Dynamic Random Networks with Node Churn

Luca Becchetti♥

joint work with

A. Clementi♦, F. Pasquale♦, L. Trevisan♣, and I. Ziccardi♠

♥Sapienza University of Rome, ♦Tor Vergata University of Rome,
♣Bocconi University, ♠University of l'Aquila

A day on Random Graphs - May 30th, 2022

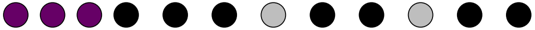
Underlying setting: Node Churn

- Time is continuous
- Nodes join and leave the system in each time unit

A set of V_t of nodes in the system at any time t

Underlying setting: Node Churn

- Time is continuous
- Nodes join and leave the system in each time unit



Nodes join and leave the network within the next time unit: $V_t \rightarrow V_{t+1}$

Underlying setting: Node Churn

- Time is continuous
- Nodes join and leave the system in each time unit

Nodes join and leave the network within the next time unit: $V_t \rightarrow V_{t+1}$

Underlying setting: Node Churn

- Time is continuous
- Nodes join and leave the system in each time unit
- Arrivals: Poisson process with rate λ;
- Departures: Nodes' lifetimes are exponential with parameter μ;
- We set $\lambda = 1$ wlog and we let $n = 1/\mu$

Lemma 1 (Pandurangan et al. 2003)

Given λ and μ such that $n = \lambda/\mu$ is sufficiently large, for every fixed real $t \geq 3n$:

$$\mathrm{P}\left(0.9n \leq |V_t| \leq 1.1n\right) \geq 1 - 2e^{-\sqrt{n}}. \tag{1}$$

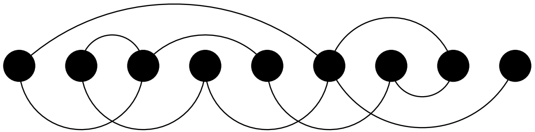
Underlying setting: Node Churn

- Time is continuous
- Nodes join and leave the system in each time unit

What if we want to maintain a network over time?

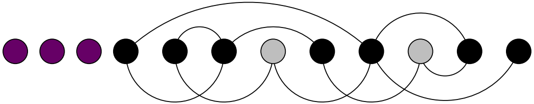
- For every t: V_t is set of nodes, e.g., clients of a P2P network;
- Set E_t of edges at time t is the result of an algorithm.

Maintaining a network



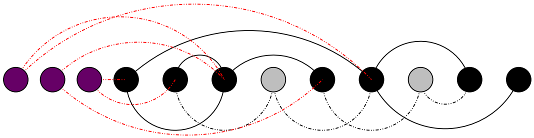
We have a network at some time t ...

Maintaining a network



New nodes arrive, some nodes depart ...

Maintaining a network



- Node u entering the network at time t: to which nodes in V_t should u connect?
- u loses an edge to a node v leaving the network at t. Should u replace the lost edge?
- **Goal:** design an algorithm to maintain a network with desired properties

Desiderata

Analytical models of Topology Dynamics with **natural** evolution rules

- **Homogenous:** All agents run the same rule at every time
- **Local:** Nodes exchange few short messages with few neighbors
- **Random:** Opportunistic/random interactions among the nodes
- **Simple:** *Natural Algorithms* (Chazelle - Comm. ACM 2012).

Desiderata

Analytical models of Topology Dynamics with **natural** evolution rules

- **Homogenous:** All agents run the same rule at every time
- **Local:** Nodes exchange few short messages with few neighbors
- **Random:** Opportunistic/random interactions among the nodes
- **Simple:** *Natural Algorithms* (Chazelle - Comm. ACM 2012).

Properties of interest:

- **Connectivity/expansion**
- **Information Spreading/flooding**

Desiderata

Analytical models of Topology Dynamics with **natural** evolution rules

- **Homogenous:** All agents run the same rule at every time
- **Local:** Nodes exchange few short messages with few neighbors
- **Random:** Opportunistic/random interactions among the nodes
- **Simple:** *Natural Algorithms* (Chazelle - Comm. ACM 2012).

Properties of interest:

- **Connectivity/expansion**
- **Information Spreading/flooding**

Technical challenge: modelling and analyzing (Random) Node Churn in simple Topology Dynamics

Dynamics /Evolving Graphs: State of the Art

Since 2008 (Avin et al - ICALP'08, Clementi et al - PODC'08), important theoretical advances in the area of Dynamic Graphs have been achieved.

- **Link-Based Graph Dynamics:**
  Probabilistic Models: Markovian Evolving Graphs
  Deterministic Models: Time-Varying Graphs
- **Node-Based Graph Dynamics:**
  Probabilistic Models: Random Walks, Random Way-Point Models
  Adversarial Models: mobile agents over grids and other graph topologies.

Dynamics /Evolving Graphs: State of the Art

Since 2008 (Avin et al - ICALP'08, Clementi et al - PODC'08), important theoretical advances in the area of Dynamic Graphs have been achieved.

