MODELING LINK RECOMMENDATIONS AS A NETWORK GROWTH MECHANISM AND THEIR IMPACT ON SOCIAL CONTAGION

Björn Komander[1,2*] **Jesus Cerquides**[1] **Jeffrey Chan** [2] **Azadeh Avali** [2]
[1] Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Spain
[2] School of Computing Technologies, RMIT University, Melbourne, Australia
{bkomander, cerquide}@iiia.csic.es
{jeffrey.chan, azadeh.alavi}@rmit.edu.au

ABSTRACT

Link recommendation algorithms significantly shape online social networks, influencing both their structural evolution and critical processes such as information and behavior spread. This paper investigates how these algorithms affect simple and complex contagion processes by modeling recommendations as additional network growth mechanisms. We introduce a synthetic network model that integrates preferential attachment, triadic closure, and choice homophily, then extend it with various link recommenders, including heuristics and graph neural networks (GNNs). Our findings show that while simple contagions exhibit relatively modest shifts under most recommenders, complex contagions are highly sensitive to clustering- and homophily-based recommendations, thriving at moderate recommendation strengths but sharply diminishing under excessive recommendation strength. These results underscore the nuanced interplay between network structure, recommendation strength, and contagion dynamics, highlighting the importance of incorporating social contagions into the design of link recommendation algorithms.

1 INTRODUCTION

The access to information and the ability to share information and behaviors are central functions of online social networks (Easley & Kleinberg, 2010; Myers et al., 2012; Rogers et al., 2008; Centola, 2010). The dynamical processes describing how information and behaviors spread online are often referred to as social contagion (Centola & Macy, 2007; Granovetter, 1978). Similar to contagious diseases, **simple contagions** describe the diffusion of information through a network, often requiring just a single exposure to an infected source (Kempe et al., 2003). In contrast, **complex contagions** involve the propagation of behaviors, movements, and social norms, typically requiring multiple exposures due to social reinforcement (Centola & Macy, 2007).

Understanding these processes is key to studying how novel information, misinformation, and social movements spread, as well as how collective behavior emerges. (Bakshy et al., 2012; Vosoughi et al., 2018; González-Bailón et al., 2011; Fink et al., 2016; Mønsted et al., 2017; Centola, 2010). Like many dynamic processes on networks, the topology of the underlying network fundamentally shapes how behavior and information propagate online. Simple contagions leverage weak ties (Bakshy et al., 2012; Granovetter, 1973) and hubs (Cencetti et al., 2023; Nematzadeh et al., 2014) to spread rapidly, while complex contagions require local clustering and wide bridges for sufficient social reinforcement (Centola & Macy, 2007; Guilbeault et al., 2018; Guilbeault & Centola, 2021).

It has been argued that online social networks increasingly favor mechanisms that promote simple contagions, optimizing for rapid information spread through weak ties and hubs, while undermining the spread of complex contagions, such as social norms and collective actions (Centola, 2018). While this argument is a matter of debate, platform design should account for both types of contagion to foster diverse and meaningful social dynamics. One increasingly influential aspect of online social network design is the use of algorithmic curation (Lazer, 2015). With advancements in ma-

chine learning, recommender systems have become a powerful tool for curating both content and connections on social platforms and therefore raise concerns about their impact on social dynamics (Pappalardo et al., 2024). Among these, link recommenders have been shown to induce profound structural changes in online social networks. For example, LinkedIn's weak-tie recommendations were shown to improve job market outcomes (Rajkumar et al., 2022), while Twitter's "Who to Follow" algorithm increased clustering and amplified the popularity of already well-connected users (Su et al., 2016). Synthetic studies further show that link recommenders can drive inequality (Espín-Noboa et al., 2022) or polarization (Santos et al., 2021), underscoring their potential impact on network evolution and dynamics.

This raises a key question: How do link recommendations influence social contagion? Since simple and complex contagions thrive in different network structures, it is crucial to examine how link recommenders reshape these structures. This involves comparing their effects to counterfactual scenarios, such as organic growth or varying recommendation strengths. However, existing studies and methodologies often fall short in fulfilling these requirements, limiting their ability to provide comprehensive insights.

Observational studies, for instance, can leverage quasi-experimental settings, such as examining network measures before and after the introduction of link recommenders (Su et al., 2016). However, these approaches are constrained by their reliance on data from a single network and the availability of suitable quasi-experiments, making it difficult to generalize findings across diverse network structures. Experimental studies (Su et al., 2020; Rajkumar et al., 2022) offer greater control and flexibility but face challenges in addressing networked interference, which complicates causal inference (Gui et al., 2015; Zhang et al., 2023) and rely on corporate collaboration, limiting their feasibility. **Synthetic experiments** offer an appealing alternative, allowing researchers to systematically test how different link recommendation algorithms reshape network structures under various conditions. However, many existing synthetic studies adopt a narrow set of initial network topologies, limiting their applicability (Espín-Noboa et al., 2022; Zhang et al., 2023; Stoica et al., 2018). They also frequently overlook counterfactual scenarios, such as comparing outcomes to organic growth or assessing varying recommendation strengths.

Contributions. In this paper, we address these challenges by:

1. **Proposing a versatile synthetic network model** that can generate a wide range of network topologies through three main growth mechanisms: *(i) preferential attachment, (ii) triadic closure*, and *(iii) choice homophily*.
2. **Modeling link recommendations as an additional growth mechanism** within the network, enabling us to systematically vary the strength of recommendations relative to organic growth.
3. **Incorporating feature-based algorithms**, specifically Graph Neural Networks (GNNs), to capture homophily-based recommendations, and comparing them with simpler heuristic algorithms.
4. Analyzing how simple and complex contagions evolve under different network structures, recommendation algorithms, and recommendation strengths, to better understand the interplay between network topology, recommendation mechanisms, and contagion dynamics.

Our findings highlight that link recommendation algorithms can substantially alter both simple and complex contagion dynamics, with **complex contagions** being particularly sensitive to scenarios involving strong clustering or homophily. These results underscore the importance of tailoring recommendation strategies to network structures and contagion types, and they emphasize the need for more nuanced models in studying algorithmic interventions in social networks.

2 Related Work

Impact of Link Recommendation on Social Networks. A widely used synthetic network model for studying the effects of link recommendations is the DPAH model (Espín-Noboa et al., 2022). The model incorporates preferential attachment for scale-free degree distributions and supports varying levels of homophily. Once the model is configured, new links are added through various link recommenders. To isolate the effects of link recommendations from changes in network density, the

model removes an existing link whenever a new node forms a connection, effectively maintaining a fixed network density. Using this approach Santos et al. (2021) found that link recommenders increase levels of affective polarization across different settings, while Ferrara et al. (2022) found significant effects on the network topology, including amplified popularity bias as well as minority group visibility disparities (Espín-Noboa et al., 2022; Ferrara et al., 2022). However, the model is limited to comparing networks entirely rewired by recommenders to their initial configurations or to networks generated by different recommenders. It does not support comparisons with organic growth or varying recommendation strengths. While the model offers some flexibility in configuring preferential attachment and homophily, it cannot generalize to other network topologies, such as clustered networks.

An improvement is the work of Zhang et al. (2023), which extends a synthetic *clustered* network model (Jackson & Rogers, 2007) to alternate between recommendation-driven and organic growth phases. This approach enables counterfactual analysis, allowing the estimation of direct and indirect causal effects of link recommendations. However, their reliance on a single network model limits the generalizability of their findings to networks with varying structural characteristics. Additionally, recommendations are modeled as fixed phases of algorithmic intervention, constraining the timing of recommendation-driven and organic growth.

