Internet Mathematics Vol. 6, No. 3: 285–313

Models of Online Social Networks

Anthony Bonato, Noor Hadi, Paul Horn, Paweł Prałat, and Changping Wang

Abstract. We present a deterministic model for online social networks (OSNs) based on transitivity and local knowledge in social interactions. In the iterated local transitivity (ILT) model, at each time step and for every existing node x, a new node appears that joins to the closed neighbor set of x. The ILT model provably satisfies a number of both local and global properties that have been observed in OSNs and other real-world complex networks, such as a densification power law, decreasing average distance, and higher clustering than in random graphs with the same average degree. Experimental studies of social networks demonstrate poor expansion properties as a consequence of the existence of communities with low numbers of intercommunity edges. Bounds on the spectral gap for both the adjacency and normalized Laplacian matrices are proved for graphs arising from the ILT model indicating such bad expansion properties. The cop and domination numbers are shown to remain the same as those of the graph from the initial time step G_0, and the automorphism group of G_0 is a subgroup of the automorphism group of graphs generated at all later time steps. A randomized version of the ILT model is presented that exhibits a tunable densification power-law exponent and maintains several properties of the deterministic model.

1. Introduction

Online social networks (OSNs) such as Facebook, MySpace, Twitter, and Flickr have become increasingly popular in recent years. In OSNs, nodes represent people online, and edges correspond to a friendship relation between them. In these complex real-world networks with sometimes millions of nodes and edges, new

1542-7951/09 $0.50 per page

nodes and edges dynamically appear over time. Parallel with their popularity among the general public is an increasing interest among the mathematical and general scientific community in the properties of online social networks, both in gathering data and statistics about the networks and in finding models simulating their evolution. Data about social interactions in online networks are more readily accessible and measurable than in offline social networks, which suggests a need for rigorous models capturing their evolutionary properties.

The small-world property of social networks, introduced in [Watts and Strogatz 76], is a central notion in the study of complex networks, with roots in [Milgram 67] on short paths of friends connecting strangers. The small-world property posits low average distance (or diameter) and high clustering, and has been observed in a wide variety of complex networks.

An increasing number of studies have focused on the small-world and other complex network properties in OSNs. An early study of an online social network at Stanford University is provided in [Adamic et al. 03], and the authors found that the network has the small-world property. Correlation between friendship and geographic location was found in [Liben-Nowell et al. 05] using data from LiveJournal. The evolution of the online networks Flickr and Yahoo!360 were studied in [Kumar et al. 06]. The authors found (among other things) that the average distance between users actually decreases over time, and that these networks exhibit power-law degree distributions. In [Golder et al. 07], the Facebook network was analyzed by studying the messaging patterns between friends with a sample of 4.2 million users. The authors also found a power-law degree distribution and the small-world property. Similar results were found in [Ahn et al. 07], which studied Cyworld, MySpace, and Orkut, and in [Mislove et al. 07], which examined data collected from four online social networks: Flickr, YouTube, LiveJournal, and Orkut. Power laws for both the in- and out-degree distributions, low diameter, and high clustering coefficient were reported in the Twitter friendship graph in [Java et al. 07]. In [Krishnamurthy et al. 08], geographic growth patterns and distinct classes of users were investigated in Twitter. For further background on complex networks and their models, see the books [Bonato 08, Caldarelli 07, Chung and Lu 06, Durrett 06].

Recent work [Leskovec et al. 05a] underscores the importance of two additional properties of complex networks above and beyond more traditionally studied phenomena such as the small-world property. A graph G with e_t edges and n_t nodes satisfies a *densification power law* if there is a constant $a \in (1, 2)$ such that e_t is proportional to n_t^a. In particular, the average degree grows to infinity with the order of the network (in contrast, for instance, to the preferential attachment model, which generates graphs with constant average degree). In [Leskovec et al. 05a], densification power laws were reported in several real-

world networks such as a physics citation graph and the Internet graph at the level of autonomous systems. Another striking property found in such networks (and also in online social networks; see [Kumar et al. 06]) is that distances in the networks (measured by either diameter or average distance) decrease with time. The usual models such as preferential attachment and copying models have logarithmically or sublogarithmically growing diameters and average distances with time. Various models (such as the forest fire [Leskovec et al. 05a] and Kronecker multiplication [Leskovec et al. 05b] models) have been proposed to simulate power-law degree distribution, densification power laws, and decreasing distances.

We present a new model, called *iterated local transitivity* (ILT), for OSNs and other complex networks that dynamically simulates many of their properties. The present article is the full version of the proceedings paper [Bonato et al. 09]. Although modeling has been done extensively for other complex networks such as the Web graph (see [Bonato 08]), models of OSNs have only recently been introduced (such as those in [Crandall et al. 08, Kumar et al. 06, Liben-Nowell et al. 05]). The central idea behind the ILT model is what sociologists call *transitivity*: if u is a friend of v, and v is a friend of w, then u is a friend of w (see, for example, [Frank 80, Scott 00, White et al. 76]). In its simplest form, transitivity gives rise to the notion of *cloning*, whereby u is joined to all of the neighbors of v. In the ILT model, given some initial graph as a starting point, nodes are repeatedly added over time that clone *each* node, so that the new nodes form an independent set. The ILT model not only incorporates transitivity, but uses only local knowledge in its evolution, in that a new node joins only to neighbors of an existing node. Local knowledge is an important feature of social and complex networks, in which nodes have only limited influence on the network topology. We stress that our approach is mathematical rather than empirical; indeed, the ILT model (apart from its potential use by computer and social scientists as a simplified model for OSNs) should be of theoretical interest in its own right.

Variants of cloning were considered earlier in duplication models for protein–protein interactions [Bebek et al. 06, Bhan et al. 02, Chung et al. 03, Pastor-Satorras et al. 03], and in copying models for the Web graph [Bonato and Janssen 09, Kumar et al. 00]. There are several differences between the duplication and copying models and the ILT model. For one, duplication models are difficult to analyze due to their rich dependence structure. While the ILT model displays a dependency structure, determinism makes it more amenable to analysis. The ILT model may be viewed as a simplified snapshot of the duplication model, whereby *all* nodes are cloned in a given time step, rather than nodes being duplicated one by one over time. Cloning all nodes at each time

step as in the ILT model leads to densification and high clustering, along with bad expansion properties (as we describe in Section 1.2).

We finish the introduction with some asymptotic notation. Let f and g be functions whose domain is some fixed subset of $\mathbb{R}$. We write $f \in O(g)$ if

$$\limsup_{t\to\infty} \frac{f(t)}{g(t)}$$

exists and is finite. We will abuse notation and write $f = O(g)$. We write $f = \Omega(g)$ if $g = O(f)$, and $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$. If $\lim_{t\to\infty} |\frac{f(t)}{g(t)}| = 0$, then $f = o(g)$ (or $g = \omega(f)$). So if $f = o(1)$, then f tends to 0.

1.1. The ILT Model

We now give a precise formulation of the model. The ILT model generates finite, simple, undirected graphs $(G_t : t \geq 0)$. *Time step* t, for $t \geq 1$, is defined to be the transition between G_{t-1} and G_t. (Note that a directed graph model will be considered in the sequel. See also Section 3.) The only parameter of the model is the initial graph G_0, which is any fixed finite *connected* graph. Assume that for a fixed $t \geq 0$, the graph G_t has been constructed. To form G_{t+1}, for each node $x \in V(G_t)$, add its *clone* x', such that x' is joined to x and all of its neighbors at time t. Note that the set of new nodes at time $t+1$ forms an independent set of cardinality $|V(G_t)|$. See Figure 1 for the graphs generated from the 4-cycle over the time steps $t = 1$, 2, 3, and 4.