- **Link-Based Graph Dynamics:**
  Probabilistic Models: Markovian Evolving Graphs
  Deterministic Models: Time-Varying Graphs
- **Node-Based Graph Dynamics:**
  Probabilistic Models: Random Walks, Random Way-Point Models
  Adversarial Models: mobile agents over grids and other graph topologies.

Crucial Constraint: Set of partipating nodes does not change.

Network Formation and Maintanance with Node Churn

Previous Analytical Work

- Dynamic-Graph Protocols with access to Central Servers and/or Random Oracles:

  Pandurangan et al. - IEEE FOCS'03
  Duchon et al. - LATIN'14

Network Formation and Maintanance with Node Churn

Previous Analytical Work

- Dynamic-Graph Protocols with access to Central Servers and/or Random Oracles:

  Pandurangan et al. - IEEE FOCS'03
  Duchon et al. - LATIN'14

- Dynamic-Graph Protocols based on Random Walks

  Cooper et Al - Combinatorics, Probability and Computing 2007
  Law and Siu - IEEE INFOCOM'03
  Augustine et al - IEEE FOCS'15

Common feature of previous work: no natural dynamics

Static setting

No node churn, static network

A very simple random graph dynamics:

Static setting

No node churn, static network

A very simple random graph dynamics:

The *d*-**Random Choice Protocol**

- **Time** $t = 0$: a set of n nodes/ agents $V_0 = V$; an empty edge set $E_0 = \emptyset$.
- **Time** $t = 1$: **Independently and u.a.r.**, each node u selects d (out-)neighbors from V and connects to each of them (discarding multi-edges). Each selected edge is added to $E_1 = E_t, \forall t \geq 1$.

Random Oracle

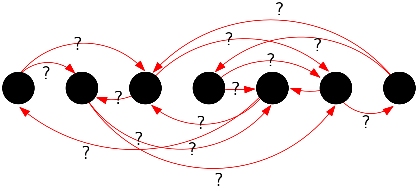
The *d*-**Random Choice Protocol** requires a simple PULL mechanism that each node can call to select one random node in the graph.

The d-Random Choice Protocol in the Static Setting

Figure: d-Random Choice Protocol

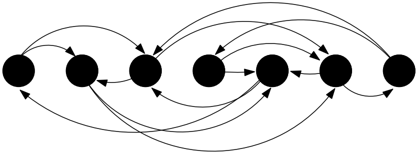
The d-Random Choice Protocol in the Static Setting

Figure: d-Random Choice Protocol



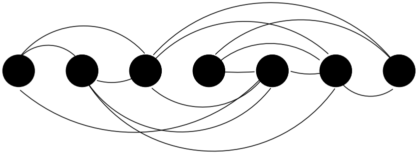
The d-Random Choice Protocol in the Static Setting

Figure: d-Random Choice Protocol



The d-Random Choice Protocol in the Static Setting

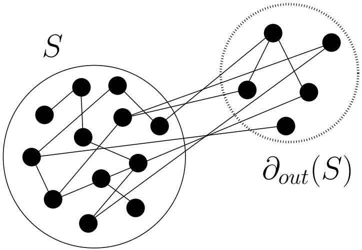
Figure: d-Random Choice Protocol



A Key notion: Vertex Expansion of a Graph

Outer boundary

Let $G = (V, E)$ be a graph of n nodes. For each $S \subseteq V$, $\partial_{out}(S)$ is the *outer boundary of* S, i.e. the set of nodes in $V - S$ with at least one neighbor in S.



A Key notion: Vertex Expansion of a Graph

Outer boundary

Let $G = (V, E)$ be a graph of n nodes. For each $S \subseteq V$, $\partial_{out}(S)$ is the *outer boundary of* S, i.e. the set of nodes in $V - S$ with at least one neighbor in S.

Vertex isoperimetric number

$$h_{out}(G) = \min_{0 \leq |S| \leq n/2} \frac{|\partial_{out}(S)|}{|S|} \tag{1}$$

Vertex expansion

Let $\varepsilon > 0$ be an arbitrary constant. Then, G is a ε-*expander* if $h_{out}(G) \geq \varepsilon$.

d-Random Choice Protocol: expansion properties

Theorem 1

For sufficiently large n, for any $d \geq 3$, the random graph $G(V, E)$ is a $\Theta(1)$-Expander, with high probability (for short, w.h.p.).

d-Random Choice Protocol: expansion properties

Theorem 1

For sufficiently large n, for any $d \geq 3$, the random graph $G(V, E)$ is a $\Theta(1)$-Expander, with high probability (for short, w.h.p.).

Proof Ingredients:

Mutually-Independent Random Choices, Standard Counting Arguments, Union Bound

d-Random Choice Protocol: expansion properties

Theorem 1

For sufficiently large n, for any $d \geq 3$, the random graph $G(V, E)$ is a $\Theta(1)$-Expander, with high probability (for short, w.h.p.).

Proof Ingredients:

Mutually-Independent Random Choices, Standard Counting Arguments, Union Bound

COROLLARY

The diameter of G and its Flooding/Rumor-Spreading Time is $O(\log n)$, w.h.p..

