The Iterated Local Directed Transitivity Model for Social Networks

Anthony Bonato$^1$, Daniel W. Cranston$^2$, Melissa A. Huggan$^1$, Trent Marbach$^1$, and Raja Mutharasan$^1$

$^1$ Ryerson University, Toronto, ON, Canada
abonato@ryerson.ca, melissa.huggan@ryerson.ca, trent.marbach@ryerson.ca, rmutharasan@ryerson.ca

$^2$ Virginia Commonwealth University, Richmond, VA, USA
dcranston@vcu.edu

Abstract. We introduce a new, deterministic directed graph model for social networks, based on the transitivity of triads. In the Iterated Local Directed Transitivity (ILDT) model, new nodes are born over discrete time-steps and inherit the link structure of their parent nodes. The ILDT model may be viewed as a directed graph analog of the ILT model for undirected graphs introduced in [4]. We investigate network science and graph-theoretical properties of ILDT digraphs. We prove that the ILDT model exhibits a densification power law, so that the digraphs generated by the models densify over time. The number of directed triads are investigated, and counts are given of the number of directed 3-cycles and transitive 3-cycles. A higher number of transitive 3-cycles are generated by the ILDT model, as found in real-world, on-line social networks that have orientations on their edges. We discuss the eigenvalues of the adjacency matrices of ILDT digraphs. We finish by showing that in many instances of the chosen initial digraph, the model eventually generates digraphs with Hamiltonian directed cycles.

1 Introduction

Real-world, complex networks contain numerous mechanisms governing link formation. Balance theory (or structural balance theory) in social network analysis cites several mechanisms to complete triads (that is, subgraphs consisting of three nodes) in social networks [9,11]. A central mechanism in balance theory is transitivity: if $x$ is a friend of $y$, and $y$ is a friend of $z$, then $x$ is a friend of $z$; see, for example, [16]. Directed networks of ratings or trust scores and models for their propagation were first considered in [10]. Status theory for directed networks, first

* The first author acknowledges funding from an NSERC Discovery grant, while the third author acknowledges support from an NSERC Postdoctoral Fellowship.
introduced in [14], was motivated by both trust propagation and balance theory. While balance theory focuses on likes and dislikes, status theory posits that a directed link indicates that the creator of the link views the recipient as having higher status. For example, on Twitter or other social media, a directed link captures one user following another, and the person they follow may be of higher social status. Evidence for status theory was found in directed networks derived from Epinions, Slashdot, and Wikipedia [14]. For other applications of status theory and directed triads in social networks, see also [12,18].

The Iterated Local Transitivity (ILT) model introduced in [3,4] and further studied in [2,17], simulates social networks and other complex networks. Transitivity gives rise to the notion of cloning, where a node $x$ is adjacent to all of the neighbors of $y$. Note that in the ILT model, the nodes have local influence within their neighbor sets. Although the model graph evolves over time, there is still a memory of the initial graph hidden in the structure. The ILT model simulates many properties of social networks. For example, as shown in [4], graphs generated by the model densify over time and exhibit bad spectral expansion. In addition, the ILT model generates graphs with the small-world property, which requires the graphs to have low diameter and dense neighbor sets.

In the present work, we introduce a directed analog of the ILT model, where nodes are added and copy the in- and out-neighbors of existing nodes. The model simulates link creation in social networks, where new actors enter the network, and directed edges are added via transitivity through the lens of status theory. For example, in link formation in a directed social network such as Twitter, a new user may reciprocally follow an existing one, then in turn follow their followers. We consider the simplified setting where new nodes copy all of the links of their parent node. Our model, and iterated models more generally [2], provide counterparts for earlier studied random models for complex networks involving copying [13] or duplication [8].

More formally, the Iterated Local Directed Transitivity (ILDT) model is deterministically defined over discrete time-steps as follows. The only parameter of this deterministic model is the initial digraph $G = G_0$. For a non-negative integer $t$, the graph $G_t$ represents the digraph at time-step $t$. Suppose that the directed graph $G_t$ has been defined for a fixed time $t \geq 0$. To form $G_{t+1}$, for each $x \in V(G_t)$, add a new node $x'$ called the clone of $x$. We refer to $x$ as the parent of $x'$, and $x'$ as the
child of x. Between x and x' we add a bidirectional arc, representing a reciprocal status (or friendship) relationship between them. For arcs (x, z) and (y, x) in G_t, we add arcs (x', z) and (y, x'), respectively, in G_{t+1}. See Figure 1. We refer to G_t as an ILDT digraph. Note that the clones form an independent set in G_{t+1}.