Stoica et al. (2018) also compare the impact of link recommendation algorithms to organic growth, showing that recommendation systems significantly amplify demographic disparities—such as the underrepresentation of women in top social hierarchies—beyond what organic growth alone produces. Their model introduces a probability parameter, p, that determines whether a given growth step is organic or recommendation-driven. Organic growth occurs when a new node is added to the network and connects to existing nodes based on preferential attachment and homophily. In contrast, during recommendation-driven growth $(1 - p)$, a randomly selected existing node forms a new edge with another existing node based on the recommender's algorithm and the incoming node is discarded. This approach neglects organic growth among existing nodes, reducing the model's flexibility in capturing interactions between organic and recommendation-driven dynamics. Merging common network growth mechanisms—such as homophily and preferential attachment (Stoica et al., 2018; Espín-Noboa et al., 2022) as well as clustering (Zhang et al., 2023) - into a unified model that accounts for the interplay between organic and recommendation-driven dynamics represents a critical gap in the existing literature.

Link Recommendation and Link Prediction Link recommendation methods often build upon models for link prediction, which estimate the likelihood of connections forming between pairs of nodes (Liben-Nowell & Kleinberg, 2003; Li et al., 2017). Among these, graph neural networks (GNNs) have gained prominence due to their ability to process relational data efficiently while incorporating both structural and feature-based information (Zhang & Chen, 2018). As GNN-based methods achieve higher performance and are increasingly adopted in real-world applications (Borisyuk et al., 2024), research has begun to explore their implications for fairness (Li et al., 2022) and biases (Wang & Derr, 2022; Subramonian et al., 2024). However, their impact on the structure and dynamics of social networks remains unclear. A key limitation is the lack of synthetic network models that integrate node features and enable the evaluation of feature-driven algorithms.

3 METHODOLOGY

3.1 NETWORK GROWTH MODEL

Understanding how networks evolve requires models that capture their fundamental growth mechanisms. A key distinction is whether growth occurs through the addition of nodes (Barabási & Albert, 1999) or through the addition of edges among existing nodes (Leskovec et al., 2007). While both processes shape online social networks, we focus on edge formation among existing nodes.

Our model operates on a set of n nodes $\mathcal{N}$, fixed over time, with new edges forming over discrete time steps $t \in \{1, 2, \ldots, T\}$. Let $\mathcal{E}_t$ denote the set of edges at time t. For simplicity, we consider undirected graphs, but the model can be extended to directed graphs. Accordingly, $\boldsymbol{A}_t \in \{0, 1\}^{n \times n}$ denotes the symmetric and unweighted adjacency matrix corresponding to graph $\mathcal{G}_t(\mathcal{N}, \mathcal{E}_t)$.

New connections are formed through three distinct mechanisms: (1) Preferential attachment (Barabási & Albert, 1999), where nodes connect to others proportional to their degree; (2) Triadic closure (Newman, 2001; Bianconi et al., 2014; Kossinets & Watts, 2009), which captures the tendency of nodes to connect with the neighbors of their neighbors; and (3) Choice homophily (McPherson et al., 2001; Kossinets & Watts, 2009; Currarini et al., 2016), where nodes preferentially connect with others sharing similar characteristics.

Instead of using discrete types or group memberships to represent homophily (Espín-Noboa et al., 2022; Stoica et al., 2018), we model it as similarity across multiple features. This approach is more compatible with integrating GNNs.

First, each node is assigned a group membership $\boldsymbol{z}$, analogous to block membership in stochastic block models. The set of groups $\mathbb{K}$ is fixed, and group membership does not change over time, $z_i \in \{1, 2, \ldots, K\}$ for all $i \in \mathcal{N}$. While we assume equal group sizes for simplicity, the model can accommodate arbitrary group size distributions, including majority-minority dynamics (Espín-Noboa et al., 2022).

To generate node features, we borrow ideas from recent advances in benchmarking GNNs with synthetic networks (Tsitsulin et al., 2022; Palowitch et al., 2022). Specifically, group memberships $\boldsymbol{z}$ are used to create features through Gaussian sampling and covariance structure modeling. Each group k is assigned a center vector μ_k, sampled from a multivariate normal distribution with a variance parameter σ_c^2, and a covariance matrix Σ_k, sampled from an inverse Wishart distribution with parameters ensuring stability and positive definiteness. Node feature vectors $\boldsymbol{x}_i$ are then generated by sampling from a multivariate normal distribution with the group-specific μ_k and Σ_k. This results in a feature matrix $\boldsymbol{X} \in \mathbb{R}^{n \times d}$, where n is the number of nodes and d is the feature dimension.

While Palowitch et al. (2022) use this process to assign features to nodes in a stochastic block model, we **invert** this process. We first assign nodes to groups, then generate features based on these group memberships. The features are subsequently used to model choice homophily and guide the evolution of the underlying network. For details on the feature generation process, see A.1.

The probability of forming an edge between two unconnected nodes u and v is determined by one of the three mechanisms:

1. **Preferential Attachment (PA)**: The probability that node u connects to node v is proportional to the relative degree of node v:
$$p_{PA}(u, v) \propto k_v, \tag{1}$$
where k_v denotes the degree of node v. If all the degrees of unconnected nodes v are zero, we set $p_{PA}(u, v) = \frac{1}{|\mathcal{N}|}$ for all v.
2. **Triadic Closure (TC)**: The probability that node u connects to node v depends on the number of their shared neighbors:
$$p_{TC}(u, v) \propto |N(u) \cap N(v)|, \tag{2}$$
where $N(u)$ denotes the neighbors of node u. If $N(u)$ is empty, we set $p_{TC}(u, v) = \frac{1}{|\mathcal{N}|}$ for all v.
3. **Choice Homophily (HP)**: To model choice homophily, we use the generated node features. The feature matrix $\boldsymbol{X}$ is normalized column-wise to range between zero and one, and scaled by $\sqrt{d}$, where d is the feature dimensionality, to standardize their contributions. Using these normalized features, we compute a pairwise Euclidean distance matrix $\boldsymbol{D}$, with $D_{uv} = \|X_u - X_v\|_2$ The probability of forming an edge between u and v decreases exponentially with distance, parameterized by a temperature constant T:
$$p_{HP}(u, v) \propto \exp\left(-\frac{D_{uv}}{T}\right). \tag{3}$$

At each time step, a node u is chosen uniformly at random. For u, the probability of forming a new edge with each unconnected node v (i.e., where $\boldsymbol{A}_{uv} = 0$ and $u \neq v$) is calculated using one of the three mechanisms: $p_{PA}(u, v)$, $p_{TC}(u, v)$, or $p_{HP}(u, v)$. The mechanism is chosen probabilistically, with weights α, β, and γ for preferential attachment, triadic closure, and choice homophily, respectively, where $\alpha + \beta + \gamma = 1$.

A node v is then sampled from this distribution to form an undirected edge with u. This process repeats until the desired number of edges is added.

3.2 LINK RECOMMENDATIONS

Given a graph $\mathcal{G}_t$ with $\mathcal{N}$ nodes and $\mathcal{E}_t$ undirected edges evolved until t, and a respective feature vector $\boldsymbol{X} \in \mathbb{R}^{d \times n}$, we aim to design a link recommender that scores the likelihood of forming a link between an unconnected node pair (u, v), where $\boldsymbol{A}_{uv} = 0$ and $u \neq v$. Let $\Gamma(u, v)$ denote this function. We then define the link recommendation probability $p_{rc}(u, v)$ as:

$$p_{rc}(u, v) = \frac{\Gamma(u, v)}{\sum_{v \in \boldsymbol{A}_{uv}=0} \Gamma(u, v)} \tag{4}$$

We incorporate the link recommender as an *additional growth mechanism* in our model, with its influence controlled by a parameter θ. At each timestep, for a given candidate node u, unconnected nodes v are scored according to one of the following probability functions:

$$p_{PA}(u, v), \quad p_{TC}(u, v), \quad p_{HP}(u, v), \quad \text{or } p_{rc}(u, v).$$

These functions are selected probabilistically, with weights $\alpha(1 - \theta)$, $\beta(1 - \theta)$, $\gamma(1 - \theta)$, and θ, respectively.