Models

Topology Dynamics

We adapt the *d*-**Random Choice Dynamics** to a Dynamic Framework where:

- Churn: nodes join/leave the network according to Poisson arrivals/exponential departure times.
- Edges incident to leaving nodes disappear;
- Topology dynamics: do active nodes replace disappeared edges?
  - Yes $\rightarrow$ edge regeneration
  - No

Models

Topology Dynamics

We adapt the d-**Random Choice Dynamics** to a Dynamic Framework where:

- Churn: nodes join/leave the network according to Poisson arrivals/exponential departure times.
- Edges incident to leaving nodes disappear;
- Topology dynamics: do active nodes replace disappeared edges?
  - Yes $\rightarrow$ edge regeneration
  - No

Caveat

- Edges are undirected but ...
- We speak of the *outgoing edges* of a node u as those edges that resulted from u's connections requests (consistently, we speak of v's *out-degree*).

Poisson Dynamics with(out) Edge Regeneration - PDGR(PDG)

A PDGR (PDG) $\mathcal{G}(\lambda, \mu, d)$ is a stochastic process $\{G_t = (V_t, E_t) : t \in \mathbb{R}^+\}$, where:

Node Churn

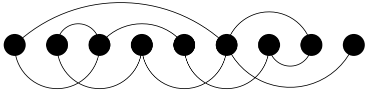
- $V_0 = \emptyset$. Nodes joining V_t follow a sequential *Poisson process with mean* λ.
- Once in V_t the *life time* of a node has *exponential distribution with parameter* μ.

Topology: *d*-Random Choice Dynamics.

- Initially, $E_0 = \emptyset$. For $t > 0$, a node joinining the network at time t selects d (out-)neighbors from V_t **independently and u.a.r.**
- If node v leaves the graph at time t, then:
  1. All edges incident to v disappear;
  2. **Regeneration:** Each node in V_t losing outgoing edges to v selects new neighbours from V_t (independently and u.a.r) to restore its out-degree to d. Hence, after any node churn, every node has exactly d out-edges.

Poisson Dynamics with edge Regeneration - PDGR

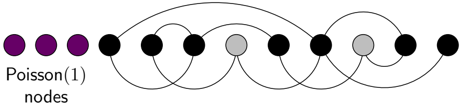
Figure: Poisson Model



t

Poisson Dynamics with edge Regeneration - PDGR

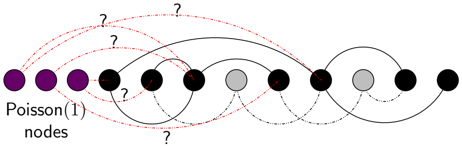
Figure: Poisson Model



$t + 1$

Poisson Dynamics with edge Regeneration - PDGR

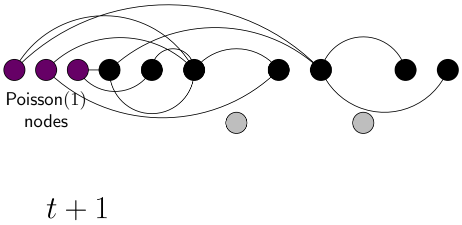
Figure: Poisson Model



$t + 1$

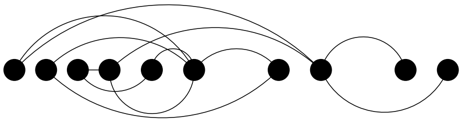
Poisson Dynamics with edge Regeneration - PDGR

Figure: Poisson Model



Poisson Dynamics with edge Regeneration - PDGR

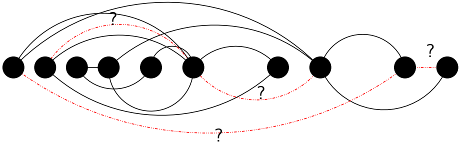
Figure: Poisson Model



$t + 1$

Poisson Dynamics with edge Regeneration - PDGR

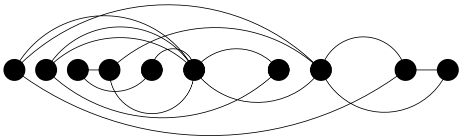
Figure: Poisson Model



$t + 1$

Poisson Dynamics with edge Regeneration - PDGR

Figure: Poisson Model



$t + 1$

Simplified streaming model used in this presentation

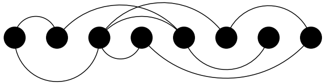
A *Streaming Dynamic Graph with(out) edge Regeneration* SDGR (SDG) $\mathcal{G}(n, d)$ is a stochastic process $\{G_t = (V_t, E_t), t \geq 1\}$ defined as follows.

- **Node Churn Events.** $V_0 = \emptyset$. At each round $t \geq 1$, a new node joins V_t and *it stays alive* up to round $t + n$, then it leaves the network. So, at every $t \geq n$, the *oldest* node v leaves the network and a *new node* u joins it, i.e., $V_t := (V_{t-1} \setminus \{v\}) \cup \{u\}$.
- **Topology: The d-Random Choice Dynamics.** E_t evolves as follows:
  i) All edges incident to **leaving** node v disappear;
  ii) The new node u selects d (out-)neighbors from N_t *independently and u.a.r.*;
  iii) **Regeneration:** Nodes in V_t that lose any out-going edges to v, select new neighbours (independently and u.a.r from V_t) to restore their (out-)degrees to d.

