On Superposition of Heterogeneous Edge Processes in Dynamic Random Graphs

Zhongmei Yao (Univerity of Dayton)
Daren B.H. Cline (Texas A&M University)
Dmitri Loguinov (Texas A&M University)

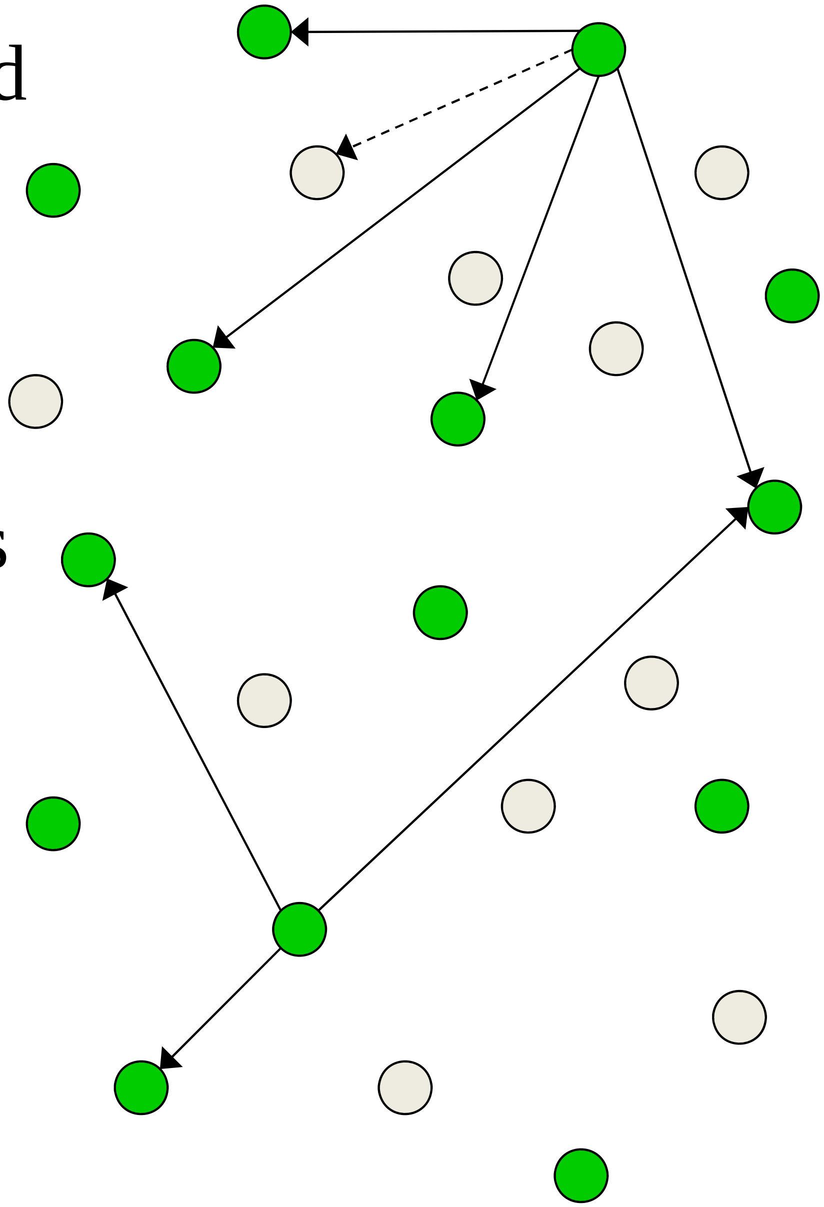
IEEE INFOCOM 2012

March 26

Agenda

- Introduction
  - Motivation and background
- General edge-creation Model
- Aggregate edge arrival process
- Wrap-up

