Modeling the Growth of Attributed Networks

ABSTRACT

We propose a network growth model based on local processes that jointly explains the emergence of key structural properties of real-world attributed directed networks: heavy-tailed indegree distribution, attribute mixing patterns, high local clustering and degree-clustering correlation. In real-world networks, individuals form edges under constraints of limited information and partial network access. However, well-known growth models that preserve multiple structural properties do not incorporate these resource constraints. Conversely, resource constrained growth models cannot jointly preserve multiple structural properties of real networks. Furthermore, most growth models disregard the effect of homophily on edge formation and global network structure.

Our Attributed Random Walk (ARW) model explains how structural & content-based properties of real-world networks jointly arise from individual preferences & edge formation under constraints of limited information and network access. In our model, each node that joins the network selects a seed node from which it initiates a biased random walk to concurrently explore the network and link to existing nodes. Our experimental results against eight well-known growth models indicate significant improvement (2.5-10x) in accurately preserving global structural properties and attribute mixing patterns of six large scale real-world networks.

KEYWORDS

Network evolution, Network growth models, Attributed networks, Homophily

1 INTRODUCTION

We develop a resource constrained model of network growth that explains the emergence of key structural properties. The problem is important for several reasons. Individuals form real-world networks acting under resource constraints and while using local information. These networks that individuals form exhibit rich structural properties. However, we lack an understanding of mechanisms that are resource constrained and which use local information explain the emergence of these structural related properties.

Classic models of network growth, make unrealistic assumptions about what agents who form edges do. Consider as a simple stylized example, the process of finding the a set of papers to cite when writing an article. In the preferential attachment model [3] of network growth, a node making m citations would pick a paper uniformly at random from *all* papers in the domain, and either cite it or copy one of its references. We would repeat this process, till we've exhausted our budget of m references. Notice that the process assumes access to the entire dataset, and that one would pick papers uniformly at random. An equivalent formulation of this copying model is to cite papers from the entire dataset in proportion to their in degree. The latter formulation assumes that agent making citations know the entire in-degree distribution. While preferential attachment models explains the emergence of the power-law degree distribution, the attachment model is an unrealistic representation of how agents make decisions on edge formation.

The problem of developing a model of network growth, where agents act under resource constraints, including access to only local information is hard. The problem lies in identifying simple mechanisms, with few parameters, where the agents only use local information and *jointly* preserve the properties related structure.

We propose a random walk based model of network growth that jointly explains the emergence of the following properties: heavy-tailed in-degree distributions, local clustering and clustering-degree relationships. In the growth model, an incoming node picks a recent node as the seed. It will link to this node with some constant linking probability. Then, it follows the outgoing link or the incoming link of this seed node and arrives at a new node. At each new node, it decides to link with the same constant linking probability. Then it has to decide whether to jump back to the seed node, or following incoming or outgoing links. The process repeats until the agent has exhausted its budget for linking. To summarize, new nodes concurrently acquire information and form edges by exploring the local neighborhoods of existing nodes, without access to the entire network.

Our main contributions are as follows:

- We propose a model of network growth using a local edge formation mechanism that incoporates the resource constraints that influence individuals' edge formation mechanisms in real-world networks.
- We propose a model that jointly explains multiple structural properties, including in-degree distribution, clustering, degree clustering relationship and edge densification.

We conducted extensive experimental results, against state of the art baselines, on large citation network datasets. We show that our growth model outperforms that best competing model in jointly and accurately preserving multiple structural properties—degree distribution, clustering and degree-clustering relationship—by a significant margin.

The rest of the paper is organized as follows. In Section 8, we describe the related work. Then, in **??**, we define key structural properties and introduce the datasets. We formally state the goal of the paper in Section 2. In **??** and Section 4, we report prominent structural characteristics of citation networks and propose a network growth model respectively. This is followed by Section 5, where we validate our model against multiple baselines.

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 under 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 be simple but expressive enough to generate networks with varying structural properties.

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 large-scale citation networks from diverse sources: research articles, utility patents and judicial cases. We focus on citation networks for three reasons. First, nodes in citation networks form all outgoing edges to existing nodes at the time of joining the network; Nodes do not form or delete edges at a later time. This allows us to analyze the edge formation mechanisms of new nodes that join the network form edges. Second, citation network datasets include the time (e.g., publication year of academic papers) at which nodes join the network. As a result, local edge formation processes and global structural properties can be better understood by studying network snapshots at different stages of the growth process. Third, the citation networks are large networks that tend to have one or more nodal attributes (e.g. category of patents) and span multiple decades. As a result, the structural and content properties of the citation networks considered are well-defined.

