Growing Attributed Networks through Local Processes

ABSTRACT

This paper proposes an attributed network growth model. Despite the knowledge that individuals use limited resources to form connections to similar others, we lack an understanding of how local and resource-constrained mechanisms explain the emergence of rich structural properties found in real-world networks. We make three contributions. First, we propose a parsimonious and accurate model of attributed network growth that jointly explains the emergence of in-degree distribution, local clustering, clustering-degree relationship and attribute mixing patterns. Second, we make use of biased random walks to develop a model that forms edges locally, without recourse to global information. Third, we account for multiple sociological phenomena—bounded rationality; structural constraints; triadic closure; attribute homophily; preferential attachment. Our experiments show that the proposed Attributed Network Growth (ARW) model accurately preserves network structure and attribute mixing patterns of six real-world networks; it improves upon the performance of eight state-of-the-art models by a statistically significant margin of 2.5–10×.

1 INTRODUCTION

We present a network growth model that explains how distinct structural properties of attributed networks can emerge from local edge formation processes. In real-world networks, individuals tend to form edges under resource constraints such as limited information and partial network access. Additionally, phenomena such as triadic closure and homophily *simultaneously* influence individuals' decisions to form connections. Over time, these decisions cumulatively shape real-world networks to exhibit rich structural characteristics: heavy-tailed in-degree distribution, skewed local clustering and homophilic mixing patterns. However, we lack an understanding of local, resource constrained mechanisms that incorporate sociological factors to explain the emergence of these rich structural characteristics.

Classic models of network growth tend to make unrealistic assumptions about how individuals form edges. Consider a simple stylized example: the process of finding a set of papers to cite when writing an article. In preferential attachment [4] or fitness [6, 11, 51] based models, a node making m citations would pick papers from the *entire* network in proportion to their in-degree or fitness respectively. This process assumes that individuals possess complete knowledge of in-degree or fitness of every node in the network. An equivalent formulation—vertex copying [26]—induces preferential attachment: for every citation, a node would pick a paper uniformly at random from *all* papers, and either cite it or copy its citations. Notice that the copying mechanism assumes individuals have complete access to the network and forms each edge independently. Although these models explain the emergence of power law degree distributions, they are unrealistic: they require global knowledge (e.g., preferential attachment requires knowledge of the global in-degree distribution) or global access (e.g., vertex copying requires random access to all nodes). Moreover, these models do not account for the fact that many networks are attributed (e.g., a paper is published at a venue; a Facebook user may use gender, political interests to describe them) and that assortative mixing is an important network characteristic [38].

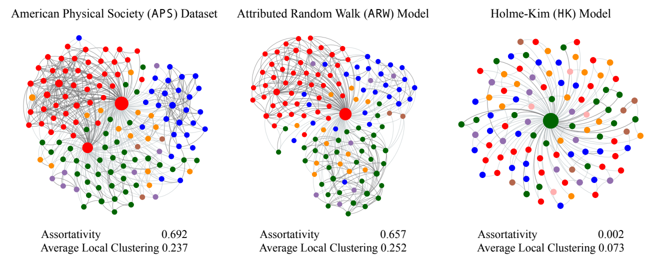

Figure 1: The figure shows how our proposed model of an Attributed Random Walk (ARW) accurately preserves local clustering and assortativity; we contrast with a non-attributed growth model [21] to underscore the importance of using attributes for network growth.

Recent papers tackle resource constraints [34, 52, 54] as well as nodal attributes [13, 18]. However, the former disregard attributes and the latter do not provide a realistic representation of edge formation under resource constraints. Furthermore, both sets of models do not jointly preserve multiple structural properties. Developing a parsimonious and accurate model of attributed network growth that accounts for observed sociological phenomena is non-trivial. Accurate network growth models are useful for synthesizing networks as well as to extrapolate existing real-world networks.

We propose an Attributed Random Walk (ARW) model that jointly explains the emergence of in-degree distributions, local clustering, clustering-degree relationship and attribute mixing patterns through a resource constrained mechanism based on random walks (see Figure 1). In particular, ARW relies entirely on local information to grow the network, without access to information of all nodes. In ARW, incoming nodes select a seed node based on attribute similarity and initiate a biased random walk: at each step of the walk, the incoming node either jumps back to its seed or chooses an outgoing link or incoming link to visit another node; it links to each visited node with some probability and halts after it has exhausted its budget to form connections. We have three primary contributions:

1. **Attributed:** We propose a parsimonious and accurate model of attributed network growth.
2. **Local information:** Our model is based on a random walk and uses local processes to form edges, without recourse to global information of the network.
3. **Unified account:** To the best of our knowledge, ARW is the first model that accounts for multiple sociological phenomena—bounded rationality; structural constraints; triadic closure; attribute homophily; preferential attachment—through an entirely local process to model global network structure and attribute mixing patterns.

ARW preserves key structural properties—in-degree distribution, clustering and indegree-clustering relationship—with high accuracy. Our experiments on six large-scale network datasets indicate that the proposed growth model outperforms eight state-of-the-art network growth models, including attributed growth models, by a statistically significant margin of 2.5–10×.

The rest of the paper is organized as follows. We begin by defining the problem statement in Section 2. In Section 3, we outline six network datasets, describe key structural properties of real-world networks and discuss insights from sociological studies. Then, in Section 4, we describe the network growth model. We follow by presenting experiments in Section 5, analysis of assortative mixing in Section 6 and discussion in Section 7. We conclude in Section 9.

2 PROBLEM STATEMENT

Consider an attributed directed network $G = (V, E, B)$, where V & E are sets of nodes & edges and each node has an attribute value $b \in B$. The goal is to develop a directed network growth model that preserves structural and attribute based properties observed in G. The growth model should be normative, accurate and parsimonious:

1. **Normative**: The model should account for normative behavior. In real-world networks, multiple sociological phenomena influence how individuals form edges under constraints of limited global information and partial network access.
2. **Accurate**: The model should preserve key structural and attribute based properties such as heavy tailed degree distribution, skewed local clustering, negatively correlated degree-clustering relationship and attribute mixing patterns.
3. **Parsimonious**: The model should have as few parameters as possible, but be expressive enough to generate networks with varying structural properties.

Next, we present extensive empirical analysis on real-world datasets to motivate our attributed random walk model.

3 EMPIRICAL ANALYSIS

In this section, we begin by describing six large-scale network datasets that we use in our analysis and experiments. Then, we describe key factors that impact edge formation and analyze global structural properties of real-world networks. Finally, we briefly discuss insights from empirical studies in sociology and common assumptions in network modeling.

3.1 Datasets

We consider six citation networks of different scales (size, time) from diverse sources: research articles, utility patents and judicial cases. We list the summary statistics and global network properties of these datasets in Table 1. Three of the six datasets are attributed networks; that is, each node has a categorical attribute value.

We focus on citation networks for two reasons. First, since nodes in citation networks form all outgoing edges to existing nodes at the time of joining the network, citation networks provide a clean basis to study edge formation mechanisms in attributed social networks. Second, citation network span long periods of time (e.g., the USSC judicial citation network span several hundred years). As a result, identifying local edge formation processes that accurately model growth for this duration is non-trivial. Next, we study the structural and content properties of these networks.

