Dynamics of Large Multi-View Social Networks: Synergy, Cannibalization and Cross-View Interplay

Yu Shi[†*] Myunghwan Kim[‡*] Shaunak Chatterjee[‡] Mitul Tiwari[§] Souvik Ghosh[‡] Rómer Rosales[‡]
[†]Dept. of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL USA
[‡§]LinkedIn Corporation, Mountain View, CA USA
[†]yushi2@illinois.edu [‡]{mukim, shchatterjee, sghosh, rrosales}@linkedin.com [§]mitultiwari@gmail.com

ABSTRACT

Most social networking services support multiple types of relationships between users, such as getting connected, sending messages, and consuming feed updates. These users and relationships can be naturally represented as a dynamic multi-view network, which is a set of weighted graphs with shared common nodes but having their own respective edges. Different network views, representing structural relationship and interaction types, could have very distinctive properties individually and these properties may change due to interplay across views. Therefore, it is of interest to study how multiple views interact and affect network dynamics and, in addition, explore possible applications to social networking.

In this paper, we propose approaches to capture and analyze multi-view network dynamics from various aspects. Through our proposed descriptors, we observe the synergy and cannibalization between different user groups and network views from LinkedIn dataset. We then develop models that consider the synergy and cannibalization per new relationship, and show the outperforming predictive capability of our models compared to baseline models. Finally, the proposed models allow us to understand the interplay among different views where they dynamically change over time.

Keywords

Multi-view networks, network dynamics, social networks.

1. INTRODUCTION

One of the most important challenges in large social networks is to help users create and curate vibrant and healthy networks of their own [31, 32]. Such vibrancy entails much more than mere structural relationships – interactions along the relationships are critical, as are the distribution of the interactions over time and relationships [5, 16]. In order to build such vibrant networks, which can be expressed in terms of interactions, relationships and their dynamics, we not only need to understand how networks evolve structurally, but also how interactions on such networks evolve with those structural changes.

In this paper, we focus our attention on studying this dynamic evolution of a multi-view network [2, 12, 13, 30]. The nodes represent users, while the multiple views fall into two categories: structural association among users such as connection and follow relation – the *relationship views* – as well as various interaction types along the relationships including message exchange, content broadcast and consumption – the *interaction views*. In particular, we believe that the values among linked users are primarily achieved by the interactions along the relationships, not just by the formation of the relationships. Given this hypothesis, by paying more attention to the study of interaction views, we aspire to help users obtain the value from their relationships.

To address the objective of building a good social network for a user, we identify a few key building blocks, which are expressive and flexible enough to (i) describe the current state and the dynamics of the user's ego network, and (ii) define what a good network is for the user. The first block represents the ties between two groups of users – the ability to specify or shape the dynamics of relationships and interactions between the two groups. The second block indicates the association between different views – how the multiple views interplay with one another. While these two blocks can be used to express both the current state and dynamics of a network, they can also be used to define instances of a "good" network. An instance of such a "good" network can be one where interactions grow monotonically with relationships.

For each of these two types of blocks, we study two primary evolution effects in more detail – *synergy* and *cannibalization* [7, 17, 23] – among users within a certain view as well as across network views. We focus on synergy and cannibalization specifically since these are the two ends of the spectrum defining the pairwise relationships for each of the two building blocks.

Synergy is defined as the mutually enhancing effect that two entities have on one another. For example, consider a user x who is currently connected to users y_1 and y_2, and exchanges messages with them 10 times a week. By adding a new connection z, x now exchanges messages 15 times a week with y_1 and y_2 (her old connections) and 5 times a week with z (her new connection). In such a scenario, x's existing connections and her new connection are said to have synergy in the message exchange view. Similarly, synergy across views is defined to be the positive influence that two views have on one another – *i.e.*, the presence of either is likely to enhance the value of the other.

Cannibalization is the very antithesis to synergy – when two entities negatively impact one another. For instance, in the previous example, if by connecting to z, x's messaging rate fell to 5 times a week with y_1 and y_2, then the connection z would have cannibalized the messaging view between x and her old connections. Cannibalization across views refers to the situation where the presence

[*]These authors contributed equally to this work.

KDD '16, August 13–17, 2016, San Francisco, CA, USA

ISBN 978-1-4503-4232-2/16/08…$15.00

DOI: http://dx.doi.org/10.1145/2939672.2939814

of either is likely to decrease the value of the other.

In addition to the pairwise relationships of synergy and cannibalization, we also study how a view evolves given available observations in other views. We call this *cross-view interplay*.

When describing a current network state, or reasoning about potential network evolution and try to drive it in a certain direction, understanding the two blocks – user groups and views – are very important. For instance, in the context of network growth, a user's existing connections and a group of potential new connections may be of the most interest. A good network might involve balancing the interactions in these two groups, *i.e.*, aiming to minimize cannibalization among them, or better still create synergy.

Similarly, a good network might concern not only the number of connections (a relationship view) but also the number of messages being exchanged (an interaction view). The different interaction views are important pillars of a user's overall experience with the network, and the ability to shape the composition among views is critical to our objective to create the aforementioned "vibrancy".

In addition to identifying these building blocks to express network state and dynamics with the purpose of building a good network, we also propose a unified formulation to model synergy, cannibalization and cross-view dynamics.

Through detailed data analysis on a large professional network, we demonstrate that synergy and cannibalization exist among user groups as well as among views. We also show that our unified model can be used to predict user activity level and achieve large performance gain compared to baseline methods.