When a recommendation mechanism is chosen (based on θ), instead of sampling a single edge, the top j edges from the candidate set are selected based on their probabilities. These selected edges are then assigned an equal probability of $\frac{1}{j}$, ensuring a fair distribution of edge formation across the top recommendations. In this work we consider $j = 3$, mimicking the limited recommendations shown to users (Santos et al., 2021; Espín-Noboa et al., 2022; Su et al., 2016)

3.3 SOCIAL CONTAGION

To evaluate changes of the network structure, we consider different models of social contagion: a simple contagion process, modeled using an independent cascade-like structure (Kempe et al., 2003; Goldenberg et al., 2001a;b), and a complex contagion process, implemented through a threshold-based mechanism (Granovetter, 1978; Centola & Macy, 2007). Both models work in discrete time steps $t \in [1, 2, \ldots, t_f]$, where t_f denotes the final timestep when no new nodes are infected. We monitor the size and rate of contagion for both models. Let n denote the number of nodes in the network and $I(t)$ the number of infected nodes at time t. The contagion size CS is then defined as $CS = \frac{I(t_f)}{N}$ and the contagion rate CR as $CR = \frac{I(t_f)}{t_f}$.

In the simple contagion model, each infected node has a fixed probability p of infecting its neighbors upon contact. The model operates in discrete time steps, where each infected node independently attempts to spread the contagion to its adjacent nodes, with the process repeating until no further infections occur. Each edge is evaluated once, therefore each infected node has a single chance to infect a specific neighbor.

The complex contagion model employs a threshold-based approach to simulate scenarios where individuals require multiple exposures before adopting the contagion. Each node is assigned a threshold value τ, drawn from a truncated normal distribution, $\mathcal{N}(\lambda, \sigma^2; a, b)$. The threshold represents the fraction of neighbors who must be infected/adopted before a node adopts the contagion. The term adoption more adequately describes a complex contagion, but to align the terminology with the simple contagion, we use infection and infected.

Specifically, a node is infected if the fraction of neighbors in the infected state is larger than the node's threshold. Let $|I_{it}|$ denote the number of infected neighbors of node i at time t and $|N_i|$ the number of neighbors of i. The node i becomes infected if $\frac{|I_{it}}{|N_i|} > \tau$. The process is controlled by the mean threshold λ.

4 EXPERIMENTS

4.1 LINK PREDICTION ALGORITHMS

Below, we describe the different variations for the link scoring function $\Gamma(u, v)$. Let $N(u)$ the set of neighbors of node u. For algorithms that don't require any learning, the implementation of $\Gamma(u, v)$ is straightforward:

1. **Adamic Adar (AA)**: The Adamic-Adar index assigns a similarity score to each node pair by summing the inverse logarithm of each shared neighbor's degree (Adamic & Adar, 2003). This method gives greater weight to shared neighbors with fewer connections, making it especially useful for identifying links where shared neighbors are uncommon but indicative of a strong connection.

$$\Gamma_{\text{AA}}(u, v) = \sum_{z \in N(u) \cap N(v)} \frac{1}{\log |N(v)|} \tag{5}$$

2. **Preferential Attachment (PA)**: The preferential attachment score is based on the idea that nodes with higher degrees are more likely to attract new links (Barabási et al., 2002; Newman, 2001; Liben-Nowell & Kleinberg, 2003). For a given pair of nodes, this score is computed as the product of their degrees, representing the tendency for high-degree nodes to form new links due to their popularity. This approach is frequently used in modeling network growth patterns where "rich get richer" dynamics are present (Stoica et al., 2018).

$$\Gamma_{PA}(u, v) = |N(u)||N(v)| \tag{6}$$

For the learning algorithms, we consider graph neural networks (GNN) for link prediction. Specifically, the link scoring function $\Gamma(u, v)$ is defined as the inner product of the node representations obtained through a GNN-based encoder. This scoring function captures the relationship strength between two nodes based on their learned representations.

$$\Gamma_{GNN}(u, v) = (h_u^l)^T h_v^L \tag{7}$$

where h_u^L is the final layer representation of node u. In this work we consider graph convolutional networks (Kipf & Welling, 2017) as a simple yet efficient GNN architecture. However our model could be extended with any GNN architecture.

1. **Graph Convolutional Network (GCN)**: The GCN aggregates features from neighboring nodes through a layer-wise propagation rule that incorporates both the adjacency matrix $\tilde{\boldsymbol{A}}$ and a degree matrix $\tilde{\boldsymbol{D}}$. The propagation is given by:

$$\boldsymbol{H}^{(l+1)} = \sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}\right)$$

   where $\boldsymbol{H}^{(l)}$ represents the node feature matrix at layer l, $\boldsymbol{W}^{(l)}$ is the weight matrix for layer l, and σ is a non-linear activation function. The GCN thus learns node representations by considering the structure of the graph, enabling the model to capture local connectivity patterns in the feature space.

We generate networks using different combinations of α, β, and γ. Each network consists of $n = 500$ nodes, evenly split into $|K| = 4$ groups, with each node having $d = 50$ features. The network evolves until $t_{\text{train}} = 2000$, resulting in $\mathcal{G}_{\text{train}}$.

We train the GCN model on $\mathcal{G}_{\text{train}}$ using all positive edges, sampling a new set of negative edges each epoch at a 1:8 positive-to-negative ratio. Training runs for 60 epochs with a learning rate of 0.001. After training, we predict scores for all potential future edges to model p_{rc}, without further retraining in subsequent evaluation steps. This process is repeated 5 times for each combination of α, β, and γ.

4.2 RECOMMENDATIONS

After training, we evolve the network for 4000 additional steps, incorporating potential recommendations with strengths $\theta \in (0, 0.5, 1)$. Every 50 steps, we evaluate the network via contagion processes, running 100 simple and complex contagion simulations and reporting the mean infection size and rate.

For simple contagion, we use an infection probability of $p = 0.15$, selecting a single infected node at random. For complex contagion, node thresholds are sampled from a truncated normal distribution with mean $\lambda = 0.23$ and standard deviation $\sigma = 0.1$, with 1% of nodes initially infected. At each evaluation step, both the initial infections and thresholds are resampled for all 100 contagion processes.

5 RESULTS

Our experiments show that the introduction of link recommendations can have significant impacts on the size and rate of both simple and complex contagion processes. However, the effects vary between simple and complex contagions, between recommender algorithms and their strength and across different network structures

5.1 STRONG HOMOPHILOUS SCENARIO

First, we examine the effects of link recommendations in a homophilous network $\mathcal{G}_{HP}$ with $\gamma = 1$, meaning the network evolves solely through choice homophily. Figure 1 shows that in such a setting, strong recommendations ($\theta = 1$) significantly reduce simple contagion size, particularly for AA and GCN. However, weakening recommendation strength ($\theta = 0.5$) mitigates this effect, making contagion dynamics resemble those under organic growth.

Importance of modeling organic growth. Figure 1 highlights the necessity of organic growth as a baseline for evaluating recommendation effects over time. Under organic growth, simple contagions start with almost no spread at $t = 0$ but expand to reach nearly all nodes by $t = 4000$. Without this baseline, one might mistakenly conclude that PA-driven recommendations enhance contagion size compared to AA and GCN. However, relative to organic growth, all recommendation algorithms actually reduce contagion size, with PA having the smallest negative impact.

Effects on Complex Contagion. For complex contagions, organic growth initially increases contagion size before it declines over time. Recommendations significantly alter this trajectory, even at moderate strength ($\theta = 0.5$):

- **AA:** At $\theta = 0.5$, AA increases contagion size, likely by enhancing local clustering, which benefits complex contagions. However, at $\theta = 1$, this effect reverses, significantly reducing contagion size. This suggests that high initial homophily combined with excessive clustering dampens complex contagion spread.
- **GCN:** At $\theta = 0.5$, GCN slightly boosts contagion size, but at $\theta = 1$, it strongly suppresses it. Similar to AA, the interaction between initial homophily and GCN's recommendations has a pronounced negative impact as θ increases.