Now, we briefly describe the datasets considered in this paper. Three of the six network datasets have nodal attribute data; That is, each node has a categorical attribute value. Table 1 provides summary statistics of the following networks:

(1) **Association of Computational Linguistics** (`ACL`) [40] is an attributed academic citation network that consists of papers published in ACL conferences, journals and workshops. The attribute value of each paper is the name of the venue where it was published.

Network	$\|V\|$	$\|E\|$	T	A	$\|A\|$
`USSC`	30,288	216,738	1754-2002	-	-
`HEP-PH`	34,546	421,533	1992-2002	-	-
`Semantic`	7,706,506	59,079,055	1991-2016	-	-
`ACL`	18,665	115,311	1965-2016	VENUE	50
`APS`	577,046	6,967,873	1893-2015	JOURNAL	13
`Patents`	3,923,922	16,522,438	1975-1999	CATEGORY	6

Table 1: Summary statistics of six network datasets: number of nodes $|V|$ and edges $|E|$, time period T, categorical attribute A and number of attribute values $|A|$.

(2) **U.S. Supreme Court Cases** (`USSC`) [14] is a judicial citation network of U.S. Supreme Court cases. There is an edge from case i to case j if and only if case i cites case j in its majority opinion.

(3) **ArXiv HEP-PH** (`HEP-PH`) [15] is an academic citation network of HEP-PH (high energy physics phenomenology) papers in the ArXiv e-print.

(4) **APS Journals** (`APS`) [1] is an attributed academic citation network maintained by the American Physical Society (`APS`). The attribute value of each paper is the `APS` journal in which it was published.

(5) **U.S Utility Patents** (`Patents`) [26] is an attributed citation network of U.S. utility patents maintained by the National Bureau of Economic Research (NBER). The attribute value of each patent is an NBER patent category.

(6) **Semantic Scholar** (`Semantic`) [2] is an academic citation network of Computer Science and Neuroscience papers, released in June 2017 by Semantic Scholar.

Next, we study the structural and content properties of these networks in subsection 3.2 and empirically validate the effectiveness of the proposed model using these network datasets in Section 5.

3.2 Observations from Network Data

Compact statistical descriptors of global network properties [33] such as degree distribution, local clustering and attribute assortativity quantify the extent to which local edge formation phenomena shape global network structure.

Preferential Attachment & Degree Distribution In the preferential attachment process [3, 43], nodes with higher degree receive links at a faster rate because incoming nodes tend to link to well-connected nodes that have more visibility. As a result, initial differences in node connectivity get reinforced over time, giving rise to a rich-get-richer effect. This phenomenon explains why the citation networks datasets exhibit heavy tailed degree distributions; It also implies that most papers receive zero or a few citations, but a small but significant percent of the nodes turn into popular hubs that receive many citations. Log-normal fits describe the indegree distribution of all network datasets, well consistent with Broido & Clauset's [9] observation that real-world networks with truly power law degree distributions are rare; The parameters of the log-normal fits are listed in table 2. Our model explains the emergence of heavy tailed indegree distributions through a *local* process that adjusts bias towards linking to well-connected nodes

Triadic Closure & Clustering In the triadic closure phenomenon [34, 42], nodes with common neighbor(s) have an increased likelihood of forming a connection. The local clustering coefficient of a node measures the prevalence of triadic closure in its neighborhood; It is the probability that two randomly chosen neighbors of the node i are connected. In directed networks, the neighborhood of a node i can refer to the set of nodes that link to i, set of nodes that i links to or the union of both sets. We define the neighborhood to be the set of all nodes that link to node i. Real-world networks tend to exhibit high average local clustering, as shown in table 2. However, average local clustering is not a representative statistic of the *skewed* local clustering distributions shown in Figure 1.

[1] https://journals.aps.org/datasets

Network Dataset	LN (μ, σ)	DPL α	Avg. LCC	AA r
USSC	(1.19, 1.18)	2.32	0.12	-
HEP-PH	(1.32, 1.41)	1.67	0.12	-
Semantic	(1.78, 0.96)	1.58	0.06	-
ACL	(1.93, 1.38)	1.43	0.07	0.07
APS	(1.62, 1.20)	1.26	0.11	0.44
Patents	(1.10, 1.01)	1.94	0.04	0.72

Table 2: Global network properties: lognormal (LN) indegree distribution mean & standard deviation (μ, σ), densification power law (DPL) exponent α, average local clustering coefficient (LCC) and attribute assortativity (AA) coefficient of six network datasets.