In Figure 2, we illustrate several time-steps of the ILDT model beginning with the directed 3-cycle.

The paper is organized as follows. In Section 2, we prove that the ILDT model exhibits a densification power law, so that the digraphs generated by the model densify over time. The number of directed triads are investigated in Theorem 2 and precise counts are given of the number of directed 3-cycles and transitive cycles. These counts are contrasted, and it is shown that the transitive cycles are more abundant (as is the case with social networks; see [14]). We include a discussion of the eigenvalues of the adjacency matrices of ILDT directed graphs. In Section 3, the eigenvalues of ILDT directed graphs are investigated. In Section 4, we explore directed cycles of larger order in ILDT graphs. We show that ILDT digraphs are acyclic if the initial digraph is such, and that for many instances of initial digraphs, the model eventually generates graphs with Hamiltonian directed cycles. We finish with open problems.

For a general reference on graph theory, the reader is directed to [19]. For background on social and complex networks, see [1,5,7]. Throughout the paper, we consider finite, directed graphs with bidirectional edges
allowed. We refer to a directed edge as an arc. For nodes $x$ and $y$ of a graph, if there is an arc between $x$ and $y$, we denote it by $(x, y)$; if there is a bidirectional arc between $x$ and $y$ we denote such an arc with usual graph theoretic notation of $xy$. When counting the number of arcs within a graph, bidirectional arcs each contribute 2 to the final count. We use $\log n$ to be the logarithm of $n$ in base 2.

2 Densification and triad counts

As we referenced in the introduction, social networks densify, in the sense that the ratio of their number of arcs to nodes tends to infinity over time [15]. We show in this section that the ILDT model always generates digraphs that densify, and we give a precise statement below.
of its densification power law. As a phenomenon specific to digraphs, we consider the differing counts of directed and transitive 3-cycles in graphs generated by the model.

The number of nodes of $G_t$ is denoted $n_t$, the number of arcs is denoted $e_t$, and the number of bidirectional arcs is denoted $b_t$. Note that $e_t$ contains two arcs for each bidirectional arc. We establish elementary but important recursive formulas for these parameters.

**Lemma 1.** Let $G_0$ be a digraph with $n_0$ nodes, $e_0$ arcs, and $b_0$ bidirectional arcs. For all $t \geq 1$, we have the following:

1. $n_t = 2^t n_0$;
2. $e_t = 3e_{t-1} + 2n_{t-1}$; and
3. $b_t = 3b_{t-1} + n_{t-1}$.

**Proof.** Item (1) follows immediately as the number of nodes doubles in each time-step. For item (2), for each node in $G_{t-1}$ with $t > 0$, after cloning there will be a bidirectional arc between each parent and their child. These count as $2n_{t-1}$-many arcs. For every arc $(x, y)$ in $G_{t-1}$, arcs $(x, y')$ and $(x', y)$ appear in $G_t$. Hence, for every arc in $G_{t-1}$, three arcs of $G_t$ are generated. Summing these two counts gives the desired expression for $e_t$. Item (3) follows analogously to (2), except that we count the bidirectional edges once; hence, there are $n_{t-1}$-many bidirectional arcs. 

We now state the densification power law for ILDT graphs. For positive integer-valued functions $f_t$ and $g_t$, we use the expression $f_t \sim g_t$ if $f_t/g_t$ tends to 1 as $t$ tends to $\infty$.

**Corollary 1.** In the ILDT model, we have that

$$\frac{e_t}{n_t} \sim \left(\frac{3}{2}\right)^t \frac{(e_0 + 2n_0)}{n_0}.$$ 

In particular, we have that $e_t \sim C \cdot (n_t)^a$, where $a = \log 3$ and $C = \frac{e_0 + 2n_0}{(n_0)^a}$.