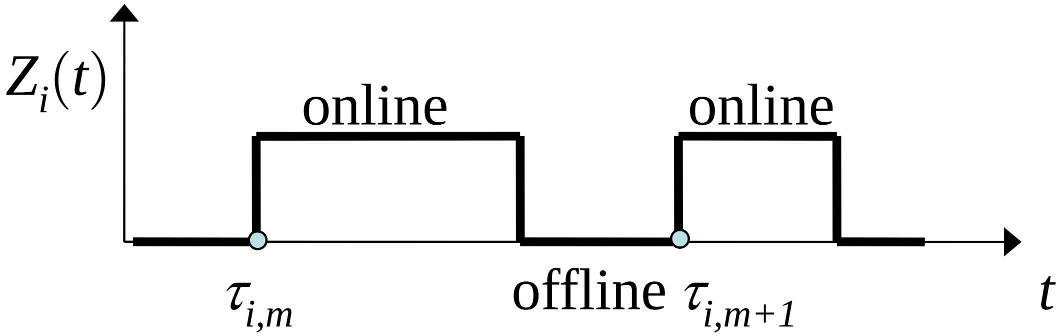
Introduction



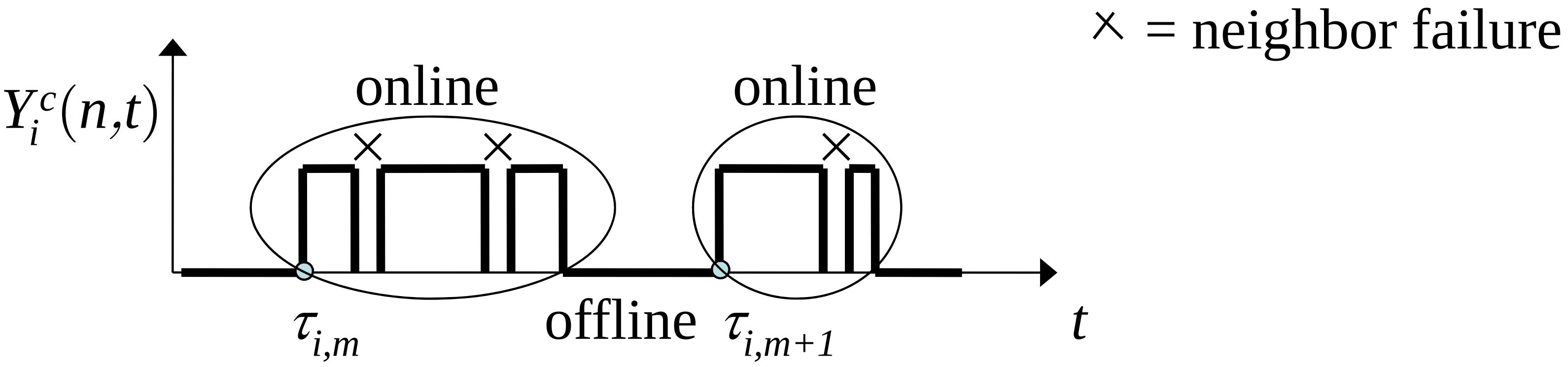
- Modern distributed systems can often be modeled as decentralized graphs
  - Nodes rely on communication services of other servers in the system
- System of n heterogeneous nodes
  - States: ON (green) and OFF (grey)
  - User on/off durations may follow different distributions
- Each user i selects k_i out-going neighbors
  - Repair links upon neighbor failure
  - Degree-irregular graphs