Implications of Recommendation Strength. This analysis highlights the importance of modeling recommendation strength. Without varying θ, one might incorrectly generalize that AA and GCN recommendations harm complex contagions at $\theta = 1$. However, reducing $\theta = 0.5$ alters the effects of these algorithms, shifting their implications for contagion processes.

5.2 VARYING NETWORK TOPOLOGY

To study the effect of link recommendations on social contagion, it is crucial to consider networks with different topologies. In this analysis, we vary triadic closure levels β and set $\gamma = 1 - \beta$. For each combination of β, algorithm, and recommendation strength θ, we track the average difference in contagion size relative to organic growth, denoted as ΔCS. Figure 2 illustrates how ΔCS varies for AA and GCN across different β values.

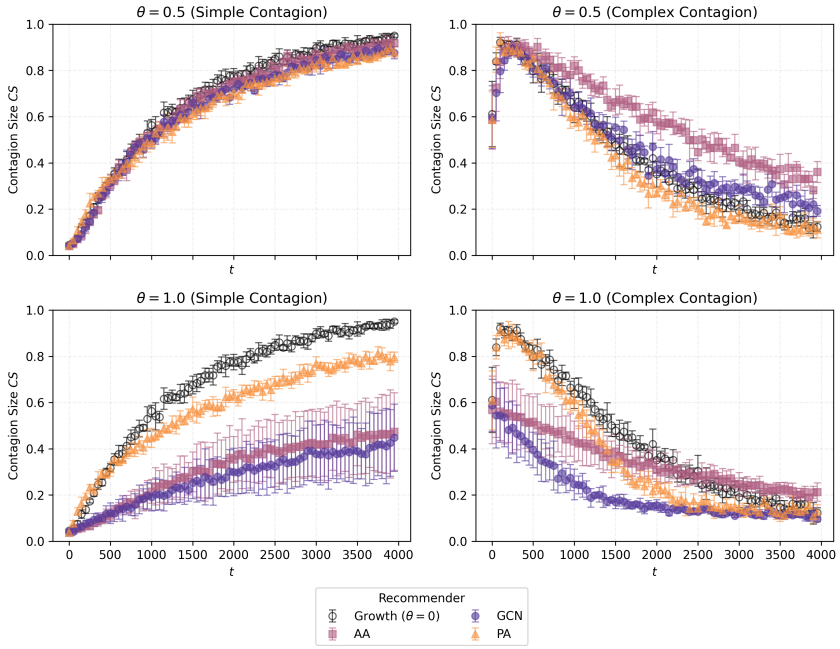


Figure 1: The impact of link recommendations on simple and complex contagion for a strong homophilous graph with $\gamma = 1$. The first row show the result for simple and complex contagions with recommendations strength $\theta = 0.5$. For the second row $\theta = 1.0$. The organic growth ($\theta = 0$) is included in all plots as a reference point. Error bars represent standard deviations calculated across 5 runs of the same parameters.

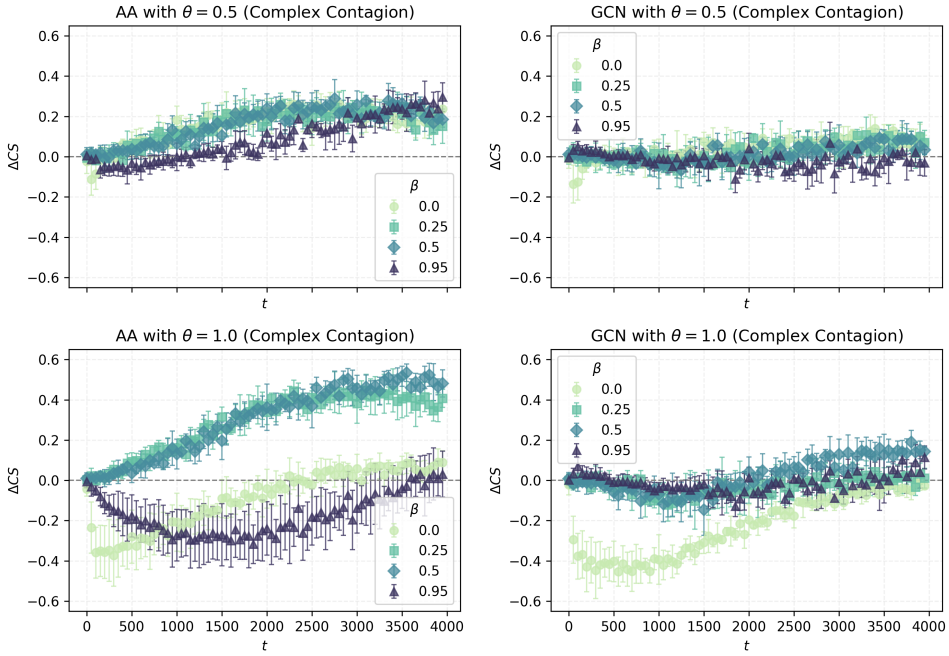
AA. Similar to the findings of the strong homophily scenario, there seems to be interplay between initial levels clustering and additional clustering. While for moderate recommendation strengths ($\theta = 0.5$), ΔCS is generally positive across all levels of β, this changes for $\theta = 1$. Specifically, for moderate triadic closure ($\beta = 0.5$), AA has strongly positive effects on ΔCS. However, for extreme clustering ($\beta = 0.95$) or strong homophily ($\beta = 0.0$), ΔCS becomes mostly negative. In strong homophily networks ($\beta = 0.0$), ΔCS tends towards zero over time, while in highly clustered networks ΔCS remains negative for longer. This implies that for moderate clustering, strong increases in clustering ($\theta = 1$) enhance complex contagion, while in highly clustered networks, such increases can actually reduce the spread of complex contagion compared to organic growth.

GCN. For moderate recommendation strength, GCN has minimal effects on ΔCS across most β values, with differences remaining close to zero. Even at $\theta = 1$, GCN's impact remains minor for most β values. However, for moderate clustering, GCN shows a moderate positive effect on ΔCS. As previously observed, in strong homophily networks, GCN's impact becomes highly negative.

Our results show that the effects of link recommendations vary significantly across network topologies, recommender algorithms, and recommendation strength relative to organic growth.

For **simple contagions**, recommendation algorithms generally have only minor effects on contagion size as the network evolves. In contrast, **complex contagions** exhibit much stronger effects, particularly for algorithms that enhance clustering (AA) or rely on graph neural networks (GCN).

Recommendation strength also plays a crucial role. While stronger recommendations typically amplify differences, in some cases, their effects reverse. In highly clustered networks, moderate recommendation strength can enhance complex contagions, but strong recommendations through AA can suppress them. Complex contagions typically thrive in clustered networks (Centola & Macy, 2007), yet excessive clustering appears to hinder their spread. Similarly, in strong homophily networks, GCN-based recommendations—which effectively capture homophily (Zhang & Chen, 2018; Palowitch et al., 2022)—initially boost contagion size at moderate strengths but strongly suppress it at higher strengths.



(a) Adamic Adar (b) GCN

Figure 2: The average contagion size difference ΔCS for varying levels of β and $\gamma = 1 - \beta$. The left column shows the differences for the Adamic Adar (AA) recommender, the right for GCN. The first row shows recommendations strength $\theta = 0.5$. For the second row $\theta = 1.0$.