**Proof.** By Lemma 1, we have that

$$e_t = 3^t e_0 + 3^{t-1} 2^1 n_0 + 3^{t-2} 2^2 n_0 + \ldots + 3^1 2^{t-1} n_0 + 2^t n_0$$

$$= 3^t e_0 + 3^{t-1} 2n_0 \left(1 - \left(\frac{3}{2}\right)^{t-1}\right)$$

$$= 3^t (e_0 + 2n_0) - 2^t (2n_0).$$
We then derive that
\[
\frac{e_t}{n_t} = \frac{3^t (e_0 + 2 \cdot n_0) - 2^t (2n_0)}{2^t n_0} \sim \left( \frac{3}{2} \right)^t \frac{(e_0 + 2n_0)}{n_0},
\]
and the result follows.

We next consider 3-node subgraph counts in the ILDT model, where we find a higher number of transitive 3-cycles relative to directed 3-cycles. A similar phenomenon was found in directed network samples in social media such as Epinions, Slashdot, and Wikipedia, where transitive 3-cycles appear much more commonly than directed 3-cycles; see \[14\]. If the initial graph has no directed 3-cycles, then our results show there are none at any time-step in the evolution of the model. For nodes \(x, y, z\), if there exists a cycle with arcs \((x, y), (y, z),\) and \((z, x)\), then we abbreviate this directed 3-cycle to \((x, y, z)\). We take the directed 3-cycles \((x, y, z)\) and \((y, z, x)\) to be the same cycle. For a 3-cycle where the arcs \((x, y), (y, z),\) and \((x, z)\) are present, we denote this transitive 3-cycle by \(xyz\). A bidirectional 3-cycle is one that consists of three bidirectional arcs. Although, strictly speaking, a bidirectional 3-cycle contains six transitive 3-cycles and two directed 3-cycles, we will distinguish these by asserting that each transitive and directed 3-cycle must contain at least one non-bidirectional arc.

**Theorem 1.** In the graph \(G_t\), let \(D_t\) be the number of directed 3-cycles, \(T_t\) be the number of transitive 3-cycles, and \(B_t\) be the number of bidirectional 3-cycles. We then have that

1. \(D_{t+1} = 4D_t;\)
2. \(T_{t+1} = 4T_t + 4(e_t - 2b_t);\) and
3. \(B_{t+1} = 4B_t + 2b_t.\)

**Proof.** For item (1), consider a directed 3-cycle in a graph \(G_t\), labeled \((a, b, c)\). Each node can be replaced by its clone to produce a new directed 3-cycle; hence, \((a', b, c), (a, b', c),\) and \((a, b, c')\) are all directed 3-cycles generated by the directed 3-cycle \((a, b, c)\) from \(G_t\). Thus, for each 3-cycle in \(G_t\), there are four directed 3-cycles in \(G_{t+1}\). Therefore, \(D_{t+1} \geq 4D_t.\)

To establish the upper bound, suppose, by way of contradiction, that there exists another directed 3-cycle in \(G_{t+1}\) which was not previously counted. Such a directed 3-cycle cannot involve only parent nodes (since
it would be from \(G_t\) and it also cannot involve two clones because clones form an independent set. Hence, it must involve two nodes from \(G_t\) and one of the clones, call it \((a, d', c)\). But since \(d'\) has all the same adjacencies as \(d\), this implies that \((a, d, c)\) is a directed 3-cycle which is in \(G_t\), a contradiction, or that \(a, d, c\) are not all distinct. In the later case, we can assume that \(a = d\), and so the directed 3-cycle includes arcs \((a, c)\) and \((c, a')\), implying that 3-cycle is bidirectional, which is not counted in \(D_{t+1}\). Hence, \(D_{t+1} = 4D_t\).

To prove item (2), observe that for every non-bidirectional arc \((x, y)\) in \(G_t\), there will be four transitive 3-cycles in \(G_{t+1}\) formed with their clones using bidirectional arcs: \(xx'y, x'xy, xyy',\) and \(xy'y\). Thus, we add \(4(e_t - 2b_t)\) to the count of \(T_{t+1}\) (note that \(e_t - 2b_t\) counts the number of non-bidirectional arcs).