Introduction – Link Dynamics

- User ON/OFF processes $\{Z_i(t)\}$, $i = 1, 2, \ldots, n$

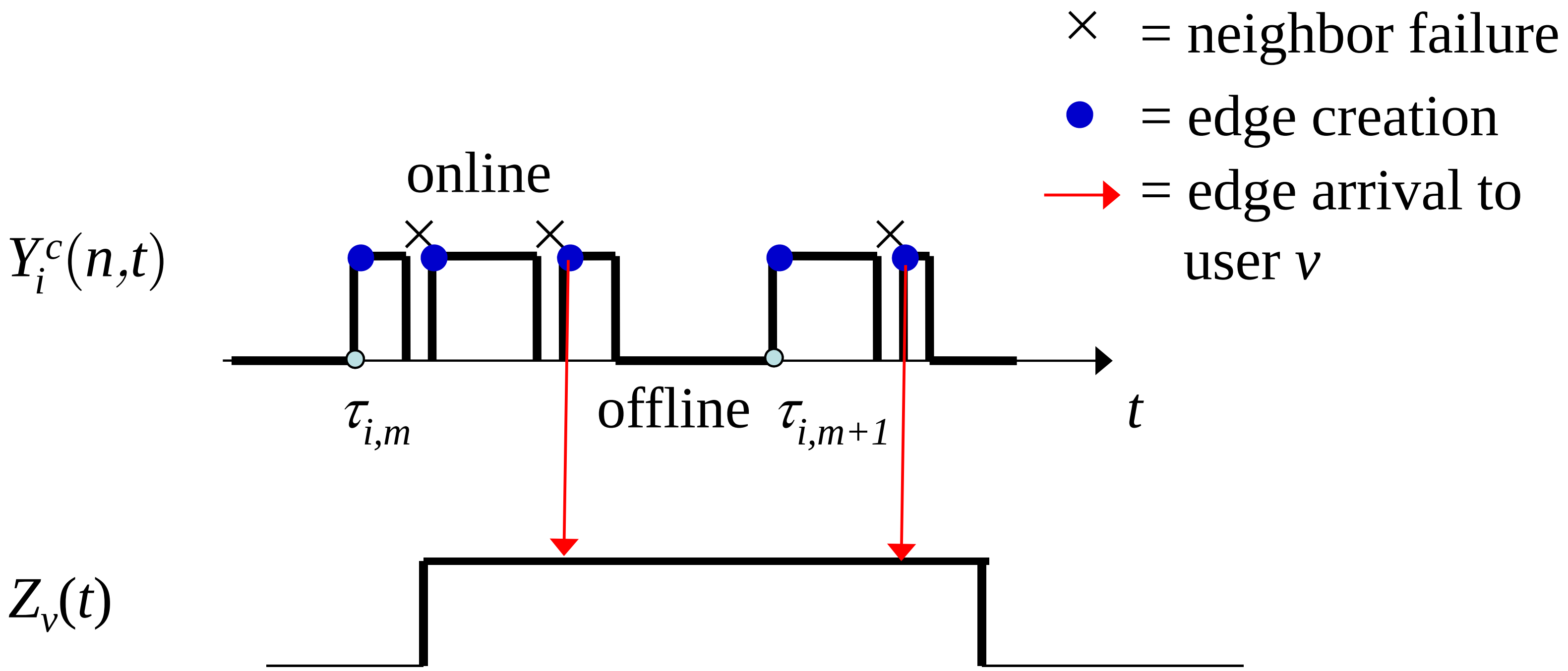


- User *i*'s link DEAD/ALIVE processes $\{Y_i^c(n,t)\}$, $c = 1, \ldots, k_i$



  - Each $\{Z_i(t)\}$ process spawns k_i link DEAD/ALIVE processes

Introduction – Edge Arrival Processes



- Let $\{\xi_{n,i}(t)\}$ be an edge-arrival process from i to v
  - Mark processes $Y_i^c(n, t)$ if user i throws edges to node v
- The superposition $\xi_n(t) = \sum_{i=1}^{n} \xi_{n,i}(t)$ is the aggregate edge arrival process from the system to v
  - More in-coming links, more likely this node will be overloaded
  - More in-coming links, smaller isolation probability

Motivation

- Previous work has analyzed numerous avenues for comprehending and improving decentralized systems
  - Graph connectivity [Gupta1998]
  - Resilience [Leonard2005, Yao2009]
  - Load balancing [Wang2007]
  - Routing mobility [Tshopp2008]
  - Improving capacity [Govindasamy2007]
- Prior studies rely on separate models
- This field has reached sufficient maturity that calls for a unifying foundation for explaining the behavior of the aggregate edge process

Related Existing Results

- The Palm-Khintchine Theorem [Heyman1982] states that the superposition process converges to Poisson in distribution if
    - Each stationary renewal process is independent from any other process;
    - Each individual process becomes sparser as n increases; and
    - The aggregate arrival rate converges to a *constant* as n increases
- The Poisson approximation on the weakly dependent superposition of sparse point processes [Chen 2006]
    - The Poisson approximation is adequate if points exhibit a locally dependent structure
- Our work is rather different
    - Due to the intricate dependency that arises in space (co-existing nodes on the graph) and time (between different lifetimes)

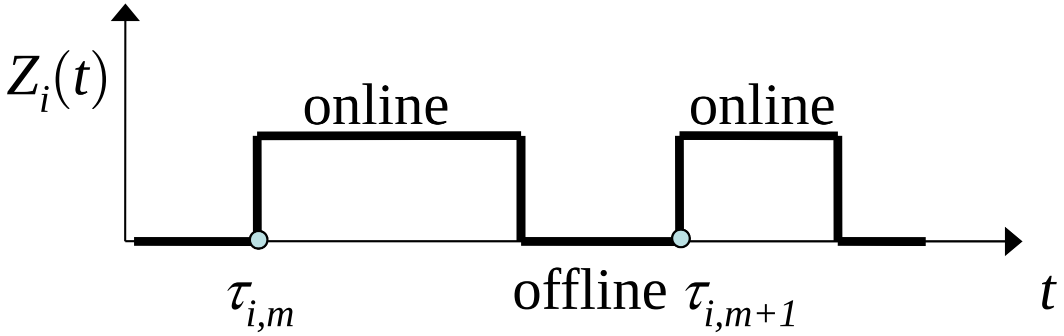
Focus of This Paper

- A complete *generic* modeling framework for understanding link dynamics
- Superposition of a large number of *dependent* edge arrival processes
- Understand *when/how* dynamic decentralized graphs develop the Poisson dynamics

Agenda

- Introduction
  - Motivation and background
- General edge-creation Model
- Aggregate edge arrival process
- Wrap-up