We write $\deg_t(x)$ for the degree of a node at time t, n_t for the order of G_t, and e_t for the number of its edges. It is straightforward to see that $n_t = 2^t n_0$. Given a node x at time t, let x' be its clone. The elementary but important recurrences governing the degrees of nodes are given as

$$\deg_{t+1}(x) = 2\deg_t(x) + 1, \tag{1.1}$$

$$\deg_{t+1}(x') = \deg_t(x) + 1. \tag{1.2}$$

1.2. Main Results

We state our main results on the ILT model, with proofs deferred to the next section. We give rigorous proofs that the ILT model generates graphs satisfying a densification power law and in many cases decreasing average distance (properties shared by the forest fire [Leskovec et al. 05a] and Kronecker multiplication [Leskovec et al. 05b] models). A randomized version of the ILT model is introduced with tunable densification power-law exponent. Properties of the ILT model not shown in the models of [Leskovec et al. 05a, Leskovec et al. 05b] exhibit higher clustering than in random graphs with the same average degree,

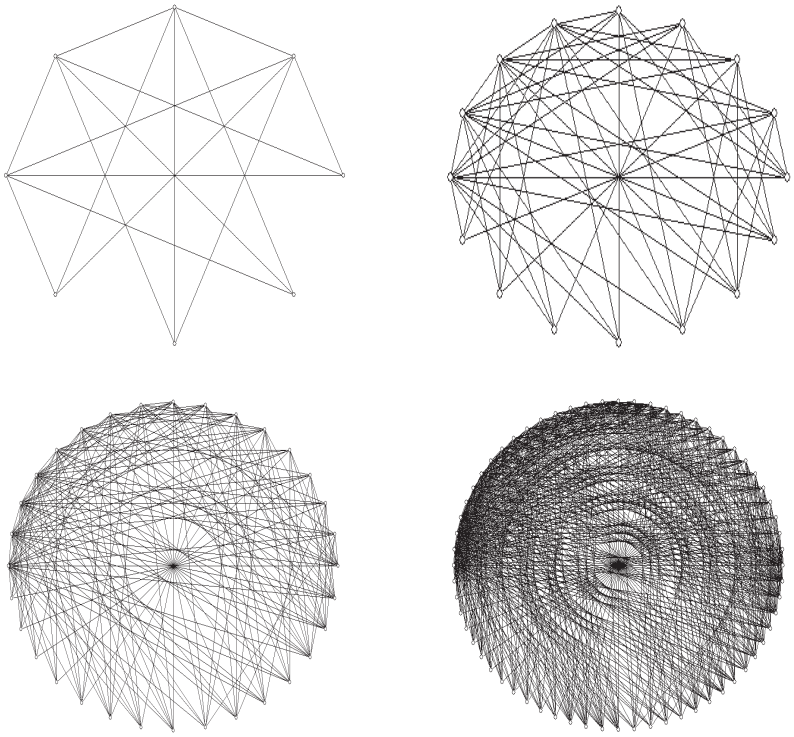

Figure 1. The evolution of the ILT model with $G_0 = C_4$, for $t = 1, 2, 3, 4$ (left to right, top to bottom).

and smaller spectral gaps for both their normalized Laplacian and adjacency matrices than in random graphs. Further, the cop and domination numbers are shown to remain the same as those of the initial graph G_0, and the automorphism group of G_0 is a subgroup of the automorphism group of graphs generated at all later times. The ILT model (unlike the models of [Leskovec et al. 05a, Leskovec et al. 05b]) does not, however, generate graphs with a power-law degree distribution. The number of nodes in the ILT model grows exponentially with time (as in the Kronecker multiplication model, but unlike the forest fire model).

We first demonstrate that the model exhibits a densification power law. Define the *volume* of G_t by

$$\mathrm{vol}(G_t) = \sum_{x \in V(G_t)} \deg_t(x) = 2e_t.$$

Theorem 1.1. *For $t > 0$, the average degree of G_t equals*

$$\left(\frac{3}{2}\right)^t \left(\frac{\mathrm{vol}(G_0)}{n_0} + 2\right) - 2.$$

Note that Theorem 1.1 supplies a densification power law with exponent $a = \frac{\log 3}{\log 2} \approx 1.58$. We think that the densification power law makes the ILT model realistic, especially in light of real-world data mined from complex networks (see [Leskovec et al. 05a]).

We study the average distances and clustering coefficient of the model as time tends to infinity. Define the *Wiener index* of G_t as

$$W(G_t) = \frac{1}{2} \sum_{x,y \in V(G_t)} d(x, y).$$

The Wiener index may be used to define the *average distance* of G_t as

$$L(G_t) = \frac{W(G_t)}{\binom{n_t}{2}}.$$

We will compute the average distance by deriving first the Wiener index. Define the *ultimate average distance of* G_0 as

$$\mathrm{UL}(G_0) = \lim_{t \to \infty} L(G_t),$$

assuming that the limit exists. Note that the ultimate average distance is a new graph parameter. We provide an exact value for $L(G_t)$ and compute the ultimate average distance for any initial graph G_0.

Theorem 1.2.

(1) For $t > 0$,

$$W(G_t) = 4^t \left(W(G_0) + (e_0 + n_0) \left(1 - \left(\frac{3}{4} \right)^t \right) \right).$$

(2) For $t > 0$,

$$L(G_t) = \frac{4^t \left(W(G_0) + (e_0 + n_0) \left(1 - \left(\frac{3}{4} \right)^t \right) \right)}{4^t n_0^2 - 2^t n_0}.$$

(3) For all graphs G_0,

$$\mathrm{UL}(G_0) = \frac{W(G_0) + e_0 + n_0}{n_0^2}.$$

Further, $\mathrm{UL}(G_0) \leq L(G_0)$ *if and only if* $W(G_0) \geq (n_0 - 1)(e_0 + n_0)$.

Note that the average distance of G_t is bounded above by $\text{diam}(G_0) + 1$ (in fact, by $\text{diam}(G_0)$ in all cases except cliques). Further, the condition in (3) for $\text{UL}(G_0) < L(G_0)$ holds for large cycles and paths. Hence, for many initial graphs G_0, the average distance decreases, a property observed in OSNs and other complex networks (see [Kumar et al. 06, Leskovec et al. 05a]).

Let $N_t(x)$ be the neighbor set of x at time t, let $G_t \restriction N_t(x)$ be the subgraph induced by $N_t(x)$ in G_t, and let $e(x,t)$ be the number of edges in $G_t \restriction N_t(x)$. For a node $x \in V(G_t)$ with degree at least 2 define

$$c_t(x) = \frac{e(x,t)}{\binom{\deg_t(x)}{2}}.$$

By convention, $c_t(x) = 0$ if the degree of x is at most 1. The *clustering coefficient* of G_t is

$$C(G_t) = \frac{\sum_{x \in V(G_t)} c_t(x)}{n_t}.$$

The clustering coefficient of the graph at time t generated by the ILT model is estimated and shown to tend to 0 more slowly than a $G(n,p)$ random graph with the same average degree.

Theorem 1.3.

$$\Omega\left(\left(\frac{7}{8}\right)^t t^{-2}\right) = C(G_t) = O\left(\left(\frac{7}{8}\right)^t t^2\right).$$

Observe that $C(G_t)$ tends to 0 as $t \to \infty$. If we let $n_t = n$ (so $t \sim \log_2 n$), then this gives that

$$C(G_t) = n^{\log_2(7/8)+o(1)}.$$

In contrast, for a random graph $G(n,p)$ with comparable average degree

$$pn = \Theta((3/2)^{\log_2 n}) = \Theta\left(n^{\log_2(3/2)}\right)$$

as G_t, the clustering coefficient is $p = \Theta(n^{\log_2(3/4)})$, which tends to zero much faster than $C(G_t)$. (For a discussion of the clustering coefficient of $G(n,p)$, see [Bonato 08, Chapter 2].)

Social networks often organize into separate clusters in which the number of intracluster links is significantly greater than that of intercluster links. In particular, social networks contain communities (characteristic of social organization), where tightly knit groups correspond to the clusters [Girvan and Newman 02]. As a result, social networks possess bad expansion properties realized by small gaps between their first and second eigenvalues [Estrada 06]. We find that the

ILT model has bad expansion properties as indicated by the spectral gap of both its normalized Laplacian and adjacency matrices.

For regular graphs, the eigenvalues of the adjacency matrix are related to several important graph properties, such as in the expander mixing lemma. The normalized Laplacian of a graph, introduced in [Chung 97], relates to important graph properties even in the case that the underlying graph is not regular (as is the case in the ILT model). Let A denote the adjacency matrix and D the diagonal adjacency matrix of a graph G. Then the normalized Laplacian of G is

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

Let $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$ denote the eigenvalues of $\mathcal{L}$. The *spectral gap* of the normalized Laplacian is

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

It is observed in [Chung et al. 04] that for random power-law graphs with some parameters (effectively in the case that $d_{\min} = c\log^2 n$ for some constant $c > 0$ and all integers $n > 0$), $\lambda \leq (1 + o(1))\frac{4}{\sqrt{d}}$, where d is the average degree.

For the graphs G_t generated by the ILT model, we observe that the spectra behave quite differently, and in fact, the spectral gap has a constant order. The following theorem suggests a significant spectral difference between graphs generated by the ILT model and random graphs. Define $\lambda(G_t)$ to be the spectral gap of the normalized Laplacian of G_t.

Theorem 1.4. *For $t \geq 1$, $\lambda(G_t) > \frac{1}{2}$.*

Theorem 1.4 represents a drastic departure from the good expansion found in random graphs, where $\lambda = o(1)$ [Chung 97, Chung et al. 04, Furedi and Komlos 81], and from the preferential attachment model [Gkantsidis et al. 03]. If G_0 has bad expansion properties, and has $\lambda_1 < 1/2$ (and thus $\lambda > 1/2$), then in fact, this trend of bad expansion continues, as shown by the following theorem.