Existing transitive 3-cycles from \(G_t\) also exist in \(G_{t+1}\), and substituting a clone for its parent will produce a new transitive 3-cycle. Thus, each original transitive 3-cycle in \(G_t\) gives four transitive 3-cycles in \(G_{t+1}\) (namely, the cycle itself, and three produced from clone substitution). It is straightforward to check that these are the only transitive 3-cycles in \(G_{t+1}\). Therefore, \(T_t + 3T_t\) contributes to the total count of transitive 3-cycles. We then have that \(T_{t+1} = 4T_t + 4(e_t - 2b_t)\).

Item (3) can be shown similarly, as each bidirectional 3-cycle of \(G_t\) will correspond to four bidirectional 3-cycles in \(G_{t+1}\). Each bidirectional arc \(xy\) in \(G_t\) will correspond to two unique bidirectional 3-cycles in \(G_{t+1}\) formed from the nodes \(x, x', y, y'\).

We now present an exact expressions for \(D_t, T_t,\) and \(B_t\).

**Theorem 2.** In the ILDT digraph \(G_t\), we have that

1. \(D_t = 4^tD_0\);
2. \(T_t = 4^tT_0 + 4(4^t - 3^t)(e_0 - 2b_0)\); and
3. \(B_t = 4^tB_0 + 2b_0(4^t - 3^t) + n_0(4^t - 2 \cdot 3^t + 2^t)\).

**Proof.** Item (1) follows from Theorem 1 (1) by induction. For (2), by Theorem 1 (2), we derive that

\[
T_t = 4^tT_0 + \sum_{i=1}^{t} 4^i(e_{t-i} - 2b_{t-i}).
\]
By the proof of Corollary 1 along with a similar argument for $b_t$, we have that

$$e_{t-i} = 3^{t-i}(e_0 + 2n_0) - 2^{t-i+1}n_0,$$

$$b_{t-i} = 3^{t-i}(b_0 + n_0) - 2^{t-i}n_0,$$

and so (1)

$$e_{t-i} - 2b_{t-i} = 3^{t-i}(e_0 - 2b_0).$$

From (2), we derive that

$$T_t = 4^t T_0 + \sum_{i=1}^{t} 4^i 3^{t-i}(e_0 - 2b_0).$$

and item (2) follows by summing the geometric series.

For item (3), we have that $B_t = 4^t B_0 + \frac{1}{2} \sum_{i=1}^{t} 4^i b_{t-i}$. From equation (1), we find (in a way analogous to $T_t$) the desired expression for $B_t$.

We consider next the ratio of $D_t$ and $T_t$, which gives more explicit estimates on the relative abundance of transitive versus directed 3-cycles. By Theorem 2, we have that

$$\frac{D_t}{T_t} = \frac{4^t D_0}{4^t T_0 + 4(4^t - 3^t)(e_0 - 2b_0)} \sim \frac{D_0}{T_0 + 4(e_0 - 2b_0)}.$$}

Hence, the ratio $\frac{D_t}{T_t}$ may be made as small as we like by choosing an appropriate initial digraph.

### 3 Eigenvalues

We next consider eigenvalues of the adjacency matrices of ILDT digraphs. Spectral graph theory is a well-developed area for undirected graphs (see [6]) but less so for directed graphs (where the eigenvalues may be complex numbers with non-zero imaginary parts). We observe that if $G_t$ has adjacency matrix $A$, then $G_{t+1}$ has the following adjacency matrix:

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

where $I$ and 0 are the appropriately sized identity and zero matrices, respectively. The following recursive formula (analogous to the one in the
ILT model) allows us to determine all the eigenvalues of ILDT graphs from the spectrum of the initial graph. We have the following theorem from [1].

**Theorem 3.** Let \( t \geq 0 \). If \( \rho \) is an eigenvalue of the adjacency matrix of \( G_t \), then the eigenvalues of the adjacency matrix of \( G_{t+1} \) are

\[
\rho \pm \sqrt{\rho^2 + 4(\rho + 1)^2}.
\]

It is of interest to consider properties of the distribution of eigenvalues arising from ILDT digraphs graphs in the complex plane. For this, we consider the special case of \( G_0 \) a directed 3-cycle, which has eigenvalues the 3rd roots of unity. The resulting eigenvalue distribution of these ILDT digraphs suggests a rich structure. We plot the eigenvalues corresponding to \( G_t \) for \( 1 \leq t \leq 5 \) in Figure 3a.