3.2 Global Network Properties

Compact statistical descriptors of global network properties [36] such as degree distribution, local clustering, and attribute assortativity quantify the extent to which local edge formation phenomena shape global network structure.

Heavy tailed degree distribution: Real-world networks tend to exhibit heavy tailed degree distributions. These distributions can emerge from the well-known preferential attachment process [4, 47], where incoming nodes connect with nodes in proportion to their degree. Over time, preferential attachment amplifies initial differences in node degree, giving rise to heavy tailed distributions. Log-normal fits, with parameters listed in Table 1, well describe the in-degree distribution of all network datasets, consistent Broido and Clauset's [10] observation that real-world networks with truly power law degree distributions are rare.

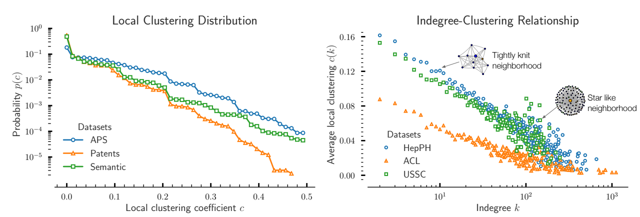

Figure 2: Local clustering in real-world networks have common characteristics: skewed local clustering distribution (left subplot) and a negatively correlated relationship between in-degree and average local clustering (right subplot).

Network	Description	$\lvert V\rvert$	$\lvert E\rvert$	T	A, $\lvert A\rvert$	LN (μ, σ)	DPL α	Avg. LCC	AA r
USSC [15]	U.S. Supreme Court cases	30,288	216,738	1754-2002	-	(1.19, 1.18)	2.32	0.12	-
HEP-PH [16]	ArXiv Physics manuscripts	34,546	421,533	1992-2002	-	(1.32, 1.41)	1.67	0.12	-
Semantic [3]	Academic Search Engine	7,706,506	59,079,055	1991-2016	-	(1.78, 0.96)	1.58	0.06	-
ACL [44]	NLP papers	18,665	115,311	1965-2016	VENUE, 50	(1.93, 1.38)	1.43	0.07	0.07
APS [1]	Physics journals	577,046	6,967,873	1893-2015	JOURNAL, 13	(1.62, 1.20)	1.26	0.11	0.44
Patents [29]	U.S. NBER patents	3,923,922	16,522,438	1975-1999	CATEGORY, 6	(1.10, 1.01)	1.94	0.04	0.72

Table 1: Summary statistics & global properties of six network datasets: $|V|$ nodes join the networks and form edges $|E|$ over time period T. In attributed networks, each node has a categorical attribute value that belongs to set A of size $|A|$. The networks exhibit lognormal (LN) in-degree distribution with mean μ and standard deviation σ, high average local clustering (LCC) & attribute assortativity (AA) coefficient and densify over time with power law (DPL) exponent α.

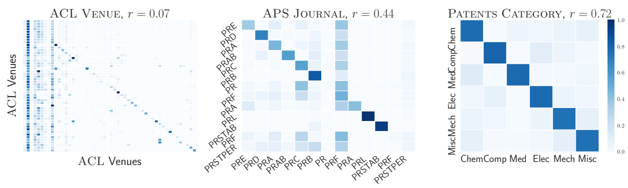

Figure 3: Attributed networks exhibit varying levels of homophily. The subplots illustrate the mixing patterns in ACL, APS and Patents w.r.t. attributes Venue ($r = 0.07$), Journal ($r = 0.44$) and Category ($r = 0.72$) respectively.

High Local Clustering: Real-world networks tend to exhibit high average local clustering (LCC), as shown in Table 1. Local clustering can arise from triadic closure [37, 46], where nodes with common neighbor(s) have an increased likelihood of forming a connection. The coefficient of node i equals the probability with which two randomly chosen neighbors of the node i are connected. In directed networks, the neighborhood of a node i can refer to the nodes that link to i, nodes that i links to or both. We define the neighborhood to be the set of all nodes that link to node i. In Figure 2, we show that (a) average local clustering is not a representative statistic of the skewed local clustering distributions and (b) real-world networks exhibit a negative correlation between node in-degree and local clustering. That is, low in-degree nodes have small, tightly knit neighborhoods and high in-degree nodes tend have large, star-shaped neighborhoods.

Homophily: Attributed networks tend to exhibit homophily [31], the phenomenon where similar nodes are more likely to be connected than dissimilar nodes. The assortativity coefficient [38] $r \in [-1, 1]$, quantifies the level of homophily in an attributed network and indicates the extent to which attribute similarity influences edge formation. Intuitively, assortativity compares the observed fraction of edges between nodes with the same attribute value to the expected fraction of edges between nodes with same attribute value if the edges were rewired randomly. In Figure 3, we show that attributed networks ACL, APS and Patents exhibit varying level of homophily with assortativity coefficient ranging from 0.07 to 0.72.

Increasing Out-degree over Time: The out-degree of nodes that join real-world networks tends to increase as functions of network size and time. This phenomenon densifies networks and can shrink effective diameter over time. Densification tends to exhibit a power law relationship [29] between the number of edges $e(t)$ and nodes $n(t)$ at time t: $e(t) \propto n(t)^{\alpha}$. Table 1 lists the densification power law (DPL) exponent α of the network datasets.

To summarize, citation networks tend to be homophilic networks that undergo accelerated network growth and exhibit regularities in structural properties: heavy tailed in-degree distribution, skewed local clustering distribution, negatively correlated degree-clustering relationship, and varying attribute mixing patterns.

3.3 Insights from Sociological Studies

Sociological studies on network formation seek to explain how individuals form edges in real-world networks.

Interplay of Triadic Closure and Homophily: Empirical studies [7, 25] that investigate the interplay between triadic closure and homophily in evolving networks indicate that *both* structural proximity and homophily are statistically significant factors that simultaneously influence edge formation. While homophilic preferences [31] induce edges between similar nodes, structural factors (e.g., network distance) act as constraints that restrict edge formation to structurally proximate nodes (e.g. friend of a friend).

Bounded Rationality: Extensive work [17, 30, 48] on individual decision making indicates that individuals are boundedly rational actors; constraints such as limited information, cognitive capacity and time impact decision making. This suggests that resource-constrained individuals that join networks are likely to employ simple rules to form edges using limited information and partial network access.

Current preferential attachment and fitness-based models [4, 14, 23, 49] make two assumptions that are at variance with these findings in the Social Sciences. First, by assuming that successive edge formations are independent, these models disregard the effect of triadic closure and structural proximity. Second, these models implicitly require incoming nodes to have complete network access (e.g., be able to connect to any node) or explicit knowledge of one or more properties (e.g., fitness, degree) of every node in the network. For example, a preferential attachment model, by making connections in proportion to degree, requires non-local information: the degree distribution of the entire network.

To summarize, insights from the Social Sciences suggest that individuals form edges under constraints of limited information and global network access. Moreover, resource-constrained edge formation, which tend to be biased towards nodes that are similar, proximate or well-connected, cumulatively modulates global properties of real-world networks.

