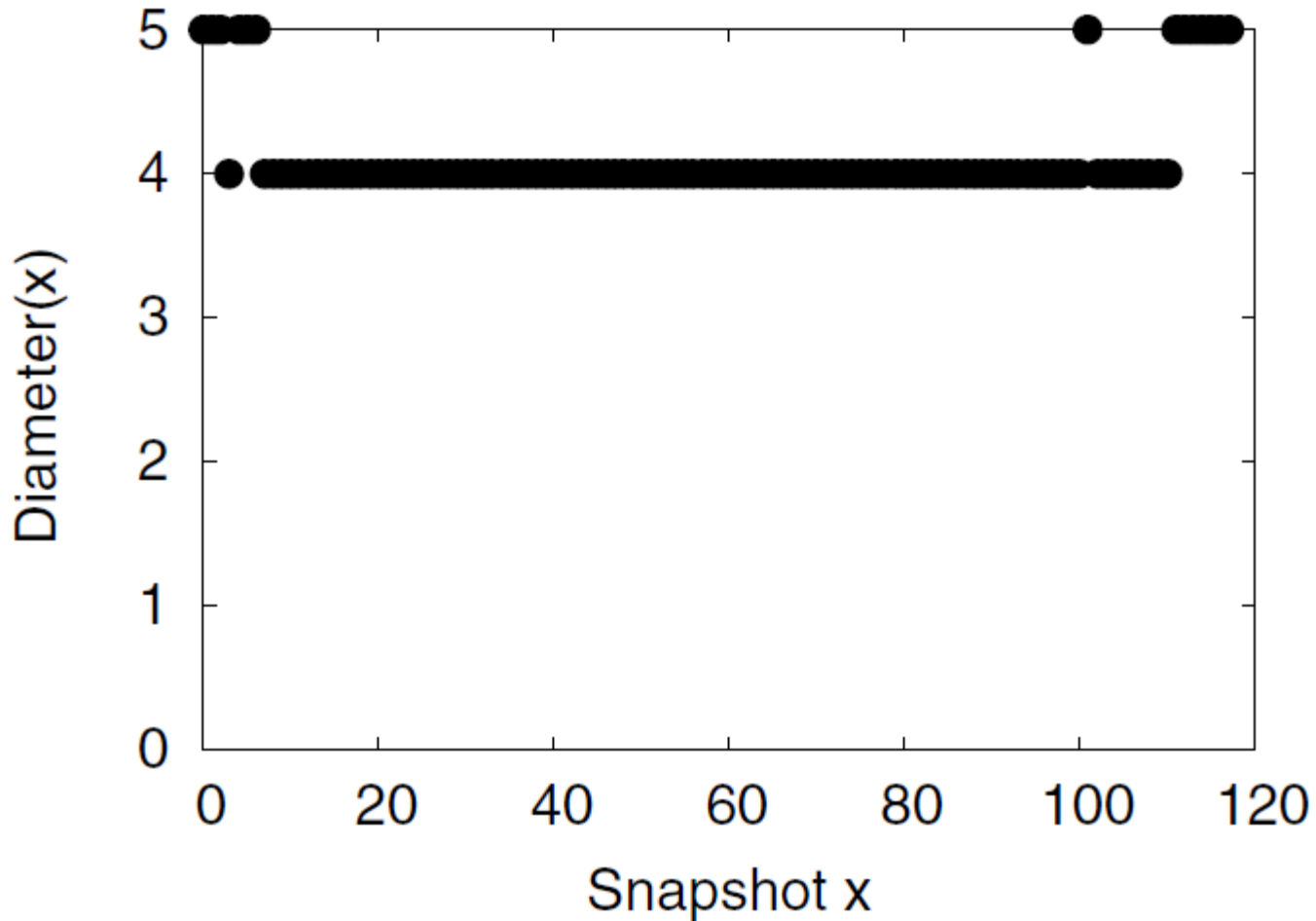 - Maximum distance between any two nodes

- Average Shortest Path
  - Shortest path of a node for all other nodes using breadth-first search

- Clustering Coefficient
  - Average node clustering coefficient
  - Node Clustering Coefficient: probability to have an edge between two of its neighbors