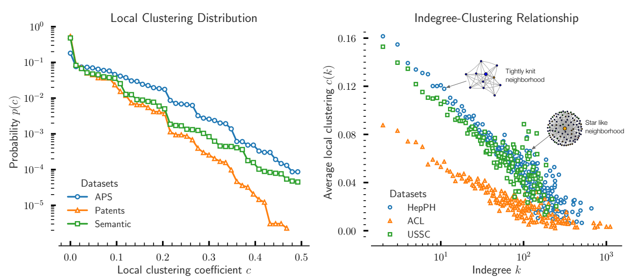

Figure 1: Local clustering in real-world networks have common characteristics: skewed local clustering distribution (left subplot) and a negatively correlated relationship between indegree and average local clustering (right subplot).

Furthermore, real-world networks exhibit a negative correlation between node indegree and local clustering. As shown in fig. 1, the average local clustering decreases as indegree increases. That is, low indegree nodes have small, tightly knit neighborhoods and high indegree nodes tend have large, star-shaped neighborhoods. We propose a model that explains how clustering in real-world networks can arise from local processes of exploration & link formation.

Homophily & Global Assortativity Real-world attributed networks tend to exhibit homophily [28], the phenomenon in which similar nodes are more likely to be connected than dissimilar nodes. The assortativity coefficient [35] $r \in [-1, 1]$, defined as the ratio between the observed modularity and the maximum possible modularity with respect to set of attribute values $B = \{b_1 \dots b_l\}$, quantifies the level of homophily in an attributed network. Intuitively, it 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. Attributed networks ACL, APS & Patents exhibit varying level of homophily, as shown in Figure 2, with assortativity coefficient ranging from 0.07 to 0.72. The magnitude of the attribute assortativity signifies the extent to which attribute similarity influences edge formation. We embed attribute based preferences at the local level lead to generate networks with varying attribute mixing patterns.

Increasing Outdegree over Time The average outdegree of nodes that join real-world networks tends to increase as functions of network size and time. This phenomenon densifies networks and shrinks their diameter over time; Leskovec et al. [26] show that

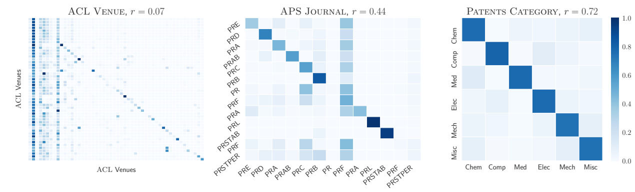

Figure 2: Attributed networks exhibit varying levels of homophily. The subplots illustrate the mixing patterns in ACL, APS and Patents w.r.t. attributes Venue ($r = 0.07$), Journal ($r = 0.44$) and Category ($r = 0.72$) respectively.

densification in many real networks exhibit a power law relationship between the number of edges $e(t)$ and nodes $n(t)$ at time t: $e(t) \propto n(t)^{\alpha}$. Table 2 lists the densification power law exponent α in the network datasets. In our proposed model, we increase the outdegree of incoming nodes at a linear or superlinear rate to account for the accelerated network growth observed in real networks.

To summarize, factors such as preferential attachment, triadic closure and homophily not only effect how individuals form connections at the local level but also explain regularities in global structural properties of real-world networks. Next, we discuss empirical studies in sociology that examine network formation and decision making.

3.3 Insights from Sociological Studies

Sociological studies on network formation seek to explain how individuals form edges in real-world networks. Empirical studies [6, 23] 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 [28] 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). Furthermore, extensive work [16, 27, 44] on individual decision making have established that individuals are *boundedly* rational actors. That is, individuals make decisions under constraints of limited information, cognitive capacity and time. This implies that individuals that join networks employ simple rules to form edges under constraints of limited information and partial network access. For example, a researcher cites academic papers without knowledge of or access to the entire literature in her or his field.

Based on these observations, a faithful characterization of edge formation in real-world networks necessitates bias towards nodes that are similar, proximate or well-connected under constraints of limited information and network access. Existing preferential attachment & fitness-based models [3, 13, 21, 45] make two assumptions that are inconsistent with the findings of aforementioned empirical studies. First, by assuming that successive edge formations are independent, these model disregard the effect of triadic closure and structural proximity. Second, they implicitly require incoming nodes to have complete network access (e.g., connect to any node) or explicit knowledge of one or more properties (e.g., fitness) of every node in the network.

To summarize, citation networks tend to be homophilic networks that undergo accelerated network growth and exhibit regularities in structural properties: heavy tailed indegree distribution, skewed local clustering distribution, negatively correlated degree-clustering relationship and varying attribute mixing patterns. These global properties are a cumulative effect of resource constrained edge formation decisions.

Next, we propose a growth model that unifies multiple sociological phenomena to explain how *local* factors that affect edge formation lead to the emergence of global structural properties observed in real-world networks.