Streaming Model: SDGR $\mathcal{G}(n, d)$

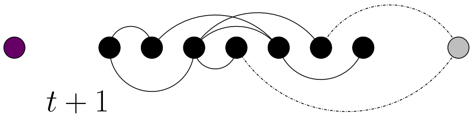
Figure: Streaming Model

t



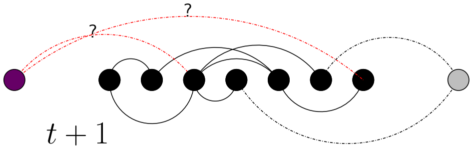
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



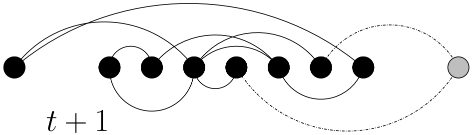
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



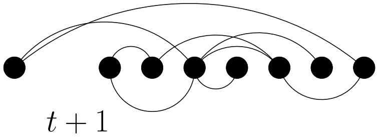
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



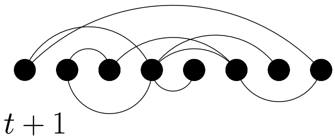
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



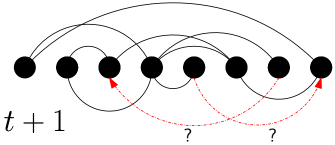
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



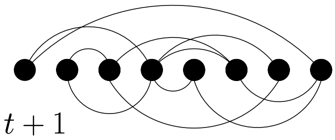
Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



Streaming Model: SDGR $\mathcal{G}(n, d)$

Figure: Streaming Model



Vertex expansion with edge regeneration

Edge regeneration affords expansion w.h.p.

Theorem 2

- Streaming Model SDGR $\mathcal{G}(n, d)$. *For any sufficiently large d (i.e. $d \geq 14$), and for any $t \geq \Omega(n)$, the snapshot $G_t(V_t, E_t)$ is a $(1/10)$-expander, with probability $1 - 1/n^{\Theta(d)}$.*

Vertex expansion with edge regeneration

Edge regeneration affords expansion w.h.p.

Theorem 2

- Streaming Model SDGR $\mathcal{G}(n, d)$. *For any sufficiently large d (i.e. $d \geq 14$), and for any $t \geq \Omega(n)$, the snapshot $G_t(V_t, E_t)$ is a $(1/10)$-expander, with probability $1 - 1/n^{\Theta(d)}$.*
- Poisson Model PDGR $\mathcal{G}(\lambda, \mu, d)$. *Let $\lambda = 1$ and $n = 1/\mu$, and let $d \geq 35$. Then, for any $t \geq \Omega(n \log n)$, the snapshot $G_t(V_t, E_t)$ is a $(1/10)$-expander, with probability $1 - 1/n^{\Theta(1)}$.*

Vertex expansion without edge regeneration

For every $t \geq n$: fraction $\Omega(e^{-d})$ isolated nodes w.h.p. but ...

Vertex expansion without edge regeneration

For every $t \geq n$: fraction $\Omega(e^{-d})$ isolated nodes w.h.p. but ...

Lemma 3 (Expansion of large subsets)

For every constant $d \geq 20$ and for every sufficiently large n, let $\{G_t = (V_t, E_t) \,:\, t \in \mathbb{N}\}$ be an SDG sampled from $\mathcal{G}(n, d)$. For every fixed $t \geq n$, w.h.p. the snapshot G_t satisfies the following:

$$\min_{S \subseteq N_t \,:\, ne^{-d/10} \leq |S| \leq n/2} \frac{|\partial_{out}(S)|}{|S|} \geq 0.1\,.$$

Vertex expansion without edge regeneration

For every $t \geq n$: fraction $\Omega(e^{-d})$ isolated nodes w.h.p. but ...

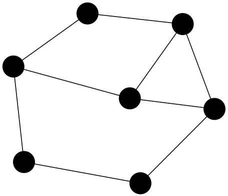
Lemma 3 (Expansion of large subsets)

For every constant $d \geq 20$ and for every sufficiently large n, let $\{G_t = (V_t, E_t) \,:\, t \in \mathbb{N}\}$ be an SDG sampled from $\mathcal{G}(n, d)$. For every fixed $t \geq n$, w.h.p. the snapshot G_t satisfies the following:

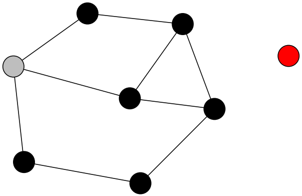
$$\min_{S \subseteq N_t \,:\, ne^{-d/10} \leq |S| \leq n/2} \frac{|\partial_{out}(S)|}{|S|} \geq 0.1\,.$$

Similar result for PDGR.