Next, we propose a growth model that explains how local processes of edge formation can lead to the emergence of global structural and attribute properties observed in real-world networks.

4 ATTRIBUTED RANDOM WALK MODEL

We propose an Attributed Random Walk (ARW) model to explain the emergence of key structural properties of real-world networks through entirely local edge formation mechanisms.

Consider a stylized example of how a researcher might go about finding relevant papers to cite. First, the researcher broadly identifies one or more relevant papers, possibly with the help of external information sources (e.g. Google Scholar). These initial set of papers act as seed nodes. Then, acting under time and information constraints, she will examine papers that cite a seed paper, as well as those papers cited by the seed. Thus she navigates a chain of backward and forward references to identify similar papers relevant to addressing that research question in which she is interested. Next, through careful analysis, she will cite a subset of these papers.

ARW grows a directed network as new nodes join the network. The mechanism is motivated by the stylized example: an incoming node selects a seed node and initiates a random walk to explore the network by navigating through neighborhoods of existing nodes. It halts the random walk after connecting to a few visited nodes.

In this section, we describe the edge formation mechanisms underlying ARW and explain how ARW provides a unified treatment of the observations from empirical data as well as Social Science studies. Then, we discuss the methods required to fit ARW to data.

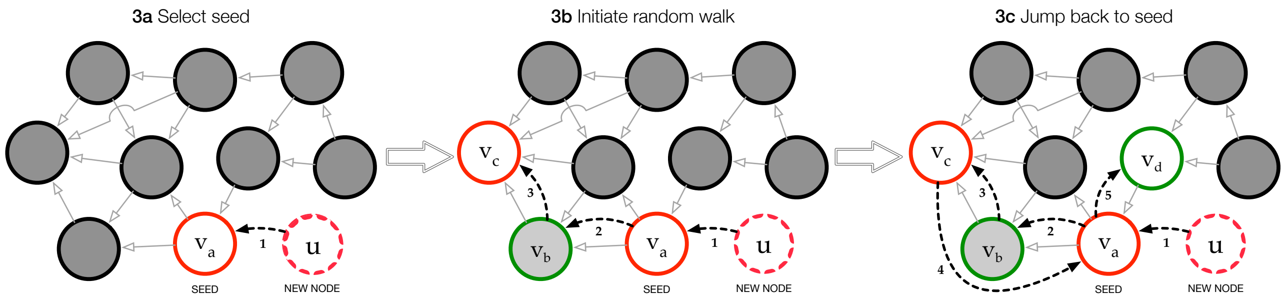

Figure 4: Edge formation in ARW: consider an incoming node u with outdegree $m = 3$ and attribute value $B(u)$ = RED ∈ {RED, GREEN}. In fig. 3a, u joins the network and selects seed v_a via SELECT-SEED. Then, in fig. 3b, u initiates a RANDOM-WALK and traverses from v_a to v_b to v_c. Finally, u jumps back to its seed v_a and restarts the walk, as shown in fig. 3c. Node u halts the random walk after linking to v_a, v_c & v_d.

4.1 Model Details

The Attributed Random Walk (ARW) model grows a directed network $\{\hat{G}_t\}_{t=1}^{T}$ in T time steps. More formally, at every discrete time step t, a new node u, with attribute value $B(u)$, joins the network $\hat{G}_t$. After joining the network, node u forms $m(t)$ edges to existing nodes.

The edge formation mechanism consists of two components: SELECT-SEED and RANDOM-WALK. As shown in Figure 4, an incoming node u with attribute value $B(u)$ that joins the network at time t first selects a seed node $S(u)$ using SELECT-SEED:

> SELECT-SEED
>
> (1) With probability $p_{\text{same}}/p_{\text{same}}+p_{\text{diff}}$, randomly select $S(u)$ from the set of existing nodes that have the same attribute value, $B(u)$.
>
> (2) Otherwise, with probability $p_{\text{diff}}/p_{\text{same}}+p_{\text{diff}}$, randomly select $S(u)$ from the set of existing nodes that do *not* have the same attribute value, $B(u)$.

SELECT-SEED accounts for homophilic preferences of incoming nodes using parameters p_{same} and p_{diff}, which incorporates attribute preferences of incoming nodes. As shown in Figure 4, after selecting the seed $S(u)$, node u initiates a random walk using RANDOM-WALK to form $m(t)$ links. The RANDOM-WALK mechanism consists of four parameters: attribute-based parameters p_{same} & p_{diff} model edge formation decisions and the jump parameter p_{jump} & out-link parameter p_{out} characterize random walk traversals:

> RANDOM-WALK
>
> (1) At each step of the walk, new node u visits node v_i.
> - If $B(u) = B(v_i)$, u links to v_i with probability p_{same}
> - Otherwise, u links to v_i with probability p_{diff}
>
> (2) Then, with probability p_{jump}, u jumps back to seed s_u.
>
> (3) Otherwise, with probability $1 - p_{\text{jump}}$, u continues to walk. It picks an outgoing edge with prob. p_{out} *or* an incoming edge with prob. $1 - p_{\text{out}}$ to visit a neighbor of v_i.
>
> (4) Steps 1-3 are repeated until u links to $m(t)$ nodes.

When attribute data is absent, ARW simplifies further, as a single link parameter p_{link} replaces both attribute-informed parameters p_{same} & p_{diff}. SELECT-SEED reduces to uniform seed node selection, whereas in RANDOM-WALK, the probability of linking to visited nodes simply equals p_{link}.

Notice that ARW has two exogenous parameters: the average outdegree $m(t)$ and attribute $B(u)$ of the node joining the network. The parameter $m(t)$ is similar to the parameter m in the classic Preferential-Attachment model [4], except that $m(t)$ is the mean-field value of out-degree m at time t in the observed network. While it is straightforward to model $m(t)$ endogenously by incorporating a densification power-law DPL exponent to ARW, we decided against it, since exogenous factors may also explain changes to $m(t)$. For example, conference venue and paper topic can influence the number of citations in a paper. Moreover, our analysis indicates that papers that join networks earlier tend to have fewer citations on average, perhaps explained by availability of fewer papers to cite. The attribute distribution $B(u)$ varies with time as new journals or venues crop up, necessitating an exogenous parameter.

Next, we explain how each parameter is necessary to conform to normative behavior of individuals in evolving networks.

4.2 ARW and Normative Behavior

The Attributed Random Walk model unifies multiple sociological phenomena into its edge formation mechanisms.

Phenomenon 1. *(Limited Resources) Individuals are boundedly rational [17, 30, 48] actors that form edges under constraints of limited information, partial network access and finite cognitive capacity.*

ARW uses a random walk to incorporate constraints of limited information and partial network access. A new node u selects a seed node from which it initiates a biased random walk. Then, u uses simple rules to connect to each visited nodes probabilistically and halts the walk after forming $m(t)$ edges, as shown in Figure 4. Random walks require information only about the 1-hop neighborhood of visited nodes, thus accounting for the constraints of limited information and partial network access.