These strong and varied effects across algorithms and network topologies highlight the **importance of considering social contagion in platform design.** However, deriving clear normative implications is challenging. Simple contagions can drive both the spread of valuable information (Bakshy et al., 2012) and misinformation (Vosoughi et al., 2018), making their amplification both beneficial and problematic. Similarly, while complex contagions are often viewed positively, they can also drive harmful behaviors (e.g., radicalization or violent movements). Given concerns that online platforms increasingly favor simple contagions (Centola, 2018), our findings underscore the need to carefully assess how recommendations affect complex contagions across different network structures.

6 Limitations

Synthetic and Real World Data While our synthetic experiments allow for systematic exploration of underlying mechanisms, they abstract away some complexities of online social networks. Validating our findings with real-world data remains a challenge due to the difficulty of directly observing social contagion processes and the limited availability of comprehensive datasets. A promising alternative is semi-synthetic validation, where real-world networks with features serve as the foundation for simulating network growth and recommendation processes (Santos et al., 2021; Wang & Kleinberg, 2024). This hybrid approach could help bridge the gap between fully synthetic models and empirical validation. Future work should also explore scaling these experiments to larger, more realistic network sizes.

Evaluating Social Contagion. Our study relies on extensive simulations of simple and complex contagions across various parameter settings. However, as network size and density increase, these simulations become computationally expensive, raising the open question of how to efficiently evaluate contagion processes in large-scale datasets. Additionally, integrating the evaluation of simple and complex contagions into a unified measure remains an open challenge.

Modeling Recommendation Mechanisms We use simplified recommendation mechanisms to model link recommendations, which provide valuable insights but lack the complexity of real-world systems. Future work should explore dynamic, adaptive GNN-based recommenders that evolve with the network over time, incorporating retraining mechanisms to study emerging biases as social networks and algorithms coevolve (Pedreschi et al., 2023).

Moreover, future studies should consider diverse recommendation acceptance policies, where nodes selectively accept or reject recommendations based on different criteria. This could lead to varied network dynamics and help construct more realistic synthetic models for understanding user behavior in recommendation-driven systems.

7 Conclusion

This work examined how link recommendation algorithms influence social contagion across various network topologies. Our findings reveal that especially complex contagions are sensitive to link recommendations. Recommendation strategies that enhance clustering or leverage graph neural networks can substantially shape the spread of complex contagions. Notably, increasing recommendation strength does not always amplify contagion; its effects critically depend on network structure. In highly clustered or strong homophily networks, moderate recommendation strengths may facilitate contagion spread, whereas excessive clustering from stronger recommendations can hinder it.

These insights highlight the importance of incorporating social contagion dynamics into platform design. Recommendation systems influence both beneficial (e.g., novel idea dissemination) and harmful (e.g., misinformation, radicalization) contagions, requiring careful calibration. Our findings demonstrate that recommendation design should account for more than just the algorithms, but also consider the underlying network structure, the composition of growth mechanisms, and the extent to which users rely on recommendations when forming new connections. Considering these dimensions can help platforms better understand the broader impact of their recommendation systems and represents a crucial step towards informed and responsible design choices.

In summary, this work advances the study of AI-driven social dynamics by illustrating the nuanced interplay between recommendation algorithms and social contagion. While grounded in synthetic simulations, our findings raise important questions about real-world systems. Future work will focus on validating these insights through semi-synthetic experiments using real-world networks, and on integrating more realistic recommender behaviors and user response models. Ultimately, this work lays the groundwork for building recommendation systems that balance connectivity with the mitigation of harmful contagion effects.

Acknowledgments

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nº 101034328.

References

Lada A Adamic and Eytan Adar. Friends and neighbors on the Web. *Social Networks*, 25(3): 211–230, July 2003. ISSN 0378-8733. doi: 10.1016/S0378-8733(03)00009-1. URL https://www.sciencedirect.com/science/article/pii/S0378873303000091.

Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, pp. 519–528, New York, NY, USA, April 2012. Association for Computing Machinery. ISBN 978-1-4503-1229-5. doi: 10.1145/2187836.2187907. URL https://doi.org/10.1145/2187836.2187907.

A. L Barabási, H Jeong, Z Néda, E Ravasz, A Schubert, and T Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3): 590–614, August 2002. ISSN 0378-4371. doi: 10.1016/S0378-4371(02)00736-7. URL https://www.sciencedirect.com/science/article/pii/S0378437102007367.

Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *Science*, 286 (5439):509–512, October 1999. doi: 10.1126/science.286.5439.509. URL https://www.science.org/doi/abs/10.1126/science.286.5439.509. Publisher: American Association for the Advancement of Science.

Ginestra Bianconi, Richard K. Darst, Jacopo Iacovacci, and Santo Fortunato. Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E*, 90(4): 042806, October 2014. doi: 10.1103/PhysRevE.90.042806. URL https://link.aps.org/doi/10.1103/PhysRevE.90.042806. Publisher: American Physical Society.

Fedor Borisyuk, Shihai He, Yunbo Ouyang, Morteza Ramezani, Peng Du, Xiaochen Hou, Chengming Jiang, Nitin Pasumarthy, Priya Bannur, Birjodh Tiwana, Ping Liu, Siddharth Dangi, Daqi Sun, Zhoutao Pei, Xiao Shi, Sirou Zhu, Qianqi Shen, Kuang-Hsuan Lee, David Stein, Baolei Li, Haichao Wei, Amol Ghoting, and Souvik Ghosh. LiGNN: Graph Neural Networks at LinkedIn. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, pp. 4793–4803, New York, NY, USA, August 2024. Association for Computing Machinery. ISBN 9798400704901. doi: 10.1145/3637528.3671566. URL https://dl.acm.org/doi/10.1145/3637528.3671566.

Giulia Cencetti, Diego Andrés Contreras, Marco Mancastroppa, and Alain Barrat. Distinguishing Simple and Complex Contagion Processes on Networks. *Physical Review Letters*, 130(24): 247401, June 2023. doi: 10.1103/PhysRevLett.130.247401. URL https://link.aps.org/doi/10.1103/PhysRevLett.130.247401. Publisher: American Physical Society.

Damon Centola. The Spread of Behavior in an Online Social Network Experiment. *Science*, 329 (5996):1194–1197, September 2010. doi: 10.1126/science.1185231. URL https://www.science.org/doi/full/10.1126/science.1185231. Publisher: American Association for the Advancement of Science.

Damon Centola. How Behavior Spreads: The Science of Complex Contagions. In *How Behavior Spreads*. Princeton University Press, June 2018. ISBN 978-1-4008-9009-5. doi: 10.23943/9781400890095. URL https://www.degruyter.com/document/doi/10.23943/9781400890095/html.

Damon Centola and Michael Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734, November 2007. ISSN 0002-9602. doi: 10.1086/521848. URL https://www.journals.uchicago.edu/doi/abs/10.1086/521848. Publisher: The University of Chicago Press.

Sergio Currarini, Jesse Matheson, and Fernando Vega-Redondo. A simple model of homophily in social networks. *European Economic Review*, 90:18–39, November 2016. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2016.03.011. URL https://www.sciencedirect.com/science/article/pii/S0014292116300642.

David Easley and Jon Kleinberg. Networks, Crowds, and Markets. *Cambridge Books*, 2010. URL https://ideas.repec.org//b/cup/cbooks/9780521195331.html. Publisher: Cambridge University Press.

Lisette Espín-Noboa, Claudia Wagner, Markus Strohmaier, and Fariba Karimi. Inequality and inequity in network-based ranking and recommendation algorithms. *Scientific Reports*, 12(1): 2012, February 2022. ISSN 2045-2322. doi: 10.1038/s41598-022-05434-1. URL https://www.nature.com/articles/s41598-022-05434-1. Publisher: Nature Publishing Group.

Antonio Ferrara, Lisette Espin-Noboa, Fariba Karimi, and Claudia Wagner. Link recommendations: Their impact on network structure and minorities. In *Proceedings of the 14th ACM Web Science Conference 2022*, WebSci '22, pp. 228–238, New York, NY, USA, June 2022. Association for Computing Machinery. ISBN 978-1-4503-9191-7. doi: 10.1145/3501247.3531583. URL https://dl.acm.org/doi/10.1145/3501247.3531583.