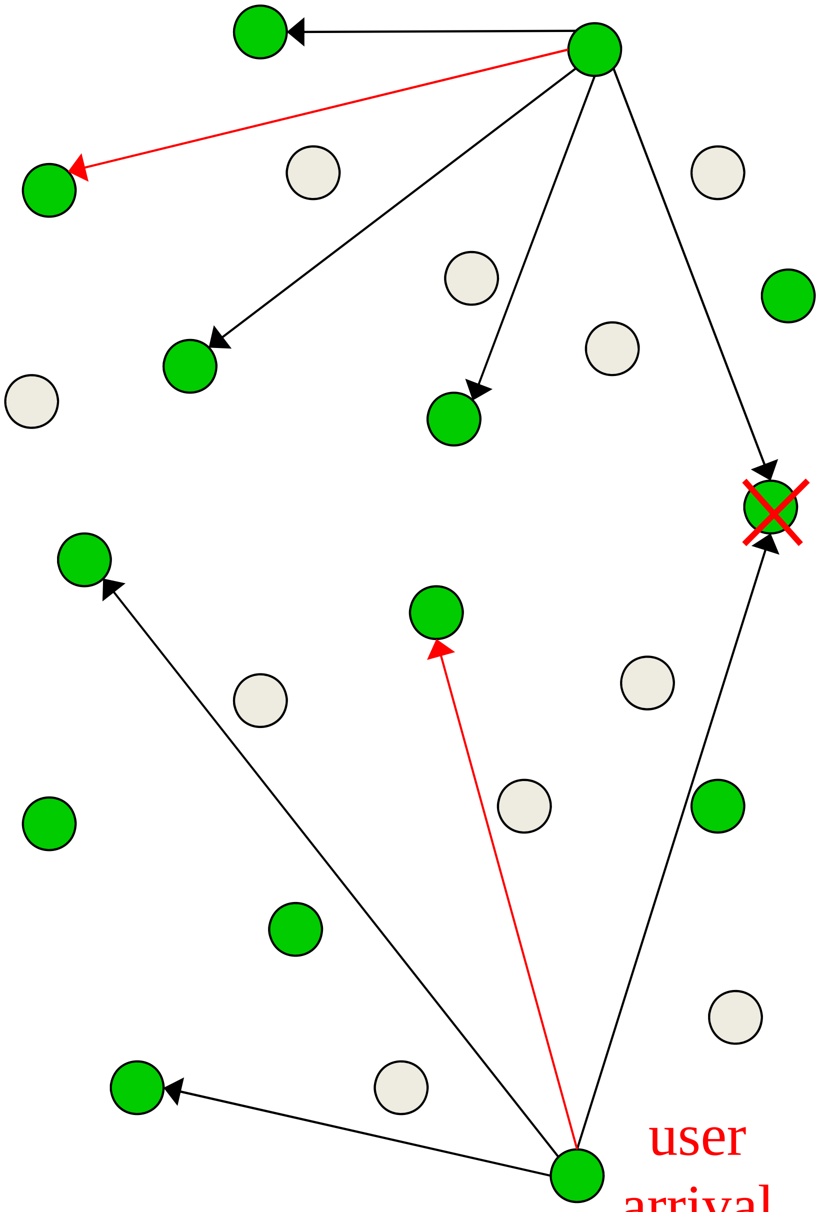
Modeling Assumptions



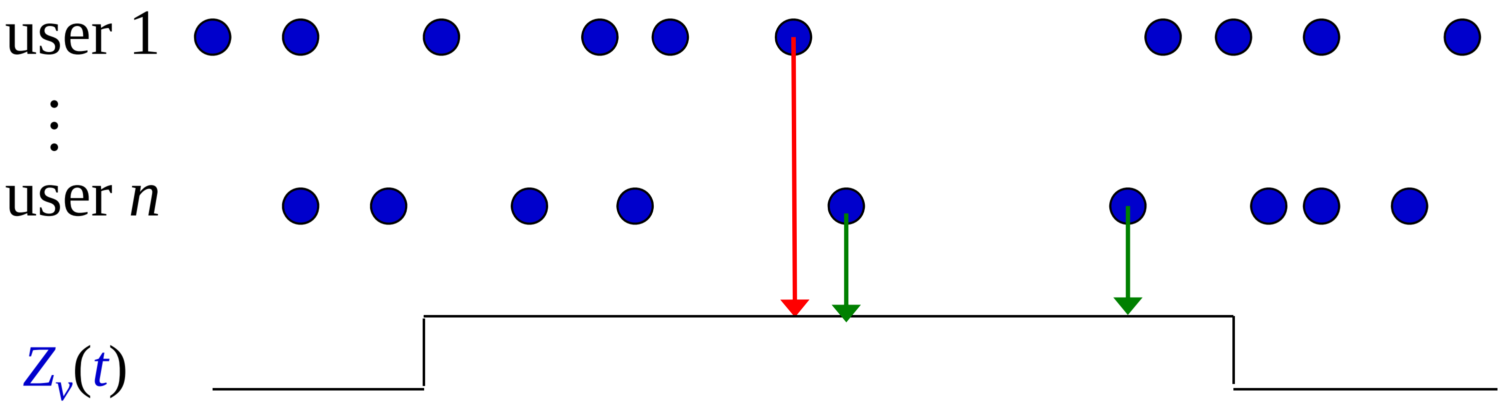
- Assumption 1 (user ON/OFF processes):
  1) Given a fixed set of user types, the user ON/OFF durations of type j respectively follow CDFs $F^{(j)}(x)$ and $G^{(j)}(x)$ with finite means
  2) Each user ON/OFF duration CDF is labeled with type j with probability p_j, where $\sum_j p_j = 1$
  3) Given that users have chosen their types, $\{Z_i(t)\}_{i=1}^{n}$ are mutually independent, stationary alternating renewal processes
- Assumption 2 (out-degree):
  – The number of outlinks k_i each user i monitors is drawn from some distribution $K(x)$ with mean k