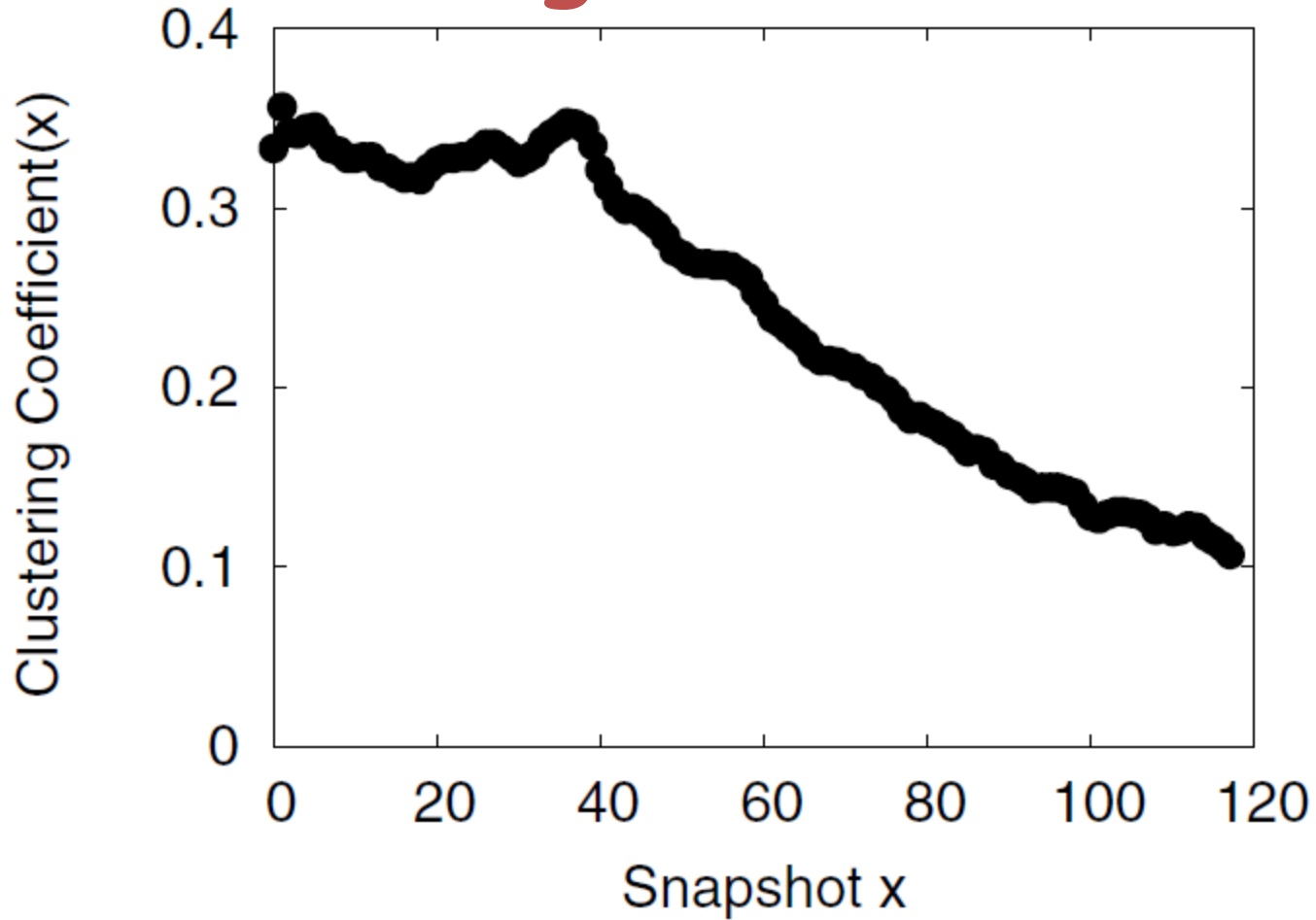- Maximum Degree: the largest degree of any node