Clay Fink, Aurora Schmidt, Vladimir Barash, John Kelly, Christopher Cameron, and Michael Macy. Investigating the Observability of Complex Contagion in Empirical Social Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 10(1):121–130, 2016. ISSN 2334-0770. doi: 10.1609/icwsm.v10i1.14751. URL https://ojs.aaai.org/index.php/ICWSM/article/view/14751. Number: 1.

Jacob Goldenberg, Barak Libai, and Eitan Muller. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters*, 12(3):211–223, August 2001a. ISSN 1573-059X. doi: 10.1023/A:1011122126881. URL https://doi.org/10.1023/A:1011122126881.

Jacob Goldenberg, Barak Libai, and Eitan Muller. Using Complex Systems Analysis to Advance Marketing Theory Development: Modeling Heterogeneity Effects on New Product Growth through Stochastic Cellular Automata. 2001b.

Sandra González-Bailón, Javier Borge-Holthoefer, Alejandro Rivero, and Yamir Moreno. The Dynamics of Protest Recruitment through an Online Network. *Scientific Reports*, 1(1):197, December 2011. ISSN 2045-2322. doi: 10.1038/srep00197. URL https://www.nature.com/articles/srep00197. Publisher: Nature Publishing Group.

Mark Granovetter. Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443, May 1978. ISSN 0002-9602. doi: 10.1086/226707. URL https://www.journals.uchicago.edu/doi/10.1086/226707. Publisher: The University of Chicago Press.

Mark S. Granovetter. The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380, May 1973. ISSN 0002-9602. doi: 10.1086/225469. URL https://www.journals.uchicago.edu/doi/abs/10.1086/225469. Publisher: The University of Chicago Press.

Huan Gui, Ya Xu, Anmol Bhasin, and Jiawei Han. Network A/B Testing: From Sampling to Estimation. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, pp. 399–409, Republic and Canton of Geneva, CHE, May 2015. International World Wide Web Conferences Steering Committee. ISBN 978-1-4503-3469-3. doi: 10.1145/2736277.2741081. URL https://dl.acm.org/doi/10.1145/2736277.2741081.

Douglas Guilbeault and Damon Centola. Topological measures for identifying and predicting the spread of complex contagions. *Nature Communications*, 12(1):4430, July 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24704-6. URL https://www.nature.com/articles/s41467-021-24704-6. Publisher: Nature Publishing Group.

Douglas Guilbeault, Joshua Becker, and Damon Centola. Complex Contagions: A Decade in Review. In Sune Lehmann and Yong-Yeol Ahn (eds.), *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, pp. 3–25. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77332-2. doi: 10.1007/978-3-319-77332-2_1. URL https://doi.org/10.1007/978-3-319-77332-2_1.

Matthew O. Jackson and Brian W. Rogers. Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review*, 97(3):890–915, June 2007. ISSN 0002-8282. doi: 10.1257/aer.97.3.890. URL https://www.aeaweb.org/articles?id=10.1257/aer.97.3.890.

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pp. 137–146, New York, NY, USA, August 2003. Association for Computing Machinery. ISBN 978-1-58113-737-8. doi: 10.1145/956750.956769. URL https://doi.org/10.1145/956750.956769.

Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks, February 2017. URL http://arxiv.org/abs/1609.02907. arXiv:1609.02907.

Gueorgi Kossinets and Duncan J. Watts. Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*, 115(2):405–450, September 2009. ISSN 0002-9602. doi: 10.1086/599247. URL https://www.journals.uchicago.edu/doi/full/10.1086/599247. Publisher: The University of Chicago Press.

David Lazer. The rise of the social algorithm. *Science*, 348(6239):1090–1091, June 2015. doi: 10.1126/science.aab1422. URL https://www.science.org/doi/abs/10.1126/science.aab1422. Publisher: American Association for the Advancement of Science.

Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, 1(1):2–es, March 2007. ISSN 1556-4681. doi: 10.1145/1217299.1217301. URL https://dl.acm.org/doi/10.1145/1217299.1217301.

Yanying Li, Xiuling Wang, Yue Ning, and Hui Wang. FairLP: Towards Fair Link Prediction on Social Network Graphs. *Proceedings of the International AAAI Conference on Web and Social Media*, 16:628–639, May 2022. ISSN 2334-0770. doi: 10.1609/icwsm.v16i1.19321. URL https://ojs.aaai.org/index.php/ICWSM/article/view/19321.

Zhepeng (Lionel) Li, Xiao Fang, and Olivia R. Liu Sheng. A Survey of Link Recommendation for Social Networks: Methods, Theoretical Foundations, and Future Research Directions. *ACM Trans. Manage. Inf. Syst.*, 9(1):1:1–1:26, October 2017. ISSN 2158-656X. doi: 10.1145/3131782. URL https://dl.acm.org/doi/10.1145/3131782.

David Liben-Nowell and Jon Kleinberg. The link prediction problem for social networks. In *Proceedings of the twelfth international conference on Information and knowledge management*, pp. 556–559, New Orleans LA USA, November 2003. ACM. ISBN 978-1-58113-723-1. doi: 10.1145/956863.956972. URL https://dl.acm.org/doi/10.1145/956863.956972.

Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(Volume 27, 2001):415–444, August 2001. ISSN 0360-0572, 1545-2115. doi: 10.1146/annurev.soc.27.1.415. URL https://www.annualreviews.org/content/journals/10.1146/annurev.soc.27.1.415. Publisher: Annual Reviews.

Seth A. Myers, Chenguang Zhu, and Jure Leskovec. Information diffusion and external influence in networks. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '12, pp. 33–41, New York, NY, USA, August 2012. Association for Computing Machinery. ISBN 978-1-4503-1462-6. doi: 10.1145/2339530.2339540. URL https://dl.acm.org/doi/10.1145/2339530.2339540.

Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, and Sune Lehmann. Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PLOS ONE*, 12(9):e0184148, September 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0184148. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0184148. Publisher: Public Library of Science.

Azadeh Nematzadeh, Emilio Ferrara, Alessandro Flammini, and Yong-Yeol Ahn. Optimal Network Modularity for Information Diffusion. *Physical Review Letters*, 113(8):088701, August 2014. doi: 10.1103/PhysRevLett.113.088701. URL https://link.aps.org/doi/10.1103/PhysRevLett.113.088701. Publisher: American Physical Society.

M. E. J. Newman. Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2):025102, July 2001. doi: 10.1103/PhysRevE.64.025102. URL https://link.aps.org/doi/10.1103/PhysRevE.64.025102. Publisher: American Physical Society.

John Palowitch, Anton Tsitsulin, Brandon Mayer, and Bryan Perozzi. GraphWorld: Fake Graphs Bring Real Insights for GNNs, July 2022. URL http://arxiv.org/abs/2203.00112. arXiv:2203.00112.

Luca Pappalardo, Emanuele Ferragina, Salvatore Citraro, Giuliano Cornacchia, Mirco Nanni, Giulio Rossetti, Gizem Gezici, Fosca Giannotti, Margherita Lalli, Daniele Gambetta, Giovanni Mauro,

Virginia Morini, Valentina Pansanella, and Dino Pedreschi. A survey on the impact of AI-based recommenders on human behaviours: methodologies, outcomes and future directions, June 2024. URL http://arxiv.org/abs/2407.01630. arXiv:2407.01630 [cs].