![Fig. 3: Eigenvalues in the complex plane of ILDT digraphs \( G_t \), where \( 1 \leq t \leq 5 \). Colors distinguish the time-steps. Figures (a) depicts the eigenvalues, (b) the normalized eigenvalues, and (c) depicts the curves \( C_t \).](image-url)
If $\rho$ is an eigenvalue of the adjacency matrix of $G_t$ and $\rho$ is large in magnitude, then there is an eigenvalue of the adjacency matrix of $G_{t+1}$ that is approximately equal to $(1 + \sqrt{5})/2 \rho$, and in a similar way there is an eigenvalue of the adjacency matrix of $G_{t+\alpha}$ that is approximately equal to $(1 + \sqrt{5})/2 \alpha \rho$. We normalize the eigenvalues corresponding to $G_t$ by dividing them by $((1 + \sqrt{5})/2)^t$ for $1 \leq t \leq 5$ in Figure 3b.

Let $C_0$ be the circle in the complex plane of radius 1 centered at the origin. By applying the function $f(z) = (z \pm \sqrt{z^2 + 4(z+1)^2})/2$ iteratively $t$ times to the points of $C_0$, we obtain a curve $C_t$ in the complex plane. If we let $G_0$ be the directed $n$-cycle, then the eigenvalues of $G_t$ lie on $C_t$. In Figure 3c, we include $C_t$ and the eigenvalues of $G_t$ after normalization by dividing them by $((1 + \sqrt{5})/2)^t$, where $G_0$ is the directed 3-cycle. The curve $C_t$ was plotted after normalization for $t \leq 30$, and there was no noticeable difference between the time-steps $t = 15$ and $t = 30$. The structure after 30 iterations is provided in Figure 4. We suspect that as $t$ approaches infinity, the normalization of $C_t$ approaches a specific, limiting curve. As a result, the normalized mapping applied to the $n$th roots of unity would approach limiting points, and these limiting points can be calculated from the curve.

![Fig. 4: The mapping $f(z)$ applied to the complex unit circle over 30 iterations, after normalization.](image)

## 4 Directed cycles

As a consequence of Theorem 2, there are no directed 3-cycles in an ILDT digraph unless there is one present in the initial graph. We generalize
The ILDT Model for Social Networks

this property in the following result. Note that our directed cycles in the theorem are oriented, and so do not include directed 2-cycles. However, we are allowed to include bidirected edges (traversed in a single-direction) as part of our directed cycles.

Theorem 4. For all \( t \geq 0 \), the digraph \( G_t \) contains an oriented directed cycle if and only if \( G_{t+1} \) contains an oriented directed cycle.

Proof. The forward implication is immediate, so we focus on the reverse implication. Let \((a_1, \ldots, a_k)\) be a directed cycle of length \( k \) in \( G_{t+1} \). Define a function \( f : V(G_{t+1}) \to V(G_t) \) that maps a clone in \( G_{t+1} \) to its parent node, and acts as the identity mapping, otherwise. If \((a_i, a_{i+1})\) is an arc in \( G_{t+1} \), then \((f(a_i), f(a_{i+1}))\) is an arc in \( G_t \). The subgraph induced by the edges of the closed directed walk \((f(a_1), \ldots, f(a_k))\) has in-degree equal to out-degree at every node (counting the multiplicity of the edges in the walk) precisely because it is a closed directed walk. Each time we visit a node on the walk we contribute one to the in-degree and one to the out-degree. Hence, \((f(a_1), \ldots, f(a_k))\) decomposes into an edge-disjoint collection of directed cycles (none of which is a directed 2-cycle, by hypothesis). Hence, \( G_t \) contains a cycle. \( \square \)

We turn next to directed Hamiltonian cycles; that is, directed cycles visiting each node exactly once. Note that while we do not expect directed Hamiltonian cycles in real-world social networks, the emergence of this property in ILDT graphs is of graph-theoretical interest in its own right. For this, we first prove the following theorem on Hamiltonian paths in ILT undirected graphs. For a graph \( G \), we use the notation \( \text{ILT}_t(G) \) for the ILT graph resulting at time \( t \) if \( G_0 = G \); analogous notation is used for ILDT graphs. We use the notation \( G[S] \) for the subgraph induced by nodes \( S \) in \( G \).