Phenomenon 2. *(Structural Constraints) Structural factors such as network distance act as constraints that limit edge formation to proximate nodes. [25]*

We incorporate structural constraints into ARW using p_{jump}, the probability with which a new node jumps back to its seed node after each step of the random walk. This implies that the probability with which the new node is at most k steps from its seed node is $(1-p_{\text{jump}})^k$; as a result, the jump parameter p_{jump} controls the extent to which new nodes' random walks explore the network to form edges.

Phenomenon 3. *(Triadic Closure) Nodes with common neighbors have an increased likelihood of forming a connection. [46]*

When attribute data is absent, ARW controls the the effect of triadic closure on link formation using p_{link}. A new node u closes a triad through its random walk traversal by linking to a visited node and its neighbor with probability proportional to p_{link}^2. Similarly, in attributed networks, the probability of triad completion equals pq, where p and q can equal p_{same} or p_{diff}, depending on the attribute values of the incoming node and the visited nodes.

Phenomenon 4. *(Attribute Homophily) Nodes that have similar attributes are more likely to form a connection. [31]*

The attribute-biased link formation parameters p_{same} and p_{diff} effectively control global attribute assortativity. When $p_{\text{same}} > p_{\text{diff}}$, nodes are more likely to connect if they share the same attribute value, thereby resulting in a homophilic network over time. Similarly, $p_{\text{same}} < p_{\text{diff}}$ and $p_{\text{same}} = p_{\text{diff}}$ make edge formation heterophilic and attribute agnostic respectively.

Phenomenon 5. *(Preferential Attachment) Nodes tend to link to high degree nodes that have more visibility. [4]*

ARW does not rely on global degree distribution. In absence of this global information, ARW can control the degree of preferential attachment by adding structural bias to the random walk traversals by varying the outlink probability p_{out}. Random walks that traverse outgoing edges only (i.e., $p_{\text{out}} = 1$) eventually visit old nodes that tend to have high in-degree. Similarly, random walks that traverse incoming edges only (i.e., $p_{\text{out}} = 0$) visit recently joined nodes that tend to have low indegree. We use p_{out} to adjust the effect of preferential attachment on edge formation.

To summarize: ARW incorporates five well-known sociological phenomena— bounded rationality; structural constraints; triadic closure; attribute homophily; preferential attachment—into a single edge formation mechanism based on random walks.

4.3 Model Fitting

We now briefly describe methods to estimate model parameters, initialize $\hat{G}$, densify $\hat{G}$ over time and sample nodes' attribute values.

Parameter Estimation. The parameter estimation task consists of finding the set of parameters values for $(p_{\text{same}}, p_{\text{diff}}, p_{\text{jump}}, p_{\text{out}})$ that best explain the structural properties of an observed network G. We use a straightforward grid search method to estimate the four parameters. Other derivative-free optimization methods such as the Nelder-Mead [35] method can be used to speed-up parameter estimation. We describe the evaluation metrics and selection criteria in Subsection 5.1.

Initialization. The edge formation mechanism in ARW is sensitive to a large number of weakly connected components (WCCs) in the initial network $\hat{G}_0$ because incoming nodes can only form edges to nodes in the same WCC. To ensure that $\hat{G}_0$ is weakly connected, we perform an undirected breadth-first search on the observed, to-be-fitted network G that starts from the oldest node and terminates after visiting 0.1% of the nodes. The initial network $\hat{G}_0$ is the small WCC induced from the set of visited nodes.

Node Out-degree. The out-degree of incoming nodes increases nonlinearly over time in real-world networks. We coarsely reflect the rate of growth in the observed network G as follows. Each incoming node u that joins $\hat{G}$ at time t corresponds to some node that joins the observed network G in year $y(t)$; the number of edges $m(t)$ that u forms is equal to the average out-degree of nodes that join G in year $y(t)$.

Sampling Attribute Values. In real networks, the distribution over the set of attribute values $P_{\text{G}}(B)$ tends to change over time. The change in the attribute distribution over time is an exogenous factor and varies for every network. Consequently, we sample the attribute value $B(u)$ of node u, that joins $\hat{G}$ at time t, from $P_{\text{G}}(B \mid \text{year} = y(t))$, the observed attribute distribution conditioned on the year of arrival of node u.

To summarize, the ARW model intuitively describes how individuals form edges under resource constraints. ARW uses four parameters —$p_{\text{same}}, p_{\text{diff}}, p_{\text{jump}}, p_{\text{out}}$— to incorporate individuals' biases towards similar, proximate and high degree nodes. Next, our experiments in Section 5 show that ARW accurately preserves multiple structural and attribute properties of real networks

5 EXPERIMENTS

In this section, we evaluate the effectiveness of ARW in preserving structural properties of network datasets described in Subsection 3.1. relative to eight state-of-the-art growth models.

5.1 Setup

In this subsection, we describe the evaluation metrics used to quantify the extent to which the following growth models preserve global structural properties of real-world networks.

State-of-the-art Growth Models. We compare ARW to eight state-of-the-art growth models representative of the key edge formation mechanisms: preferential attachment, fitness, triangle closing and random walks. Two of the eight models account for attribute homophily and preserve attribute mixing patterns, as listed below:

(1) **Dorogovtsev-Mendes-Samukhin model** [14] (DMS) is a preferential attachment model in which the probability of linking to a node is proportional to the sum of its in-degree and "initial attractiveness."
(2) **Kim-Altmann model** [23] (KA) is a fitness-based model that defines fitness as the product of degree and attribute similarity. It can generate *attributed* networks with assortative mixing and heavy tailed degree distribution.
(3) **Relay Linking model** [49] (RL) propose a set of preferential attachment models that use relay linking to explain the change in node popularity over time. We use the iterated preferential relay-cite (IPRC) variant, which best fits real-world network properties.
(4) **Holme-Kim model** [21] (HK) is a preferential attachment model which uses a triangle-closing mechanism to generate scale-free, clustered networks.
(5) **Social Attribute Network model** [18] (SAN) generates scale-free, attributed networks with high clustering using attribute-augmented preferential attachment and triangle closing mechanisms.
(6) **Herera-Zufiria model** [45] (SK) is a random walk model that tunes the length of random walks to generate clustered networks with power law degree distributions.
(7) **Saramaki-Kaski** [20] (HZ) is a random walk model that generates scale-free networks with tunable average local clustering.