The rest of the paper is structured as follows. We cover some related work in Section 2, before introducing some notation and definitions in Section 3. In the next section, we propose techniques for make observation on multi-view dynamic networks and describe a unified model for network dynamics. Section 5 discusses how to use the formulation to recommend connections to a user, to help him build a good network. Detailed experiments and analysis results are presented in Section 6, before we summarize our contribution and identify future directions.

2. RELATED WORK

Multi-view networks have been extensively studied in recent years. This is greatly motivated by the fact that many real-world networks are naturally comprised of different types of relations or views. However, most of the existing study on multi-view networks aim at exploiting the multiplicity of views to boost performance in traditional tasks, such as clustering [19, 20, 24, 36], classification [33, 35], and dense subgraph mining [15, 28, 34]. The above methods leverage the redundancy offered by multiple views, but do not directly discuss the influence that views can exert onto each other.

Another line of research on multi-view networks involves modeling cross-view interrelations. In this area, [11, 27] proposed to discover such interrelations by analyzing correlation between link existence and network statistics. Other related work focuses on jointly modeling multiple network views using latent space models [12, 13, 30]. By projecting all network views onto a low-dimensional latent space, this approach is able to capture associations across different views. However, so far such associations ignore network dynamics, an important aspect of our work. The study in [2] introduces a problem setting related to our work, where the correlation in a multi-view network is modeled at different time-stamps and relationships (multiple views). In that study, the network at different time-stamps is treated just as a network with different additional views. As a result, network view's contribution to network dynamics is not directly analyzed.

On the other hand, the social network evolution has been studied mostly in the single-view setting. In accordance with the principle of homophile [26] from psychology, which postulates that people tend to interact if they have similar characteristics, [22] suggests the topology of a single-view network is correlated to the dynamics of the network. In this study, the use of pairwise topological measures alone is shown to outperform more direct measures in predicting future network interactions. In addition, [18] further shows that topological information such as triadic and focal closure correlates with the change of network strength. The relationship between network dynamics and more involved structures in a single-view dynamic network has also been studied in a real world social dataset [29]. The evolution of single-view social networks with consideration of their structural balance has also been discussed [9, 25]. A large amount of research has been conducted for decades on modeling the generation of single-view network [1, 3, 4, 8]. The work presented in this paper goes beyond these through the incorporation of multi-views.

3. MULTI-VIEW NETWORK

To better understand the state of one's social network, we need to holistically view one's relationships in the network as well as the interactions he or she has over such relationships. For this holistic study, here we represent all of these relationships and interactions by the form of *Multi-view Network* where each view indicates one kind of relationship network or interaction network.

3.1 Definition

First, we formally define the multi-view network as follows. For a given set U of nodes, a multi-view network with a set $\mathcal{V}$ of views is defined as a set of weighted graphs on the shared node set U. Each view $v \in \mathcal{V}$ corresponds to a weighted graph $(U, E^{(v)})$, where $E^{(v)}$ consists of edge (x, y) and the corresponding strength $\lambda^{(v)}(x, y)$ for $x, y \in U$.

While both relationships and interactions can be represented by weighted graphs, these two kinds have different characteristics in practice. While relationship networks grow as people come to know other people, interaction networks can change more dynamically. For example, in the online social network site like LinkedIn or Facebook, connection or friend network keeps growing as you know more people, but the frequency of message exchanges may vary a lot over time. Furthermore, interactions typically happen around certain relationships – news articles and photos are usually shared and consumed between connected users.

Given these characteristics, for practical purposes we classify network views into *relationship views* and *interaction views*. We denote the set of relationship views by $\mathcal{R}$ and denote the set of interaction views by $\mathcal{I}$. By definition, $\mathcal{R} \cap \mathcal{I} = \emptyset$ and $\mathcal{R} \cup \mathcal{I} = \mathcal{V}$. For simplicity, we limit our focus to $|\mathcal{R}| = 1$ throughout the paper. That is, we consider the case that there exists only one view of the relationship networks, which represents some systemically defined relationships such as LinkedIn connection network or Twitter follow network. In the rest of the paper, we refer two users linked in this unique relationship view by *related users*.

3.2 Dynamics of Network View

In general, for a given pair of users x and y, an edge strength $\lambda^{(v)}(x, y)$ in each view v can change. Such change can be induced by dynamism over time such as special events and seasonal effects. Besides, the edge strength $\lambda^{(v)}(x, y)$ can also change along with the growth of x's relationship network. For example, if a user forms a connection with the other users who frequently post great articles, this user may spend less time reading the articles posted by users

who were connected before. Hence, the change in the relationship view potentially affects all the interaction views. To set the notation, let $\lambda^{(v)}_{N_x}(x, y) \in \mathbb{R}_{\geq 0}$ be the *edge strength* between x and y[1] in network view v, at the time N_x are all users related to x. We further define

$$\Lambda^{(v)}_{N_x}(x, S) \coloneqq \sum_{y \in S} \lambda^{(v)}_{N_x}(x, y)$$

as the *aggregated edge strength* between user x and a set S of users in network view v.

In general, we address that $\Lambda^{(v)}_{N_x}(x, S) \neq \Lambda^{(v)}_{N_x \cup M}(x, S)$ for $S \subset N_x$ and $M \not\subset N_x$. Therefore, the study of the synergy and cannibalization between new (M) and existing related users (S) in each interaction view is crucial for understanding the multi-view network holistically. Once we have insights on this phenomenon, such insights can be used for recommending appropriate users with consideration of potential influence on each interaction view.