Theorem 1.5. *Suppose G_0 has at least two nodes, and for $t > 0$ let $\lambda_1(t)$ be the second eigenvalue of G_t. Then we have that*

$$\lambda_1(t) < \lambda_1(0).$$

Note that Theorem 1.5 implies that $\lambda_1(1) < \lambda_1(0)$, and this implies that the sequence $\{\lambda_1(t) : t \geq 0\}$ is strictly decreasing. This follows because G_t is

constructed from G_{t-1} in the same manner as G_1 is constructed from G_0. If G_0 is K_1, then there is no second eigenvalue, but G_1 is K_2. Hence in this case, the theorem implies that $\{\lambda_1(t) : t \geq 1\}$ is strictly decreasing.

Let $\rho_0(t) \geq |\rho_1(t)| \geq \cdots$ denote the eigenvalues of the adjacency matrix G_t. If A is the adjacency matrix of G_t, then the adjacency matrix of G_{t+1} is

$$M = \begin{pmatrix} A & A+I \\ A+I & 0 \end{pmatrix},$$

where I is the identity matrix of order n_t. We note the following recurrence for the eigenvalues of the adjacency matrix of G_t.

Theorem 1.6. *If ρ is an eigenvalue of the adjacency matrix of G_t, then*

$$\frac{\rho \pm \sqrt{\rho^2 + 4(\rho+1)^2}}{2}$$

are eigenvalues of the adjacency matrix of G_{t+1}.

We leave the reader to check that the eigenvectors of G_t can be written in terms of the eigenvectors of G_{t-1}. As in the Laplacian case, we show that there is a small spectral gap of the adjacency matrix.

Theorem 1.7. *Let $\rho_0(t) \geq |\rho_1(t)| \geq \cdots \geq |\rho_n(t)|$ denote the eigenvalues of the adjacency matrix of G_t. Then*

$$\frac{\rho_0(t)}{|\rho_1(t)|} = \Theta(1).$$

That is, $\rho_1(t) \geq c|\rho_0(t)|$ for some constant $c > 0$. Theorem 1.7 is in contrast to the fact that in $G(n,p)$ random graphs, $|\rho_1| = o(\rho_0)$ (see [Chung 97]).

In a graph G, a set S of nodes is a *dominating* set if every node not in S has a neighbor in S. The *domination number* of G, written $\gamma(G)$, is the minimum cardinality of a dominating set in G. We use S to represent a dominating set in G where each node not in S is joined to some node of S. A graph parameter bounded below by the domination number is the so-called cop (or search) number of a graph. In the game cops and robbers, there are two players, a set of s *cops* (or *searchers*) $\mathcal{C}$, where $s > 0$ is a fixed integer, and the *robber* $\mathcal{R}$. The cops begin the game by occupying a set of s nodes of a simple, undirected, and finite graph G. While the game may be played on a disconnected graph, without loss of generality, assume that G is connected (since the game is played independently on each component and the number of cops required is the sum over all components).

The cops and robber move in *rounds* indexed by nonnegative integers. Each round consists of a cop's move followed by a robber's move. More than one cop is allowed to occupy a node, and the players may *pass*, that is, remain on their current nodes. A *move* in a given round for a cop or the robber consists of a pass or moving to an adjacent node; each cop may move or pass in a round. The players know each other's current locations; that is, the game is played with *perfect information*. The cops win and the game ends if at least one of the cops can eventually occupy the same node as the robber; otherwise, $\mathcal{R}$ wins. Since placing a cop on each node guarantees that the cops win, we may define the *cop number*, written $c(G)$, as the minimum cardinality of the set of cops needed to win on G. While this node pursuit game played with one cop was introduced in [Nowakowski and Winkler 83, Quilliot 78], the cop number was first introduced in [Aigner and Fromme 84]. For a survey of results on cops and robbers, see [Hahn 07].

We prove that the domination and cop numbers of G_t depend only on the initial graph G_0. Theorem 1.8 shows that even as the graph becomes large as t progresses, the same number of nodes as that needed at time 0 to dominate the graph will be needed at time t.

Theorem 1.8. *For all $t \geq 0$, $\gamma(G_t) = \gamma(G_0)$ and $c(G_t) = c(G_0)$.*

In Theorem 1.8, we prove that the cop number remains the same for G_t. This implies that no matter how large the graph G_t becomes, the robber can be captured by the same number of cops used at time 0. In terms of OSNs, Theorem 1.8 suggests that users in the network can easily spread and track information (such as gossip) no matter how large the graph becomes.

An *automorphism* of a graph G is an isomorphism from G to itself; the set of all automorphisms forms a group under the operation of composition, written $\mathrm{Aut}(G)$. We say that an automorphism $f_t \in \mathrm{Aut}(G_t)$ *extends* to $f_{t+1} \in \mathrm{Aut}(G_{t+1})$ if

$$f_{t+1} \restriction V(G_t) = f_t;$$

that is, the restriction of the map f_{t+1} to $V(G_t)$ equals f_t. We show that symmetries from $t = 0$ are preserved at time t. This provides further evidence that the ILT model retains a memory of the initial graph from time 0.

Theorem 1.9. *For all $t \geq 0$, $\mathrm{Aut}(G_0)$ embeds in $\mathrm{Aut}(G_t)$.*

As shown in Theorem 1.1, the ILT model has a fixed densification exponent equal to $\log 3/\log 2$. We consider a randomized version of the model that allows

for this exponent to become tunable. To motivate the model, in OSNs some new users are friends outside of the OSN. Such users immediately seek each other out as they join the OSN and become friends there. The stochastic model ILT(p) is defined as follows. Define H_0 to be K_1. A sequence $(H_t : t \in \mathbb{N})$ of graphs is generated such that for all t, H_t is an induced subgraph of H_{t+1}. At time $t+1$, first clone all the nodes of H_t as in the deterministic ILT model. Let n be the number of new nodes that are added at time $t+1$. (Note that n is a function of t and is not a new parameter.) To form H_{t+1}, add edges independently between the new nodes with probability $p = p(n)$. Hence, the new nodes form a random graph $G(n, p)$.

Several properties of the ILT model are inherited by the ILT(p) model. For example, as we are adding edges to the graphs generated by the ILT model, the average distance may only decrease, and the clustering coefficient may only increase. The following theorem proves that ILT(p) generates graphs following a densification power law with exponent $\log(3+\delta)/\log 2$, where $0 \leq \delta \leq 1$. For T a positive integer representing time, we say that an event holds *asymptotically almost surely* (*a.a.s.*) if the probability that it holds tends to 1 as T tends to infinity.

Theorem 1.10. *Let* $0 \leq \delta \leq 1$*, and define*

$$p(n) = \frac{\delta n^{\frac{\log(3+\delta)}{\log 2}}}{n^2}. \tag{1.3}$$

Then a.a.s.,

$$\mathrm{vol}(H_T) = (1 + o(1))(3+\delta)^T.$$

Hence, by choosing an appropriate p, the densification power-law exponent in graphs generated by the ILT(p) model may achieve any value in the interval $[\log 3/\log 2, 2]$. We also prove that for the normalized Laplacian, the ILT(p) model maintains a large spectral gap.

Theorem 1.11. *Asymptotically almost surely,*

$$\lambda(H_T) = \Omega(1).$$

2. Proofs of Results

This section is devoted to the proofs of the theorems outlined in Section 1.

2.1. Proof of Theorem 1.1

We now consider the number of edges and average degree of G_t, and prove the following densification power law for the ILT model. Define the *volume* of G_t by

$$\mathrm{vol}(G_t) = \sum_{x \in V(G_t)} \deg_t(x) = 2e_t.$$

The proof of Theorem 1.1 follows directly from the following lemma, since the average degree of G_t is $\mathrm{vol}(G_t)/n_t$.

Lemma 2.1. *For $t > 0$,*

$$\mathrm{vol}(G_t) = 3^t \mathrm{vol}(G_0) + 2n_0(3^t - 2^t).$$

In particular,

$$e_t = 3^t(e_0 + n_0) - n_t.$$

Proof. By (1.1) and (1.2) we have that

$$\begin{aligned}
\mathrm{vol}(G_{t+1}) &= \sum_{x \in V(G_t)} \deg_{t+1}(x) + \sum_{x' \in V(G_{t+1}) \setminus V(G_t)} \deg_{t+1}(x') \\
&= \sum_{x \in V(G_t)} (2\deg_t(x) + 1) + \sum_{x \in V(G_t)} (\deg_t(x) + 1) \\
&= 3\mathrm{vol}(G_t) + n_{t+1}. \qquad (2.1)
\end{aligned}$$

Hence by (2.1) for $t > 0$,

$$\begin{aligned}
\mathrm{vol}(G_t) &= 3\mathrm{vol}(G_{t-1}) + n_t = 3^t \mathrm{vol}(G_0) + n_0 \sum_{i=0}^{t-1} 3^i 2^{t-i} \\
&= 3^t \mathrm{vol}(G_0) + 2n_0(3^t - 2^t),
\end{aligned}$$

where the third equality follows by summing a geometric series. □

2.2. Proof of Theorem 1.2

In computing distances in the ILT model, the following lemma is helpful.