Edge Creation Processes

- Each user arrival triggers k_i simultaneous edge-creation events
- Each user departure causes edge-replacement by all its in-degree neighbors
  - Red links shown in this figure



Edge Creation Processes 2

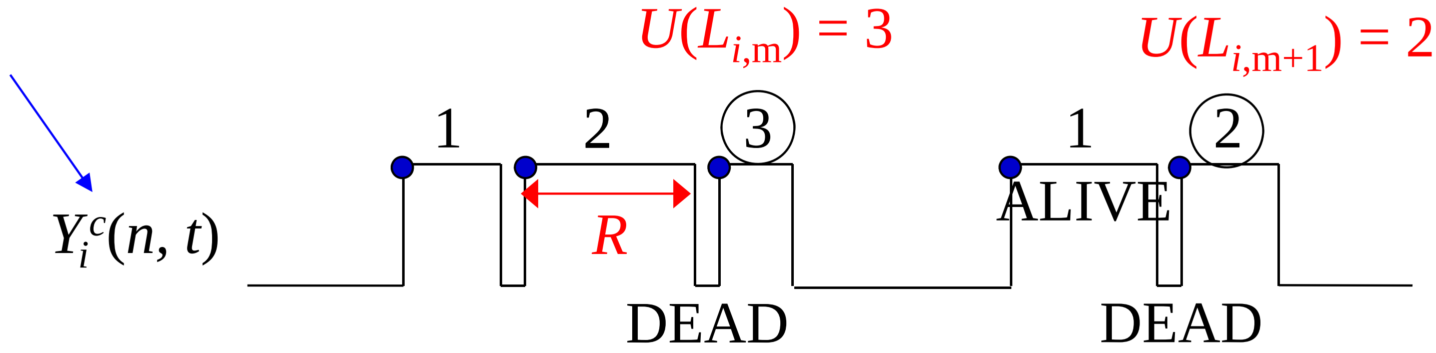


- Edge creation processes are dependent
  - Multiple users may concurrently connect to the same neighbor
  - Each out-link may point to a peer v again after v re-appears in the system
- User i's current selection depends on the history observed by i
  - As a result, the model for the number of users available at each selection time is intricate

Edge Creation Processes 3

$i = 1, \ldots, n$

$c = 1, \ldots, k_i$



- The $\{U(s)\}$ counts # of selections in an interval of length s
  - For $n \rightarrow \infty$ and uniform selection, $\{U(s)\}$ converges to a pure renewal process with cycle length $R \sim H(x)$

$$\lim_{n\to\infty} P(R < x) = l^{-1} \int_0^x (1 - F(u))du$$

mean user lifetime

aggregate user lifetime distribution

- The mean number of edges created by each i in $[0, t]$:

arrival rate of user i

$$\lim_{n\to\infty} E[W_i(n,t)|i's\ type, k_i] = k_i \lambda_i t\, E[U(L_i)]$$

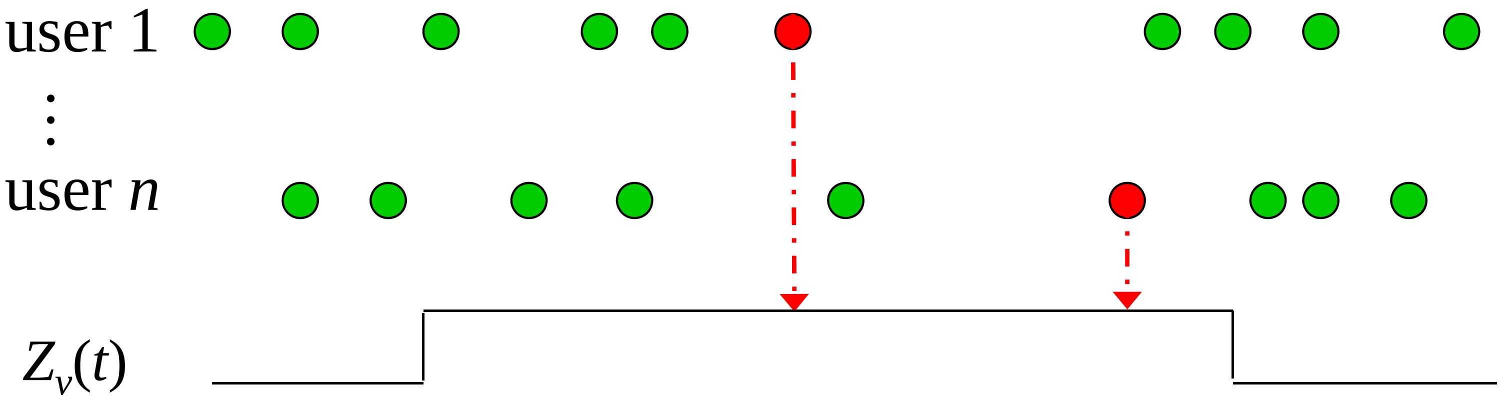
# of selections per link within i's lifetime

initial out-degree of user i

Agenda

- Introduction
  - Motivation and background
- General edge-creation Model
- Aggregate edge arrival process
- Wrap-up

Main Theorem



- Define $\xi_{n,i}(t)$ to be the edge arrival process from i to v:

# of edges that i generates in [0, t]

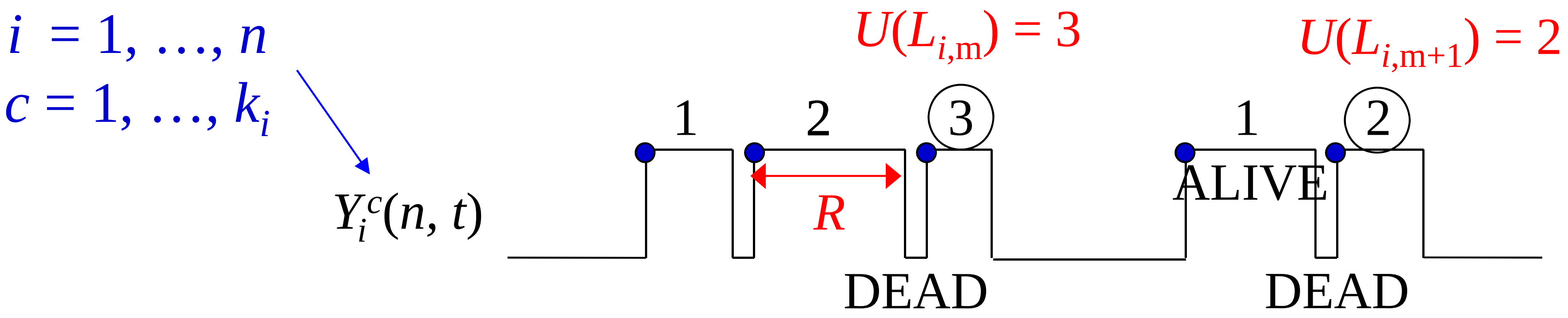
$$\xi_{n,i}(t) := \sum_{z=1}^{W_i(n,t)} I_{i,z}^{v}$$

indicator that i connects to v as its z-th selection

- Theorem 1: Under Assumptions 1-2 and uniform selection, conditioned on Z_v, the superposition $\sum_{i=1}^{n} \xi_{n,i}(t)$ converges in distribution as $n \rightarrow \infty$ to a non-homogeneous Poisson process with local rate $\gamma Z_v(t)$