4 ATTRIBUTED RANDOM WALK MODEL

We propose the Attributed Random Walk (ARW) model to explain the emergence of key structural properties of real-world networks through *entirely local* edge formation mechanisms. ARW grows a directed network over time as new nodes join the network. The mechanism that incoming nodes use to form edges intuitively corresponds to how we expect researchers to conduct a literature survey and cite relevant work. First, the researcher broadly identifies one or more *relevant* papers, possibly with the help of external information sources. Then, under time and information constraints, the individual navigates a chain of references to identify *similar* papers that either support or address the problem in hand. Next, through careful analysis, she decides to cite a subset of these papers. Similarly, 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 first describe the edge formation mechanisms underlying ARW. Then, we explain how ARW intuitively incorporates multiple sociological phenomena. Finally, we briefly discuss the methods required to fit the model to data.

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 outdegree of incoming nodes increases over time to reflect the nonlinear growth and densification of real networks.

The edge formation mechanism consists of two components: Select-Seed and Random-Walk. A new 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_a, randomly select $S(u)$ from the set of existing nodes that have attribute value $B(u)$.

(2) Otherwise, with probability $1 - p_a$, randomly select $S(u)$ from the set of existing nodes that do *not* have attribute value $B(u)$.

Seed-Select accounts for homophilic preferences of incoming nodes using attribute parameter p_a. As shown in fig. 3, 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 - α & p_a parameterize edge formation decisions and p_j & p_o 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 $\alpha \cdot p_a$
- Otherwise, u links to v_i with probability $\alpha \cdot (1 - p_a)$

(2) Then, with probability p_j, u jumps back to seed s_u.

(3) Otherwise, with probability $1 - p_j$, u continues to walk. It picks an outgoing edge with probability p_o *or* an incoming edge with probability $1 - p_o$ 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, the attribute parameter p_a is not required. Then, Seed-Select simply selects an existing node uniformly at random and the probability of edge formation in Random-Walk is equal to the rate parameter α only.

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

4.2 ARW & Normative Behavior

The Attributed Random Walk model unifies multiple well-known sociological phenomena into its edge formation mechanisms Seed-Select & Random-Walk.

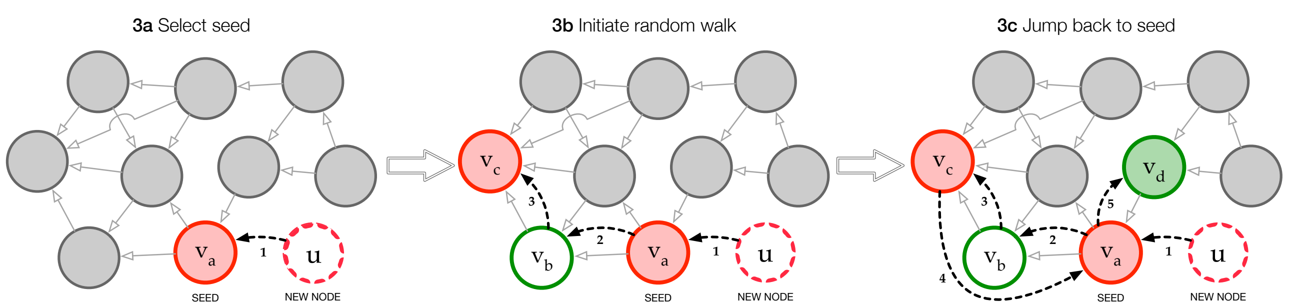

Figure 3: Edge formation in ARW: consider an incoming node u with outdegree $m = 3$ and attribute value $B(u) = \text{RED} \in \{\text{RED}, \text{GREEN}\}$. In fig. 3a, u joins the network and selects seed v_a via Select-Seed. Then, in fig. 3b, u initiates a Random-Walk and traverses from v_a to v_b 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.

Phenomenon 1. *(Limited Resources) Individuals are boundedly rational [16, 27, 44] actors that form edges under constraints of limited information, partial network access and finite cognitive capacity.*

In ARW, we use random walks 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 a few edges, as shown in fig. 3. Random walks inherently account for limited information and partial network access as they only require information about the 1-hop neighborhood of visited nodes.

Phenomenon 2. *(Structural Constraints) Structural factors such as network distance act as constraints that limit edge formation to proximate nodes. [23]*

We incorporate structural constraints into ARW using the jump parameter p_j. The jump parameter p_j is the probability 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 $1-p_j^k$; As a result, the jump parameter p_j 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. [42]*

We control the effect of triadic closure on edge formation using the rate parameter α. A new node u uses a random walk to link to each visited node with probability proportional to α. As a result, the probability with which node u closes a triad by linking to a visited node and its neighbor is proportional to α^2.