Lemma 2.2. *Let x and y be nodes in G_t with $t > 0$. Then*

$$d_{t+1}(x', y) = d_{t+1}(x, y') = d_{t+1}(x, y) = d_t(x, y)$$

and

$$d_{t+1}(x', y') = \begin{cases} d_t(x, y) & \text{if } xy \notin E(G_t), \\ d_t(x, y) + 1 = 2 & \text{if } xy \in E(G_t). \end{cases}$$

Proof. We prove that $d_{t+1}(x, y) = d_t(x, y)$. The proofs of the other equalities are analogous and so are omitted. Since in the ILT model we do not delete any edges, the distance cannot increase after a "cloning" step occurs. Hence, $d_{t+1}(x, y) \leq d_t(x, y)$. Now suppose for a contradiction that there is a path P' connecting x and y in G_{t+1} with length $k < d_t(x, y)$. Hence, P' contains nodes not in G_t. Choose such a P' with the least number of nodes, say $s > 0$, not in G_t. Let z' be a node of P' not in G_t, and let the neighbors of z' in P' be u and v. Then $z \in V(G_t)$ is joined to u and v. Form the path Q' by replacing z' by z. But then Q' has length k and has $s - 1$ nodes not in G_t, which supplies a contradiction. □

We now turn to the proof of Theorem 1.2. We prove only item (1), noting that items (2) and (3) follow from (1) by computation. We derive a recurrence for $W(G_t)$ as follows. To compute $W(G_{t+1})$, there are five cases to consider: distances within G_t, and distances of the forms $d_{t+1}(x, y')$, $d_{t+1}(x', y)$, $d_{t+1}(x, x')$, and $d_{t+1}(x', y')$. The first three cases contribute $3W(G_t)$ by Lemma 2.2. The fourth case contributes n_t. The final case contributes $W(G_t) + e_t$ (the term e_t comes from the fact that each edge xy contributes $d_t(x, y) + 1$).

Thus

$$W(G_{t+1}) = 4W(G_t) + e_t + n_t = 4W(G_t) + 3^t(e_0 + n_0).$$

Hence

$$\begin{aligned} W(G_t) &= 4^t W(G_0) + \sum_{i=0}^{t-1} 4^i \left(3^{t-1-i}\right)(e_0 + n_0) \\ &= 4^t W(G_0) + 4^t(e_0 + n_0)\left(1 - \left(\frac{3}{4}\right)^t\right). \end{aligned}$$

Diameters are constant in the ILT model. We record this as a strong indication of the (ultra) small-world property in the model.

Lemma 2.3. *For all graphs G_0 different from a clique,*

$$\operatorname{diam}(G_t) = \operatorname{diam}(G_0),$$

and $\operatorname{diam}(G_t) = \operatorname{diam}(G_0) + 1 = 2$ *when G_0 is a clique.*

Proof. This follows directly from Lemma 2.2. □

2.3. Proof of Theorem 1.3

We introduce the following dependency structure that will help us classify the degrees of nodes. Given a node $x \in V(G_0)$ we define its *descendant tree at time* t, written $T(x,t)$, to be a rooted binary tree with root x whose leaves are all of the nodes at time t. To define the $(k+1)$th row of $T(x,t)$, let y be a node in the kth row (y corresponds to a node in G_k). Then y has exactly two descendants on row $k+1$: y itself and y'. In this way, we may identify the nodes of G_t with a length-t binary sequence corresponding to the descendants of x, using the convention that a clone is labeled 1. We refer to such a sequence as the *binary sequence for* x at time t. We need the following technical lemma.

Lemma 2.4. *Let $S(x,k,t)$ be the nodes of $T(x,t)$ with exactly k zeros in their binary sequence at time t. Then for all $y \in S(x,k,t)$,*

$$2^k(\deg_0(x)+1)+t-k-1 \le \deg_t(y) \le 2^k(\deg_0(x)+t-k+1)-1.$$

Proof. The degree $\deg_t(y)$ is minimized when y is identified with the binary sequence beginning with k zeros: $(0,\dots,0,1,1,\dots,1)$. In this case,

$$\begin{aligned}\deg_t(y) &= 2(2(\cdots(2(2\deg_0(x)+1)+1)\cdots)+1)+1+(t-k)\\ &= 2^k(\deg_0(x)+1)+t-k-1.\end{aligned}$$

The degree $\deg_t(y)$ is maximized by the binary sequence ending with k zeros: $(1,1,\dots,1,0,\dots,0)$. Then

$$\begin{aligned}\deg_t(y) &= 2(2(\cdots(2(\deg_0(x)+t-k)+1)\cdots)+1)+1\\ &= 2^k(\deg_0(x)+t-k+1)-1.\end{aligned}$$

□

It can be shown (using Lemma 2.4) that the number of nodes of degree at least j at time t, denoted by $N_{(\ge j)}$, satisfies

$$\sum_{i=\log_2 j}^{t}\binom{t}{i} \le N_{(\ge j)} \le \sum_{i=\max\{\log_2 j-\log_2 t-O(1),0\}}^{t}\binom{t}{i}.$$

Indeed, when a vertex is identified with the binary sequence with $i \ge \log_2 k$ zeros, then the degree is at least k. We have $\binom{t}{i}$ such sequences. On the other hand, if the binary sequence has $i \le \log_2 k - \log_2 t - O(1)$ zeros, then the corresponding vertex has degree smaller than k. In particular, $N_{(\ge j)} = \Theta(n_t)$ for $j \le \sqrt{n_t}$, and therefore, the degree distribution of G_t does not follow a power law. Since $\binom{t}{j}$ nodes have degree around 2^j, the degree distribution has "binomial-type"

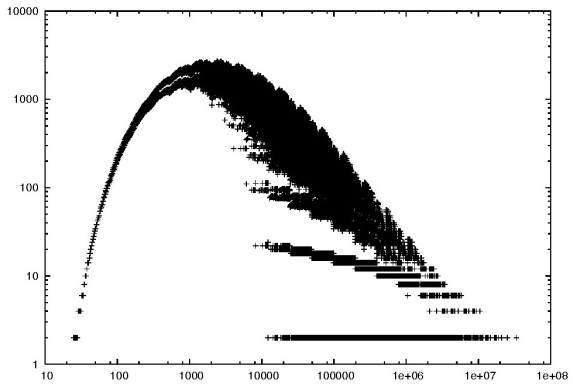

Figure 2. A log-log plot of the degree distribution for G_{25} with $G_0 = K_1$.

behavior. As an example of the degree distribution of a graph generated by the ILT model, see Figure 2.

We now prove the following lemma. Recall that $e(x,t)$ is the number of edges in $G_t \restriction N_t(x)$.

Lemma 2.5. *For all $x \in V(G_t)$ with k zeros in their binary sequence, we have that*

$$\Omega(3^k) = e(x,t) = O(3^k t^2).$$

We note that the constants hidden in the $\Omega(\cdot)$ and $O(\cdot)$ notations (both in the statement of the lemma and in the proof below) do not depend on k or on t.

Proof of Lemma 2.5. For $x \in V(G_t)$ we have that

$$\begin{aligned} e(x,t+1) &= e(x,t) + \deg_t(x) + \sum_{i=1}^{\deg_t(x)} (1 + \deg_{G_t \restriction N_t(x)}(x)) \\ &= 3e(x,t) + 2\deg_t(x). \end{aligned}$$

For x', we have that

$$e(x', t+1) = e(x,t) + \deg_t(x).$$

Since there are k zeros and $e(x,2)$ is always positive for all initial graphs G_0, $e(x,t) \geq 3^{k-2} e(x,2) = \Omega(3^k)$, and the lower bound follows.

For the upper bound, a general binary sequence corresponding to x is of the form

$$(1, \ldots, 1, 0, 1, \ldots, 1, 0, 1, \ldots, 1, 0, 1, \ldots, 1, 0, 1, \ldots, 1)$$

with the 0's in positions i_k $(1 \le i \le k)$. Consider a path in the descendant tree from the root of the tree to node x. By Lemma 2.4, the node on the path in the ith row $(i < i_j)$ has (at time i) degree $O(2^{j-1}t)$.