Significance level p < 0.001 p < 0.01

Group	Model	A: Indegree Distribution (KS Stat) USSC	HepPH	Semantic	ACL	APS	Patents	B: Local Clustering Distribution (KS Stat) USSC	HepPH	Semantic	ACL	APS	Patents	C: Indegree & Clustering Relationship (WRE) USSC	HepPH	Semantic	ACL	APS	Patents	Assortativity $\lvert r - \hat{r} \rvert < \epsilon$
Preferential attachment	DMS	0.03	0.03	0.05	0.09	0.04	0.02	0.80	0.82	0.56	0.63	0.83	0.50	1.00	1.00	1.00	1.00	1.00	1.00	✗
Preferential attachment	KA	0.11	0.19	0.22	0.26	0.13	0.06	0.80	0.82	0.56	0.63	0.82	0.50	1.00	1.00	1.00	1.00	1.00	1.00	✓
Preferential attachment	RL	0.12	0.12	0.17	0.15	0.07	0.15	0.79	0.82	0.56	0.62	0.83	0.50	0.99	1.00	1.00	0.99	1.00	1.00	✗
Triangle closing	HK	0.11	0.19	0.22	0.26	0.13	0.05	0.39	0.55	0.15	0.08	0.52	0.05	0.59	0.74	0.08	0.25	0.73	0.17	✗
Triangle closing	SAN	0.12	0.18	0.19	0.24	0.11	0.05	0.12	0.05	0.12	0.16	0.05	0.19	0.13	0.14	0.34	0.31	0.15	1.28	✓
Random walk	FF	0.16	0.17	0.14	0.12	0.46	0.32	0.53	0.54	0.33	0.69	0.19	0.40	1.64	1.74	0.54	4.11	0.15	0.73	✗
Random walk	SK	0.19	0.22	0.25	0.27	0.13	0.13	0.15	0.29	0.26	0.34	0.34	0.11	0.14	0.46	0.74	0.41	0.51	0.38	✗
Random walk	HZ	0.18	0.22	0.23	0.26	0.13	0.13	0.08	0.29	0.10	0.07	0.34	0.03	0.18	0.45	0.21	0.22	0.51	0.04	✗
	ARW	0.07	0.06	0.07	0.09	0.07	0.08	0.08	0.04	0.05	0.05	0.05	0.09	0.14	0.10	0.05	0.13	0.08	0.08	✓

Figure 5: Modeling network structure. We assess the extent to which network models fit key structural properties of six real-world networks. Tables 5A, 5B and 5C measure the accuracy of eight models in fitting the in-degree distribution, local clustering distribution, in-degree & clustering relationship respectively and global attribute assortativity. Existing models tend to underperform because they either disregard the effect of factors such as triadic closure and/or homophily or are unable to generate networks with varying structural properties. Our model, ARW, jointly preserves all three properties accurately and often performs considerably better than existing models: the cells are shaded gray or dark gray if the proposed model ARW performs better at significance level $\alpha = 0.01$ ( ) or $\alpha = 0.001$ ( ) respectively.

(8) **Forest Fire model** [29] (FF) is a recursive random walk model that preserves decreasing diameter over time, heavy-tailed degree distribution and high clustering.

Ensuring Fair Comparison. To ensure fair comparison, we modify existing models in three ways. First, for models that do not have an explicitly defined initial graph —DMS, SAN, KA— we use the initialization method used for ARW, described in Subsection 4.3. Second, we extend models that use constant node outdegree m, by increasing outdegree over time $m(t)$, using the method described in Subsection 4.3. Third, we adjust models that generate undirected networks to create directed edges and generate directed networks Harshay: TODO: parameter fitting of baselines, undirected to directed explain better

Evaluation. We evaluate the network model fit by comparing four key global network properties of G and $\hat{G}$: degree distribution, local clustering distribution, degree-clustering relationship and attribute assortativity. We use the Kolmogorov-Smirnov (KS) statistic to compare the univariate degree & local clustering distributions. We compare the bivariate degree-clustering relationship in G and $\hat{G}$ using Weighted Relative Error (WRE). The evaluation metric WRE aggregates the relative error between the average local clustering $c(k)$ and $\hat{c}(k)$ of nodes with in-degree k in G and $\hat{G}$ respectively; The weight of each relative error term equals the fraction of nodes with in-degree k in G.

Jointly preserving multiple structural properties is a multi-objective optimization problem; Model parameters that accurately preserve the degree distribution (i.e. low KS statistic) may not preserve the clustering distribution. Therefore, for each model, we use grid search to select the model parameters that minimizes the ℓ^2-norm of the evaluation metrics. We note that the sensitivity of the Forest Fire model requires a manually guided grid search method. Since the metrics have different scales, we normalize the metrics before computing the ℓ^2-norm to prevent unwanted bias towards any particular metric.

5.2 Results

Now, we evaluate the performance of ARW relative to eight well-known existing models on the datasets introduced in Subsection 3.1. Figure 5 tabulates the evaluation metrics for every pair of model and dataset. These metrics measure the accuracy with which the fitted models preserve key global network properties: degree distribution, local clustering distribution, and indegree-clustering relationship.

To evaluate the performance of network models, we first fit every model to each network dataset G as described in Subsection 5.1. Thereafter, we compare the structural properties of network dataset G and network $\hat{G}$ generated by the fitted model using evaluation metrics introduced in Subsection 5.1. We average out fluctuations in $\hat{G}$ over 100 runs.

We use one-sided permutation tests [19] to evaluate the relative performance of ARW. If ARW performs better than a model on a dataset with significance level $\alpha = 0.01$ or $\alpha = 0.001$, the corresponding cells in Figure 5 are shaded gray ( ) or dark gray ( ) respectively. We also group models that have similar edge formation mechanisms by color-coding the corresponding rows in Figure 5. We use green ticks in Figure 5 to annotate models that preserve assortativity up to two decimal places.

Figure 5 shows that existing models fail to jointly preserve multiple structural properties in an accurate manner. This is because existing models either disregard important mechanisms such as triadic closure and homophily or are not flexible enough to generate networks with varying structural properties.

Preferential attachment models: DMS, RL and KA preserve in-degree distributions but disregard clustering. DMS outperforms other models in accurately modeling degree distribution (Figure 5A) because its "initial attractiveness" parameter can be tuned to adjust preference towards low degree nodes. Unlike KA, however, DMS cannot preserve global assortativity. However, by assuming that successive edge formations are independent, both models disregard triadic closure and local clustering. (Figure 5B & Figure 5C).

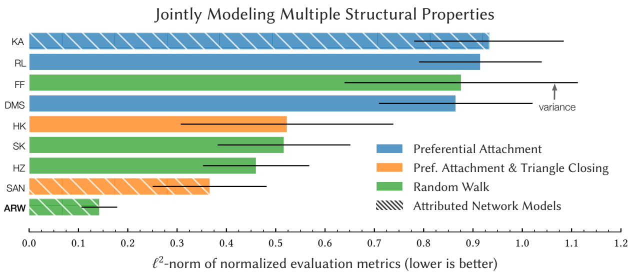

Figure 6: Jointly modeling multiple network properties: ARW outperforms existing network models in jointly preserving key structural properties—in-degree distribution, local clustering distribution and degree-clustering relationship— by a margin of 2.5x-10x.

Triangle Closing Models: HK and SAN are preferential attachment models that use triangle closing mechanisms to generate scale-free networks with high average local clustering. While triangle closing leads to considerable improvement over DMS and KA in modeling local clustering, HK and SAN are not flexible enough to preserve local clustering in all datasets (see Figure 5B & Figure 5C).

Existing random walk models: FF, SK, and HZ cannot accurately preserve structural properties of real-world network datasets. The recursive approach in FF considerably overestimates local clustering. because nodes perform a probabilistic breadth-first search and link to *all* visited/burned nodes. SK and HZ can control local clustering to some extent, as nodes perform a single random walk and link to each visited node with tunable probability μ. However, both models lack control over the in-degree distribution. Furthermore, existing random walk models disregard attribute homophily and do not account for attribute mixing patterns.

