Growing Attributed Networks through Local Processes

Harshay Shah, Suhansanu Kumar, Hari Sundaram
University of Illinois at Urbana-Champaign
{hrshah4,skumar56,hs1}@illinois.edu

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 an interpretable 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. We explore the parameter space of the proposed Attributed Network Growth (ARW) to show each model parameter intuitively modulates network structure. Our experiments show that ARW accurately preserves network structure and attribute mixing patterns of six real-world networks; it improves upon the performance of eight well-known models by a significant margin of 2.5–10×.

CCS CONCEPTS

• **Information systems** → **Web applications**; *Data mining*; *Web mining*; • **Applied computing** → *Sociology*.

KEYWORDS

Network growth; Network Structure; Attributed networks

ACM Reference Format:
Harshay Shah, Suhansanu Kumar, Hari Sundaram. 2019. Growing Attributed Networks through Local Processes. In *Proceedings of the 2019 World Wide Web Conference (WWW'19), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3308558.3313640

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 despite limited information and partial network access. Moreover, 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 properties: 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 rich structural properties.

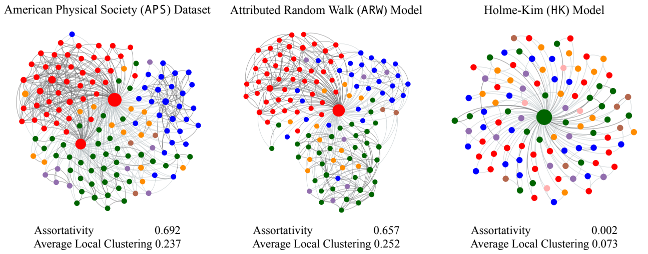

American Physical Society (APS) Dataset
Attributed Random Walk (ARW) Model
Holme-Kim (HK) Model
Assortativity 0.692
Average Local Clustering 0.237
Assortativity 0.657
Average Local Clustering 0.252
Assortativity 0.002
Average Local Clustering 0.073

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 [20] to underscore the importance of using attributes for network growth.

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 [3] or fitness [5, 10, 50] 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). Additionally, 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].

Recent papers tackle resource constraints [35, 51, 53] as well as nodal attributes [12, 17]. 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 an interpretable 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.

WWW '19, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6674-8/19/05.
https://doi.org/10.1145/3308558.3313640

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 an interpretable 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. We analyze the parameter space of the model to show how each parameter intuitively controls one or more key structural properties. 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) **Interpretability**: The model should be expressive enough to generate networks with varying structural properties, while having as few parameters as possible.

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

3 EMPIRICAL ANALYSIS

We begin by describing six large-scale network datasets that we use in our analysis and experiments. Then, we describe global network properties, insights from empirical studies in the Social Science and common assumptions in network modeling. Finally, we discuss the role of structural proximity in edge formation.

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.

Network	Description	$\lvert V \rvert$	$\lvert E \rvert$	T	A, $\lvert A \rvert$	LN (μ, σ)	DPL α	Avg. LCC	AA r
USSC [14]	U.S. Supreme Court cases	30,288	216,738	1754-2002	-	(1.19, 1.18)	2.32	0.12	-
HEP-PH [15]	ArXiv Physics manuscripts	34,546	421,533	1992-2002	-	(1.32, 1.41)	1.67	0.12	-
Semantic [2]	Academic Search Engine	7,706,506	59,079,055	1991-2016	-	(1.78, 0.96)	1.58	0.06	-
ACL [43]	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.

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 [3, 46], where incoming nodes connect with nodes in proportion to their degree. Log-normal fits, with parameters listed in Table 1, well describe the in-degree distribution of all network datasets, consistent Broido and Clauset's [9] observation that scale-free, real-world networks are rare

High Local Clustering: Real-world networks exhibit high local clustering (LCC), as shown in Table 1. Local clustering can arise from triadic closure [37, 45], 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 in-degree and 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 [32], 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. 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.

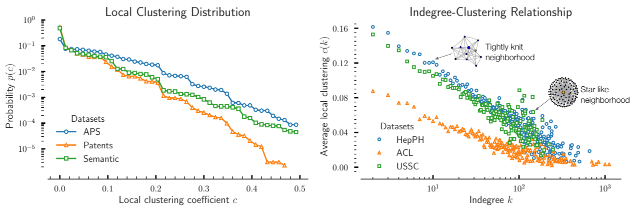


Figure 2: Local clustering in real-world networks have common characteristics: skewed local clustering distribution (left subplot) and a negatively correlated relationship between in-degree and average local clustering (right subplot).