Hence the number of edges we estimate is $O(t^2)$ until the $(i_1 - 1)$th row, increases to $3O(t^2) + O(2^1 t)$ in the next row, and increases to $3O(t^2) + O(2^1 t^2)$ in the $(i_2 - 1)$th row. By induction, we have that

$$\begin{aligned} e(x,t) &= 3(\cdots(3(3O(t^2) + O(2^1 t^2)) + O(2^2 t^2))\cdots) + O(2^k t^2) \\ &= O(t^2)3^k \sum_{i=0}^{k} \left(\frac{2}{3}\right)^j = O(3^k t^2). \end{aligned}$$

□

We now prove our result on clustering coefficients.

Proof of Theorem 1.3.. For $x \in V(G_t)$ with k zeros in its binary sequence, by Lemmas 2.4 and 2.5 we have that

$$c_t(x) = \Omega\left(\frac{3^k}{(2^k t)^2}\right) = \Omega\left(\left(\frac{3}{4}\right)^k t^{-2}\right)$$

and

$$c_t(x) = O\left(\frac{3^k t^2}{(2^k)^2}\right) = O\left(\left(\frac{3}{4}\right)^k t^2\right).$$

Hence, since we have $n_0\binom{t}{k}$ nodes with k zeros in its binary sequence,

$$C(G_t) = \frac{\sum_{k=0}^{t} n_0\binom{t}{k}\Omega\left(\left(\frac{3}{4}\right)^k t^{-2}\right)}{n_0 2^t} = \Omega\left(\frac{t^{-2}\left(1+\frac{3}{4}\right)^t}{2^t}\right) = \Omega\left(\left(\frac{7}{8}\right)^t t^{-2}\right).$$

In a similar fashion, it follows that

$$C(G_t) = \frac{\sum_{k=0}^{t} n_0\binom{t}{k} O\left(\left(\frac{3}{4}\right)^k t^2\right)}{n_0 2^t} = O\left(\left(\frac{7}{8}\right)^t t^2\right).$$

□

2.4. Proofs of Theorems 1.4, 1.5, 1.6, and 1.7

We present proofs of the spectral properties of the ILT model. For ease of notation, let $\lambda(t) = \lambda(G_t)$.

Proof of Theorem 1.4. We use the expander mixing lemma for the normalized Laplacian (see [Chung 97]). For sets of nodes X and Y we use the notation $\mathrm{vol}(X)$ for the volume of the subgraph induced by X, and $e(X,Y)$ for the number of edges with one end in each of X and Y.

Lemma 2.6. *For all sets $X \subseteq G$,*

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

We observe that G_t contains an independent set (that is, a set of nodes with no edges) with volume $\mathrm{vol}(G_{t-1}) + n_{t-1}$. Let X denote this set, that is, the new nodes added at time t. Then by (2.1) it follows that

$$\mathrm{vol}(\bar{X}) = \mathrm{vol}(G_t) - \mathrm{vol}(X) = 2\mathrm{vol}(G_{t-1}) + n_{t-1}.$$

Since X is independent, Lemma 2.6 implies that

$$\lambda(t) \geq \frac{\mathrm{vol}(X)}{\mathrm{vol}(\bar{X})} = \frac{\mathrm{vol}(G_{t-1}) + n_{t-1}}{2\mathrm{vol}(G_{t-1}) + n_{t-1}} > \frac{1}{2}.$$ □

Proof of Theorem 1.5. Before we proceed with the proof of Theorem 1.5, we begin by stating some notation and a lemma. For a given node $u \in V(G_t)$, we let $\tilde{u} \in V(G_0)$ denote the node in G_0 of which u is a descendant. Given $uv \in E(G_0)$, we define

$$\mathcal{A}_{uv}(t) = \{xy \in E(G_t) : \tilde{x} = u, \tilde{y} = v\},$$

and for $v \in E(G_0)$, we set

$$\mathcal{A}_v(t) = \{xy \in E(G_t) : \tilde{x} = \tilde{y} = v\}.$$

We use the following lemma, for which the proof of items (1) and (2) follow from Lemma 2.1. The final item contains a standard form of the Raleigh quotient characterization of the second eigenvalue; see [Chung 97].

Lemma 2.7. *(1) For $uv \in E(G_0)$,*

$$|\mathcal{A}_{uv}(t)| = 3^t.$$

(2) For $v \in V(G_0)$,

$$|\mathcal{A}_v(t)| = 3^t - 2^t.$$

(3) Define

$$\bar{d} = \frac{\sum_{v \in V(G_t)} f(v) \deg_t(v)}{\mathrm{vol}(G_t)}.$$

Then

$$\lambda_1(t) = \inf_{\substack{f: V(G_t) \to \mathbb{R}, \\ f \neq 0}} \frac{\sum\limits_{uv \in E(G_t)} (f(u) - f(v))^2}{\sum\limits_{v} f^2(v) \deg_t(v) - \bar{d}^2 \mathrm{vol}(G_t)}. \tag{2.2}$$

Note that in item (3), $\bar{d}$ is a function of f. Now let $g : V(G_0) \to \mathbb{R}$ be the harmonic eigenvector for $\lambda_1(0)$, so that

$$\sum_{v \in V(G_0)} g(v) \deg_0(v) = 0$$

and

$$\lambda_1(0) = \frac{\sum_{uv \in E(G_0)} (g(u) - g(v))^2}{\sum_{v \in V(G_0)} g^2(v) \deg_0(v)}.$$

Furthermore, we choose g scaled so that $\sum_{v \in V(G_0)} g^2(v) \deg_0(v) = 1$. This is the standard version of the Raleigh quotient for the normalized Laplacian from [Chung 97], so such a g exists as long as G_0 has at least two eigenvalues, which it does by our assumption that $G_0 \ncong K_1$. Our strategy in proving the theorem is to show that lifting g to G_1 provides an effective bound on the second eigenvalue of G_1 using the form of the Raleigh quotient given in (2.2).

Define $f : G_t \to \mathbb{R}$ by $f(x) = g(\tilde{x})$. Then note that

$$\begin{aligned}
\sum_{xy \in E(G_t)} (f(x) - f(y))^2 &= \sum_{\substack{xy \in E(G_t), \\ \tilde{x} = \tilde{y}}} (f(x) - f(y))^2 + \sum_{\substack{xy \in E(G_t) \\ \tilde{x} \neq \tilde{y}}} (f(x) - f(y))^2 \\
&= \sum_{uv \in E(G_0)} \sum_{xy \in \mathcal{A}_{uv}} (g(u) - g(v))^2 \\
&= 3^t \sum_{uv \in E(G_0)} (g(u) - g(v))^2.
\end{aligned}$$

By Lemma 2.7(1) and (2) it follows that

$$\begin{aligned}
\sum_{x \in V(G_t)} f^2(x) \deg_t(x) &= \sum_{x \in V(G_t)} \sum_{xy \in E(G_t)} f^2(x) \\
&= \sum_{u \in V(G_0)} \sum_{\substack{xy \in E(G_t), \\ \tilde{x} = u}} g^2(u) \\
&= \sum_{u \in V(G_0)} g^2(u) \left(\sum_{vu \in E(G_0)} \sum_{xy \in \mathcal{A}_{uv}} 1 + 2|\mathcal{A}_u| \right) \\
&= 3^t \sum_{u \in V(G_0)} g^2(u) \deg_0(u) + 2(3^t - 2^t) \sum_{u \in V(G_0)} g^2(u) \\
&= 3^t + 2(3^t - 2^t) \sum_{u \in G_0} g^2(u).
\end{aligned}$$

By Lemma 2.1 and proceeding as above, noting that $\sum_{v\in V(G_0)} g(v)\deg_0(v) = 0$, we have that

$$\begin{aligned}
\bar{d}^2\mathrm{vol}(G_t) &= \frac{\left(\sum\limits_{x\in V(G_t)} f(x)\deg_t(x)\right)^2}{\mathrm{vol}(G_t)} \\
&= \frac{\left(2(3^t-2^t)\sum\limits_{u\in V(G_0)} g(u)\right)^2}{\mathrm{vol}(G_t)} \\
&= \frac{4\cdot 3^{2t}\left(1-\left(\frac{2}{3}\right)^t\right)^2\left(\sum\limits_{u\in V(G_0)} g(u)\right)^2}{3^t\left(\mathrm{vol}(G_0)+2n_0\left(1-\left(\frac{2}{3}\right)^t\right)\right)} \\
&\le \frac{4\cdot 3^t\left(1-\left(\frac{2}{3}\right)^t\right)^2\sum\limits_{u\in V(G_0)} g^2(u)}{\bar{D}+2\left(1-\left(\frac{2}{3}\right)^t\right)},
\end{aligned}$$

where $\bar{D}$ is the average degree of G_0, and the last inequality follows from the Cauchy–Schwarz inequality.