3.3 Cross-View Interplay

In addition to synergy and cannibalization, the edge strength in each view of network may change over time, affecting and being affected by the other views. For example, connecting to someone who posts a lot of good articles on the site can lead to more connections because of increased engagement. On the other hand, posting a lot of articles may limit the time to spend on reading the other people's articles. Understanding this cross-view interplay will also provide values similarly to the cross-user interrelation. Hence, in the following section we will not only introduce descriptors enabling observation on synergy and cannibalization, but will also model these descriptors in the dynamic networks to obtain the insights about the cross-view interplay.

4. MODELING MULTI-VIEW NETWORK DYNAMICS

We aim to study the dynamics of each network view and the interplay between different views, along with the growth of relationship network in the multi-view scenario. We develop a simple but effective model to fulfill these objectives that allows a scalable inference algorithm.

4.1 Descriptors

First, we introduce two descriptors to track the dynamics of a network view when a new relationship is established. In our model, we consider the time granularity of the epochs when individual relationships are formed. As notations in discussing dynamics of one particular user, we always refer this user as the *source user*. When this particular *source user* forms a new relationship, we refer the user on the other end of the relationship as a *destination user*.

The first descriptor is *total evolutionary rate*

$$r^{(v)}_{N_x}(x; z) \coloneqq \frac{\Lambda^{(v)}_{N_x \cup \{z\}}(x, N_x \cup \{z\})}{\Lambda^{(v)}_{N_x}(x, N_x)},$$

which is the ratio of the total strength of source user x, in network view v, before and after forming the new relationship with z. For example, x exchanges 14 messages a week with existing connections. After getting connected to z, x now exchanges 17 messages a week with z and existing connections altogether. Then $r^{(v)}_{N_x}(x; z) = \frac{17}{14}$.

[1]For notational convenience, we describe only the symmetric case.

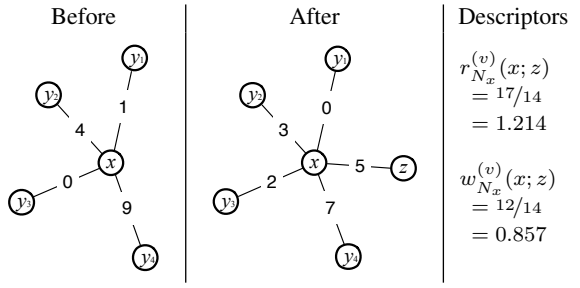


Table 1: A toy example on the change of edge strength being depicted by two descriptors, total evolutionary rate (r) and restricted evolutionary rate (w), as new relationship forms.

The second descriptor is *restricted evolutionary rate*

$$w^{(v)}_{N_x}(x; z) \coloneqq \frac{\Lambda^{(v)}_{N_x \cup \{z\}}(x, N_x)}{\Lambda^{(v)}_{N_x}(x, N_x)},$$

which is the ratio of the aggregated edge strength of x restricted on its existing neighbors, in network view v, before and after new relationship establishment. For example, x exchanges 14 messages a week with existing connections. After getting connected to z, x now exchanges only 12 messages a week with existing connections. Then $w^{(v)}_{N_x}(x; z) = \frac{12}{14}$.

Table 1 provides an illustrative example for these two descriptors. The leftmost figure represents the state of edge strength on this view. In the middle figure, user x forms relationship with a new user z, and the edge strength changes as well. These two figures display that $\Lambda^{(v)}_{N_x}(x, N_x) = 14$, $\Lambda^{(v)}_{N_x \cup \{z\}}(x, N_x) = 12$, and $\Lambda^{(v)}_{N_x \cup \{z\}}(x, N_x \cup \{z\}) = 17$. Therefore, we can compute $r^{(v)}_{N_x}(x; z) = \frac{17}{14}$ and $w^{(v)}_{N_x}(x; z) = \frac{12}{14}$.

4.2 Observation on Strength

As represented in both total evolutionary rate (r) and restricted evolutionary rate (w), the key measurement to compute the descriptors is the aggregated edge strength (Λ). In practice, a typical method of measuring the edge strength (λ) is to track the events related to a certain network view and measure its frequency. For instance, a network view representing message communication between a pair of users defines its edge strength $\lambda^{(msg)}$ based on the number of messages sent in a period of time. To obtain more robust measurements, aggregation on a long period of time such as a month is preferred as opposed to short granularity such as one hour.

However, as depicted in Figure 1, new relationships are formed with multiple users, $\tilde{N}^T_x$, over the aggregation period T. As a result, we cannot always directly measure aggregated edge strength (Λ) every time a new relationship forms. That is, it is not trivial to extract descriptors, r and w. As its alternative, we measure the ratio of the aggregated edge strength of x restricted on the existing relationships before and after a period of time, called *accumulated restricted evolutionary rate* over time period T

$$W^{(v)}_{N_x}(x; T) \coloneqq \frac{\Lambda^{(v)}_{N_x \cup \tilde{N}^T_x}(x, N_x)}{\Lambda^{(v)}_{N_x}(x, N_x)}. \quad (1)$$

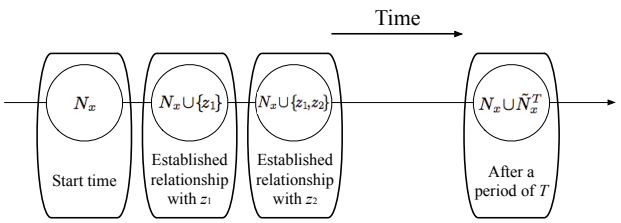


Figure 1: For each user x, we use all his related users to indicate corresponding timestamp. We do not distinguish timestamps between the instances of two consecutive relationship establishments.