Phenomenon 4. *(Attribute Homophily) Nodes that have similar attributes are more likely to form a connection. [28]*

We incorporate attribute homophily into the edge formation process via attribute parameter p_a. New node u links to each visited node v with probability $\alpha \cdot p_a$ if they share the same attribute value. Otherwise, u connects to v with probability $\alpha \cdot (1-p_a)$. The attribute parameter p_a effectively controls the global assortativity coefficient.

Phenomenon 5. *(Preferential Attachment) Nodes tend to link to high degree nodes that have more visibility. [3]*

In the absence of global information, we induce preferential attachment in ARW by adding structural bias to random walk traversals. We utilize the positive correlation between node age and node degree to adjust bias towards visiting old nodes that tend to have high degree. Indeed, random walks that traverse outgoing edges only eventually visit old nodes that tend to have high indegree. Similarly, random walks that traverse incoming edges only visit recently joined nodes that tend to have low indegree. We use out parameter p_o, the probability with which nodes choose outgoing edges in their random walks, to adjust the effect of preferential attachment on edge formation.

ARW unifies five well-known sociological phenomena into a single edge formation mechanism based on random walks. Random walks inherently account for limited information and partial network access. Furthermore, the jump parameter p_j, attribute parameter p_a, rate parameter α and out parameter p_o incorporate the effect of structural constraints, homophily, triadic closure and preferential attachment respectively.

4.3 Model Fitting

We now briefly describe methods to estimate model parameters, initialize $\hat{G}$ at time $t=0$, densify $\hat{G}$ over time and sample incoming nodes' attribute values.

Parameter Estimation. The parameter estimation task consists of finding the set of parameters values for (α, p_a, p_j, p_o) 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 [32] method can be used to quicken parameter estimation.

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 Outdegree. The outdegree of incoming nodes increases over time in real-world networks. We incorporate this phenomenon in ARW to coarsely reflect the rate of growth in the observed network G. 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 outdegree of nodes that join G in year $y(t)$.

Sampling Attribute Values. In real networks $G=(V,E,B)$, the distribution over the set of attribute values $P_G(B)$ changes over time. For instance, the attribute distribution over journals in the APS citation network changes over time as old journals decay in popularity and new journals gain traction. The change in the attribute distribution over time is an exogenous factor and varies for every network. To incorporate this phenomenon into ARW, we sample the attribute value $B(u)$ of node u, that joins $\hat{G}$ at time t, from $P_G(B \mid \text{year} = y(t))$, the observed attribute distribution conditioned on the corresponding year of node u.

To summarize, the Attributed Random Walk (ARW) model intuitively describes how individuals form edges under resource constraints. ARW uses four parameters — α, p_a, p_j, p_o — to incorporate individuals' biases towards similar, proximate and high degree nodes. Next, our experiments in section 5 show that ARW accurately preserves *multiple* structural properties of real networks

5 MODELING NETWORK STRUCTURE

In this section, we evaluate the effectiveness of our model in preserving structural properties of six real-world networks described in subsection 3.1. Our experiments compare ARW to eight well-known growth models. In Subsection 5.1, we describe existing growth models and the evaluation metrics used in the experiments. In Subsection 5.2 , we discuss our experimental results.

5.1 Experiment 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.

Existing Growth Models. We compare ARW to eight well-known growth models that are representative of the key edge formation mechanisms; Two of the eight models account for attribute homophily and preserve attribute mixing patterns.

(1) **Dorogovtsev-Mendes-Samukhin model** [13] (DMS) is a preferential attachment model in which the probability of linking to a node is proportional to the sum of its indegree and "initial attractiveness."
(2) **Kim-Altmann model** [21] (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** [45] (RL) propose a set of preferential attachment models that use relay linking to explain the change in node popularity over time. [2]
(4) **Holme-Kim model** [19] (HK) is a preferential attachment model which uses a triangle-closing mechanism to generate scale-free, clustered networks.
(5) **Social Attribute Network model** [17] (SAN) generates scale-free, attributed networks with high clustering using attribute-augmented preferential attachment and triangle closing mechanisms.
(6) **Herera-Zufiria model** [41] (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** [18] (HZ) is a random walk model that generates scale-free networks with tunable average local clustering.
(8) **Forest Fire model** [26] (FF) is a recursive random walk model that preserves decreasing diameter over time, heavy-tailed degree distribution and high clustering.

To ensure a fair comparison, we modify these models in three ways. First, models that do not have an explicitly defined initial graph use the initialization method described in Subsection 4.3. Second, we extend models that use constant node outdegree by increasing outdegree over time using the method described in Subsection 4.3. Third, we adjust models that generate undirected networks to create directed edges and thereby generated directed networks.