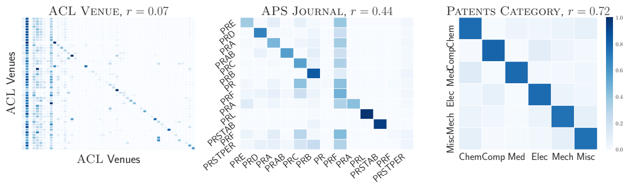


Figure 3: Attributed networks exhibit varying levels of homophily. The subplots illustrate the mixing patterns in ACL, APS and Patents w.r.t. attributes Venue ($r = 0.07$), Journal ($r = 0.44$) and Category ($r = 0.72$) respectively.

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 [6, 25] that analyze the interplay between triadic closure and homophily indicate that *both* structural proximity and homophily are statistically significant factors that simultaneously influence edge formation. Homophilic preferences [32] induce edges between similar nodes, whereas structural factors such as network distance limit edge formation to proximate nodes (e.g. friend of a friend).

Bounded Rationality: Extensive work [16, 30, 47] on decision making shows 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 [3, 13, 48] make two assumptions that are at variance with these findings. 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. 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 sociological studies indicate that edge formation in real-world networks comprises biases towards nodes that are similar, well-connected or structurally proximate. Next, we analyze the role of structural proximity in edge formation.

3.4 Proximity-biased Edge Formation

We investigate the effect of structural proximity on edge formation in real-world networks. Prior work [25] shows that the probability of edge formation in social networks decreases as a function of network distance. Indeed, triadic closure explains how individuals form additional edges to proximate nodes (e.g. friend of friend) over time. However, we lack a concrete understanding of the extent to which structural proximity influences edge formation in bibliographic networks, wherein incoming nodes form all edges at the time of joining the network. In Figure 4, we show how high structural proximity among incoming (shown in blue) node's (shown in red) connections in the Hep-PH dataset hints at edge formation processes biased towards proximate nodes in the same local neighborhood.

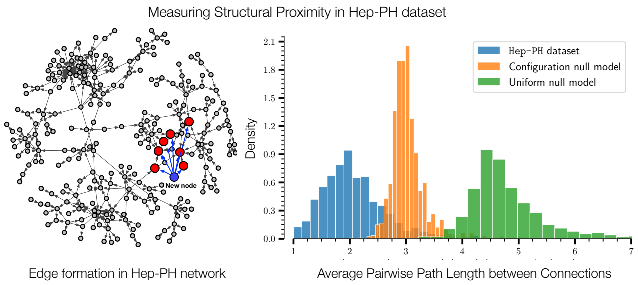


Figure 4: Proximity-biased edge formation. The diagram and proximity distributions collectively indicate how edge formation in real-world networks are biased towards structurally proximate nodes in the same locality.

We rely on network snapshots and node arrival sequence to estimate a statistic based on path length that measures structural proximity between nodes' connections. Consider an incoming node u that forms edges to nodes in $N(u)$. To measure the proximity between node u's connections, we compute the average pairwise shortest path distance between the connections in the network snapshot immediately preceding node u's arrival.

The right subplot in Figure 4 compares the proximity statistic distribution of the Hep-PH dataset to two null models: uniform and configuration. In the uniform model, incoming nodes form connections to existing nodes uniformly at random, whereas the configuration model randomly rewire all edges in Hep-PH while preserving the out-degree and in-degree distributions. We first observe that the connections of incoming nodes in the uniform null model are structurally distant from each other on average. Although the presence of hubs in the configuration model considerably decreases the distance between nodes' connections, it does not explain why the majority of connections in Hep-PH are either connected directly or via an intermediate node. The disparity between the observed and null distributions suggests that structural proximity between connections is intrinsic to edge formation in real-world networks.

To summarize, empirical analyses and insights from the Social Sciences motivate the need to model how resource-constrained edge formation processes collectively shape well-defined global network properties of large-scale networks over time.

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 (e.g. Google Scholar). These initial set of papers act as seed nodes. Then, acting under time and information constraints, she will examine papers cited by the seed and papers that cite the seed. Thus, she navigates a chain of backward and forward references to identify similar, relevant papers. Next, through careful analysis, she will cite a subset of these papers. Similarly, users in online social networks might form new friendships by navigating their social circle (e.g., friends of friends) to find similar others.

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, explain how ARW unifies multiple sociological phenomena, discuss model interpretability and summarize methods required to fit ARW to network data.