Similarly, we define *accumulated total evolutionary rate* over time period T as

$$R_{N_x}^{(v)}(x;T) := \frac{\Lambda_{N_x \cup \tilde{N}_x^T}^{(v)}(x, N_x \cup \tilde{N}_x^T)}{\Lambda_{N_x}^{(v)}(x, N_x)}.$$

4.3 Discovering Synergy and Cannibalization

Once we have these two accumulated descriptors for each network view, we can obtain insight about synergy and cannibalization among user groups and across network views. For instance, we are able to quantify the interrelation between two user groups – existing related users and new related users – on that network view. If $W_{N_x}^{(v)}(x;T) < 1$, then *cannibalization* happens, *i.e.*, the new relationship establishment draws the reduction of the aggregated edge strength restricted on existing related users ($\Lambda^{(v)}(x, N_x)$). If $W_{N_x}^{(v)}(x;T) > 1$, then *synergy* is introduced by the new relationships, *i.e.,* $\Lambda^{(v)}(x, N_x)$ is boosted by the new relationships. Furthermore, by studying the correlation of the descriptors in different network views we can represent and observe the cross-view interrelation. We present detailed approaches to identifying synergy and cannibalization within real-world cases in Section 6.2.

4.4 Model for Evolutionary Rate

With aforementioned typical situation on the aggregated edge strength (Λ), we develop the models on the restricted evolutionary rate (w) and total evolutionary rate (r), with the following objectives. First, we aim to infer the evolutionary rates through learning the models. Second, the models should be able to capture the cross-view interplay. Third, the models must be scalable so as to be applied to large multi-view social networks. In this section, we present the detailed model for w.

Considering the three objectives, we assume the log-normality in the generative process of w. We will validate this assumption later in Section 6.3. Furthermore, we decompose each w into two parts: the systemic part determined by certain covariates from each network view and the stochastic part represented by the log-normal randomness. Hence, we formulate the restricted evolutionary rate as follows:

$$w_{N_x}^{(v)}(x;z) \sim \ln\mathcal{N}\left(g_v(x,z), \sigma_v^2\right) \tag{2}$$

where g is a function of network view v, source user x and destination user z, which takes information from the ego networks of x and of z in multi-view $\mathcal{V}$. For simplicity, we assume a constant variance σ_v^2 for given view v. In this formulation, the cross-view interplay along with each new relationship can be represented by parameters in the multi-view function g. We can then obtain insights about the cross-view interplay via learning the parameters of the function g.

However, the scalable inference algorithm highly depends on the form of the multi-view function g. To tackle this, we further simplify the model by making it additive with respect to each view $\tilde{v} \in \mathcal{V}$ as well as each covariate $\phi_{x,z}^{(\tilde{v})}$ on each view $\tilde{v}$:

$$g_v(x,z) = \sum_{\tilde{v}\in\mathcal{V}} g_v^{(\tilde{v})}(x,z) \quad \text{s.t.} \quad g_v^{(\tilde{v})}(x,z) = \sum_{\phi} g_v^{(\tilde{v},\phi)}(\phi_{x,z}^{(\tilde{v})}), \tag{3}$$

where each covariate $\phi_{x,z}^{(\tilde{v})}$ in network view $\tilde{v}$ depends on x and z. This additive formulation permits a scalable model inference, as described in the next section.

4.5 Inference on Evolutionary Rate

As previously mentioned, in many practical situations, our observations are limited to accumulated evolutionary rate, such as $W_{N_x}^{(v)}(x;T)$. Thus, we need to build the bridge from *accumulated evolutionary rate* to the more fine-grained *evolutionary rate*.

To decompose $W_{N_x}^{(v)}(x;T)$ into finer granularity, we denote $\tilde{N}_x^T = \{z_1, z_2, ..., z_{|\tilde{N}_x^T|}\}$, where z_k is the k-th new related user, and denote $M_k := \{z_1, z_2, ..., z_k\}$, $k = 0, 1, ..., |\tilde{N}_x^T|$. While aggregated edge strength itself changes as new relationship forms, for any short time period, we assume $|\tilde{N}_x^T| \ll |N_x|$ and subsequently $|M_k| \ll |N_x|$, and we do not distinguish the difference in the evolution of edge strength introduced by one relationship among similar conditions

$$\frac{\Lambda_{N_x \cup M_{k-1} \cup \{z_k\}}^{(v)}(x, N_x)}{\Lambda_{N_x \cup M_{k-1}}^{(v)}(x, N_x)} = \frac{\Lambda_{N_x \cup \{z_k\}}^{(v)}(x, N_x)}{\Lambda_{N_x}^{(v)}(x, N_x)},$$

from which we derive, given T is short compared to user lifetime,

$$W_{N_x}^{(v)}(x;T) = \prod_{k=1}^{|\tilde{N}_x^T|} w_{N_x}^{(v)}(x;z_k). \tag{4}$$

Assuming $\{w_{N_x}^{(v)}(x;z_k)\}_{k=1}^{|\tilde{N}_x^T|}$ are independent, we substitute Formula (2) into (4) as follows:

$$W_{N_x}^{(v)}(x;T) \sim \ln\mathcal{N}\left(\sum_{k=1}^{|\tilde{N}_x^T|} g_v(x,z_k), |\tilde{N}_x^T| \cdot \sigma_v^2\right). \tag{5}$$

