The Iterated Local Model for Social Networks

Erin Meger

Mount Allison University
Joint work with: Anthony Bonato, Huda Chuangpishit, Sean English, and Bill Kay

Combinatorics Seminar
UQAM 17 April 2020

Outline

Complex Networks

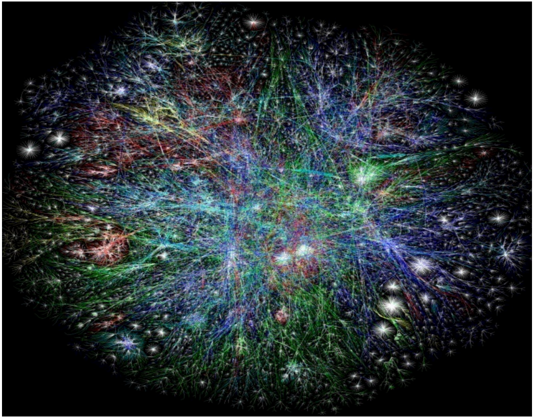


Figure: The Web Graph

Complex Networks

Five main properties:

1. Large-scale
2. Evolving over time
3. Power law degree distribution
4. Small world property
5. Densification

Power Law Degree Distribution

Degree Distribution: $\{N_{k,G} : 0 \leq k \leq n\}$

$$N_{k,G} = |\{x \in V(G) : \deg_G(x) = k|$$

Power Law: for $1 < \beta \in \mathbb{R}$, and interval of $k \in \mathbb{N}$

$$\frac{N_{k,G}}{n} \approx k^{-\beta}.$$

Small World Property

The average distance is

$$L(G) = \frac{\sum_{u,v \in V(G)} d(u,v)}{\binom{|V(G)|}{2}}$$

The clustering coefficient of G is defined as follows:

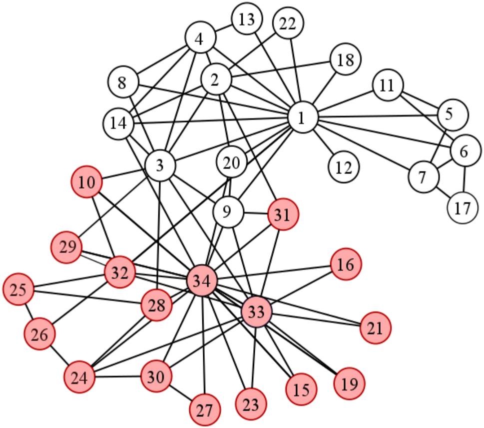
$$C(G) = \frac{1}{|V(G)|} \sum_{x \in V(G)} C_x(G), \quad \text{where} \quad C_x(G) = \frac{\left| E\left(G[N_G(x)]\right) \right|}{\binom{\deg(x)}{2}}.$$

Densification

A sequence of graphs $\{G_t : t \in \mathbb{N}\}$ densifies over time if

$$\lim_{t \to \infty} \frac{|E(G_t)|}{|V(G_t)|} \to \infty$$

Zachary Karate Club



Preferential Attachment Model

Preferential Attachment Model (2002, Barabasi, Albert)

Input: G_0, $m \in \mathbb{N}$

Iterate:
Add a new node v, with m neighbours.

Preferential Attachment Model

Preferential Attachment Model (2002, Barabasi, Albert)

Input: G_0, $m \in \mathbb{N}$

Iterate:
Add a new node v, with m neighbours.
Each neighbour u_i is chosen independently using *preferential attachment*, that is the edge $u_i v$ is added with probability

$$p_i = \frac{\deg(u_i)}{\sum_{u_j \in V(G)} \deg(u_j)}$$

Preferential Attachment Model

ACL Preferential Attachment Model (2008, Aiello, Chung, Lu)

Input: $p \in [0, 1]$

$$p_i = \frac{\deg(u_i)}{\sum_{u_j \in V(G)} \deg(u_j)}$$

Vertex Step: add a new node v, and add the edge $u_i v$ with probability p_i as above

Preferential Attachment Model

ACL Preferential Attachment Model (2008, Aiello, Chung, Lu)