4.1 Model Description

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 5, an incoming node u with attribute value $B(u)$ that joins the network at time t first selects a seed node using Select-Seed:

Select-Seed

(1) With probability $p_{\text{same}}/p_{\text{same}}+p_{\text{diff}}$, randomly select a seed node from 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 a seed node from 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 incorporate attribute preferences of incoming nodes. As shown in Figure 5, after selecting the seed 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. A single link parameter p_{link} replaces both attribute parameters p_{same} & p_{diff}. Select-Seed reduces to uniform seed selection and in Random-Walk, the probability of linking to visited nodes equals p_{link}.

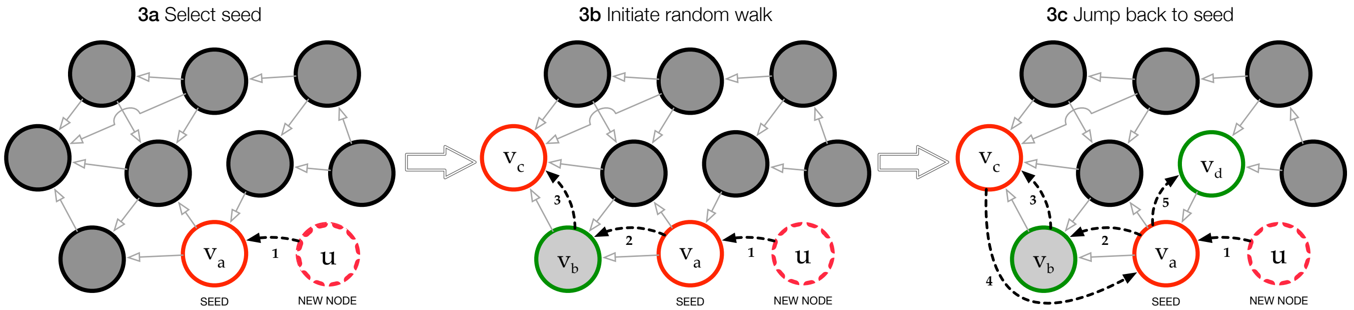


Figure 5: Edge formation in ARW: consider an incoming node u with outdegree $m = 3$ and attribute value $B(u) = \textsc{red} \in \{\textsc{red}, \textsc{green}\}$. In fig. 3a, u joins the network and selects seed v_a via Select-Seed. Then, in fig. 3b, u initiates a Random-Walk and traverses from v_a 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.

Note that ARW has two exogenous parameters: the out-degree $m(t)$ and attribute $B(u)$ of incoming nodes. The attribute distribution varies with time as new attribute values (e.g., journals) crop up, necessitating an exogenous parameter. The parameter $m(t)$ 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, exogenous factors (e.g., venue, topic) may influence node out-degree.

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 [16, 30, 47] actors that form edges under constraints of limited information, partial network access and finite cognitive capacity.*

ARW uses random walk traversals 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 5. Random walks require information only about the 1-hop neighborhood of a few visited nodes, thereby 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, p_{jump} controls the extent to which 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. [45]*

When attribute data is absent, ARW controls the effect of triadic closure on link formation using p_{link}. This is because a new node u closes a triad through its random walk by linking to both, a visited node and its neighbor, with probability proportional to p^2_{link}. 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 u and the visited nodes.

Phenomenon 4. *(Attribute Homophily) Nodes that have similar attributes are more likely to form a connection. [32]*

The attribute parameters p_{same} and p_{diff} modulate 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. [3]*

ARW controls preferential attachment by adding structural bias to the random walk traversal using outlink parameter p_{out}, instead of relying on the global degree distribution. 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. As a result, 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 Interpretability

ARW parameters intuitively shape key structural properties: in-degree distribution, local clustering, path length and attribute assortativity.

In order to understand how global network properties vary as functions of ARW parameters, we explore the parameter space of the model. As described in Subsection 4.1, ARW uses two parameterizations to model networks with or without attribute data. We analyze network structure and attribute assortativity using $(p_{\text{link}}, p_{\text{jump}}, p_{\text{out}})$ and $(p_{\text{same}}, p_{\text{diff}}, p_{\text{jump}}, p_{\text{out}})$ respectively.

Figure 6 illustrates how in-degree and local clustering depend on p_{out} and p_{link}. Increasing p_{out} steers random walks towards older nodes that tend to have higher in-degree. Over time, as more nodes join the network, initial differences in degree amplify, resulting in heavy-tailed distributions. In Figure 6, we observe that increasing p_{out} from 0.2 to 0.8 shifts probability mass from average degree nodes (B) to hubs (C) and low degree nodes (A). As a result, p_{out} controls the extent to which hubs skew the in-degree distribution. Similarly, local clustering increases as a function of p_{link} because

p_{link} implicitly controls the rate at which new nodes close triads by linking to adjacent nodes in their random walks.