In the following lemma with \( G_0 \) chosen as \( K_1 \), we label the node of \( G_0 \) as 0, and its unique child in \( G_1 \) as 1.

Lemma 2. Fix \( t \geq 1 \) and let \( G_t = \text{ILT}_t(K_1) \). For every clone \( v \in V(G_t) \), there is a Hamiltonian path in \( G_t \) from \( v \) to 0.

Proof. We use induction on \( t \geq 1 \). The base case \( t = 1 \) is straightforward, since \( G_1 \cong K_2 \). For the induction step, we assume \( t \geq 2 \). We label the node of \( G_0 \) as 0, and its unique child in \( G_1 \) as 1. Note that we can partition \( V(G_t) \) into \( V_0 \) and \( V_1 \) such that 0 \( \in \) \( V_0 \), 1 \( \in \) \( V_1 \) and \( G_t[V_0] \cong \).
\[ G_t[V_1] \cong G_{t-1} \] To see this, consider the ILDT process as starting with each of the nodes 0 and 1 independently. For \( i = 1, 2 \), the set \( V_i \) consists of all clones over subsequent time-steps starting with the initial vertex \( i \).

First suppose that \( v \in V_1 \), and choose an arbitrary \( w \in V_0 \). By the induction hypothesis, there exists a Hamiltonian path \( P_0 \) in \( G_t[V_0] \) with endpoints \( w \) and 0. Similarly, there exists a Hamiltonian path \( P_1 \) in \( G_t[V_1] \) with endpoints \( v \) and 1. Let \( P = P_1 + w + P_0 \); now \( P \) is the desired Hamiltonian path in \( G_t \) from \( v \) to 0.

Suppose instead that \( v \in V_0 \), and choose an arbitrary \( w \in V_1 \). By the induction hypothesis, there exists a Hamiltonian path \( P_0 \) in \( G_t[V_0] \) with endpoints \( v \) and 0. Similarly, there exists a Hamiltonian path \( P_1 \) in \( G_t[V_1] \) with endpoints \( w \) and 1. Let \( x \) be the neighbor of 0 on \( P_0 \). Let \( P = P_0 - x0 + x1 + P_1 + w0 \). Now \( P \) is the desired Hamiltonian path in \( G_t \) from \( v \) to 0. \(\square\)

\[ \text{Fig. 5: The two cases in the proof of Lemma 2, wavy lines represent paths.} \]

In a digraph \( D \), a closed spanning walk \( \mathcal{C} \) (that respects the orientations of \( D \)) is **nice** if for each edge \( v_i v_j \in E(\mathcal{C}) \) either (i) \( v_i v_j \) is the last edge departing \( v_i \) on \( \mathcal{C} \) or (ii) \( v_i v_j \) is the first edge entering \( v_j \) on \( \mathcal{C} \); possibly both (i) and (ii) hold for some edges of \( \mathcal{C} \). The **max frequency**, written \( s(\mathcal{C}) \), of a nice walk \( \mathcal{C} \) is the largest number of times that any node appears in \( \mathcal{C} \).

**Theorem 5.** If \( D \) is a digraph with a nice walk \( \mathcal{C} \) and \( t \in \mathbb{Z}^+ \) such that \( 2^{t-1} \geq s(\mathcal{C}) \), then \( D_t \) has a directed Hamiltonian cycle.

**Proof.** Let \( D_t = \text{ILDT}_t(D) \). We construct a Hamiltonian cycle in \( D_t \) by the algorithm below. We assume that the nodes of \( D \) are \( \{v_1, \ldots, v_n\} \).
(each $v_i$ may appear many times on $C$) and that the nodes of $D_t$ are partitioned into $V_1, \ldots, V_n$ (where $V_i$ consists of $v_i$ and all its descendants; that is, nodes that resulted by iterated cloning of $v_i$). Intuitively, we use $C$ to ensure that our walk $H$ visits each $V_i$ at least once and we use each $P_i$ to ensure that we visit all remaining vertices of $V_i$ the last time that $C$ visits $v_i$, in $D$.