Evaluation. A network model fit should generate a network $\hat{G}$ that preserves the global network structure of the observed network G. We evaluate the 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 indegree k in G and $\hat{G}$ respectively; The weight of each relative error term equals the fraction of nodes with indegree k in G.

[2]We use the iterated preferential relay-cite (IPRC) variant, which best fits real-world network properties

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

5.2 Experiment Results

Our experiment results test the efficacy of ARW in *jointly* modeling multiple structural properties relative to well-known models outlined in subsection 5.1. We evaluate the network models on network datasets outlined in subsection 3.1.

To evaluate the performance of network models, we first fit every model to each network dataset G. Then, we compare the structural properties of network dataset G and network $\hat{G}$ generated by the fitted model using metrics outlined in subsection 5.1. We evaluate multiple instances of $\hat{G}$ to average out fluctuations and acquire data to conduct statistical tests.

Table 4 lists the evaluation metrics for every pair of model and dataset; The metrics measure the accuracy with which these models preserve key global network properties: degree distribution, local clustering distribution and indegree-clustering relationship. We do not explicitly compare the extent to which these models preserve attribute assortativity because the attribute related model parameters can be tuned to obtain arbitrary precision. Instead, models that preserve assortativity up to two decimal places — KA, SAN and ARW — have green ticks (✓) in table 4. We use permutation testing to evaluate the relative performance of our model 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 table 4 are shaded gray or dark gray boxes respectively.

A common characteristic of existing models outlined in subsection 5.1 is that they fail to accurately preserve *multiple* structural properties. This is because existing models disregard important mechanisms such as triadic closure & homophily or are not flexible enough to generate networks with varying structural properties.

Preferential attachment models DMS, RL and KA preserve heavy tailed degree distributions but disregard clustering. DMS outperforms other models in accurately modeling degree distribution (table 4A) because its "initial attractiveness" parameter can be tuned to adjust preference towards low degree nodes. Unlike KA, however, DMS cannot preserve global assortativity. By assuming that successive edge formations are independent, both models disregard the effect of triadic closure and do not preserve local clustering. (tables 4B & 4C).

HK and SAN are preferential attachment models that use triangle closing mechanisms to generate networks with high average local clustering and heavy tailed degree distributions. Note that HK and KA fit degree distributions with the same KS statistic (table 4A) because they lack parameters that can generate varying degree distributions. While triangle closing leads to considerable improvement over DMS and KA in modeling local clustering, HK and SAN are not flexibe

Significance level 0.001 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	✗
	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	✓
	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	✗
Pref. Attachment & 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	✗
	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	✗
	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	✗
	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 4: Modeling network structure. We assess the extent to which network models fit key structural properties of six real-world networks. Tables A, B and C measure the accuracy of eight models in fitting the indegree distribution, local clustering distribution, indegree-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.

enough to preserve local clustering in *all* datasets (see tables 4B & 4C).

Existing random walk models FF, SK and HZ are not flexible enough to accurately preserve network structure observed in real networks datasets. The recursive approach in FF, wherein nodes perform a probabilistic breadth-first search and link to *all* visited nodes, considerably overestimates local clustering. In SK and HZ, nodes perform a single random walk and link to each visited node with some probability μ; The parameter μ indirectly controls the effect on triadic closure on edge formation and leads to some improvement over FF in preserving local clustering distribution (table 4B) and indegree-clustering relationship (table 4). However, the improvements are not substantial when compared to the performance of our model ARW. Furthermore, existing random walk models disregard attribute homophily and do not model attribute mixing patterns.

The experiment results in table 4 validate the effectiveness of the proposed model ARW in *jointly* preserving multiple global network properties. ARW can generate networks with varying degree distribution by adjusting nodes' preference towards high degree nodes using out parameter p_o. As a result, ARW accurately preserves degree distribution (table 4A), often significantly better than all models except DMS. Similarly, ARW matches the local clustering distribution (table 4B) and indegree-clustering relationship (table 4C) observed in real-world networks with high accuracy; This is because the jump parameter p_j and link parameter p_l in ARW control the effect of triadic closure on edge formation. Edge formation in ARW depends on attribute similarity via attribute parameter p_a, which can be tuned to match the attribute assortativity coefficient of attributed network datasets up to arbitrary precision.