We use contour plots to visualize how $(p_{\text{diff}}, p_{\text{same}})$ and $(p_{\text{link}}, p_{\text{jump}})$ alter attribute assortativity and average path length. As shown in the left subplot of Figure 7, $p_{\text{same}} - p_{\text{diff}}$ tunes the extent to which attributes influence edge formation. Increasing $p_{\text{same}} - p_{\text{diff}}$ increases attribute assortativity by amplifying nodes' propensity to link to similar nodes, which subsequently increases the fraction of edges between similar nodes. More importantly, when $p_{\text{same}} - p_{\text{diff}}$ remains constant, increasing $(p_{\text{same}}, p_{\text{diff}})$ raises local clustering without altering attribute assortativity. In the right subplot, we observe that increasing p_{out} while decreasing p_{jump} shortens the average path length. This is because low values of p_{jump} do not restrict incoming nodes to the local neighborhood of their seed nodes, thereby allowing incoming nodes to visit and form edges to nodes that are structurally distant. Additionally, increasing p_{out} results in greater number of hubs, which in turn act as intermediate nodes to connect nodes via short path lengths.

Thus, ARW unifies multiple sociological phenomena at the local level as well as intuitively controls key global network properties.

4.4 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 preserve the structural properties of an observed network G. We use a straightforward grid search method to estimate the four parameters using evaluation metrics and selection criterion described in Subsection 5.1.

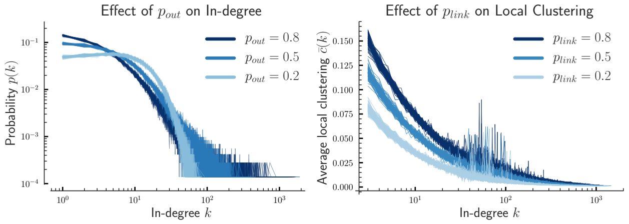


Figure 6: Effect of p_{out} and p_{link} on in-degree and local clustering. The left and right subplots show how increasing p_{out} and p_{link} yield in-degree distributions with heavier tails and neighbors with higher clustering by adding structural bias towards well-connected and proximate nodes respectively.

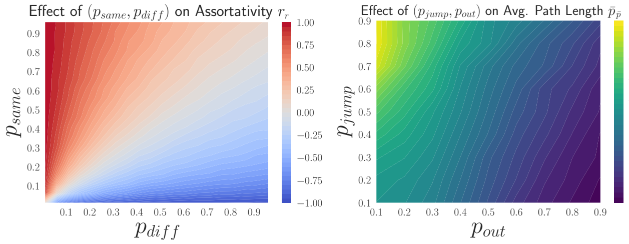


Figure 7: Effect of $(p_{\text{same}}, p_{\text{diff}})$ and $(p_{\text{link}}, p_{\text{jump}})$ on attribute assortativity and average path length. In the left, we observe that increasing $p_{\text{same}} - p_{\text{diff}}$ leads to higher assortativity by making edge formation more homophilic. Similarly, decreasing p_{jump} while increasing p_{out} shortens average path length by enabling incoming nodes to form connections in disparate regions of the network.

Initialization. ARW is sensitive to a large number of weakly connected components (WCCs) in initial network $\hat{G}_0$ because incoming nodes 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 halts 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. Node out-degree increases non-linearly over time in real-world networks. We coarsely mirror the growth rate of 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. The distribution over nodal 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. Therefore, 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, ARW intuitively describes how individuals form edges under resource constraints. ARW uses four parameters $-p_{\text{same}}$, p_{diff}, p_{jump}, $p_{\text{out}}-$ to incorporate individuals' biases towards similar, proximate and high degree nodes. Next, we discuss our experiments on the performance of ARW in accurately preserving multiple structural and attribute properties of real networks.

5 MODELING NETWORK STRUCTURE

In this section, we evaluate ARW's performance in preserving real-world network structure relative to well-known growth models.

5.1 Setup

In this subsection, we introduce eight representative growth models and describe evaluation metrics used to fit models to the datasets.