As each $W_{N_x}^{(v)}(x;T)$ is observed, the likelihood is derived after applying (3) as

$$\mathcal{L} = -\sum_{x\in U} \frac{\left(\log W_{N_x}^{(v)}(x;T) - \sum_{k=1}^{|\tilde{N}_x^T|} \sum_{\tilde{v}\in\mathcal{V},\phi} g_v^{(\tilde{v},\phi)}(\phi_{x,z_k}^{(\tilde{v})})\right)^2}{2 \cdot |\tilde{N}_x^T| \cdot \sigma_v^2} + \text{const.}$$

To infer function $g_v^{(\tilde{v},\phi)}$ determined by certain covariates, we adopt maximum likelihood estimation (MLE) [14], which is equivalent to minimizing the weighted sum of squared residuals in our case

$$\min_{g_v^{(\tilde{v},\phi)}} \sum_{x\in U} \frac{\left(\log W_{N_x}^{(v)}(x;T) - \sum_{k=1}^{|\tilde{N}_x^T|} \sum_{\tilde{v}\in\mathcal{V},\phi} g_v^{(\tilde{v},\phi)}(\phi_{x,z_k}^{(\tilde{v})})\right)^2}{|\tilde{N}_x^T|}. \tag{6}$$

4.6 Functions of Covariates

For the form of each covariate function $g_v^{(\tilde{v},\phi)}$, we use the piecewise linear function on each covariate $\phi^{(\tilde{v})}$. This formulation maintains the scalability of the inference algorithm, while supporting flexibility in the covariate function beyond monotonic relationship.

We partition the range of each covariate $\phi^{(\tilde{v})}$ into multiple segments, denote $b_i^{(\tilde{v},\phi)}$ the lower boundary of the i-th segment, and let the i-th feature be

$$f_i^{(\tilde{v},\phi)}(\phi^{(\tilde{v})}) := (\phi^{(\tilde{v})} - b_i^{(\tilde{v},\phi)})\mathbb{1}_{[\phi^{(\tilde{v})} \geq b_i^{(\tilde{v},\phi)}]}, \quad (7)$$

where $\mathbb{1}_{[\cdot]}$ is the indicator function. For each covariate $\phi^{(\tilde{v})}$, combining features derived from all segments yields a feature mapping $\phi^{(\tilde{v})} \mapsto \mathbf{f}^{(\tilde{v},\phi)}(\phi^{(\tilde{v})})$. Further concatenate these mappings for all covariates in all network views, we get a feature function $(x, z) \mapsto \mathbf{f}(x, z)$. With this feature function on covariates derived across all network views, $g_v(x, z)$ can be expressed by $\mathbf{c}_v \cdot \mathbf{f}(x, z)$, where $\mathbf{c}_v$ is the coefficient vector to learn. By substituting this representation of g into the Optimization Problem (6), we have

$$\min_{\mathbf{c}_v} \sum_{x \in U} \frac{1}{|\tilde{N}_x^T|} \left(\log W_{N_x}^{(v)}(x; T) - \mathbf{c}_v \cdot \sum_{k=1}^{|\tilde{N}_x^T|} \mathbf{f}(x, z_k) \right)^2, \quad (8)$$

which turns out to be a linear least squares problem and has closed-form solution [6], and we call this Restricted Evolutionary Model (REM). Similarly for $r_{N_x}^{(v)}(x; z)$, we derive the Total Evolutionary Model (TEM)

$$\min_{\tilde{\mathbf{c}}_v} \sum_{x \in U} \frac{1}{|\tilde{N}_x^T|} \left(\log R_{N_x}^{(v)}(x; T) - \tilde{\mathbf{c}}_v \cdot \sum_{k=1}^{|\tilde{N}_x^T|} \mathbf{f}(x, z_k) \right)^2. \quad (9)$$

Lastly, note that the inference algorithm is scalable due to the aforementioned existence of closed-form solutions.

5. APPLICATION IN RELATIONSHIP RECOMMENDATION

Social networking services help users to expand their networks by recommending potential relationships. One way of making such recommendations is to recommend relationships that are likely to eventually be established (henceforth referred as *Strategy I*). For example, LinkedIn may recommend a professional the user might wish to reach out to, while Twitter may recommend an account the user would want to follow.

However, the establishment of relationships may not necessarily be the end goal of the recommendation. For example, a desirable outcome from forming a new relationship can be that this relationship leads to more interactions and engagement. Should this be the case, the users may benefit more from their relationships and the service providers may receive more commercial value by having higher user activity level. Hence, another way of relationship recommendation is to recommend relationships that are likely to result in most interaction (henceforth referred as *Strategy II*).

Note that any type of interaction can be a network view over all users. We can therefore model the dynamics of edge strength for concerned types with our proposed multi-view network framework.

Formally, for given concerned interaction type v and user x, we need to predict $\Lambda_{N_x \cup \{z\}}^{(v)}(x, N_x \cup \{z\})$, the *aggregated edge strength* after x forms a new relationship with the user z. With the Restricted Evolutionary Model, the prediction can be realized by

$$\begin{aligned} &\hat{\Lambda}_{N_x \cup \{z\}}^{(v)}(x, N_x \cup \{z\}) \\ &\quad = \Lambda_{N_x}^{(v)}(x, N_x)\hat{w}_{N_x}^{(v)}(x; z) + \hat{\lambda}_{N_x \cup \{z\}}^{(v)}(x, z) \\ &\quad = \Lambda_{N_x}^{(v)}(x, N_x)e^{\mathbf{c}_v \cdot \mathbf{f}(x,z)} + \hat{\lambda}_{N_x \cup \{z\}}^{(v)}(x, z), \end{aligned} \quad (10)$$

where $\hat{w}_{N_x}^{(v)}(x; z)$ is derived from the log-normal model in Formula (2), and $\hat{\lambda}_{N_x \cup \{z\}}^{(v)}(x, z)$ is the predicted edge strength between x and z, which can be obtained by regressing on features of user pair (x, z). Besides, with the Total Evolutionary Model, the prediction would become

$$\hat{\Lambda}_{N_x \cup \{z\}}^{(v)}(x, N_x \cup \{z\}) = \Lambda_{N_x}^{(v)}(x, N_x)e^{\tilde{\mathbf{c}}_v \cdot \mathbf{f}(x,z)}. \quad (11)$$