Attributed Random Walk model: The results in Figure 5 clearly indicate the effectiveness of ARW in jointly preserving multiple global network properties. ARW can generate networks with tunable in-degree distribution by adjusting nodes' bias towards high degree nodes using p_{out}. As a result, ARW accurately preserves in-degree distributions (Figure 5A), often significantly better than all models except DMS. Similarly, ARW matches the local clustering distribution (Figure 5B) and in-degree & clustering relationship (Figure 5C) with high accuracy using p_{jump} and p_{link}. Similarly, ARW preserves attribute assortativity using the attribute parameters p_{same} and p_{diff}. Barring one to two datasets, ARW preserves all three properties significantly better ($\alpha < 0.001$) than existing random walk models. Figure 6 shows that ARW improves upon the average ℓ^2-norm of the second best performing model SAN by a margin of approximately 2.5x.

To summarize, ARW unifies multiple sociological phenomena into a single mechanism to jointly preserve key structural properties of real-world networks.

6 MODELING LOCAL MIXING PATTERNS

The global assortativity coefficient quantifies the average propensity of links to occur between similar nodes. However, global assortativity is not a representative summary statistic of heterogeneous mixing patterns observed in large-scale networks [42]. Furthermore, it does not quantify anomalous mixing patterns and fails to measure how mixing varies across a network.

We use local assortativity [42] to measure varying mixing patterns in an attributed network $G = (V, E, B)$ with attribute values $B = \{b_1 \ldots b_l\}$. Unlike global assortativity that counts all edges between similar nodes, local assortativity of node i, $r_l(i)$, captures mixing pattern in the local neighborhood of node i by using a locality biased weight distribution w_i. The distribution w_i reweighs edges between similar nodes based on how local they are to node i. As Peel et al. [42] indicate, there are multiple ways to define node i's weight distribution w_i other than the prescribed personalized pagerank weight distribution, which is prohibitively expensive to compute for all nodes in large graphs. We define w_i as a uniform distribution over $N_1(i)$, the set of nodes that are at most 1 hop away from node i to allow for a highly efficient local assortativity calculation. More formally, the local assortativity coefficient $r_l(i)$ of node i, with outdegree $m(i)$ and attribute value $b(i)$ is defined as follows:

$$r_l(i) = \frac{\overbrace{\frac{1}{|N(i)|} \sum_{j \in N(i)}^{m(j)>0} \sum_{k \in V} \frac{\mathcal{I}\{(j,k) \in E \wedge b(j) = b(k)\}}{m(i)}}^{\texttt{observed}} - \overbrace{\sum_{b \in B} e_b \cdot e_b}^{\texttt{random}}}{\underbrace{1}_{\texttt{max(observed)}} - \underbrace{\sum_{b \in B} e_b \cdot e_b}_{\texttt{random}}}$$

Intuitively, $r_l(i)$ compares the observed fraction of edges between similar nodes in the local neighborhood of node i (observed) to the expected fraction if the edges are randomly rewired (random).

As shown in Figure 7, local assortativity distributions of ACL, APS and Patents reveal anomalous, skewed and heterophilic local mixing patterns that are not inferred via global assortativity. Our

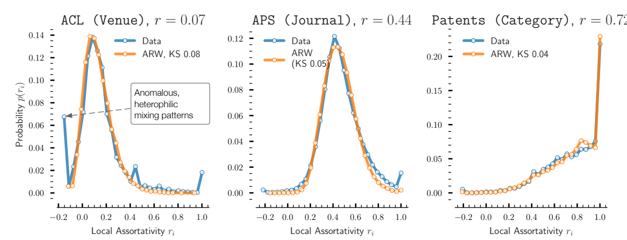

Figure 7: Local assortativity distributions of attributed networks ACL, APS and Patents reveal anomalous, skewed and heterophilic local mixing patterns. ARW accurately preserves local assortativity, but does not account for anomalous mixing patterns.

model ARW can preserve diverse local assortativity distributions with high accuracy even though nodes share the same attribute parameters p_{same} and p_{diff}. This is because, in addition to sampling attributes conditioned on time, ARW incorporates multiple sources of stochasticity through its edge formation mechanism. As a result, incoming nodes with fixed homophilic preferences can position themselves in neighborhoods with variable local assortativity by (a) selecting a seed node in a region with too few (or too many) similar nodes or (b) exhausting all its links before visiting similar (or dissimilar) nodes. We note that ARW is not expressive enough to model anomalous mixing patterns; richer mechanisms such as sampling p_{same} or p_{diff} from a mixture of Bernoullis are necessary to account for anomalous mixing patterns.

7 DISCUSSION

In this section, we discuss a major weakness of triangle closing mechanisms and limitations & potential modifications of our model ARW.

7.1 Dissecting the Triangle Closing Mechanism

A set of network models (e.g., SAN [18] & HK [21]) use triangle closing mechanisms to generate networks with varying average local clustering. However, our experimental results in Subsection 5.2 show that models that rely on triangle closing cannot model local clustering distribution or bivariate degree-clustering relationship accurately. To understand why, we examine the degree-clustering relationship in the APS network, in Figure 8.

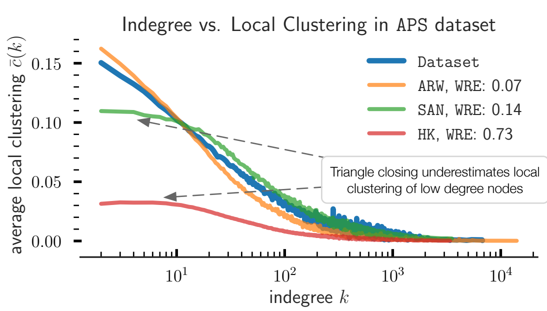

Figure 8: Triangle closing mechanisms used in SAN HK fail to model average local clustering of low in-degree nodes. In contrast, to accurately preserve local clustering, ARW uses random walks to visit low in-degree nodes and close triangles in their neighborhoods .

Figure 8 reveals that models based on triangle closing mechanisms, SAN and HK, considerably underestimate the local clustering of nodes that have low in-degree. This is because incoming nodes in SAN and HK tend to close triangles in the neighborhood of high in-degree nodes to which they connect via preferential attachment. Local clustering plateaus as in-degree decreases because triangle closing along with preferential attachment fail to form connections in neighborhoods of low in-degree nodes. In contrast, ARW accurately models the degree-clustering relationship because incoming nodes initiate random walks and close triangles in neighborhoods of low in-degree seed nodes chosen via Select-Seed.