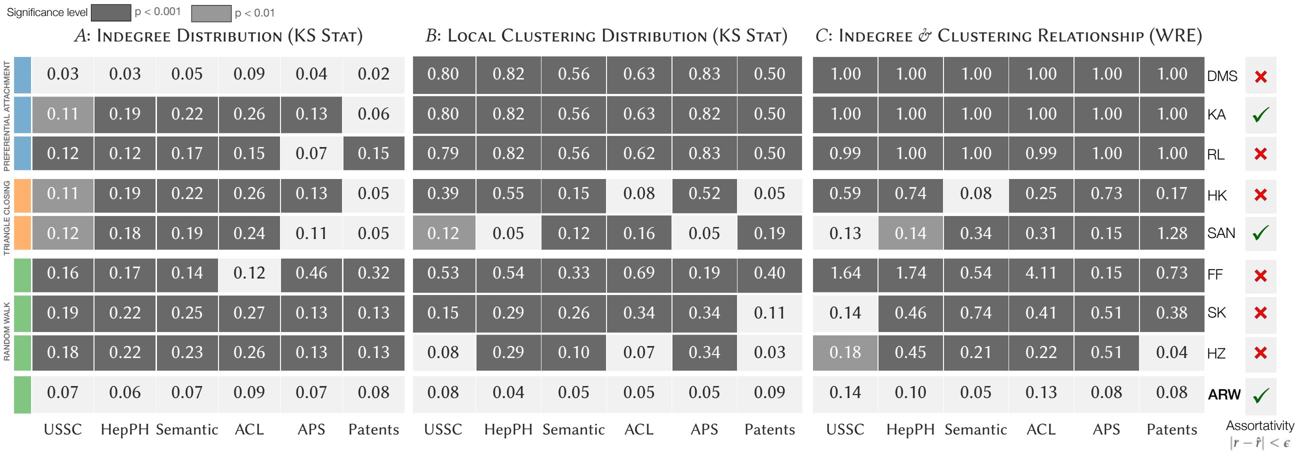
State-of-the-art Growth Models. We compare ARW to eight state-of-the-art growth models representative of the key edge formation mechanisms: preferential attachment, fitness, triangle closing and random walks. Two of the eight models account for attribute homophily and preserve attribute mixing patterns,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

(1) Dorogovtsev-Mendes-Samukhin model [13] (DMS) is a preferential attachment model that generates directed scale-free graphs. In this model, the probability of linking to a node is proportional to the sum of its in-degree and "initial attractiveness."

(2) Relay Linking model [48] (RL) comprises preferential attachment models for directed networks that use relay linking to model node popularity over time. We use the iterated preferential relay-cite (IPRC) variant, which best fits real-world network properties.

(3) Kim-Altmann model [23] (KA) is a fitness-based model that defines fitness as the product of degree and attribute similarity. It generates attributed networks with assortative mixing and power law degree distribution. To generate directed networks, we modify KA to form directed edges to nodes in proportion to their in-degree.

(4) Holme-Kim model [20] (HK) is a preferential attachment model that generates scale-free, clustered, undirected networks using a



Significance level ▇ p < 0.001 ▇ p < 0.01

	A: Indegree Distribution (KS Stat)						*B*: Local Clustering Distribution (KS Stat)						*C*: Indegree & Clustering Relationship (WRE)							Assortativity $|r - \hat{r}| < \epsilon$
	USSC	HepPH	Semantic	ACL	APS	Patents	USSC	HepPH	Semantic	ACL	APS	Patents	USSC	HepPH	Semantic	ACL	APS	Patents		
Preferential Attachment	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	DMS	✗
	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	KA	✓
	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	RL	✗
Triangle Closing	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	HK	✗
	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	SAN	✓
Random Walk	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	FF	✗
	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	SK	✗
	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	HZ	✗
	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	**ARW**	✓

Figure 8: 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.

triangle-closing mechanism. To generate directed networks, we modify HK to form directed edges to nodes in proportion to their in-degree and close triangles in their undirected 1-hop neighborhood.

(5) Social Attribute Network model [17] (SAN) generates scale-free, clustered, attributed networks via attribute-augmented preferential attachment and triangle closing processes. We modify SAN to create directed edges and thereby produce directed networks.

(6) Herera-Zufiria model [44] (SK) is a random walk model that generates scale-free, undirected networks with tunable average clustering. In order to generate directed networks, we allow the random walk mechanism in SK to traverse edges in any direction.

(7) Saramaki-Kaski [19] (HZ) is a random walk model that generates scale-free networks with tunable average local clustering. To generate directed networks, we modify HZ to allow its random walk mechanism to traverse edges in any direction.

(8) Forest Fire model [29] (FF) is a recursive random walk model that can generate directed networks with shrinking diameter over time, heavy-tailed degree distributions and high clustering.