Figure 5 illustrates the performance of network models in jointly modeling degree distribution, local clustering distribution and indegree-clustering relationship. Preferential attachment models KA, DMS and RL perform poorly because they do not preserve clustering. HK and SAN perform better than KA, DMS and RL because of edge formation mechanisms that close triangles to preserve clustering and

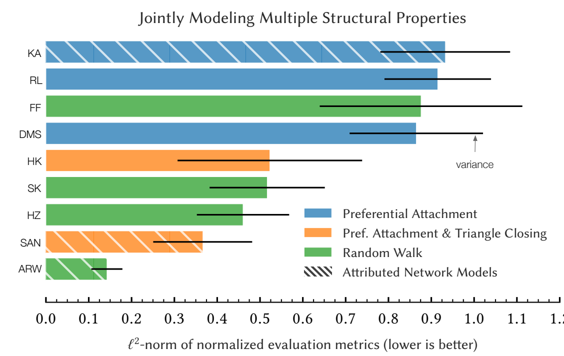

Figure 5: Jointly modeling multiple network properties: ARW outperforms existing network models in jointly preserving key structural properties—indegree distribution, local clustering distribution and degree-clustering relationship— by a margin of 2.5-10x.

its relationship with degree to some extent. The proposed model ARW outperforms existing random walk models HZ, SK and FF by a considerable margin. As shown in fig. 5, ARW improves upon the average ℓ^2-norm of the second best performing model, SAN by a margin of 2.5x.

To summarize, ARW unifies multiple factors that influence edge formation into a single mechanism. As a result, it can jointly preserve multiple structural properties of real-world networks with high accuracy. Next, we discuss limitations of the global assortativity coefficient and analyze local mixing pattterns of real-world attributed networks.

6 MODELING LOCAL MIXING PATTERNS

The global assortativity coefficient quantifies the level of homophily or heterophily in an attributed network. It sheds light on the average propensity of links to occur between similar nodes by capturing the attribute mixing pattern across the entire network. However, global assortativity is not a representative summary statistic of

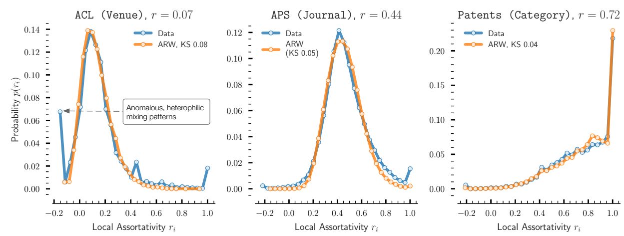

Figure 6: Local attribute mixing patterns of homophilic networks ACL, APS and Patents reveal anomalous, skewed and even heterophilic local mixing patterns. ARW preserves the observed local assortativity distributions with high accuracy, but does not account for nodes with extreme heterophilic or homophilic preferences.

heterogeneous mixing patterns observed in large-scale networks. It does not quantify anomalous mixing patterns and fails to measure how mixing varies across a network.

We use local assortativity [39] to measure varying mixing patterns in an attributed network $G = (V, E, B)$ with attribute values $B = \{b_1 \ldots b_h\}$. 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. Peel et al. [39] prescribe a personalized pagerank weight distribution, which is prohibitively expensive to compute for all nodes in large graphs; Large network datasets necessitate efficient weighting schemes. Therefore, we define w_i as a uniform distribution over $N(i)$, the set of nodes that are at most 1 hop away from node i. 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:

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

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

The local assortativity distributions of ACL, APS and Patents reveal anomalous, skewed and heterophilic local mixing patterns that are not easily inferred via the global assortativity, as shown in fig. 6. ARW can preserve diverse local assortativity distributions with high accuracy even though nodes share the same attribute preference parameter p_a. This is because ARW incorporates multiple sources of stochasticity through its edge formation mechanism. As a result, incoming nodes with fixed homophilic preferences can end up having 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. However, ARW is not expressive enough to accurately model anomalous mixing patterns. Richer mechanisms such as sampling p_a from a mixture of Bernoulli distributions are necessary to account for anomalous mixing patterns.

7 DISCUSSION

In this section, we discuss the insufficiency of the well-known triangle closing mechanism, importance of measuring distributional network properties and the limitations of our model ARW.

7.1 Dissecting the Triangle Closing Mechanism

Network models (e.g., SAN [17] & HK [19]) commonly 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 the local clustering distribution or bivariate degree-clustering relationship accurately. To understand why, we examine the degree-clustering relationship in the APS network:

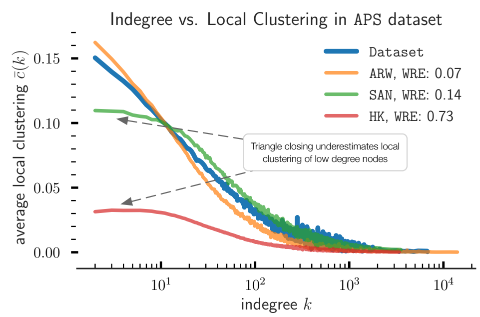

Figure 7: Triangle closing mechanisms in SAN and HK fail to model average local clustering of low indegree nodes. In contrast, the random walk mechanism in ARW visits low indegree nodes and "closes triangles" in their neighborhoods to preserve local clustering with high accuracy.