Input: $p \in [0, 1]$

$$p_i = \frac{\deg(u_i)}{\sum_{u_j \in V(G)} \deg(u_j)}$$

Vertex Step: add a new node v, and add the edge $u_i v$ with probability p_i as above

Edge Step: add a new edge rs, where both nodes r, s are chosen by preferential attachment

Preferential Attachment Model

ACL Preferential Attachment Model (2008, Aiello, Chung, Lu)

Input: $p \in [0, 1]$

$$p_i = \frac{\deg(u_i)}{\sum_{u_j \in V(G)} \deg(u_j)}$$

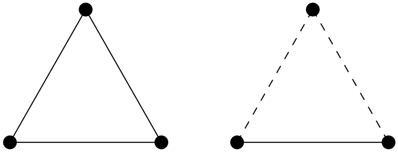
Vertex Step: add a new node v, and add the edge $u_i v$ with probability p_i as above

Edge Step: add a new edge rs, where both nodes r, s are chosen by preferential attachment

With probability p take a vertex step, and with probability $1 - p$ take an edge step.

Structural Balance Theory

Representing adversarial relationships with $(-)$ and friendly relationships with $(+)$, Structural Balance Theory says triads seek a positive product of edge signs, called closure.



ILT

Iterated Local Transitivity Model (ILT) (2009, Bonato, Hadi, Horn, Prałat, Wang)

Input: G_0

To form G_t at time t clone each $x \in V(G_{t-1})$ by adding a new node x' such that

$$N_{G_t}(x') = N_{G_{t-1}}[x]$$

Example of ILT

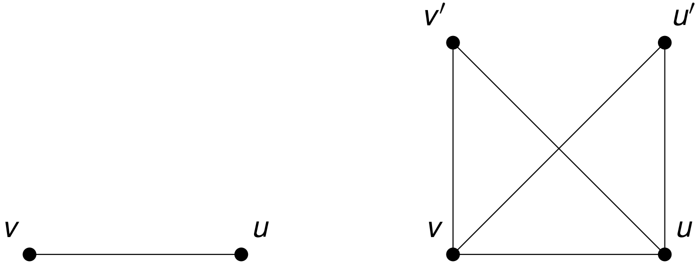


Figure: Example of one time step of ILT with $G_0 = K_2$.

Iterated Local Anti-Transitivity

Iterated Local Anti-Transitivity Model (ILAT) (2017, Bonato, Infeld, Pokhrel, Prałat)

Input: G_0

To form G_t at time t anti-clone each $x \in V(G_{t-1})$ by adding new node x^* such that

$$N_{G_t}(x^*) = V(G_{t-1}) \backslash N_{G_{t-1}}[x]$$

Example of ILAT

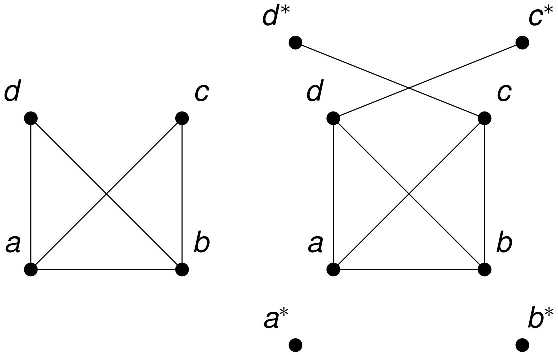


Figure: Example of one time step ILAT.

ILM

Iterated Local Model (ILM) (2019+, Bonato, Chuangpishit, English, Kay, M.)

Input: G_0 and $S = \{b_i\}_{i \in \mathbb{N}}$, where $b_i \in \{0, 1\}$

To form $\mathrm{ILM}_{t,S}(G_0)$ at time t:

- if $b_t = 1$ add a clone x' for each $x \in V(G_{t-1})$ with

$$N_{G_t}(x') = N_{G_{t-1}}[x]$$

- if $b_t = 0$ add an anti-clone x^* for each $x \in V(G_{t-1})$ with

$$N_{G_t}(x^*) = V(G_{t-1}) \setminus N_{G_{t-1}}[x]$$

Example of ILM

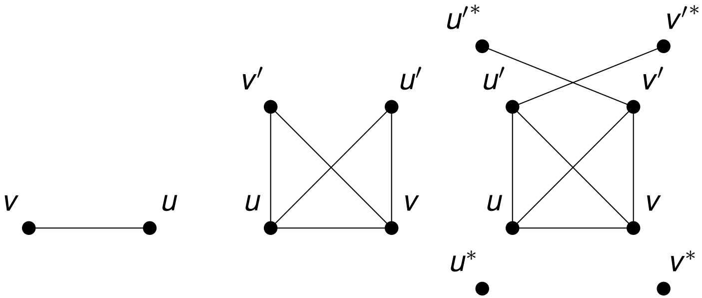


Figure: Example of ILM using $G_0 = K_2$ and $S = \{1, 0, ...\}$

Size, Evolution, and Densification

Theorem (2019+, BCEKM) Given any graph, G_0, and any binary sequence, S, with at least one zero, then at time step t

$$|E(\mathrm{ILM}_{t,S}(G))| = \Theta\left(2^{t+\beta}\left(\frac{3}{2}\right)^{t-\beta}\right) = \Theta\left(2^{\beta}\left(\frac{3}{2}\right)^{t-\beta} n_t\right)$$

Where τ is the first index such that $s_\tau = 0$, and β is the largest index such that $s_\beta = 0$.

Low Diameter

Theorem (2019+, BCEKM) Given $G \neq K_1$ be a graph that is not the disjoint union of two cliques, and a sequence with at least two zeroes, then

$$\mathrm{diam}(\mathrm{ILM}_{t,S}(G)) = 3$$

Low Diameter

Lemma $2 \leq \text{diam}(G) = \text{diam}(\text{LT}(G))$ and
$2 \leq \text{radius}(G) = \text{radius}(\text{LT}(G))$.

Proof For any $u, v \in V(G)$ with $uv \notin E(G)$

$\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u, v)$ and $\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u', v')$

Low Diameter

Lemma $2 \leq \text{diam}(G) = \text{diam}(\text{LT}(G))$ and $2 \leq \text{radius}(G) = \text{radius}(\text{LT}(G))$.

Proof For any $u, v \in V(G)$ with $uv \notin E(G)$

$\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u, v)$ and $\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u', v')$

Similarly $\text{dist}_{LT(G)}(u, v') = \text{dist}_G(u, v)$

Low Diameter

Lemma $2 \leq \text{diam}(G) = \text{diam}(\text{LT}(G))$ and $2 \leq \text{radius}(G) = \text{radius}(\text{LT}(G))$.

Proof For any $u, v \in V(G)$ with $uv \notin E(G)$

$\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u, v)$ and $\text{dist}_G(u, v) = \text{dist}_{LT(G)}(u', v')$

Similarly $\text{dist}_{LT(G)}(u, v') = \text{dist}_G(u, v)$

When $uv \in E(G)$, $\text{dist}_{LT(G)}(u', v') = 2$

Low Diameter

Proof Sketch

- LAT(G) has radius at least 3 since dist$(x, x^*) \geq 3$
- Lemma: If $\gamma(G) \geq 3$ then diam(LAT(G)) ≤ 3

Low Diameter

Proof Sketch

- LAT(G) has radius at least 3 since dist$(x, x^*) \geq 3$
- Lemma: If $\gamma(G) \geq 3$ then diam(LAT(G)) ≤ 3
- We only need to consider $\gamma(G) = 2$
- Find x, y whose closed neighborhoods partition the vertex set

Low Diameter

Proof Sketch

- LAT(G) has radius at least 3 since dist$(x, x^*) \geq 3$
- Lemma: If $\gamma(G) \geq 3$ then diam(LAT(G)) ≤ 3
- We only need to consider $\gamma(G) = 2$
- Find x, y whose closed neighborhoods partition the vertex set
- Pairs of vertices in $N[x]$ or in $N[y]$ are distance 2.