By (2.2) we have that

$$\begin{aligned}
\lambda_1(t) &\le \frac{\sum\limits_{xy\in E(G_t)}(f(x)-f(y))^2}{\sum\limits_{x\in V(G_t)} f^2(x)\deg_t(x)+\bar{d}^2\mathrm{vol}(G_t)} \\
&\le \frac{3^t\sum\limits_{uv\in E(G_0)}(g(u)-g(v))^2}{3^t+2\cdot 3^t\left(1-\left(\frac{2}{3}\right)^t\right)\left(\sum_{u\in V(G_0)} g^2(u)\right)-\frac{4\cdot 3^t\left(1-\left(\frac{2}{3}\right)^t\right)^2\sum\limits_{u\in V(G_0)} g^2(u)}{\bar{D}+2\left(1-\left(\frac{2}{3}\right)^t\right)}} \\
&= \frac{\lambda_1(0)}{1+2\left(1-\left(\frac{2}{3}\right)^t\right)\left(\sum\limits_{u\in V(G_0)} g^2(u)\right)\left(1-\frac{2\left(1-\left(\frac{2}{3}\right)^t\right)}{\bar{D}+2\left(1-\left(\frac{2}{3}\right)^t\right)}\right)} \\
&< \lambda_1(0),
\end{aligned}$$

where the strict inequality follows from the fact that $\bar{D}\ge 1$, since G_0 is connected and $G_0 \ncong K_1$. □

Proof of Theorem 1.6. We denote vectors using boldface. We first assume that $\rho \ne -1$. Hence, $\rho_+, \rho_- \ne 0$. Let $\mathbf{u}$ be an eigenvector of $A = A(G_t)$ such that $A\mathbf{u} = \rho\mathbf{u}$.

Let $\beta = \frac{(\rho+1)}{\rho}$, and let

$$\mathbf{v} = \begin{pmatrix} \mathbf{u} \\ \beta\mathbf{u} \end{pmatrix}.$$

Then we have that

$$M\mathbf{v} = \begin{pmatrix} A & A+I \\ A+I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \beta\mathbf{u} \end{pmatrix} = \begin{pmatrix} \rho\mathbf{u} + (\rho+1)\beta\mathbf{u} \\ (\rho+1)\mathbf{u} \end{pmatrix}.$$

Now $\beta\rho = \rho + 1$, and so $(\rho+1)\mathbf{u} = \beta\rho\mathbf{u}$. The condition

$$\rho = \rho + \beta(\rho+1) = \rho + \frac{(\rho+1)^2}{\rho}$$

is equivalent to ρ solving

$$x - \rho - \frac{(\rho+1)^2}{x} = 0.$$

Hence $M\mathbf{v} = \rho\mathbf{v}$, as desired.

Now let $\rho = -1$. In this case, $\rho_- = -1$. Let

$$\mathbf{v} = \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix},$$

where $\mathbf{0}$ is the appropriately sized zero vector. Thus,

$$M\mathbf{v} = \begin{pmatrix} A & A+I \\ A+I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} -\mathbf{u} \\ \mathbf{0} \end{pmatrix}.$$

Hence $M\mathbf{v} = \rho_-\mathbf{v}$, as desired. In the case that $\rho_+ = 0$ and $\rho = -1$, let

$$\mathbf{v} = \begin{pmatrix} \mathbf{0} \\ \mathbf{u} \end{pmatrix},$$

and so $M\mathbf{v} = \rho_+\mathbf{v}$. □

Proof of Theorem 1.7. Without loss of generality, we assume that G_0 is not the trivial graph K_1; otherwise, G_1 is K_2, and we may start from there. Thus, in particular, we can assume $\rho_0(0) \geq 1$.

We first observe that by Theorem 1.6,

$$\rho_0(t) \geq \left(\frac{1+\sqrt{5}}{2}\right)^t \rho_0(0).$$

By Theorem 1.6 and by taking a branch of descendants from the largest eigenvalue, it follows that

$$|\rho_1(t)| \geq \frac{2(\sqrt{5}-1)}{(1+\sqrt{5})^2}\left(\frac{1+\sqrt{5}}{2}\right)^t \rho_0(0).$$

Hence, to prove the theorem, it suffices to show that

$$\rho_0(t) \leq c\left(\frac{1+\sqrt{5}}{2}\right)^t \rho_0(0).$$

Observe that, again by Theorem 1.6 and taking the largest branch of descendants from the largest eigenvalues,

$$\rho_0(t) = \rho_0(0)\prod_{i=0}^{t-1}\left(\frac{1+\sqrt{5+\frac{8}{\rho_0(i)}+\frac{4}{\rho_0^2(i)}}}{2}\right) \leq \rho_0(0)\prod_{i=0}^{t-1}\left(\frac{1+\sqrt{5+\frac{6}{\rho_0(i)}}}{2}\right).$$

Thus,

$$\begin{aligned}
\frac{2^t\rho_0(t)}{(1+\sqrt{5})^t} &\leq \rho_0(0)\prod_{i=0}^{t-1}\frac{1+\sqrt{5+\frac{6}{\rho_0(i)}}}{1+\sqrt{5}} \\
&\leq \rho_0(0)\prod_{i=0}^{t-1}\left(1+\frac{\sqrt{5}}{1+\sqrt{5}}\frac{6}{5\rho_0(i)}\right) \\
&\leq \rho_0(0)\exp\left(\frac{6\sqrt{5}}{5(1+\sqrt{5})}\sum_{i=0}^{t-1}\rho_0(i)^{-1}\right) \\
&\leq \rho_0(0)\exp\left(\frac{6\sqrt{5}}{5(1+\sqrt{5})\rho_0(0)}\sum_{i=0}^{\infty}\left(\frac{2}{1+\sqrt{5}}\right)^{-i}\right) = \rho_0(0)c.
\end{aligned}$$

In all, we have proved that for constants c and d,

$$c\left(\frac{1+\sqrt{5}}{2}\right)^t \rho_0(0) \geq \rho_0(t) \geq |\rho_1(t)| \geq d\left(\frac{1+\sqrt{5}}{2}\right)^t \rho_0(t).$$

□

2.5. Proofs of Theorems 1.8 and 1.9

We give the proofs for the results on the cop number, domination number, and automorphism group of the ILT model.

Proof of Theorem 1.8. We prove that for $t \geq 0$, $\gamma(G_{t+1}) = \gamma(G_t)$. It then follows that $\gamma(G_t) = \gamma(G_0)$. When a dominating node $x \in V(G_t)$ is cloned, its clone x' will be dominated by x. The clone y' of a nondominating node $y \in V(G_t)$ will be joined to a dominating node, since y is joined to one. Hence, a dominating set in G_t is a dominating set in G_{t+1}, and so $\gamma(G_{t+1}) \leq \gamma(G_t)$. If S' is a dominating set in G_{t+1}, then form S by replacing (if necessary) nodes $x' \in S'$ by nodes x. Since S dominates G_t, it follows that $\gamma(G_t) \leq \gamma(G_{t+1})$.

We next show that $c(G_{t+1}) = c(G_t)$. Let $c = c(G_t)$. Assume that c cops play in G_{t+1}, so that whenever $\mathcal{R}$ is on $x' \in V(G_{t+1}) \setminus V(G_t)$, the cops $\mathcal{C}$ play as if $\mathcal{R}$ were on $x \in V(G_t)$. Either $\mathcal{C}$ captures $\mathcal{R}$ on x', or using their winning strategy in G_t, the cops move to x with $\mathcal{R}$ on x'. The cops then win in the next round. Hence,

$$c(G_{t+1}) \leq c(G_t).$$

If $b = c(G_{t+1}) < c$, then we prove that $c(G_t) \leq b$, which is a contradiction. Suppose that $\mathcal{R}$ and $\mathcal{C}$ play in G_t. At the same time this game is played, let the set of b cops $\mathcal{C}'$ play with their winning strategy in G_{t+1}, under the assumption that $\mathcal{R}$ remains in G_t. Each time a cop in $\mathcal{C}'$ moves to a cloned node x', move the corresponding cop in $\mathcal{C}$ to x. Since x and x' are joined and share the exact same neighbors in G_{t+1}, $\mathcal{C}$ may win in G_t with $b < c$ cops. □

Proof of Theorem 1.9. We first prove the following lemma.

Lemma 2.8. *Each $f_0 \in \mathrm{Aut}(G_0)$, extends to $f_t \in \mathrm{Aut}(G_t)$.*

Proof. Given $f_0 \in \mathrm{Aut}(G_0)$, we prove by induction on $t \geq 0$ that f_0 extends to $f_t \in \mathrm{Aut}(G_t)$. The base case is immediate. Assuming that f_t is defined, let

$$f_{t+1}(x) = \begin{cases} f_t(x) & \text{if } x \in V(G_t), \\ (f_t(y))' & \text{where } x = y' \ . \end{cases}$$

Let x, y be distinct nodes of $V(G_t)$. It is straightforward to see that f_{t+1} is a bijection. We show that $xy \in E(G_{t+1})$ if and only if $f_{t+1}(x)f_{t+1}(y) \in E(G_{t+1})$. This will prove that $f_{t+1} \in \mathrm{Aut}(G_t)$, since f_{t+1} extends f_t.