Ensuring Fair Comparison. To ensure fair comparison, we modify existing models in three ways. First, for DMS, SAN, KA do not have an explicitly defined initial graph, so we use initialization method used for ARW, described in subsection 4.4. Second, we extend models that use constant node outdegree m by increasing outdegree over time $m(t)$ using the method described in subsection 4.4. In the absence of model-specific parameter estimation methods, we use grid search to estimate the parameters of every network model, including ARW, using evaluation metrics and selection criterion described below.

Evaluation Metrics. We evaluate the network model fit by comparing four structural properties of G & $\hat{G}$: degree distribution, local clustering distribution, degree-clustering relationship and attribute assortativity. We use Kolmogorov-Smirnov (KS) statistic to compare in-degree & local clustering distributions. We compare the degree-clustering relationship in G and $\hat{G}$ using Weighted Relative Error (WRE), which 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 relative error between $c(k)$ and $\hat{c}(k)$ is weighted in proportion to the number 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, the selection criterion for the grid search parameter estimation method chooses the model parameters that minimizes the ℓ^2-norm of the aforementioned evaluation metrics. Since the metrics have different scales, we normalize the metrics before computing the ℓ^2-norm to prevent unwanted bias towards any particular metric. We note that the parameter sensitivity of the Forest Fire (FF) model necessitates a manually guided grid search method.

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 8 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 these models, we first fit each model to all network datasets G in Subsection 3.1. Thereafter, we compare the structural properties of network dataset G and network $\hat{G}$ generated by the fitted model using evaluation metrics in Subsection 5.1. We average out fluctuations in $\hat{G}$ over 100 runs.

We use one-sided permutation tests [18] 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 8 are shaded gray (▇) or dark gray (▇) respectively. We also group models that have similar edge formation mechanisms by color-coding the corresponding rows

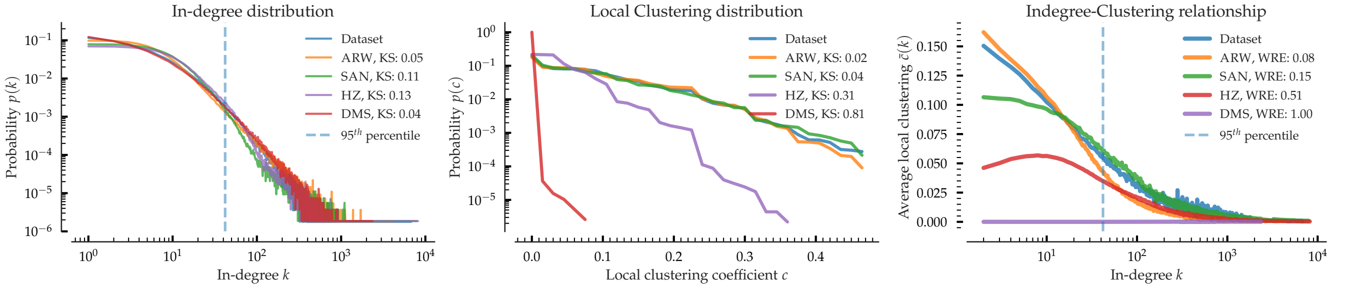


Figure 9: Performance of ARW in accurately preserving key global structural properties of the APS network dataset relative to state-of-the-art, representative network models. Existing models such as DMS and HK cannot preserve high local clustering. Moreover, the triangle closing mechanism in SAN incurs high Weighted Relative Error (WRE) because it cannot explain why low in-degree nodes have high local clustering. ARW outperforms existing network models in jointly preserving all three structural properties, in addition to attribute mixing patterns.

in Figure 8. We use green ticks in Figure 8 to annotate models that preserve assortativity up to two decimal places.

Figure 8 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 8A) 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 8B & Figure 8C).

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 8B & Figure 8C).

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: Figure 8 clearly indicates 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 8A), often significantly better than all models except DMS. Similarly, ARW matches the local clustering distribution (Figure 8B) and in-degree & clustering relationship (Figure 8C) 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.

To summarize, ARW unifies five sociological phenomena into a single mechanism to jointly preserve real-world network structure.

6 MODELING LOCAL MIXING PATTERNS

The global assortativity coefficient quantifies the average propensity of links between similar nodes. However, global assortativity is not a representative summary statistic of heterogeneous mixing patterns observed in large-scale networks [41]. Furthermore, it does not quantify anomalous mixing patterns and fails to measure how mixing varies across a network.