**Initialization:** Pick an arbitrary node $v_i$ on $C$. Choose a clone $w \in V_i$. Start $H$ at $w$. Let $P_i$ be a Hamiltonian path in $V_i$ that starts at $w$ and ends at 0 (in $V_i$), by Lemma 2 (with vertex 0 defined as in Lemma 2).

Always $v_i, v_j$ refers to vertices of $D$ and vertices of $D_t$ are denoted by $w$, $x$, or 0.

**Iteration:** Assume that $H$ currently ends at some clone $w \in V_i$ (for some $i$) and path $P_i$ is defined, possibly from the initialization. If we have followed all edges of $C$, then halt and output $H$. Otherwise, let $e = v_i v_j$ denote the next edge of $C$. (1) If $e$ is the last edge leaving $v_i$ on $C$, then follow $P_i$ from $w$ to 0 in $V_i$; otherwise, move to the next node on $P_i$. (2) If $P_j$ is undefined (we have not yet visited $v_j$ on $C$), then follow an edge to an arbitrary clone $x$ in $V_j$. In this case, define $P_j$ to be a Hamiltonian path in $V_j$ with endpoints $x$ and 0; such a $P_j$ exists by Lemma 2. If $P_j$ is defined, then follow an edge from the current node of $H$ to the next node on $P_j$ (in $V_j$). (When $e$ is the final edge of $C$, we return to the node of $V_j$ where we started, which finishes $H$.)

This completes the algorithm for constructing $H$ from $C$. We prove its correctness in two steps. First, we show that the algorithm completes, then it constructs the desired Hamiltonian cycle $H$. Second, we show that the algorithm does indeed complete. Suppose the algorithm completes successfully. Since $C$ is a spanning walk, it visits every node $v_i \in V(D)$. Thus, $H$ visits every $V_i$. The final time that $H$ visits a $V_i$ it visits every remaining node on $P_i$. Thus, $H$ visits every node in $\bigcup V_i = V(D_t)$; that is, $H$ is spanning in $D_t$.

Now we must show that the algorithm completes successfully. Every time $H$ leaves a $V_i$ it does so from a non-clone, and every time $H$ returns to a $V_j$ it returns to a clone (that has not been visited before). The number of clones in each $V_j$ is $2^{t-1}$, so this is possible precisely because $s(C) \geq 2^{t-1}$. Now we need to check that $D_t$ has the necessary edges between $V_i$ and $V_j$. Since $C$ is nice, each edge $e = v_i v_j \in E(C)$ satisfies that either (i) $v_i v_j$ is the last edge leaving $v_i$ on $C$ or (ii) $v_i v_j$ is the first edge entering $v_j$ on $C$. In (i), $H$ leaves $V_i$ from node 0, which has edges
to every node of $V_j$. In (ii), any edge from a non-clone of $V_i$ to a clone of $V_j$ suffices, since we will define $P_j$ as starting from this clone of $V_j$ (and since $C$ has never before visited $v_j$).

The following result on the Hamiltonicity of the ILT model was first proven in [2], and we give an alternative proof as a corollary of Theorem 5.

**Corollary 2.** If $G$ is a connected undirected graph and $t = \log |V(G)|$, then ILT$_t(G)$ is Hamiltonian.

**Proof.** We form a digraph $D$ from $G$ by replacing each undirected edge $vw$ with arcs $(v,w)$ and $(w,v)$. We construct a nice spanning closed walk of $D$ and apply Theorem 5. Choose an arbitrary node $v \in D$ and form $C$ by recording each edge followed in a depth-first traversal of $D$ (including what we call back-tracking edges). Consider an edge $vw \in E(C)$. If $w$ has never been visited before, then $vw$ satisfies (ii) in the definition of nice. If $w$ has been visited before, then it is straightforward to check that $vw$ satisfies (i) in the definition (precisely because $C$ arose from a depth-first traversal of $D$).