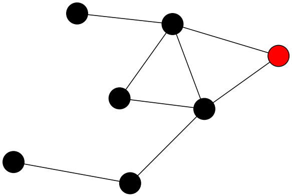
Flooding (SDGR)



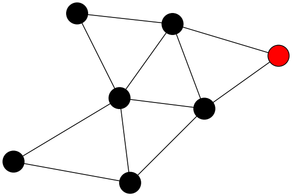
Flooding (SDGR)



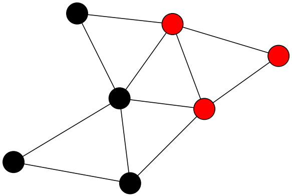
Flooding (SDGR)



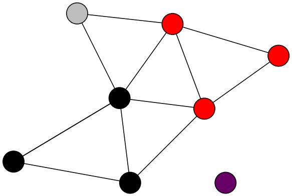
Flooding (SDGR)



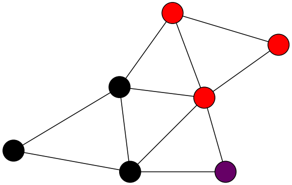
Flooding (SDGR)



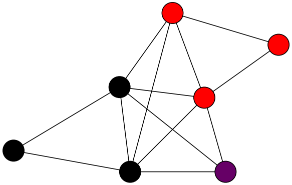
Flooding (SDGR)



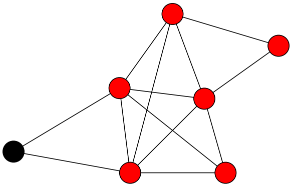
Flooding (SDGR)



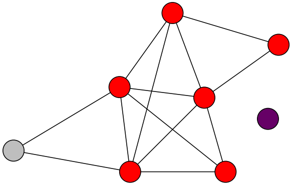
Flooding (SDGR)



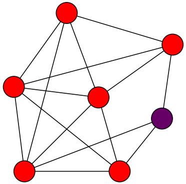
Flooding (SDGR)



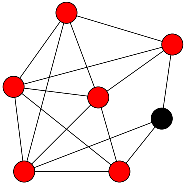
Flooding (SDGR)



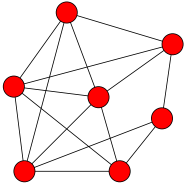
Flooding (SDGR)



Flooding (SDGR)



Flooding (SDGR)



Flooding in the Streaming Model

Set of informed nodes

Consider a SDGR $\mathcal{G}(n,d) = \{G_t = (V_t, E_t),\ t \geq 0\}$. Let s be the informed node joining the graph at round t_0 and let $I_0 = \{s\} \subseteq V_{t_0}$
Then, at each round $t \geq t_0$, after applying the d-**Random Choice Dynamics**, define I_t iteratively as follows:

$$I_t = \left(I_{t-1} \bigcup I'_t\right) \bigcap V_t\,, \text{ where } I'_t = \{v \in V_{t-1} | \exists u \in I_{t-1} : (u,v) \in E_{t-1}\}$$

Flooding with Edge Regeneration

Flooding completes in $\mathcal{O}(\log n)$ rounds whp

Theorem 4

- Streaming Model SDGR $\mathcal{G}(n, d)$. *For any sufficiently large* d *(i.e.* $d \geq 14$*), and for any* $t \geq \Omega(n)$*. Then, if an informed node joins in step* t*, after* $O(\log n)$ *time steps, all nodes of the network will be informed, w.h.p.*

Flooding with Edge Regeneration

Flooding completes in $\mathcal{O}(\log n)$ rounds whp

Theorem 4

- Streaming Model SDGR $\mathcal{G}(n, d)$. *For any sufficiently large d (i.e. $d \geq 14$), and for any $t \geq \Omega(n)$. Then, if an informed node joins in step t, after $O(\log n)$ time steps, all nodes of the network will be informed, w.h.p.*
- Poisson Model PDGR $\mathcal{G}(\lambda, \mu, d)$. *Let $\lambda = 1$ and $n = 1/\mu$, and let $d \geq 35$. Then, for any $t \geq \Omega(n \log n)$, if an informed node joins at step t, after $O(\log n)$* flooding steps, *all nodes of the network will be informed, w.h.p.*

Flooding without Edge Regeneration

Flooding can take long to complete $\rightarrow \Omega_d(m)$ rounds whp ...

Flooding without Edge Regeneration

Flooding can take long to complete $\to \Omega_d(m)$ rounds whp ...but most nodes informed quickly most of the times:

Theorem 5

For constant d sufficiently large, for every sufficiently large n and for every fixed $t_0 \geq n$, there exists $\tau = \mathcal{O}(\log n / \log d + d)$, such that the flooding over SDG $\mathcal{G}(n, d)$ starting at t_0 satisfies the following:

$$\mathrm{P}\left(|I_{t_0+\tau}| \geq (1 - e^{-d/10})n\right) \geq 1 - e^{-\Omega(d)},$$

Streaming Model SDGR Technical proofs

Expansion of $G_t = (V_t, E_t)$: main issues and key steps

Technical Issue. Differences in life times of nodes in V_t induce i) correlations among edges in E_t and ii) non uniform edge probabilities $\rightarrow$ Edges incident to **old** nodes are more likely to belong to E_t.