Dino Pedreschi, Luca Pappalardo, Emanuele Ferragina, Ricardo Baeza-Yates, Albert-Laszlo Barabasi, Frank Dignum, Virginia Dignum, Tina Eliassi-Rad, Fosca Giannotti, Janos Kertesz, Alistair Knott, Yannis Ioannidis, Paul Lukowicz, Andrea Passarella, Alex Sandy Pentland, John Shawe-Taylor, and Alessandro Vespignani. Human-AI Coevolution, June 2023. URL https://ui.adsabs.harvard.edu/abs/2023arXiv230613723P. Publication Title: arXiv e-prints ADS Bibcode: 2023arXiv230613723P.

Karthik Rajkumar, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral. A causal test of the strength of weak ties. *Science*, 377(6612):1304–1310, September 2022. doi: 10.1126/science.abl4476. URL https://www.science.org/doi/10.1126/science.abl4476. Publisher: American Association for the Advancement of Science.

Everett M. Rogers, Arvind Singhal, and Margaret M. Quinlan. Diffusion of Innovations. In *An Integrated Approach to Communication Theory and Research*. Routledge, 2 edition, 2008. ISBN 978-0-203-88701-1. Num Pages: 17.

Fernando P. Santos, Yphtach Lelkes, and Simon A. Levin. Link recommendation algorithms and dynamics of polarization in online social networks. *Proceedings of the National Academy of Sciences*, 118(50):e2102141118, December 2021. doi: 10.1073/pnas.2102141118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2102141118. Publisher: Proceedings of the National Academy of Sciences.

Ana-Andreea Stoica, Christopher Riederer, and Augustin Chaintreau. Algorithmic Glass Ceiling in Social Networks: The effects of social recommendations on network diversity. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, pp. 923–932, Republic and Canton of Geneva, CHE, April 2018. International World Wide Web Conferences Steering Committee. ISBN 978-1-4503-5639-8. doi: 10.1145/3178876.3186140. URL https://dl.acm.org/doi/10.1145/3178876.3186140.

Jessica Su, Aneesh Sharma, and Sharad Goel. The Effect of Recommendations on Network Structure. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pp. 1157–1167, Republic and Canton of Geneva, CHE, April 2016. International World Wide Web Conferences Steering Committee. ISBN 978-1-4503-4143-1. doi: 10.1145/2872427.2883040. URL https://dl.acm.org/doi/10.1145/2872427.2883040.

Jessica Su, Krishna Kamath, Aneesh Sharma, Johan Ugander, and Sharad Goel. An Experimental Study of Structural Diversity in Social Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 14:661–670, May 2020. ISSN 2334-0770. doi: 10.1609/icwsm.v14i1.7332. URL https://ojs.aaai.org/index.php/ICWSM/article/view/7332.

Arjun Subramonian, Levent Sagun, and Yizhou Sun. Networked Inequality: Preferential Attachment Bias in Graph Neural Network Link Prediction, June 2024. URL http://arxiv.org/abs/2309.17417. arXiv:2309.17417 [cs].

Anton Tsitsulin, Benedek Rozemberczki, John Palowitch, and Bryan Perozzi. Synthetic Graph Generation to Benchmark Graph Learning, April 2022. URL http://arxiv.org/abs/2204.01376. arXiv:2204.01376.

Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, March 2018. doi: 10.1126/science.aap9559. URL https://www.science.org/doi/10.1126/science.aap9559. Publisher: American Association for the Advancement of Science.

Yanbang Wang and Jon Kleinberg. On the Relationship Between Relevance and Conflict in Online Social Link Recommendations. *Advances in Neural Information Processing Systems*, 36, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/73d6c3e4b214deebbbf8256e26d2cf45-Abstract-Conference.html.

Yu Wang and Tyler Derr. Degree-Related Bias in Link Prediction. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 757–758, November 2022. doi: 10.1109/ICDMW58026.2022.00103. URL https://ieeexplore.ieee.org/abstract/document/10031056?casa_token=l9ovWRDVNNYAAAAA:IUIlnR_K6RBVrt10K36nQ55m4RqLNq957lzAGuBab949O89NkcTq74EGjmuxFHEj-2ePRAe5wTk. ISSN: 2375-9259.

Han Zhang, Shangen Lu, Yixin Wang, and Mihaela Curmei. Delayed and Indirect Impacts of Link Recommendations. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '23, pp. 545–557, New York, NY, USA, June 2023. Association for Computing Machinery. ISBN 9798400701924. doi: 10.1145/3593013.3594021. URL https://dl.acm.org/doi/10.1145/3593013.3594021.

Muhan Zhang and Yixin Chen. Link Prediction Based on Graph Neural Networks. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper/2018/hash/53f0d7c537d99b3824f0f99d62ea2428-Abstract.html.

A Appendix

You may include other additional sections here.

A.1 Detailed Node Feature Generation Process

We generate node features using the group or block memberships z_i of nodes in the following steps:

Each node is assigned a group membership $z_i \in \{1, \ldots, K\}$, defining which community it belongs to.

Sampling Steps. 1. **Group Centers:** For each group k, we sample a center vector $\mu_k \in \mathbb{R}^D$:

$$\mu_k \sim \mathcal{N}(\mathbf{0}, \sigma_c^2 \mathbf{I}_D),$$

where $\mathbf{I}_D$ is the D-dimensional identity matrix, and σ_c^2 controls the spread of the group centers around the origin.

2. **Covariance Matrices:** For each group k, we sample a covariance matrix $\mathbf{\Sigma}_k$ from the inverse Wishart distribution:

$$\mathbf{\Sigma}_k \sim \mathcal{W}^{-1}(\nu, \sigma_K^2 \mathbf{I}_D),$$

where σ_K^2 determines the spread within each cluster, and ν is chosen to ensure stability (typically $\nu \geq D + 2$).

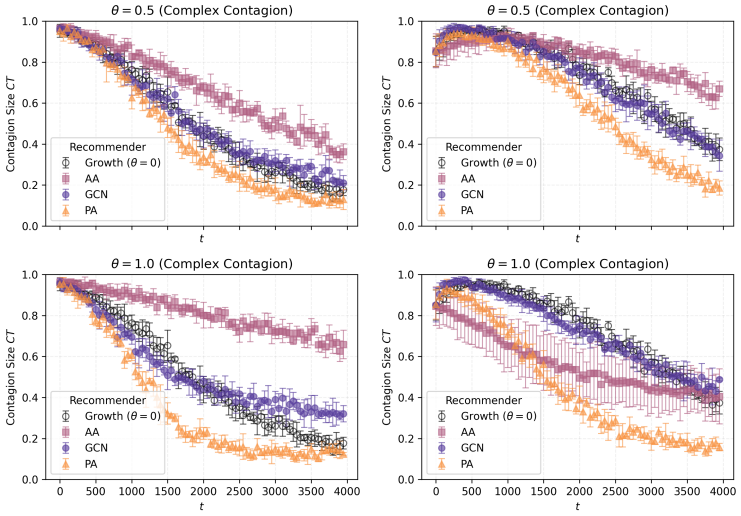
3. **Node Features:** For each node i in group k, we generate a feature vector $\mathbf{x}_i$ by sampling from a multivariate normal distribution:

$$\mathbf{x}_i \sim \mathcal{N}(\mu_k, \mathbf{\Sigma}_k).$$

Final Feature Matrix. The resulting feature matrix $\boldsymbol{X} \in \mathbb{R}^{N \times D}$, where N is the total number of nodes, is constructed by stacking the feature vectors $\mathbf{x}_i$ for all nodes.

A.2 Additional Results

Figure 3 shows changes in infection rates across two different network topologies. In figure 3a, the network is generated with $\beta = 0.5$ and $\gamma = 0.5$. Here PA always decreases the average infection size of a complex contagion, with stronger effects as the strength of recommendations increases. In contrast AA always increases the average infection sizes of a complex contagion with stronger effects as the strength of recommendation increases. However as we change $\beta = 0.95$ and $\gamma = 0.05$, the initial positive effect of AA for $\theta = 0.5$ turns strongly negative for $\theta = 1$. PA remains strongly negative.