Low Diameter

Proof Sketch

- LAT(G) has radius at least 3 since dist$(x, x^*) \geq 3$
- Lemma: If $\gamma(G) \geq 3$ then diam(LAT(G)) ≤ 3
- We only need to consider $\gamma(G) = 2$
- Find x, y whose closed neighborhoods partition the vertex set
- Pairs of vertices in $N[x]$ or in $N[y]$ are distance 2.
- If $u \in N[x]$ has a neighbour in $N[y]$

Low Diameter

Proof Sketch

- LAT(G) has radius at least 3 since $\text{dist}(x, x^*) \geq 3$
- Lemma: If $\gamma(G) \geq 3$ then $\text{diam}(\text{LAT}(G)) \leq 3$
- We only need to consider $\gamma(G) = 2$
- Find x, y whose closed neighborhoods partition the vertex set
- Pairs of vertices in $N[x]$ or in $N[y]$ are distance 2.
- If $u \in N[x]$ has a neighbour in $N[y]$
- Otherwise, we get this picture using a counting argument and case analysis

Clustering

Theorem (2019+, BCEKM) Given a sequence with bounded gaps between zeroes, and k a constant such that there are no gaps of length k,

$$C(\mathrm{ILM}_{t,S}(G)) \geq (1 + o(1))\frac{1}{2^{2k+4}}.$$

Clustering

Theorem (2019+, BCEKM) Given a sequence with bounded gaps between zeroes, and k a constant such that there are no gaps of length k,

$$C(\mathrm{ILM}_{t,S}(G)) \geq (1 + o(1))\frac{1}{2^{2k+4}}.$$

The clustering coefficient is bounded away from zero.

Clustering

Let β_1 and β_2 be the two most recent 0 steps

Clustering

Let β_1 and β_2 be the two most recent 0 steps

If v is not a clone in G_{β_1}, then $\deg_{\beta_1}(v) = \frac{n_{\beta_1}}{2} - 1$, and then

$$\deg_{\beta_1+1}(v) = 2\deg_{\beta_1}(v) + 1 = \frac{n_{\beta_1+1}}{2} - 1.$$

Clustering

Let β_1 and β_2 be the two most recent 0 steps

If v is not a clone in G_{β_1}, then $\deg_{\beta_1}(v) = \frac{n_{\beta_1}}{2} - 1$, and then

$$\deg_{\beta_1+1}(v) = 2\deg_{\beta_1}(v) + 1 = \frac{n_{\beta_1+1}}{2} - 1.$$

Since all entries between β_1 an β_2 are 1, we have, inductively that

$$\deg_{\beta_2-1}(v) = \frac{n_{\beta_2-1}}{2} - 1.$$

Clustering

Define two sets:

$$X_{\beta_2} = N_{\beta_2}(v) \cap (V(G_{\beta_2}) \setminus V(G_{\beta_2 - 1}))$$

$$Y_{\beta_2} = (V(G_{\beta_2}) \setminus V(G_{\beta_2 - 1})) \setminus N_{\beta_2}(v)$$

.
With $|X_{\beta_2}| = |Y_{\beta_2}| = n_{\beta_2}/4$.

Clustering

Define two sets:

$$X_{\beta_2} = N_{\beta_2}(v) \cap (V(G_{\beta_2}) \setminus V(G_{\beta_2 - 1}))$$

$$Y_{\beta_2} = (V(G_{\beta_2}) \setminus V(G_{\beta_2 - 1})) \setminus N_{\beta_2}(v)$$

.
With $|X_{\beta_2}| = |Y_{\beta_2}| = n_{\beta_2}/4$.

For transitive steps, create X_{β_2+1} and Y_{β_2+1} by adding the set of all respective clones to each set

Note $X_{\beta_2+1} \subseteq N_{\beta_2+1}(v)$ and $Y_{\beta_2+1} \cap N_{\beta_2+1}(v) = \emptyset$

Clustering

After an anti-transitive step, each vertex in X_{t-1} will be adjacent to each clone of vertices in Y_{t-1}.

Clustering

After an anti-transitive step, each vertex in X_{t-1} will be adjacent to each clone of vertices in Y_{t-1}.

Thus, $N_t(v)$ contains at least $|X_{t-1}| \cdot |Y_{t-1}| = n_t^2/64$ edges.

Since $\deg_t(v) = \frac{n_t}{2} - 1$, we have that

$$c_t(v) \geq \frac{n_t^2/64}{\binom{n_t/2-1}{2}} = (1+o(1))\frac{1}{8}.$$

Clustering

Note this is holds for all vertices $v \in V(G_{\beta_1 - 1})$.