As annotated in fig. 7, models based on triangle closing mechanisms, SAN and HK, considerably underestimate the local clustering of nodes that have low indegree. This is because incoming nodes in SAN and HK tend to close triangles in the neighborhood of high indegree nodes to which they connect via preferential attachment; Local clustering plateaus as indegree decreases because triangle closing along with preferential attachment fail to form connections in neighborhoods of low indegree nodes. In contrast, ARW accurately models the degree-clustering relationship because incoming nodes initiate random walks in neighborhoods of seeds nodes that tend to have low indegree.

7.2 Measurement of Global Network Properties

Despite their widespread usage, summary statistics of global network properties such as global assortativity and average clustering have limited representative power. Unlike point estimates, distributional properties reveal variance, skewness and anomalies in network data. Notably, understanding local processes via distributional network properties guided the development of ARW, which consists of entirely *local* processes that do not rely on global information (e.g. fitness values of all nodes). For instance, the *skewed* clustering distribution and the relationship between clustering and degree necessitated the jump parameter p_j in our model. The structural constraints imposed by the jump parameter amplify the effect of triadic closure and preserve high clustering observed in neighborhoods of low degree nodes. To summarize, we believe that the analysis and evaluation of *distributional* network properties is crucial to accurately model network structure.

7.3 ARW Limitations

We discuss two limitations of our work. First, ARW does not preserve the average path length distribution of real-world networks. This is because the random walk mechanism is inherently local and does not form long-range connections to bridge distant regions in the network. Preliminary experiments on forming "structural bridges" by initiating multiple random walks for every node indicate a trade-off between modeling small average path length and high local clustering. Second, we only consider citation network datasets in order to study edge formation mechanisms of incoming nodes that form all edges at once. We can adapt ARW to other kinds of networks: attributed random walks that pause and resume intermittently can jointly model edge formation processes between new and existing nodes in social networks; Similarly, metapath based random walks can model interactions between nodes of different types in heterogeneous information networks. In our future work, we plan to study the emergence of higher-order clustering [49] over time and the effect of homophily on the formation of temporal motifs [38] via local processes such as ARW.

In this section, we first discussed the weaknesses of triangle closing mechanisms and the importance of distributional network properties. Then, we briefly described simple methods to extend ARW and address current limitations of our model.

8 RELATED WORK

network growth models seek to explain a subset of structural properties observed in real networks. Well-known network growth models can be broadly categorized by their edge formation mechanism(s):

Preferential Attachment & Fitness In preferential attachment and fitness-based models [4, 5, 10, 29], 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; Node fitness is defined as a dimensionless measure of node attractiveness. For instance, linear preferential attachment functions [3, 13, 24] lead to power law degree distributions and small diameter [8] and attachment functions of degree & node age [47] can preserve realistic temporal dynamics. Extensions of preferential attachment [31, 48, 50] 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 [19, 22, 25] 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 [12, 17, 20, 51] 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 [46], random walk models are inherently local. Models [7] in which new nodes only link to terminal nodes of short random walks generate networks with power law degree distributions [11] and small diameter [30] but do not preserve clustering. Models such as SK [41] and HZ [18], 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 [26] 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. [37] use preferential & uniform attachment to model the decreasing power law exponent of real-world, undirected networks in which average degree increases over time. Singh et al. [45] (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. To the best of our knowledge, ARW is the first model that unifies multiple sociological phenomena into an entirely local process to model network structure *and* attribute mixing patterns. We point the reader to extensive surveys [1, 36] of network growth models for more information.

9 CONCLUSION

In this paper, we model resource-constrained network growth model in which nodes use a random walk process to form edges under constraints of limited information and network access constraints. The problem is important because edge formation in real networks is usually a local process. In typical network growth scenarios, nodes in the network either have limited information about the other nodes in the network or the system allows access to only restricted portion of the existing network. It therefore becomes imperative to model how the local processes of link formation gives rise to network characteristics. In this work, we show that multiple structural properties of real networks can arise from the local process of exploration and link formation. Our results indicate significant improvement over the next best competing model HZ [18] by a significant margin.

REFERENCES

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

[2] 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*.

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

[4] 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.

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

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

[7] 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.

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

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

[10] 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.

[11] 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.

[12] 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.

[13] 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).

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

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

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

[17] 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.

[18] 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.

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

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

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

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

[23] 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

[24] 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.

[25] 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.

[26] 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.

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

[28] 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.

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

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

[31] 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.

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

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

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

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

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

[37] 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.

[38] 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.

[39] 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.

[40] 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

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

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

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

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

[45] 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.

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

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

[48] 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.

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

[50] 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.

[51] 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.