The case for $x, y \in V(G_t)$ is immediate, since $f_t \in \mathrm{Aut}(G_t)$. Next, we consider the case for $x \in V(G_t)$ and $y' \in V(G_{t+1})$. Now $xy' \in E(G_{t+1})$ if and only if

$$f_{t+1}(x)f_{t+1}(y') = f_t(x)(f_t(y))' \in E(G_{t+1}).$$

Note that $x'y' \notin E(G_{t+1})$ for all $x', y' \in V(G_{t+1}) \setminus V(G_t)$. But $f_{t+1}(x')f_{t+1}(y') \notin E(G_{t+1})$ by definition of G_{t+1}. □

We now prove that for all $t \geq 0$, $\mathrm{Aut}(G_t)$ is isomorphic to a subgroup of $\mathrm{Aut}(G_{t+1})$. The proof of Theorem 1.9 then follows from this fact by induction on t. Define

$$\phi : \mathrm{Aut}(G_t) \to \mathrm{Aut}(G_{t+1})$$

by

$$\phi(f)(x) = \begin{cases} f(x) & \text{if } x \in V(G_t), \\ (f(y))' & \text{if } x = y' \in V(G_{t+1}) \setminus V(G_t). \end{cases}$$

Note that $\phi(f)(x)$ is injective, since $f \neq g$ implies that $\phi(f) \neq \phi(g)$ by the definition of ϕ.

We prove that for all $x \in V(G_{t+1})$ and $f, g \in \mathrm{Aut}(G_t)$,

$$\phi(fg)(x) = \phi(f)\phi(g)(x).$$

If $x \in V(G_t)$, then

$$\phi(fg)(x) = fg(x) = \phi(f)\phi(g)(x).$$

If $x \notin V(G_t)$, then say $x = y'$, with $y \in V(G_t)$. We then have that

$$\phi(fg)(x) = (fg(y))' = (\phi(f)\phi(g)(y))' = \phi(f)(g(y))' = \phi(f)\phi(g)(x). \qquad \square$$

2.6. Proofs of Theorems 1.10 and 1.11

We give the proofs for the results on the randomized ILT model, ILT(p). Without loss of generality, we assume that $0 < p < 1$.

Proof of Theorem 1.10. By the definition of the ILT(p) model, we obtain the following conditional expectation:

$$\mathbb{E}(\mathrm{vol}(H_{t+1}) \mid \mathrm{vol}(H_t)) = 3\mathrm{vol}(H_t) + n_{t+1} + n_t(n_t - 1)p(n_t).$$

At the beginning of the process, we cannot control the random variable $\mathrm{vol}(H_t)$; it may be far from its expectation. However, if t is large enough, a number of additional edges added in a random process may be controlled, and $\mathrm{vol}(H_t)$ eventually approaches its expected value. Let

$$t_0(T) = \frac{4 \log \log T}{\log(3 + \delta)} \tag{2.3}$$

be the time from which we can control the process (note that $t_0(T)$ tends to infinity with T). Now suppose that

$$\mathrm{vol}(H_{t_0}) = (3 + \delta)^{t_0}(1 + A(t_0)).$$

The function $A(t_0)$ measures how far $\mathrm{vol}(G_{t_0})$ is from its expectation; we do not give an explicit formula for this, but the bounds $-1 \le A(t_0) \le (\frac{4}{3+\delta})^{t_0}$ apply (deterministically; note that -1 corresponds to the empty graph, while $(\frac{4}{3+\delta})^{t_0}$ corresponds to a complete graph). We first demonstrate that for any t (where $t_0(T) \le t \le T$), with probability at least $(1 - T^{-2})^t$,

$$\mathrm{vol}(H_t) = (1 + o(1))(3 + \delta)^t \left(1 + \left(\frac{3}{3+\delta}\right)^{t-t_0} A(t_0)\right). \tag{2.4}$$

We prove (2.4) by induction on t. The base case, $t = t_0$, trivially holds. For the inductive step, assume that (2.4) holds for $t_0 = t_0(T) \le t < T$ (with probability at least $(1-T^{-2})^t$). We want to show that (2.4) holds for $t+1$ (with probability at least $(1-T^{-2})^{t+1}$). Using (2.3) and (1.3), we have that the expected number of *random edges* added at time $t+1$ (that is, edges added between new nodes) is

$$\begin{aligned}\mathbb{E}X &= 2^t(2^t - 1)p(2^t) = (1 - (1/2)^t)\delta(3 + \delta)^t \\ &\ge (1 + o(1))\delta(3 + \delta)^{t_0} \ge (1 + o(1))\delta \log^4 T.\end{aligned}$$

Using the Chernoff bound

$$\mathbb{P}(|X - \mathbb{E}X| \ge \varepsilon \mathbb{E}X) \le 2\exp(-\varepsilon^2 \mathbb{E}X/3)$$

with $\varepsilon = 1/\log T$, we derive that the number of random edges is not concentrated with probability at most

$$2\exp\left(-\frac{\varepsilon^2 \mathbb{E}X}{3}\right) \le 2\exp\left(-\frac{\delta \log^2 T}{4}\right) \le T^{-2}.$$

Thus, with probability at least $(1 - T^{-2})^{t+1}$, we have that

$$\begin{aligned}\mathrm{vol}(H_{t+1}) &= 3\mathrm{vol}(H_t) + 2^{t+1} + (1 + O(\log^{-1} T))\delta(3 + \delta)^t \\ &= (1 + o(1))(3 + \delta)^t \left(3 + 3\left(\frac{3}{3+\delta}\right)^{t-t_0} A(t_0) + \delta\right) \\ &= (1 + o(1))(3 + \delta)^{t+1} \left(1 + \left(\frac{3}{3+\delta}\right)^{t+1-t_0} A(t_0)\right).\end{aligned}$$

By the bounds on $A(t_0)$ it follows that

$$\left(\frac{3}{3+\delta}\right)^{T-t_0} A(t_0) = \exp\left(-\Omega(T) + O(t_0)\right) = o(1).$$

Therefore, the assertion holds with probability at least

$$(1 + T^{-2})^T = \exp((1 + o(1))T^{-1}) = 1 + o(1).$$

□

Proof of Theorem 1.11. Let

$$X = V(H_T) \setminus V(H_{T-1})$$

and

$$\bar{X} = V(H_T) \setminus X = V(H_{T-1}).$$

By computation it follows that a.a.s.,

$$\begin{aligned}
\mathrm{vol}(X) &= (1+o(1))(1+\delta)(3+\delta)^{T-1}, \\
\mathrm{vol}(\bar{X}) &= (1+o(1))2(3+\delta)^{T-1}, \\
\mathrm{vol}(H_T) &= (1+o(1))(3+\delta)(3+\delta)^{T-1},
\end{aligned}$$

and

$$e(X,X) = (1+o(1))(3+\delta)^{T-1}.$$

Thus, by Lemma 2.6 we have that a.a.s.,

$$\begin{aligned}
\lambda(T) &\geq (1+o(1))\frac{|3+\delta-(1+\delta)^2|}{2(1+\delta)} \\
&= (1+o(1))\frac{2-\delta-\delta^2}{2(1+\delta)} = \Omega(1).
\end{aligned}$$

□

3. Conclusion and Further Work

We have introduced the ILT model for OSNs and other complex networks, whereby the network is cloned at each time step. We have proved that the ILT model generates graphs with a densification power law, in many cases decreasing average distance (and in all cases, the average distance and diameter are bounded above by constants independent of time), with higher clustering than random graphs with the same average degree, and with smaller spectral gaps for both their normalized Laplacian and adjacency matrices than in random graphs. The cop and domination numbers were shown to remain the same as those for the graph from the initial time step G_0, and the automorphism group of G_0 is a subgroup of the automorphism group of graphs generated at all later times. A randomized version of the ILT model was introduced with tunable densification power-law exponent.

As we noted after the statement of Lemma 2.4, the ILT model does not generate graphs with a power-law degree distribution, and neither does the ILT(p) model. An interesting problem is to design and analyze a randomized version of the ILT model satisfying the properties displayed in the ILT model as well as generating power-law graphs. Such a randomized ILT model should with high

probability generate power-law graphs with topological and spectral properties similar to those of graphs from the deterministic ILT model.