If p_{xz} stands for the probability a user x establishes the relationship with a user z given z is recommended to x, then according to *Strategy I* the user recommended foremost to x would be the z that maximizes p_{xz}. However, in case of *Strategy II*, where interaction is further considered, the unconnected user recommended foremost to x would be the z that maximizes the *predicted aggregated edge strength given z is recommended to x*, *i.e.*,

$$\hat{z} = \arg\max_z p_{xz} \cdot \hat{\Lambda}_{N_x \cup \{z\}}^{(v)}(x, N_x \cup \{z\}),$$

where $\hat{\Lambda}_{N_x \cup \{z\}}^{(v)}(x, N_x \cup \{z\})$ is computed by Equation (10) or Equation (11).

Note that in Equation (10), we choose not to directly approximate $\hat{\Lambda}_{N_x \cup \{z\}}^{(v)}(x, N_x)$ with the observed $\Lambda_{N_x}^{(v)}(x, N_x)$, and instead calibrate the latter with $\hat{w}_{N_x}^{(v)}(x; z)$, because aggregated edge strength restricted on the existing relationships N_x change as new relationships form, which is confirmed by experiment results reported in section 6.2.

6. EXPERIMENTS

In this section, we analyze the LinkedIn multi-view network using the proposed methods and validate the methods through prediction tasks. This section mainly consists of three parts. First, using the two accumulated descriptors, R and W, we observe some cases of synergy and cannibalization between existing and new related users as well as among network views. Second, we quantitatively verify our proposed models through prediction tasks, which also serve as the most important piece of brick in the application discussed in Section 5. Last, given the verified model, we provide more case study on the cross-view interplay at the temporal granularity that each individual relationship is created.

6.1 Data Description and Data Processing

LinkedIn, the largest professional online social network site, has over 430 million users and provides lots of interactions among users. Here we use the LinkedIn data by focusing on three phases ($T = 28$ days) of observations: Phase 0 (Aug. 4–31), Phase 1 (Sep. 1–28), and Phase 2 (Sep. 29–Oct. 26) in 2014. Phase 1 and 2 are used for most analysis on the dynamics of multi-view network, while Phase 0 data is mainly used to compute some covariates (*e.g.*, number of new connections formed between Phase 0 and 1) for the proposed model. We then sample over 1 million users who had at least one interaction in either Phase 1 or 2 and made at least one connection in Phase 2.

Network views. Multiple types of interactions exist among users in LinkedIn network, which compose a multi-view network. In particular, connections in LinkedIn are naturally regarded as its relationship view. We mainly consider the following network views, which include one relationship view and three interaction views:

- connection (*conn*): binary representation of undirected relationships.
- profile_view (*p_v*): the action of viewing the profile of another user.

- feed (*feed*): the action of consuming posts or updates from another user via the newsfeed.
- active_connection (*act_conn*): connection along which transient interactions have occurred in a given time period, where transient interactions (*trans_interactions*) include message exchange, profile view, feed consumption, endorsement, *etc.*

Note that *connection* and *active_connection* are binary whereas *profile_view* and *feed* are weighted interactions. The edge strength between a pair of users is defined over the number of the corresponding interaction events in each Phase (*i.e.*, for 28 days). For given user x, any user connected to x is also referred as a connection of x when context is clear. Since the problem is modeled assuming symmetry in edge strength for now, we do not distinguish between viewing others and being viewed by others. The same treatment is also applied to other interaction types if applicable.

Covariates and logarithmic transformation In Section 4.4∼4.6, we introduced the way of modeling descriptors with covariates ($\phi^{(\tilde{v})}$) and of making piece-wise linear functions on each covariate. Here we show the covariates we use for our analysis and experiments.

- Node strength (Σv): the sum of edge strength associated with a certain user in each network view v.
- Incremental node strength (Δv): the change of node strength between Phase 0 and 1 per network view v.

Also, as these covariates are used to model the evolutionary rate, $r^{(v)}_{N_x}(x;z)$ and $w^{(v)}_{N_x}(x;z)$, when each new connection z is made for a given user x, we use both source user x's covariates and destination user z's covariates. Therefore, the total number of covariates is 4 per view.

To embed these covariates as features in the proposed model as described in Equation (7), we use 5 segments per covariate based on quantile values. However, as degree distributions in relationships and interactions are typically heavy-tailed [10, 21], we use the transformed values of covariates based on the following transformation function:

$$\psi : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \text{sign}(x) \log_2(|x| + 1)\,.$$

Overall, we have 4 network views, 4 covariates per network view, and 5 segments per covariate; that is, we use 80 features for the proposed model.

6.2 Synergy and Cannibalization

In this section, we first study the synergy and cannibalization between two user groups – one's existing connections and her new connections – before analyzing such relationship across the network views. Such analyses are made possible by the proposed accumulated descriptors, total evolutionary rate (R) and restricted evolutionary rate (W).