We use local assortativity [41] to measure varying mixing patterns in an attributed network $G = (V, E, B)$ with attribute values $B = \{b_1 ... b_k\}$. Unlike global assortativity that counts all edges between similar nodes, local assortativity of node i, $r_{\text{local}}(i)$, captures attribute mixing patterns in the neighborhood of node i using a proximity-biased weight distribution w_i. The distribution w_i reweighs edges between similar nodes based on proximity to node i. As Peel et al. [41] 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 $N_2(i)$, the two-hop local neighborhood of node i, to allow for a highly efficient local assortativity calculation. Intuitively, $r_{\text{local}}(i)$ compares the observed fraction of edges between

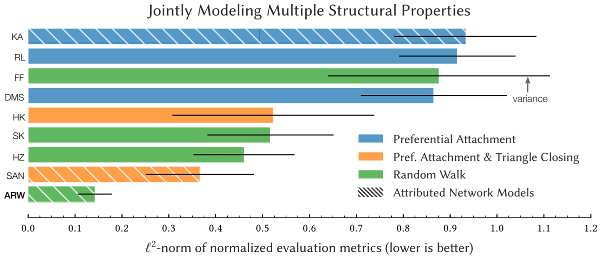


Figure 10: ARW outperforms existing network models in jointly preserving key structural properties—in-degree distribution, local clustering distribution and degree-clustering relationship— by a significant margin of 2.5x-10x.

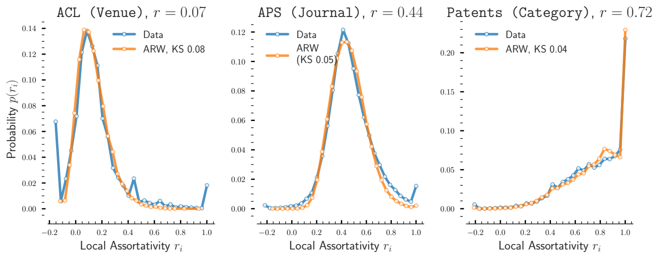


Figure 11: Local assortativity distributions of attributed networks ACL, APS and Patents reveal anomalous, skewed and heterophilic local mixing patterns. ARW accurately preserves local assortativity, but does not account for anomalous mixing patterns.

similar nodes in the local neighborhood of node i to the expected fraction if the edges are randomly rewired.

As shown in Figure 11, local assortativity distributions of ACL, APS and Patents reveal anomalous, skewed and heterophilic local mixing patterns that are not inferred via global assortativity. Our 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 weaknesses of triangle closing mechanisms, the effect of out-degree on network diameter and limitations & potential modifications of our model ARW.

7.1 Dissecting the Triangle Closing Mechanism

A set of network models (e.g., SAN [17] & HK [20]) 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 explain local clustering distribution or bivariate degree-clustering relationship accurately. To understand why, we examine the degree-clustering relationship in the APS network in Figure 12.

Figure 12 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 Effect of Out-degree on Network Diameter

Extensive analyses [21, 29, 31] on evolving real-world networks reveal two key temporal properties: network densification and diameter shrinkage over time. Growth models can be adjusted to densify networks over time by allowing node out-degree to increase super-linearly as a function of network size. However, we lack a concrete understanding of existing edge formation mechanisms' inability to preserve diameter shrinkage. Through our analysis, we observe that the out-degree sequence of incoming nodes in network models has a significant impact on effective diameter over time.

Figure 13 illustrates the effective diameter of network models fitted to Hep-PH as a function of node out-degree sequence and time. By increasing the out-degree $m(t)$ over time using the method described in subsection 4.4, network models representative of key edge formation mechanisms—ARW, FF, DMS & SAN—generate networks that exhibit diameter shrinkage. In particular, FF [1] and ARW mirror the observed rate at which the effective diameter shrinks over time. However, when the out-degree $\bar{m} = n^{-1} \sum_i m(i)$ of incoming nodes is constant, fitted networks, including Forest Fire (FF), cannot preserve shrinking diameter; the effective diameter of the fitted models remain consistently lower than that of Hep-PH.

Increasing out-degree over time can effectively incorporate diameter shrinkage in all representative network models. This phenomenon is best understood through the simple Uniform null model,

[1] FF inherently increases out-degree over time because incoming nodes "burn" through the network for duration in proportion to the network size.