7.2 ARW Limitations

We discuss three limitations of our model ARW. First, we only consider citation network datasets, as nodes (i.e., papers) form all edges at the time of joining. This allows us to analyze edge formation in the absence of confounding edge processes such as edge deletion and edge creation between existing nodes. We plan to extend ARW to handle social networks, where individuals can form edges at any time. One way is to incorporate random walks that pause and resume intermittently, thus allowing for older nodes to connect with more recent arrivals. Similarly, meta-path based random walks could potentially model interactions between nodes of different types in heterogeneous information networks. Second, the out-degree $m(t)$ of incoming nodes in ARW rely on the observed, temporal out-degree sequence that might be unavailable in network datasets without fine-grained temporal information. In this case, ARW can be adapted to rely on the prescribed range of densification power law exponent [29] and observed, average outdegree. Third, ARW does not account for mixing patterns of numerical nodal attributes such as publication year and age; we plan to conduct additional analysis on mixing patterns of numerical attributes in real-world networks to extend ARW in this direction.

8 RELATED WORK

Network growth models seek to explain a subset of structural properties observed in real networks We point the reader to extensive surveys [2, 39] of network growth models for more information. Note that, unlike growth models, statistical models of network replication [28, 43] do not model how networks grow over time and are not relevant to our work. Below, we discuss relevant and recent work on modeling network growth.

Preferential Attachment & Fitness: In preferential attachment and fitness-based models [5, 6, 11, 32], a new node u links to an existing node v with probability proportional to the attachment function $f(k_v)$, a function of either degree k_v or fitness ϕ_v of node v. For instance, linear preferential attachment functions [4, 14, 26] lead to power law degree distributions and small diameter [9] and attachment functions of degree & node age [51] can preserve realistic temporal dynamics. Extensions of preferential attachment [34, 52, 54] that incorporate resource constraints disregard network properties other than power law degree distribution and small diameter. Additional mechanisms are necessary to explain network properties such as clustering and attribute mixing patterns.

Triangle Closing: A set of models [21, 24, 27] incorporate triadic closure using triangle closing mechanisms, which increase average local clustering by forming edges between nodes with one or more common neighbors. However, as explained in Subsection 7.1, models based on preferential attachment and triangle closing do not preserve the local clustering of low degree nodes.

Attributed network models: Attribute network growth models [13, 18, 22, 55] account for the effect of attribute homophily on edge formation and preserve mixing patterns. Existing models can be broadly categorized as (a) fitness-based model that define fitness as a function of attribute similarity and (b) microscopic models of network evolution that require complete temporal information about edge arrivals & deletion. Our experiment results in Subsection 5.2 show that well-known attributed network models SAN and KA preserve assortative mixing patterns, degree distribution to some extent, but not local clustering and degree-clustering correlation.

Random walk models: First introduced by Vazquez [50], random walk models are inherently local. Models [8] in which new nodes only link to terminal nodes of short random walks generate networks with power law degree distributions [12] and small diameter [33] but do not preserve clustering. Models such as SK [45] and HZ [20], in which new nodes probabilistically link to each visited nodes incorporate triadic closure but are not flexible enough to preserve skewed local clustering of real-world networks, as shown in Subsection 5.2. We also observe that recursive random walk models such as FF [29] preserve temporal properties such as shrinking diameter but considerably overestimate local clustering and degree-clustering relationship of real-world networks. Furthermore, existing random walk models disregard the effect of homophily and do not model attribute mixing patterns.

Recent Work: Pálovics et al. [40] use preferential and uniform attachment to model the decreasing power law exponent of real-world, undirected networks in which average degree increases over time. Singh et al. [49] (RL) augment preferential attachment to explain the shift in popularity of nodes over time via the concept of relay linking. Both models do not incorporate mechanisms to preserve clustering, attribute mixing patterns, and resource constraints that affect how individuals form edges in real-world networks.

To summarize, existing models do not explain how resource constrained and local processes *jointly* preserve multiple global network properties of attributed networks.

9 CONCLUSION

In this paper, we proposed a novel, parsimonious model of attributed network growth. ARW grows a directed network in the following manner: an incoming node selects a seed node based on attribute similarity, initiates a biased random walk to explore the network by navigating through neighborhoods of existing nodes, and halts the random walk after connecting to a few visited nodes. To the best of our knowledge, ARW is the first model that unifies multiple sociological phenomena—bounded rationality; structural constraints; triadic closure; attribute homophily; preferential attachment—into a single local process to model global network structure *and* attribute mixing patterns. Our experiments on six large-scale citation networks show that ARW outperforms relevant and recent existing models by a statistically significant factor of 2.5–10×.

We plan to extend the ARW model in three ways: understanding the emergence of higher-order clustering [53] through local processes, modeling the effect of homophily on the formation of temporal motifs [41] and extending ARW to model undirected, social networks.

REFERENCES

[1] [n. d.]. APS Datasets for Research. ([n. d.]). https://journals.aps.org/datasets

[2] Réka Albert and Albert-László Barabási. 2002. Statistical mechanics of complex networks. *Reviews of modern physics* 74, 1 (2002), 47.

[3] Waleed Ammar, Dirk Groeneveld, Chandra Bhagavatula, Iz Beltagy, Miles Crawford, Doug Downey, Jason Dunkelberger, Ahmed Elgohary, Sergey Feldman, Vu Ha, Rodney Kinney, Sebastian Kohlmeier, Kyle Lo, Tyler Murray, Hsu-Han Ooi, Matthew Peters, Joanna Power, Sam Skjonsberg, Lucy Lu Wang, Chris Wilhelm, Zheng Yuan, Madeleine van Zuylen, and Oren Etzioni. 2018. Construction of the Literature Graph in Semantic Scholar. In *NAACL*.

[4] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512.

[5] Michael Bell, Supun Perera, Mahendrarajah Piraveenan, Michiel Bliemer, Tanya Latty, and Chris Reid. 2017. Network growth models: A behavioural basis for attachment proportional to fitness. *Scientific Reports* 7 (2017), 42431.

[6] Ginestra Bianconi and Albert-László Barabási. 2001. Bose-Einstein condensation in complex networks. *Physical review letters* 86, 24 (2001), 5632.

[7] Per Block and Thomas Grund. 2014. Multidimensional homophily in friendship networks. *Network Science* 2, 2 (2014), 189–212.

[8] Avrim Blum, TH Hubert Chan, and Mugizi Robert Rwebangira. 2006. A random-surfer web-graph model. In *2006 Proceedings of the Third Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*. SIAM, 238–246.

[9] Béla Bollobás and Oliver Riordan. 2004. The diameter of a scale-free random graph. *Combinatorica* 24, 1 (2004), 5–34.

[10] Anna D Broido and Aaron Clauset. 2018. Scale-free networks are rare. *arXiv preprint arXiv:1801.03400* (2018).

[11] Guido Caldarelli, Andrea Capocci, Paolo De Los Rios, and Miguel A Munoz. 2002. Scale-free networks from varying vertex intrinsic fitness. *Physical review letters* 89, 25 (2002), 258702.

[12] Prasad Chebolu and Páll Melsted. 2008. PageRank and the random surfer model. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 1010–1018.

[13] Maurício L de Almeida, Gabriel A Mendes, G Madras Viswanathan, and Luciano R da Silva. 2013. Scale-free homophilic network. *The European Physical Journal B* 86, 2 (2013), 38.