Streaming Model SDGR Technical proofs

Expansion of $G_t = (V_t, E_t)$: main issues and key steps

Technical Issue. Differences in life times of nodes in V_t induce i) correlations among edges in E_t and ii) non uniform edge probabilities $\rightarrow$ Edges incident to **old** nodes are more likely to belong to E_t.

Lemma 6

Let $k \leq t-1$ and let u be the node with age $k+1$. Then, if node $v \in V_t$ was born before u, probability that a specific out-going edge from u has destination v is

$$\frac{1}{n-1}\left(1+\frac{1}{n-1}\right)^k . \qquad (2)$$

If v was born after u, above probability is always $\leq \frac{1}{n-1}$.

Good News. Since $k \leq n$, Eq. (2) is $\leq \Theta(1/n)$.

Streaming Model SDGR Technical proofs

Theorem 7

Let n be sufficiently large and $d \geq 21$. Then, for any $t \geq n$, the snapshot G_t of a SDGR $\mathcal{G}(n, d)$ is a vertex expander with parameter $\varepsilon \geq 0.1$, w.h.p.

Proof Strategy

We consider two cases:
Case 1. Small subsets, i.e., $|S| \leq n/4$,
Case 2. Large subsets, i.e., $n/4 \leq |S| \leq n/2$,

Remark

In both cases, the S expansion is obtained by only looking at the **out-going** edges of set S, i.e., those edges determined by the d random choices of each node in S.

Streaming Model SDGR Technical proofs

Lemma (Case 2)

For every pair of vertex subsets (S, T) with $|S| \leq n/4$ and $|T| = 0.1|S|$, such that $S \cap T = \emptyset$, the event "*all the out-neighbors of* S *are in* T", i.e. $\partial_{out}(S) \subseteq T$, does happen with negligible probability, i.e., with probability $O(1/n^{\Theta(1)})$.

Streaming Model SDGR Technical proofs

Lemma (Case 2)

For every pair of vertex subsets (S, T) with $|S| \le n/4$ and $|T| = 0.1|S|$, such that $S \cap T = \emptyset$, the event *"all the out-neighbors of S are in T"*, i.e. $\partial_{out}(S) \subseteq T$, does happen with negligible probability, i.e., with probability $O(1/n^{\Theta(1)})$.

Proof

For any S and any $T \subseteq N_t - S$, we define the event $A_{S,T} = \{\partial_{out}(S) \subseteq T\}$ So, we have that

$$\Pr\left(\min_{n/4 \le |S| \le n/2} \frac{|\partial_{out}(S)|}{|S|} \le 0.1\right) \le \sum_{\substack{n/4 \le |S| \le n/2 \\ |T| = 0.1|S|}} \Pr\left(A_{S,T}\right) . \tag{3}$$

The next step is to upper bound $\Pr\left(A_{S,T}\right)$.

Streaming Model SDGR Technical proofs

Lemma (Case 2)

$\Pr(A_{S,T})$ is upper bounded by the probability that each outgoing edge of each node in S has destination in $S \cup T$.
From Lemma 6, since $k \le n-1$, the probability that any outgoing edge of u has destination some node v is at most $e/(n-1)$.
Since $\partial_{out}(S) \subseteq T$ implies that every outgoing edge of $u \in S$ has destination in $S \cup T$ we have:

$$\Pr(A_{S,T}) \le \left(\frac{\mathrm{e}}{\mathrm{n}-1} \cdot |S \cup T|\right)^{d|S|} . \tag{4}$$

So, from (3) and (4), for any $d \ge 21$, and standard calculus,

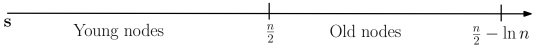
$$\Pr\left(\min_{1\le|S|\le n/4} \frac{|\partial_{out}(S)|}{|S|} \le 0.1\right) \le \sum_{s=1}^{n/4} \binom{n}{s}\binom{n-s}{0.1s}\left(\frac{1.1s\cdot e}{n-1}\right)^{ds} \le \frac{1}{n^4} . \tag{5}$$

Alternating paths argument

Node demographics

With respect to node s joining at round t_0:

- **Young nodes:** $age \leq n/2$;
- **Old nodes:** $n/2 < age \leq n/2 - \ln n$;
- **Very old nodes:** $age > \ln n$.

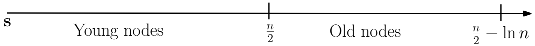


Alternating paths argument

Node demographics

With respect to node s joining at round t_0:

- **Young nodes:** $age \leq n/2$;
- **Old nodes:** $n/2 < age \leq n/2 - \ln n$;
- **Very old nodes:** $age > \ln n$.