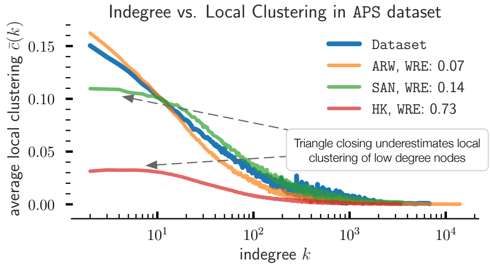


Figure 12: Triangle closing mechanisms used in SAN HK fail to model average local clustering of low in-degree nodes. In contrast, to accurately preserve local clustering, ARW uses random walks to visit low in-degree nodes and close triangles in their neighborhoods .

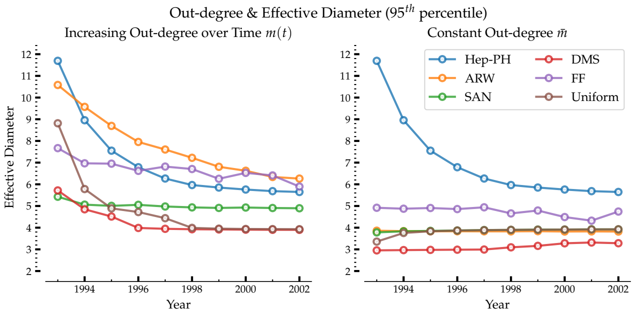


Figure 13: Effect of out-degree on effective diameter. The left subplot shows that increasing out-degree over time leads to diameter shrinkage for all models. The right subplot shows that constant out-degree sequence does not account for diameter shrinkage and consistently underestimates effective diameter of the Hep-PH network.

in which incoming nodes form edges to existing nodes chosen uniformly at random. In the Uniform model, nodes with higher out-degree have a greater probability of linking to existing nodes that are structurally distant from each other. Consequently, over time, incoming nodes with higher out-degree are more likely to bridge distant regions of the network, reduce path length between existing nodes and subsequently shrink effective diameter.

To summarize, our analysis indicates how increasing out-degree over time enables existing models that rely on different edge formation processes to account for diameter shrinkage.

7.3 ARW Limitations

We discuss three limitations of ARW. First, we consider only bibliographic network datasets in which nodes 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 potential way is to incorporate random walks that pause and resume intermittently, thus allowing for older nodes to connect with more recent arrivals. Second, the out-degree $m(t)$ of incoming nodes in ARW rely on the observed out-degree sequence, which might be unavailable in datasets without fine-grained temporal data. In this case, ARW can be adapted to rely on the prescribed range of densification exponent α_{DPL} [29] in real-world networks. Since $e(t) = m(t)n(t)$, the power law relationship $e(t) \propto n(t)^{\alpha_{\text{DPL}}}$ between number of edges $e(t)$ and nodes $n(t)$ at time t implies that out-degree $m(t)$ must be proportional to $n(t)^{\alpha_{\text{DPL}}-1}$. Third, ARW focuses on modeling networks in which nodes have a single attribute. The difficulty in incorporating multiple attributes into the edge formation mechanism rests on how we measure attribute similarity. If two nodes are similar only when all their attribute values are identical, we can simply create a new categorical attribute that encodes all multiple attribute combinations and then directly apply ARW. Additional analysis is necessary to identify definitions of attribute similarity that best describe how multiple attributes influence individuals' edge formation processes.

8 RELATED WORK

Network growth models seek to explain a subset of structural properties observed in real networks. Note that, unlike growth models, statistical models of network replication [28, 42] 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 [4, 5, 10, 33], 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 [3, 13, 26] lead to power law degree distributions and small diameter [8] and attachment functions of degree & node age [50] can preserve realistic temporal dynamics. Extensions of preferential attachment [35, 51, 53] 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 [20, 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: These models [12, 17, 22, 54] 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 [49], 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 [34] but do not preserve clustering. Models such as SK [44] and HZ [19], 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. [39] 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. [48] (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 simple, interpretable 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. We explored the parameter space of the model to show how each parameter intuitively controls one or more key structural properties. Our experiments on six large-scale citation networks showed 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: modeling undirected, social networks, understanding the emergence of higher-order clustering [52] and modeling the effect of homophily on the formation of temporal motifs [40]

REFERENCES

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

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

[19] 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.

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

[21] Haibo Hu and Xiaofan Wang. 2009. Evolution of a large online social network. *Physics Letters A* 373, 12-13 (2009), 1105–1110.

[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] Mary McGlohon, Leman Akoglu, and Christos Faloutsos. 2011. Statistical properties of social networks. In *Social network data analytics*. Springer, 17–42.

[32] 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.

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

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

[35] 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.

[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] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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

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

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

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

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

[48] 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.

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

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

[51] 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.

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

[53] 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.

[54] 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.