(a) $\alpha = 0$ $\beta = 0.5$ and $\gamma = 0.5$. (b) $\alpha = 0$, $\beta = 0.95$ and $\gamma = 0.05$.

Figure 3: The impact of link recommendations on complex contagion processes for different levels of clustering. Figure 3a (left) corresponds to a network grown with equal parts of triadic closure β and homophily γ. Figure 3b corresponds to a network with high clustering ($\beta = 0.95$). The first rows shows medium recommendations strengths ($\theta = 0.5$), the second pure recommendations ($\theta = 1.0$). The errorbars indicate standard deviations across 5 runs of different networks with the same growth parameters. For mixed clustering and increase in recommendation strength can leads to stronger positive effect in the case of AA, and stronger negative effects in the case of PA. However for the network with high clustering, an increase in recommendation strength turns the initial positive effect of AA from positive at $\theta = 0.5$ to negative at $\theta = 1$.

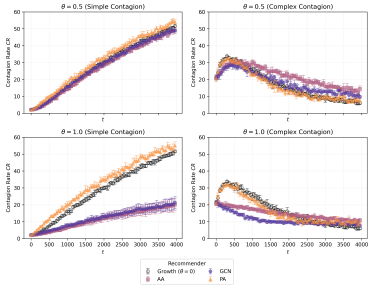
A.3 Results for Infection Rates

If we consider the infection rate instead of the infection size, the differences between algorithms become partly clearer. PA across all topologies generally increases the rate of infection for simple contagions and decreases for complex contagions. For simple contagions, the effect of AA and GCN mostly align, with expection of the strong clustering case with $\theta = 1$, where the infection rate of AA is significantly lower. For complex contagion AA has positive effects besides the strong clustering case and the strong homophilous case with $\theta = 1$. The changes of complex contagion infection rates of the GCN recommender mostly align with the organic growth. Exceptions are for high levels of γ. For $\gamma = 0.75$ or $\gamma = 1$, and recommendation strength $\theta = 1$, GCN decreases infection rates significantly. This difference is further illustrated in figure 5, which summarizes above findings.

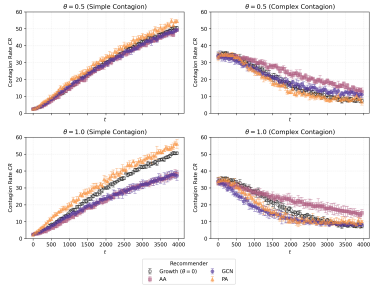
A.4 Impact of Recommendations on Network Structure

To understand *why* the impact of link recommendations on social contagions differs across network topologies, recommender algorithms and strengths as well as simple and complex contagions, a first step is to consider changes in the network structure. In similar fashion to figure 2, we compare the difference in average clustering and average shortest path length across different β, γ combinations in figure 6. Differences refer to organic growth ($\theta = 0$). A first observation is that the differences in graph measures are much stronger than the differences in contagion sizes might suggest. For instance differences in average clustering are up to 0.7 high while the average shortest path length can increase up to 2. We find that the average clustering in figure 6a increases across all β and θ configurations for AA. The increases are stronger for $\theta = 1$. Similarly for GCN the average clustering increases with recommendations. However, the effects are not as strong as for AA. For $\beta = 0.95$ the increases of clustering are the smallest, with partially negative effects.

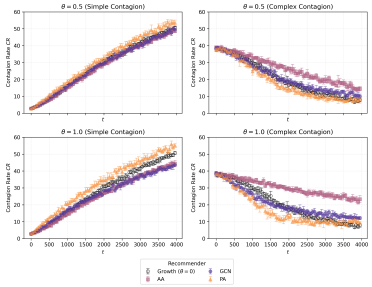
For average shortest path length, we find that AA significantly increases the length of average shortest paths. However these effects are only visible for recommendation strength $\theta = 1$. Moreover,



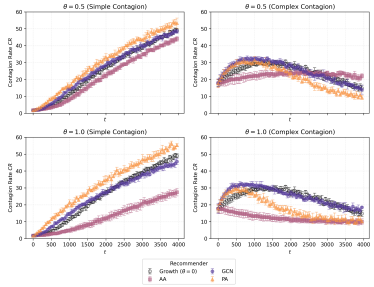
(a) $\alpha, \beta = 0, \gamma = 1$



(b) $\alpha, \beta = 0.25, \gamma = 0.75$

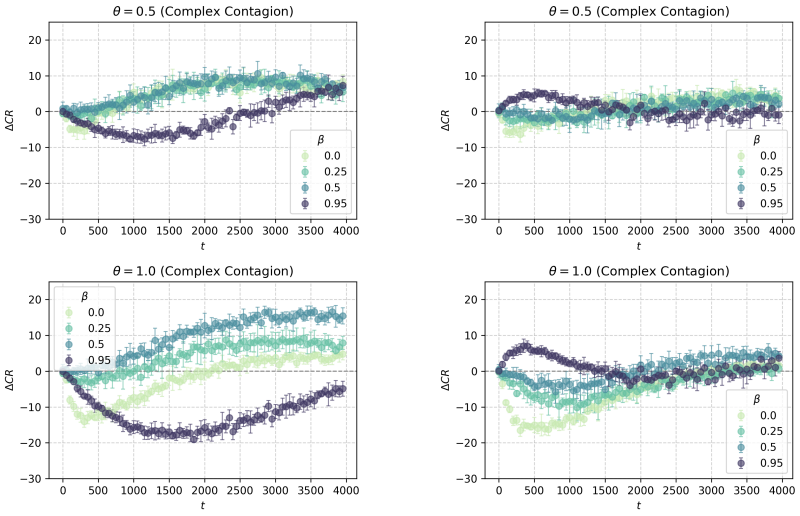


(c) $\alpha, \beta = 0.5, \gamma = 0.5$



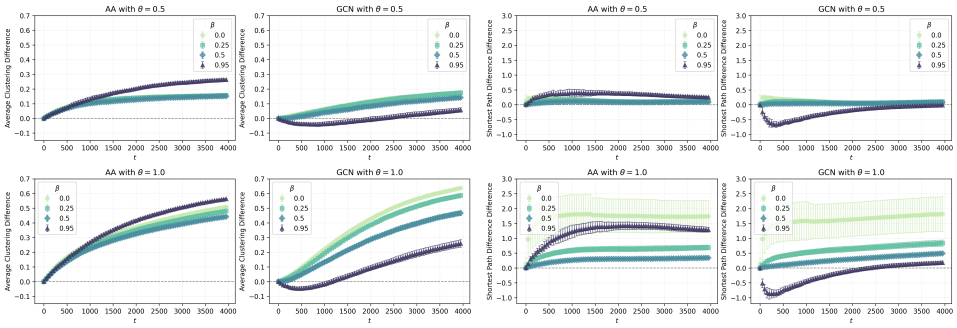
(d) $\alpha, \beta = 0.95, \gamma = 0.05$

Figure 4: The effect of link recommendations on **the rate of infection** for both simple and complex contagion processes. We consider the strong homophilous case 4a, low clustering 4b, medium clustering 4c and strong clustering 4d.



(a) Adamic Adar (b) GCN

Figure 5: Contagion Rate differences ΔCS for varying levels of β. The first row presents differences for $\theta = 0.5$, the second for $\theta = 1$.



(a) Average Clustering (b) Average Shortest Path Length

Figure 6: Differences of graph measures for varying levels of β and $\gamma = 1 - \beta$. 6a left shows the difference of average clustering and 6b differences in average shortest path length. For each subfigure, the left column shows the differences for Adamic Adar (AA) recommender, the right for GCN. The first row shows recommendations strength $\theta = 0.5$. For the second row $\theta = 1.0$.

these differences are more pronounced for the high homphily and the high clustering network. The same holds for GCN, however there are some initial decreases in shortest path length for the highly clustered network.