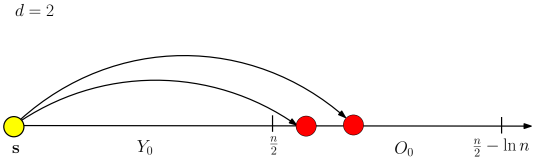


Key ingredients: i) we only consider outgoing edges from young to old nodes; ii) we neglect edges from young nodes to very old ones; iii) we neglect nodes arriving after s.
Remark: i), ii) and iii) $\Longrightarrow$ we establish a subset of the edges that will exist in the window $[t_0, t_0 + \ln n]$.

Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.

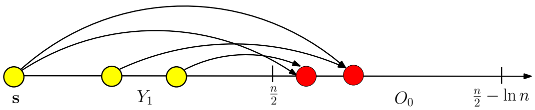


Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.

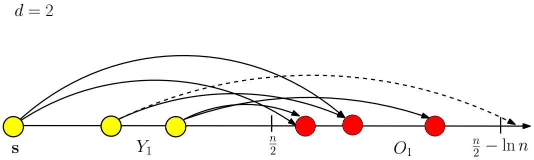
$d = 2$



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

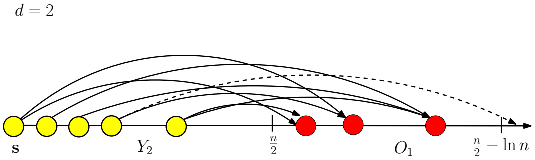
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

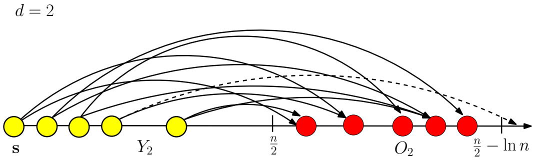
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

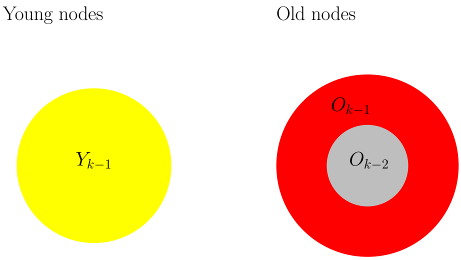
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

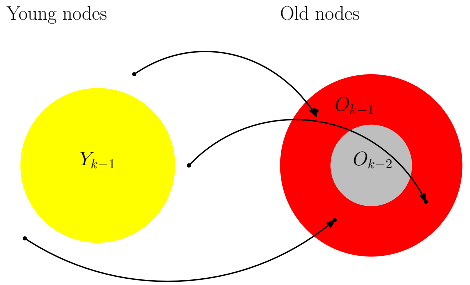
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

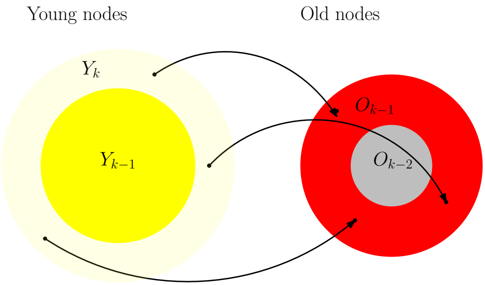
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

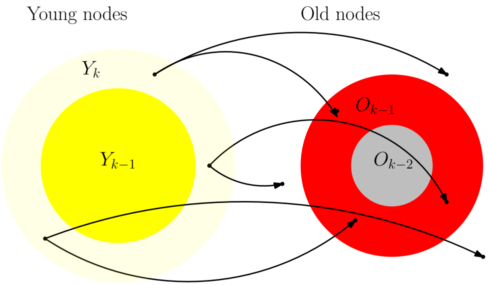
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

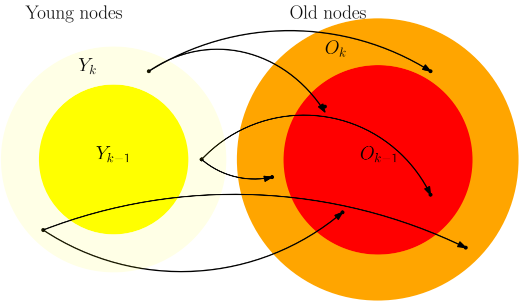
- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Alternating paths argument

We iteratively add edges. At the end of the k-th phase:

- Y_k: subset of young nodes that can be reached from s along an alternating path of length $2k$;
- O_k: subset of old nodes that can be reached from s along an alternating path of length $2k+1$.



Proof: informing $\Omega(n)$ nodes

Theorem 8

Assume s joins the network at step t_0. There is $\tau = \mathcal{O}(\log n/d)$, such that:

$$P\left(|I_{t_0+\tau}| \geq \frac{2n}{d}\right) \geq 1 - e^{\Omega(d)}.$$

Proof.

Assume $|Y_{k-1}| \leq n/d$ and $|O_{k-1}| \leq n/d$ for $k \geq 1$:

Claim 8.1

$$P\left(|Y_k - Y_{k-1}| > \frac{d}{20}y \mid |O_{k-1} - O_{k-2}| \geq y\right) \geq 1 - e^{-yd/100}$$

Proof: informing $\Omega(n)$ nodes

Theorem 8

Assume s joins the network at step t_0. There is $\tau = \mathcal{O}(\log n/d)$, such that:

$$\mathrm{P}\left(|I_{t_0+\tau}| \geq \frac{2n}{d}\right) \geq 1 - e^{\Omega(d)}.$$

Proof.