Diameter

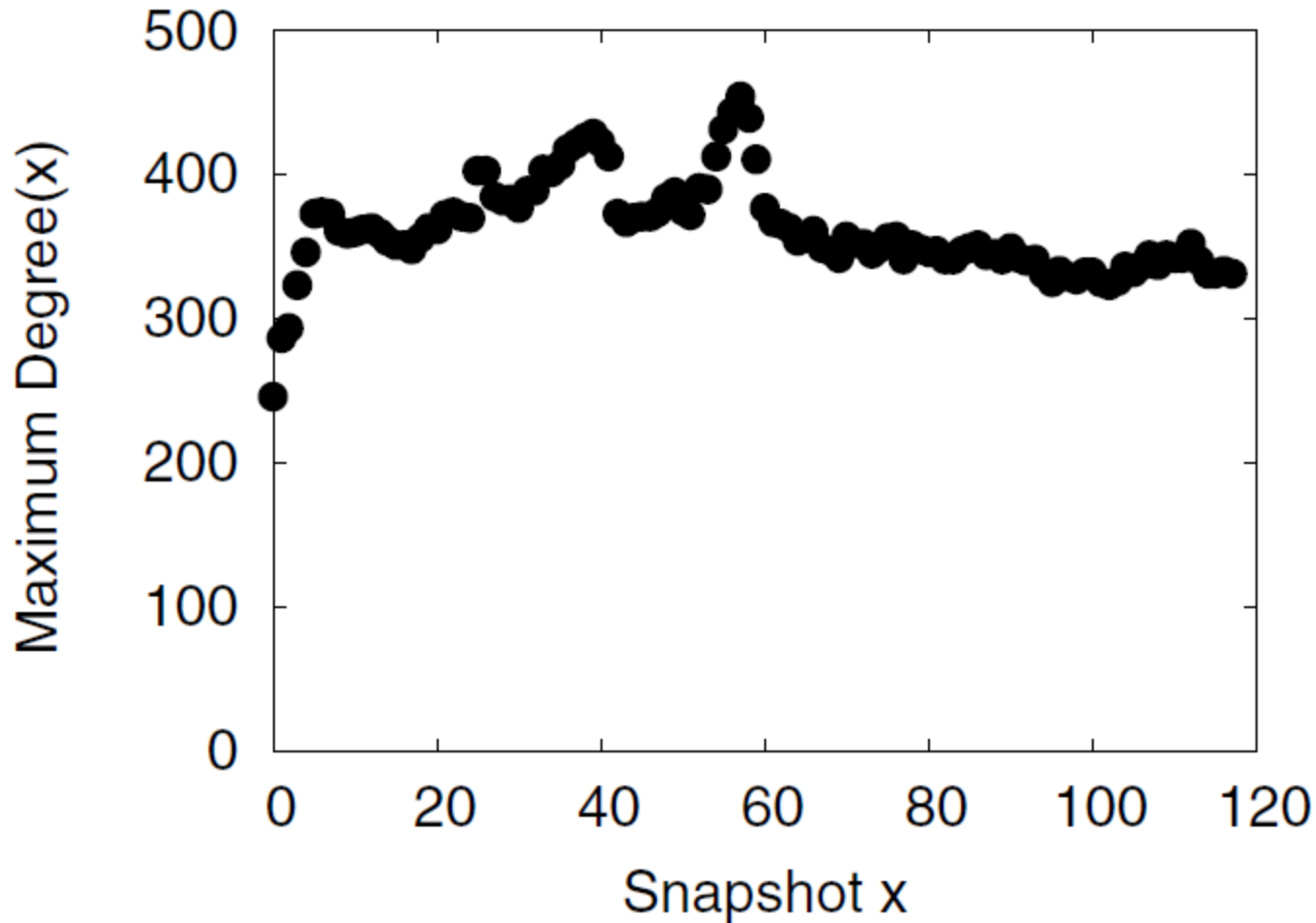It tends to remain stable between 4 and 5