constant

Proof Overview

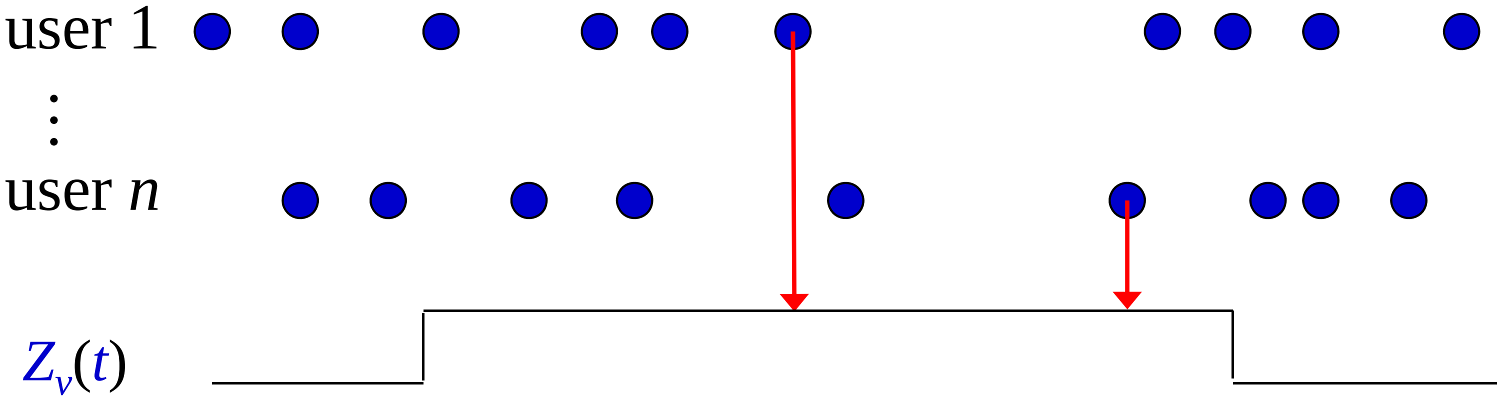


- The aggregate edge arrival rate γ to user v when v is alive converges to

$$\gamma = \lim_{n\to\infty} E\left[\frac{\sum_{i=1, i\neq v}^{n} W_i(n,t)}{t} \cdot \frac{1}{\#\ of\ live\ users}\right] = \frac{k+\theta}{l}$$

  - The edge arrival rate is the sum of the mean number of new edges k and the mean number of replacement edges θ generated per user lifetime l

Proof Overview 2



- Remaining tasks are to show [Resnick87]:
  - Continuity: the probability that no point occurs exactly at time t is 1
  - Mean convergence:

$$\forall t > 0 : \lim_{n\to\infty} E\left[\sum_{i=1,i\neq v}^{n} \xi_{n,i}(t)|Z_v\right] = \gamma \int_0^t Z_v(u)du$$

  - Probability convergence:

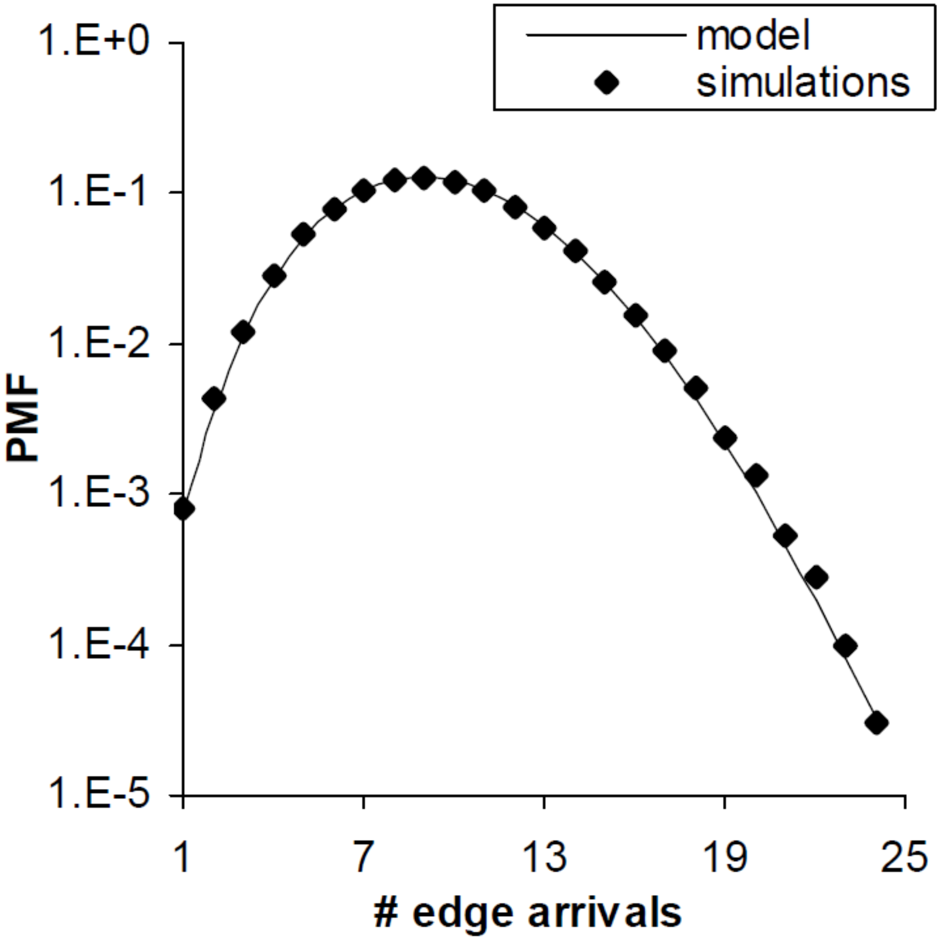
$$\forall t > 0 : \lim_{n\to\infty} P\left(\left(\sum_{i=1,i\neq v}^{n} \xi_{n,i}(t)\right) = 0|Z_v\right) = \exp\left(-\gamma \int_0^t Z_v(u)du|Z_v\right)$$