There are $n_{\beta_1 - 1} \geq n_{t-2k-1} = \frac{n_t}{2^{2k+1}}$ such vertices,

$$C(G_t) \geq \frac{(1 + o(1))\frac{1}{8} \cdot \frac{n_t}{2^{2k+1}}}{n_t} = (1 + o(1))\frac{1}{2^{2k+4}},$$

□

Spectral Gap

A the adjacency matrix and D the diagonal degree matrix of a graph G. The *normalized Laplacian* of G is

$$\mathcal{L} = I - D^{-1/2} A D^{-1/2}.$$

With eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$

Spectral Gap

A the adjacency matrix and D the diagonal degree matrix of a graph G. The *normalized Laplacian* of G is

$$\mathcal{L} = I - D^{-1/2} A D^{-1/2}.$$

With eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$

The *spectral gap* of the normalized Laplacian is defined as

$$\lambda = \max\{|\lambda_1 - 1|, |\lambda_{n-1} - 1|\}.$$

Spectral Gap

Expander Mixing Lemma

If G is a graph with spectral gap λ, then, for all sets $X \subseteq V(G)$,

$$\left| e(X,X) - \frac{(\mathrm{vol}(X))^2}{\mathrm{vol}(G)} \right| \leq \lambda \frac{\mathrm{vol}(X)\mathrm{vol}(\overline{X})}{\mathrm{vol}(G)}.$$

Where $\mathrm{vol(X)} = \sum_{v \in X} \deg v$ for $X \subseteq V(G)$.

Spectral Gap

After a transitive step, let $X = V(G_t) \setminus V(G_{t-1})$

Since X is an independent set, $e(X, X) = 0$. We derive that

$$\mathrm{vol}(G_t) = 6e_{t-1} + 2n_{t-1}$$

$$\mathrm{vol}(X) = 2e_{t-1} + n_{t-1}$$

$$\mathrm{vol}(\overline{X}) = \mathrm{vol}(G_t) - \mathrm{vol}(X) = 4e_{t-1} + n_{t-1}$$

Spectral Gap

After a transitive step, let $X = V(G_t) \setminus V(G_{t-1})$

Since X is an independent set, $e(X, X) = 0$. We derive that

$$\mathrm{vol}(G_t) = 6e_{t-1} + 2n_{t-1}$$

$$\mathrm{vol}(X) = 2e_{t-1} + n_{t-1}$$

$$\mathrm{vol}(\overline{X}) = \mathrm{vol}(G_t) - \mathrm{vol}(X) = 4e_{t-1} + n_{t-1}$$

$$\begin{aligned} \lambda_t &\geq \frac{(\mathrm{vol}(X))^2}{\mathrm{vol}(G_t)} \cdot \frac{\mathrm{vol}(G_t)}{\mathrm{vol}(X)\mathrm{vol}(\overline{X})} \\ &= \frac{\mathrm{vol}(X)}{\mathrm{vol}(\overline{X})} \\ &= \frac{2e_{t-1} + n_{t-1}}{4e_{t-1} + n_{t-1}} > 1/2 \end{aligned}$$

Spectral Gap

After an antitransitive step, Let $X = V(G_t) \backslash V(G_{t-1})$

$$\mathrm{vol}(G_t) = 2n_{t-1}^2 - 2e_{t-1} - 2n_{t-1}$$

$$\mathrm{vol}(X) = n_{t-1}^2 - 2e_{t-1} - n_{t-1}$$

$$\mathrm{vol}(\overline{X}) = \mathrm{vol}(G_t) - \mathrm{vol}(X) = n_{t-1}^2 - n_{t-1}.$$

$$\begin{aligned}\lambda_t &\geq \frac{\mathrm{vol}(X)}{\mathrm{vol}(\overline{X})} \\ &= \frac{n_{t-1}^2 - 2e_{t-1} - n_{t-1}}{n_{t-1}^2 - n_{t-1}} \\ &= 1 - \frac{2e_{t-1}}{n_{t-1}^2 - n_{t-1}} = \frac{1}{4} - o(1)\end{aligned}$$

Spectral Gap

Therefore, we have that

$$2e_{t-1} \leq 2\binom{n_{t-1}}{2} - 2\binom{n_{t-1}/2}{2} = \frac{3}{4}n_{t-1}^2 - \frac{1}{2}n_{t-1},$$

□

Induced Subgraphs

Theorem (2019+, BCEKM) If F is a graph, then there exists some constant $t_0 = t_0(F)$ such that for all $t \geq t_0$, all graphs G, and all binary sequences S, F is an induced subgraph of $\mathrm{ILM}_{t,S}(G)$.