Average Shortest Path
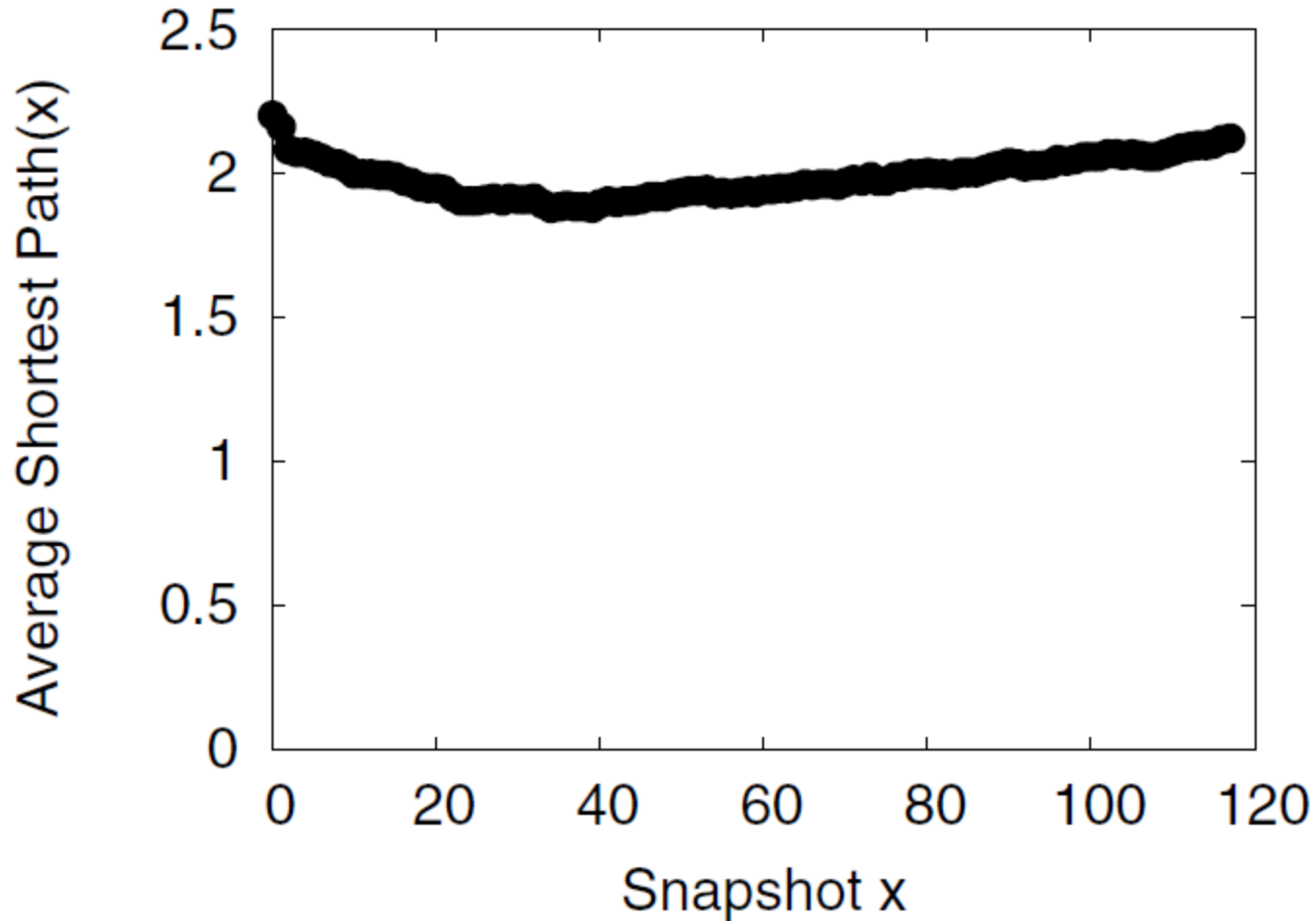
It tends to remain stable ≈ 2

Clustering Coefficient

It decreases due to new partnerships between nodes over time

Maximum Degree

It remains stable between 300 and 400 partnerships

Conclusions and Future Work

- Conclusions
  - Three centrality profiles
  - Over a transmission:
    - Nodes tend to remain with the same centrality profile, despite the change in partnerships
    - Network tends to remain stable (exception: clustering coefficient)
  - In general: little dynamism

- Future Work
  - Validate findings in other applications (PPLive)
  - Build realistic P2P live streaming simulation environments

Thanks!

Kênia Carolina – keniacarolina@dcc.ufmg.br

Alex Borges - alex.borges@ufjf.edu.br

Jussara Almeida - jussara@dcc.ufmg.br

Ana Paula Silva - anapaula.silva@ufjf.edu.br

Humberto Marques-Neto – humberto@pucminas.br

Sérgio Campos – scampos@dcc.ufmg.br

Centrality Profiles: Overview

		High Centrality (HC)	Intermediate Centrality (IC)	Low Centrality (LC)
	% Nodes	4,76%	32,69%	62,53%
Degree	Average	282,83	257,99	86,94
	CV	0,17	0,2	0,82
Betweenness	Average	3312,52	1212,52	129,45
	CV	0,45	0,3	1,6
Closeness	Average	0,005	0,008	0,005
	CV	1,25	1,2	1,16

Results for one experiment
(representative of all experiments)

Network Properties Over a Transmission

	Average	CV
Diameter	4,11	0,07
Average Shortest Path	1,98	0,03
Clustering Coefficient	0,24	0,34
Maximum Degree	361,47	0,08