Assume $|Y_{k-1}| \leq n/d$ and $|O_{k-1}| \leq n/d$ for $k \geq 1$:

Claim 8.2

$$\mathrm{P}\left(|O_k - O_{k-1}| \geq \frac{d}{20}x \mid |Y_k - Y_{k-1}| \geq x\right) \geq 1 - e^{-dx/100}.$$

Proof: informing $\Omega(n)$ nodes

Claim 8.1.

$$P\left(|Y_k - Y_{k-1}| > \frac{d}{20}y \mid |O_{k-1} - O_{k-2}| \geq y\right) \geq 1 - e^{-yd/100}$$

Proof.

Consider $v \in Y - Y_{k-1}$.

- $Z_v = 1$ iff $\exists u \in O_{k-1} - O_{k-2} : v \to u$ using at least of the first $1, \ldots, d/2$ links
- $P\left(Z_v = 1 \mid |O_{k-1} - O_{k-2}| \geq y\right) \geq 1 - \left(1 - \frac{y}{n}\right)^{\frac{d}{2}}$
- The Z_v's are independent
- Apply Chernoff's bound

Proof: informing $\Omega(n)$ nodes

Claim 8.2.

$$\mathrm{P}\left(|O_k - O_{k-1}| \geq \frac{d}{20}x \mid |Y_k - Y_{k-1}| \geq x\right) \geq 1 - e^{-dx/100}\,.$$

Proof.

At a high level:

- $|O_k - O_{k-1}| = \sum_{u \in O - O_{k-1}} A_u$
- $A_u = 1$ iff $\exists v \in Y_k - Y_{k-1} : v \to u$ using one of its last $d/2$ links
- The A_u'a are *negatively correlated*
- Bounded correlation $\to$ use variant of Azuma's inequality

Proof: informing almost every node

Theorem 9

a constant $\tau_2 = \Theta(d)$ exists such that, for $\tau_1 = \mathcal{O}(\log n / \log d)$ we have:

$$\mathrm{P}\left(\left| I_{t_0+\tau_1+\tau_2} \right| \geq (1 - e^{-d/10}) n \right) \geq 1 - e^{-\Omega(d)} .$$

Proof.

Proof uses expansion of large sets in the "no-regeneration" case:

Lemma 10

For every $t \geq n$, the following holds whp:

$$\mathrm{P}\left(\min_{S \subseteq N_t \,:\, n e^{-d/10} \leq |S| \leq n/2} \frac{|\partial_{out}(S)|}{|S|} \geq 0.1 \right) \leq \frac{1}{n^4}.$$

Proof: informing almost every node

Theorem 9

a constant $\tau_2 = \Theta(d)$ *exists such that, for* $\tau_1 = \mathcal{O}(\log n / \log d)$ *we have:*

$$\mathrm{P}\left(|I_{t_0+\tau_1+\tau_2}| \geq (1 - e^{-d/10})n\right) \geq 1 - e^{-\Omega(d)}\,.$$

Proof.

Proof of Lemma 10: Consider $ne^{-d/10} \leq |S| \leq n/2$ and $T = 0.1|S|, S \cap T = \emptyset$.

1. Union bound over all possible pairs S and T;
2. Exponentially many pairs, but $\mathrm{P}\left(\partial_{out}(S) \subseteq T\right)$ for specific pair (S, T) is exponentially small.

Similar to Case 1 of Theorem 7. □

From SDG(R) to PDG(R)

Some subtle changes occur

Main Challenges

- Number of nodes at any moment "stable" but not fixed
- Nodes can come into existence and disappear at any time
- In the Poisson model, we assume it takes one unit to deliver a message across an edge
- What if the edge disappears in the interim?

From SDG(R) to PDG(R)

Some subtle changes occur

Main Challenges

- Number of nodes at any moment "stable" but not fixed
- Nodes can come into existence and disappear at any time
- In the Poisson model, we assume it takes one unit to deliver a message across an edge
- What if the edge disappears in the interim?

Refining the analysis

- Only consider instants t at which a change in V_t occurs
- The resulting process is discrete Markov chain
- Though the intuitions remain the same, proofs sometimes need substantial revisiting

Open Questions and the End

- Expected degree is constant for every t
- Maximum degree is $\mathcal{O}(\log n)$ (could still be non-constant)

Major Open Question:

Design and Analysis of **Natural** Graph Dynamics in the presence of Node Churn that yield **Bounded-Degree Expanders**, w.h.p.

Open Questions and the End

- Expected degree is constant for every t
- Maximum degree is $\mathcal{O}(\log n)$ (could still be non-constant)

Major Open Question:

Design and Analysis of **Natural** Graph Dynamics in the presence of Node Churn that yield **Bounded-Degree Expanders**, w.h.p.

THANKS!!!!!!