Network view (v)	$R^{(v)}_{N_x}(x;T)$	$W^{(v)}_{N_x}(x;T)$	$W^{(v)}_{N_{x_0}}(x_0;T)$
profile_view	1.028	0.626	0.830
feed	1.029	1.015	0.989
trans_interactions	1.023	0.672	0.818

Table 2: Geometric means of accumulated descriptors of all users (x) and of users who form no new connections (x_0).

Geometric mean reveals interrelation across user groups. In each of the aforementioned network views, we report the geometric mean (GM[·]) of $R^{(v)}_{N_x}(x;T)$ and of $W^{(v)}_{N_x}(x;T)$, over the time period between Phase 1 and 2 in Table 2, which can be directly computed from aggregated edge strength (Λ) specified in Section 6.1. The reported quantities are all significantly different from 1 (with $p < 0.0001$). Note that $R^{(conn)}_{N_x}(x;T) = \frac{|N_x \cup \tilde{N}^T_x|}{|N_x|}$. Therefore, we do not discuss the view of *connection* here. Geometric mean is a natural choice as each descriptor indicates the ratio metric.

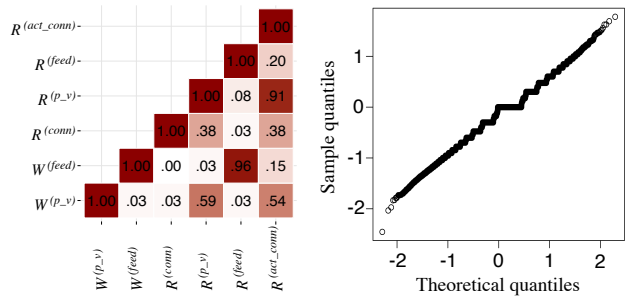
For the purpose of comparison, we also study users, denoted by x_0, who were active in Phase 1 but did not form any new connection during Phase 2. In this case, $R^{(v)}_{N_{x_0}}(x_0;T) = W^{(v)}_{N_{x_0}}(x_0;T)$ due to the invariance in the connections. Thus, we only show GM$[W^{(v)}_{N_{x_0}}(x_0;T)]$ in the rightmost column of Table 2.

Each network view grows. For all network views, geometric mean of the accumulated total evolutionary rate, GM$[R^{(v)}_{N_x}(x;T)]$, is greater than 1. This implies that the sampled users generally become more active after 28 days in the sense that they are involved in more interactions. In contrast, from the rightmost column, the user who did not establish new connections actually shrank their interactions. This agrees with our intuition that when users expand their connection network, they are more engaged in the social networking site, *i.e.*, having more interactions. In order to engage users in more interactions, we should encourage them to expanse their connection network.

Cannibalization in *profile_view*. The accumulated restricted evolutionary rate (W) in *profile_view* tends to be smaller than 1, as shown in Table 2, GM$[W^{(p_v)}_{N_x}(x;T)] = 0.626 < 1$. In other words, while a user tends to have more overall profile views, those happening with existing connections actually drop as time elapses. One can think that the shrinkage might have happened spontaneously. However, spontaneous shrinkage alone can not fully explain such shrinkage. In fact, compared with the rightmost column, this shrinkage (GM$[W] = 0.626$) is lower than 0.830 – the case when there is no interference by new connections. The lower GM$[W]$ for users who added new connections indeed demonstrates cannibalization exists between existing connections and new connections with respect to *profile_view*. We understand this phenomenon in the way that users spend more time exploring the profiles of the newly connected users by sacrificing time spent on the existing connections.

Synergy in *feed*. Unlike *profile_view*, the accumulated restricted evolutionary rate (W) in *feed* tends to be greater than 1, as in Table 2, GM$[W^{(feed)}_{N_x}(x;T)] = 1.015 > 1$. That is, not only overall *feed* interaction has increased, *feed* interaction with existing connections has grown as well. Note that while W for $feed$ tends to be greater than 1, it is still not as great as accumulated total evolutionary rate (GM$[R] = 1.029$), which indicates both new and existing connections do contribute to aggregated edge strength (Λ). We therefore deduce that the *feed* interaction with existing connections and that with new connections behave synergetically.

This could be explained by the following scenario. If new connections of a user bring in interesting articles into the *feed*, then the given user could gain increased satisfaction at the *feed* experience and become more engaged. As a result, this user might be willing to put more time and energy on LinkedIn *feed* interaction. In this case, existing connections and new connections would not fight for the user's time and energy. Instead, the growth of one would boost the growth of the other by making feed interaction a fun thing for this user.



(a) (b)

Figure 2: (a) Pearson's correlation coefficients between pairs of accumulated descriptors in log scale. (b) Q-Q plot on the logarithmic value of restricted evolutionary rate (w) for *profile_view*. The strong linear trend demonstrates that the logarithmic value of this descriptor is along with normal distribution.

Further evidence from correlation of descriptors. We further consolidate the synergy and cannibalization observation over *profile_view* and *feed* by studying cross-view correlation. Specifically, we compute the Pearson's correlation coefficients of $R^{(v)}_{N_x}(x;T)$ and $W^{(v)}_{N_x}(x;T)$ in log scale (abbreviated as $\log R^{(v)}$ and $\log W^{(v)}$). These Pearson's correlation coefficients can be found in Figure 2a. We can observe that, $\log R^{(feed)}$ and $\log W^{(feed)}$ are highly correlated (0.96). That is, the higher growth a user enjoys in overall *feed* interaction, the more likely this user would also have higher growth in *feed* interaction restricted over existing connections. This high correlation is in line with our previous conclusion that the *feed* interactions grow synergetically for existing connections and new connections.