Proof Overview 3

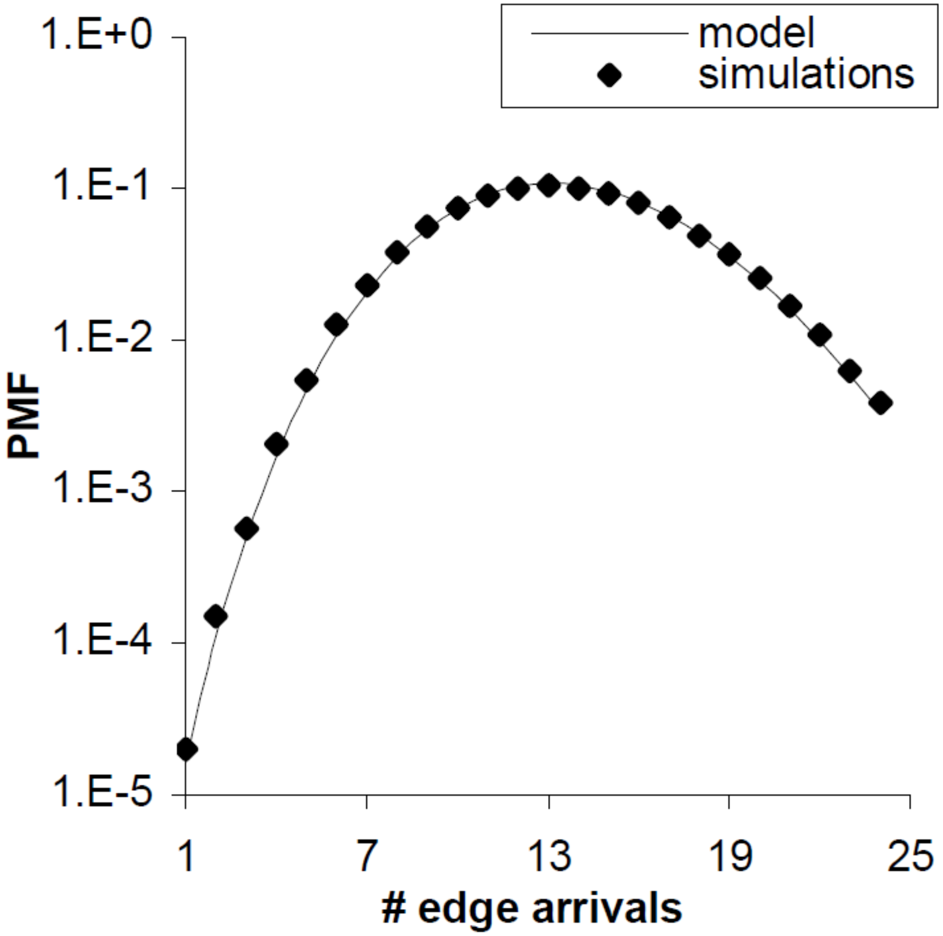
- Intuitive thinking
  - Under Assumptions 1-2 and uniform selection, as *n* increase, the pool of available users for selection becomes larger
  - The probability that each user *i* selects *any* other peer *more than once* in [0, t] becomes smaller
- To bound the above probability, we first must show that moments of collection $\{W_i(n, t)\}_{n>1}$ exist for all *n*
  - Lemma 3 in the paper
- Currently, model is intractable under other neighbor selection strategies

Simulations

- t



Pareto lifetimes with
shape parameter = 3



Pareto lifetimes with
shape parameter = 1.5

Agenda

- Introduction
    - Motivation and background
- General edge-creation Model
- Aggregate edge arrival process
- Wrap-up

Wrap-up

- A generic modeling framework for understanding user join/departure and edge arrival
- Closed-form results on the edge-arrival process to each user
- Open problems:
  - Non-uniform selection
  - Non-stationary user churn