Certain OSNs such as Twitter are directed networks, in which users may either be friends with other users (represented by undirected edges), or follow them (represented by a directed edge pointing to the follower). Hence, a more accurate model for such networks would be directed, and we will consider a directed version of the ILT model in the sequel.

References

[Adamic et al. 03] L. A. Adamic, O. Buyukkokten, and E. Adar. "A Social Network Caught in the Web." *First Monday* 8:6 (2003). Available at http://firstmonday.org/issues/issue8_6/adamic/index.html.

[Ahn et al. 07] Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong. "Analysis of Topological Characteristics of Huge Online Social Networking Services." In *Proceedings of the 16th International Conference on World Wide Web*, pp. 835–844. New York: ACM Press, 2007.

[Aigner and Fromme 84] M. Aigner and M. Fromme. "A Game of Cops and Robbers." *Discrete Applied Mathematics* 8 (1984), 1–12.

[Bebek et al. 06] G. Bebek, P. Berenbrink, C. Cooper, T. Friedetzky, J. Nadeau, and S. C. Sahinalp. "The Degree Distribution of the Generalized Duplication Model." *Theoretical Computer Science* 369 (2006), 234–249.

[Bhan et al. 02] A. Bhan, D. J. Galas, and T. G. Dewey. "A Duplication Growth Model of Gene Expression Networks. *Bioinformatics* 18 (2002), 1486–1493.

[Bonato 08] A. Bonato. *A Course on the Web Graph*, American Mathematical Society Graduate Studies Series in Mathematics. Providence: American Mathematical Society, 2008.

[Bonato and Janssen 09] A. Bonato and J. Janssen. "Infinite Limits and Adjacency Properties of a Generalized Copying Model. *Internet Mathematics* 4:2–3 (2009), 199–223.

[Bonato et al. 09] A. Bonato, N. Hadi, P. Horn, P. Pralat, and C. Wang. "Dynamic Models of On-Line Social Networks." In *Algorithms and Models for the Web-Graph: 6th International Workshop, WAW 2009, Barcelona, Spain, February 12–13, 2009, Proceedings*, Lecture Notes in Computer Science 5427, pp. 127–142. Berlin: Springer, 2009.

[Caldarelli 07] G. Caldarelli. *Scale-Free Networks*. Oxford: Oxford University Press, 2007.

[Chung 97] F. Chung, *Spectral Graph Theory*. Providence: American Mathematical Society, 1997.

[Chung and Lu 06] F. Chung and L. Lu, *Complex Graphs and Networks*. Providence: American Mathematical Society, 2006.

[Chung et al. 03] F. Chung, L. Lu, T. Dewey, and D. Galas. "Duplication Models for Biological Networks." *Journal of Computational Biology* 10 (2003), 677–687.

[Chung et al. 04] F. Chung, L. Lu, and V. Vu. "The Spectra of Random Graphs with Given Expected Degrees." *Internet Mathematics* 1:3 (2004), 257–275.

[Crandall et al. 08] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. "Feedback Effects between Similarity and Social Influence in Online Communities." In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 160–168. New York: ACM Press, 2008.

[Durrett 06] R. Durrett. *Random Graph Dynamics*. New York: Cambridge University Press, 2006.

[Ebel et al. 03] H. Ebel, J. Davidsen, and S. Bornholdt. "Dynamics of Social Networks." *Complexity* 8 (2003), 24–27.

[Estrada 06] E. Estrada. "Spectral Scaling and Good Expansion Properties in Complex Networks." *Europhys. Lett.* 73 (2006), 649–655.

[Frank 80] O. Frank. "Transitivity in Stochastic Graphs and Digraphs." *Journal of Mathematical Sociology* 7 (1980), 199–213.

[Furedi and Komlos 81] Z. Furedi and J. Komlos. "The Eigenvalues of Random Symmetric Matrices." *Combinatorica* 1 (1981), 233–241.

[Hahn 07] G. Hahn. "Cops, Robbers and Graphs." *Tatra Mountain Mathematical Publications* 36 (2007), 163–176.

[Girvan and Newman 02] M. Girvan and M. E. J. Newman. "Community Structure in Social and Biological Networks." *Proceedings of the National Academy of Sciences* 99 (2002), 7821–7826.

[Gkantsidis et al. 03] C. Gkantsidis, M. Mihail, and A. Saberi. "Throughput and Congestion in Power-Law Graphs." In *Proceedings of the 2003 ACM SIGMETRICS International Conference on Measurement Modeling of Computer Systems*, pp. 148-159. New York: ACM Press, 2003.

[Golder et al. 07] S. Golder, D. Wilkinson, and B. Huberman. "Rhythms of Social Interaction: Messaging within a Massive Online Network." In *Communities and Technologies 2007: Proceedings of the Third Communities and Technologies Conference, Michigan State University 2007*, edited by Charles Steinfield, Brian T. Pentland, Mark Ackerman, and Noshir Contractor, pp. 41–66. New York: Springer, 2007.

[Java et al. 07] A. Java, X. Song, T. Finin, and B. Tseng. "Why We Twitter: Understanding Microblogging Usage and Communities." In *Proceedings of the Joint 9th WebKDD and 1st SNA-KDD Workshop on Web Mining and Social Network Analysis*, pp. 56–65. New York: ACM Press 2007.

[Krishnamurthy et al. 08] B. Krishnamurthy, P. Gill, and M. Arlitt. "A Few Chirps about Twitter." In *Proceedings of The First ACM SIGCOMM Workshop on Online Social Networks*, pp. 19–24. New York: ACM Press, 2008.

[Kumar et al. 00] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. "Stochastic Models for the Web Graph." In *Proceedings of the 41th Annual Symposium on Foundations of Computer Science*, pp. 57–65. Washington, DC: IEEE, 2000.

[Kumar et al. 06] R. Kumar, J. Novak, and A. Tomkins. "Structure and Evolution of Online Social Networks." In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 611–617. New York: ACM Press, 2006.

[Leskovec et al. 05a] J. Leskovec, J. Kleinberg, and C. Faloutsos. "Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations." In *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 177–187. New York: ACM Press, 2005.

[Leskovec et al. 05b] J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos. "Realistic, Mathematically Tractable Graph Generation and Evolution, Using Kronecker Multiplication." In *Knowledge Discovery in Databases: PKDD 2005: 9th European Conference on Principles and Practice of Knowledge Discovery in Databases, Porto, Portugal, October 3–7, 2005, Proceedings*, Lecture Notes in Computer Science 3721, pp. 133-145. New York: Springer, 2005.

[Liben-Nowell et al. 05] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. "Geographic Routing in Social Networks." *Proceedings of the National Academy of Sciences* 102 (2005), 11623–11628.

[Milgram 67] S. Milgram. "The Small World Problem." *Psychology Today* 2 (1967), 60–67.

[Mislove et al. 07] A. Mislove, M. Marcon, K. Gummadi, P. Druschel, and B. Bhattacharjee. "Measurement and Analysis of Online Social Networks." In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pp. 29–42. New York: ACM Press, 2007.

[Nowakowski and Winkler 83] R. Nowakowski and P. Winkler. "Vertex-to-Vertex Pursuit in a Graph." *Discrete Mathematics* 43 (1983), 235–239.

[Quilliot 78] A. Quilliot. "Jeux et Pointes Fixes sur les Graphes." PhD thesis, Université de Paris VI, 1978.

[Pastor-Satorras et al. 03] R. Pastor-Satorras, E. Smith, and R. V. Sole. "Evolving Protein Interaction Networks through Gene Duplication." *J. Theor. Biol.* 22 (2003), 199–210.

[Scott 00] J. P. Scott, *Social Network Analysis: A Handbook*. London: Sage Publications Ltd., 2000.

[Watts and Strogatz 76] D. J. Watts, and S. H. Strogatz. "Collective Dynamics of 'Small-World' Networks." *Nature* 393 (1998), 440–442.

[White et al. 76] H. White, S. Harrison, and R. Breiger. "Social Structure from Multiple Networks, I: Blockmodels of Roles and Positions." *American Journal of Sociology* 81 (1976), 730–780.

Anthony Bonato, Department of Mathematics, Ryerson University, Toronto, ON, Canada, M5B 2K3 (abonato@ryerson.ca)

Noor Hadi, Department of Mathematics, Wilfrid Laurier University, Waterloo, NS, Canada, N2L 3C5 (hadi4130@wlu.ca)

Paul Horn, Department of Mathematics and Computer Science, Emory University, Atlanta, GA, U.S.A., 30322 (phorn@mathcs.emory.edu)

Paweł Prałat, Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310 (pralat@math.wvu.edu)

Changping Wang, Department of Mathematics, Ryerson University, Toronto, ON, Canada, M5B 2K3 (cpwang@ryerson.ca)

Received June 30, 2009; accepted June 3, 2010.