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$
- Iteratively delete edges of the clique to form F

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$
- Iteratively delete edges of the clique to form F
- For each $uv \notin E(F)$ replace with $u'v'$

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$
- Iteratively delete edges of the clique to form F
- For each $uv \notin E(F)$ replace with $u'v'$

Show $ILM_{2r,S}(G)$ contains $ILT_r(K_1)$, by induction.

- One transitive step will increase the r for the induced copy of ILT

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$
- Iteratively delete edges of the clique to form F
- For each $uv \notin E(F)$ replace with $u'v'$

Show $ILM_{2r,s}(G)$ contains $ILT_r(K_1)$, by induction.

- One transitive step will increase the r for the induced copy of ILT
- Two anti-transitive steps will similarily increase the induced copy of ILT

Finite Subgraphs

Proof (sketch):

Show $ILT_k(K_1)$ contains induced copy of F

- Find a t with a clique of size $|V(F)|$
- Iteratively delete edges of the clique to form F
- For each $uv \notin E(F)$ replace with $u'v'$

Show $ILM_{2r,s}(G)$ contains $ILT_r(K_1)$, by induction.

- One transitive step will increase the r for the induced copy of ILT
- Two anti-transitive steps will similarly increase the induced copy of ILT
- In any binary sequence of length $2r$ there exist r 0s or r 1s.

Hamiltonicity

Theorem For $G \neq K_1$ and S a binary sequence with at least two non-consecutive zeros, then $\mathrm{ILM}_{t,S}(G) = G_t$ is Hamiltonian.

Hamiltonicity

Theorem For $G \neq K_1$ and S a binary sequence with at least two non-consecutive zeros, then $\text{ILM}_{t,S}(G) = G_t$ is Hamiltonian.

Definition Let $\zeta(G)$ be the first value such that $ILT_{\zeta(G)}$ is Hamiltonian, and let ζ_n be the maximum over all graphs of order n.

Hamiltonicity

Theorem For $G \neq K_1$ and S a binary sequence with at least two non-consecutive zeros, then $\mathrm{ILM}_{t,S}(G) = G_t$ is Hamiltonian.

Definition Let $\zeta(G)$ be the first value such that $ILT_{\zeta(G)}$ is Hamiltonian, and let ζ_n be the maximum over all graphs of order n.

Theorem For all $n \geq 3$

$$\log_2(n-1) \leq \zeta_n \leq \lceil \log_2(n-1) \rceil + 1$$

Hamiltonicity

Proof (sketch):

- Using $\Delta(G_t) = \frac{n_t}{2} - 1$, the compliment of G_t is Hamiltonian by Dirac's theorem

Hamiltonicity

Proof (sketch):

- Using $\Delta(G_t) = \frac{n_t}{2} - 1$, the compliment of G_t is Hamiltonian by Dirac's theorem
- Four clone vertices form a clique in the HC of $\overline{G_t}$

Hamiltonicity

Proof (sketch):

- Using $\Delta(G_t) = \frac{n_t}{2} - 1$, the compliment of G_t is Hamiltonian by Dirac's theorem
- Four clone vertices form a clique in the HC of $\overline{G_t}$
- Find two cycles that partition the vertex set and perform an edge switch

Other Structural Properties

For the model with certain restrictions on the input sequence and graph:

- $\chi(G) + t - 1 \leq \chi\left(\text{ILM}_{t,S}(G)\right) \leq \chi(G) + t$
- $\gamma(\text{ILM}_{t,S}(G)) \leq 3$

Future Directions

- Graph Limits
- Domination number in remaining cases
- Randomization of the model
- Improve Clustering, still unknown for ILAT

Thank You