[14] Sergey N Dorogovtsev, Jose Ferreira F Mendes, and Alexander N Samukhin. 2000. Structure of Growing Networks: Exact Solution of the Barabási–Albert's Model. *arXiv preprint cond-mat/0004434* (2000).

[15] James H Fowler and Sangick Jeon. 2008. The authority of Supreme Court precedent. *Social networks* 30, 1 (2008), 16–30.

[16] Johannes Gehrke, Paul Ginsparg, and Jon Kleinberg. 2003. Overview of the 2003 KDD Cup. *ACM SIGKDD Explorations Newsletter* 5, 2 (2003), 149–151.

[17] Gerd Gigerenzer and Daniel G Goldstein. 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological review* 103, 4 (1996), 650.

[18] Neil Zhenqiang Gong, Wenchang Xu, Ling Huang, Prateek Mittal, Emil Stefanov, Vyas Sekar, and Dawn Song. 2012. Evolution of social-attribute networks: measurements, modeling, and implications using google+. In *Proceedings of the 2012 Internet Measurement Conference*. ACM, 131–144.

[19] Phillip Good. 2013. *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media.

[20] Carlos Herrera and Pedro J Zufiria. 2011. Generating scale-free networks with adjustable clustering coefficient via random walks. In *Network Science Workshop (NSW), 2011 IEEE*. IEEE, 167–172.

[21] Petter Holme and Beom Jun Kim. 2002. Growing scale-free networks with tunable clustering. *Physical review E* 65, 2 (2002), 026107.

[22] Fariba Karimi, Mathieu Génois, Claudia Wagner, Philipp Singer, and Markus Strohmaier. 2017. Visibility of minorities in social networks. *arXiv preprint arXiv:1702.00150* (2017).

[23] Kibae Kim and Jörn Altmann. 2017. Effect of homophily on network formation. *Communications in Nonlinear Science and Numerical Simulation* 44 (2017), 482–494.

[24] Konstantin Klemm and Victor M Eguiluz. 2002. Highly clustered scale-free networks. *Physical Review E* 65, 3 (2002), 036123.

[25] Gueorgi Kossinets and Duncan J. Watts. 2009. Origins of Homophily in an Evolving Social Network. *Amer. J. Sociology* 115 (2009), 405–450. http://www.journals.uchicago.edu/doi/abs/10.1086/599247

[26] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D Sivakumar, Andrew Tomkins, and Eli Upfal. 2000. Stochastic models for the web graph. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*. IEEE, 57–65.

[27] Jure Leskovec, Lars Backstrom, Ravi Kumar, and Andrew Tomkins. 2008. Microscopic evolution of social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 462–470.

[28] Jure Leskovec, Deepayan Chakrabarti, Jon Kleinberg, Christos Faloutsos, and Zoubin Ghahramani. 2010. Kronecker graphs: An approach to modeling networks. *Journal of Machine Learning Research* 11, Feb (2010), 985–1042.

[29] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. 2005. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 177–187.

[30] Barton L Lipman. 1995. Information processing and bounded rationality: a survey. *Canadian Journal of Economics* (1995), 42–67.

[31] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.

[32] Matúš Medo, Giulio Cimini, and Stanislao Gualdi. 2011. Temporal effects in the growth of networks. *Physical review letters* 107, 23 (2011), 238701.

[33] Abbas Mehrabian and Nick Wormald. 2016. Itś a small world for random surfers. *Algorithmica* 76, 2 (2016), 344–380.

[34] Stefano Mossa, Marc Barthelemy, H Eugene Stanley, and Luis A Nunes Amaral. 2002. Truncation of power law behavior in scale-free network models due to information filtering. *Physical Review Letters* 88, 13 (2002), 138701.

[35] John A Nelder and Roger Mead. 1965. A simplex method for function minimization. *The computer journal* 7, 4 (1965), 308–313.

[36] Mark Newman. 2010. *Networks: an introduction*. Oxford university press.

[37] Mark EJ Newman. 2001. Clustering and preferential attachment in growing networks. *Physical review E* 64, 2 (2001), 025102.

[38] Mark EJ Newman. 2002. Assortative mixing in networks. *Physical review letters* 89, 20 (2002), 208701.

[39] Mark EJ Newman. 2003. The structure and function of complex networks. *SIAM review* 45, 2 (2003), 167–256.

[40] Róbert Pálovics and András A Benczúr. 2017. Raising graphs from randomness to reveal information networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM, 23–32.

[41] Ashwin Paranjape, Austin R Benson, and Jure Leskovec. 2017. Motifs in temporal networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM, 601–610.

[42] Leto Peel, Jean-Charles Delvenne, and Renaud Lambiotte. 2018. Multiscale mixing patterns in networks. *Proceedings of the National Academy of Sciences* 115, 16 (2018), 4057–4062.

[43] Joseph J Pfeiffer III, Sebastian Moreno, Timothy La Fond, Jennifer Neville, and Brian Gallagher. 2014. Attributed graph models: Modeling network structure with correlated attributes. In *Proceedings of the 23rd international conference on World wide web*. ACM, 831–842.

[44] Dragomir R. Radev, Pradeep Muthukrishnan, Vahed Qazvinian, and Amjad Abu-Jbara. 2013. The ACL anthology network corpus. *Language Resources and Evaluation* (2013), 1–26. https://doi.org/10.1007/s10579-012-9211-2

[45] Jari Saramäki and Kimmo Kaski. 2004. Scale-free networks generated by random walkers. *Physica A: Statistical Mechanics and its Applications* 341 (2004), 80–86.

[46] Georg Simmel. 1950. *The sociology of georg simmel*. Vol. 92892. Simon and Schuster.

[47] Herbert A Simon. 1955. On a class of skew distribution functions. *Biometrika* 42, 3/4 (1955), 425–440.

[48] Herbert A Simon. 1972. Theories of bounded rationality. *Decision and organization* 1, 1 (1972), 161–176.

[49] Mayank Singh, Rajdeep Sarkar, Pawan Goyal, Animesh Mukherjee, and Soumen Chakrabarti. 2017. Relay-linking models for prominence and obsolescence in evolving networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1077–1086.

[50] Alexei Vazquez. 2000. Knowing a network by walking on it: emergence of scaling. *arXiv preprint cond-mat/0006132* (2000).

[51] Dashun Wang, Chaoming Song, and Albert-László Barabási. 2013. Quantifying long-term scientific impact. *Science* 342, 6154 (2013), 127–132.

[52] Li-Na Wang, Jin-Li Guo, Han-Xin Yang, and Tao Zhou. 2009. Local preferential attachment model for hierarchical networks. *Physica A: Statistical Mechanics and its Applications* 388, 8 (2009), 1713–1720.

[53] Hao Yin, Austin R Benson, and Jure Leskovec. 2018. Higher-order clustering in networks. *Physical Review E* 97, 5 (2018), 052306.

[54] Jianyang Zeng, Wen-Jing Hsu, and Suiping Zhou. 2005. Construction of scale-free networks with partial information. *Lecture notes in computer science* 3595 (2005), 146.

[55] Elena Zheleva, Hossam Sharara, and Lise Getoor. 2009. Co-evolution of social and affiliation networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1007–1016.