We conjecture that for every strongly connected digraph $D$ there exists an integer $t$ such that ILT$_t(D)$ has a Hamiltonian cycle. In a sense, this conjecture is best possible, since if ILT$_t(D)$ is Hamiltonian for some $t$, then $D$ must be strongly connected. Namely, if there exist $v_i, v_j \in V(D)$ such that $D$ has no directed path from $v_i$ to $v_j$, then ILT$_t(D)$ has no directed path from $V_i$ to $V_j$, so ILT$_t(D)$ is not Hamiltonian. We suspect that this conjecture can be proved by somehow modifying the proof of Theorem 5.

5 Conclusion and further directions

We introduced and analyzed the Iterated Local Directed Transitivity (ILDT) model for social networks, motivated by status theory, transitivity in triads, and the ILT model in the undirected case [4]. We proved that the ILDT model, as in social networks, generates graphs that densify over time. A count of the directed, transitive, and bidirectional 3-cycles was given, and it was shown that the 3-transitive cycles count may be far more abundant by choice of the initial graph of the model. We studied
the eigenvalues of the adjacency matrices of ILDT graphs, with a discussion of the limiting distribution of eigenvalues of the directed 3-cycle. We concluded our results with an analysis of directed cycles in ILDT graphs and proved that in many instances of the initial graph, ILDT graphs have Hamiltonian cycles.

Given our limited space, we did not explore distance properties of the model, although we expect the model should generate small-world graphs, as is the case for ILT graphs. In the full version of the paper, it would be interesting to analyze the clustering coefficient, domination number, and degree distribution of ILDT graphs. The eigenvalues of ILDT graphs are worthy of further study, both in their limiting distribution in the complex plane and regarding their spectral expansion.

References

1. A. Bonato, *A Course on the Web Graph*, American Mathematical Society Graduate Studies Series in Mathematics, Providence, Rhode Island, 2008.
2. A. Bonato, H. Chuangpishit, S. English, B. Kay, E. Meger, The iterated local model for social networks, Accepted to *Discrete Applied Mathematics*.
3. A. Bonato, N. Hadi, P. Prałat, C. Wang, Dynamic models of on-line social networks, In: *Proceedings of WAW’09*, 2009.
4. A. Bonato, N. Hadi, P. Horn, P. Prałat, C. Wang, Models of on-line social networks, *Internet Mathematics* 6 (2011) 285–313.
5. A. Bonato, A. Tian, Complex networks and social networks, invited book chapter in: *Social Networks*, editor E. Kranakis, Springer, Mathematics in Industry series, 2011.
6. F.R.K. Chung, *Spectral Graph Theory*, American Mathematical Society, Providence, Rhode Island, 1997.
7. F.R.K. Chung, L. Lu, *Complex Graphs and Networks*, American Mathematical Society, Providence, Rhode Island, 2006.
8. F.R.K. Chung, G. Dewey, D.J. Galas, L. Lu, Duplication models for biological networks, *Journal of Computational Biology* 10 (2003) 677–688.
9. D. Easley, J. Kleinberg, *Networks, Crowds, and Markets Reasoning about a Highly Connected World*, Cambridge University Press, 2010.
10. R.V. Guha, R. Kumar, P. Raghavan, A. Tomkins, Propagation of trust and distrust, In: *Proceedings of WWW*, 2004.
11. F. Heider, *The Psychology of Interpersonal Relations*, John Wiley & Sons, 1958.
12. J. Huang, H. Shen, L. Hou, X. Cheng, Signed graph attention networks, In: *Proceedings of Artificial Neural Networks and Machine Learning ICANN*, 2019.
13. R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, E. Upfal, Stochastic models for the web graph, In: *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science*, 2000.
14. J. Leskovec, D.P. Huttenlocher, J.M. Kleinberg, Signed networks in social media, In: *Proceedings of the ACM SIGCHI*, 2010.
15. J. Leskovec, J. Kleinberg, C. Faloutsos, Graphs over time: densification Laws, shrinking diameters and possible explanations, In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2005.
16. J.P. Scott, Social Network Analysis: A Handbook, Sage Publications Ltd, London, 2000.
17. L. Small, O. Mason, Information diffusion on the iterated local transitivity model of online social networks, Discrete Applied Mathematics 161 (2013) 13381344.
18. D. Song, D.A. Meyer, A model of consistent node types in signed directed social networks, In: Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2014.
19. D.B. West, Introduction to Graph Theory, 2nd edition, Prentice Hall, 2001.