As for *profile_view*, the correlation between $\log R^{(p\_v)}$ and $\log W^{(p\_v)}$ is much less significant (0.59). Together with previous observation on GM$[W^{(p_v)}_{N_x}(x;T)]$, we conclude that cannibalization exists between existing connections and new connections in terms of *profile_view*.

Synergy among network views. We thus far study the synergy and cannibalization between existing connections and new connections. To examine the synergistic or cannibalizing dynamics across the network views, we pose a holistic observation using proposed descriptors in different views by calculating their correlation coefficients. In Figure 2a, we present Pearson's correlation coefficients in log scale between each pair of $\log R^{(conn)}$, $\log R^{(p\_v)}$, $\log R^{(feed)}$, $\log R^{(act\_conn)}$, $\log W^{(p\_v)}$ and $\log W^{(feed)}$.

Synergy exists between *connection* and *profile_view* due to their observed positively correlated total evolutionary rate (R). In other words, users who are frequently adding new connections are also likely to view increasingly more profiles, and vice versa. This can be explained by, on the one hand, users who frequently add new connections will have a increasingly bigger pool of profiles available to view. On the other hand, by viewing profiles, users can explore other professionally relevant users mentioned in these profiles. These professionally relevant users can then be potential new connections.

To the contrary, such cross-view synergistic relationships do not exist between *connection* and *feed*, which corresponds to an insignificant correlation coefficient (0.03). We surmise the reason could be, while new connections always have profiles available for viewing, not all new connections would generate interesting feeds. As a result, merely adding many new connections does not necessarily guarantee further engagement in *feed* interaction. The other way round, unlike profiles, feeds seldom mention users and hence cannot serve as a significant source of potential connections. Together with previous observation that cross-user synergy exists between new connections and existing connections with respect to *feed* interaction, we conclude that just making connections may not necessarily be beneficial for *feed* engagement but getting connected to users who can provide quality feeds would further boost the *feed* engagement.

6.3 Model Validation

In the previous section, we observed the instances of synergy and cannibalization between user groups as well as among network views. Toward the fine granularity of cross-view interplay in the dynamic setting, we proposed models that allow us to infer such interplay incurred by one single new connection. In this section, we verify the fundamental log-normal assumption and quantitatively validate our models through a prediction task, which serve as the ground for the application in Section 5.

Log-normality validation. The log-normality assumption made in Formula (2) plays a key role in building the scalable inference algorithm. However, as the performance on the model inference is sensitive to how close our assumption is to the real-world data, here we validate our log-normality assumption through Q-Q plot.

For empirical distribution of $w^{(v)}_{N_x}(x;z)$, we sample LinkedIn users who are active in both Phase 1 and 2 as well as make exactly 1 new connection during Phase 2. For these users, their restricted evolutionary rate, $w^{(v)}_{N_x}(x;z)$'s in the network view *profile_view*, are derivable from data since only one instead of multiple new connections are made during the given period. We draw the Q-Q plot of observed $w^{(v)}_{N_x}(x;z)$'s against theoretical log-normal distribution in Figure 2b. The linear trend in this plot validates that the stochastic part of our model follows log-normal distribution. Therefore, our choice of log-normal model is reasonable.

Predicting aggregated edge strength. Aggregated edge strength (Λ) can practically represent the engagement level of a user on the given interaction. Through prediction task on Λ, we can achieve two things. First, we envision the predictive capability of the proposed model, and thereby validate the foundation of the recommendation product brought up in Section 5. Second, by quantitatively validating the proposed models through this prediction task, we support the arguments about the qualitative analysis in the following section.

In the prediction task, we aim to predict two quantities for each test user in Phase 2: aggregated edge strength over all connections, $\Lambda^{(v)}_{N_x\cup\tilde{N}^T_x}(x, N_x\cup\tilde{N}^T_x)$, and aggregated edge strength over existing connections, $\Lambda^{(v)}_{N_x\cup\tilde{N}^T_x}(x, N_x)$. These two quantities are both important because knowing both metrics will allow us to design the product balancing between existing and new connections. All information available for these prediction tasks is from Phase 0 and 1. We randomly partition users into two parts, 70% as training set and 30% as test set.

Here we present the two strategies: REM and TEM, as introduced in Section 4.6. The REM solves the Optimization Problem (8) and learn the coefficients of the features by regressing on $W^{(v)}_{N_x}(x;T)$. On the other hand, the TEM solves the Optimization Problem (9) and learn the coefficients of the features by regressing on $R^{(v)}_{N_x}(x;T)$. Hence, given the state of each user x and new connection z's state on each network view, REM and TEM infer $\hat{w}^{(v)}_{N_x}(x;z)$ and $\hat{r}^{(v)}_{N_x}(x;z)$, respectively, both of which depict fine grained dynamics whenever one single new connection forms.

Once we obtain the inferred restricted evolutionary rate $\hat{w}^{(v)}_{N_x}(x;z)$, REM makes prediction through Equation (1) and (4) by

$$\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=\Lambda^{(v)}_{N_x}(x,N_x)\prod_i \hat{w}^{(v)}_{N_x}(x;z_i),$$
$$\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)+\sum_i \hat{\lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,z_i),$$

where $\{z_1,z_2,...,z_{|\tilde{N}^T_x|}\}$ is the set of all new connections x made during the observation period, and $\hat{\lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,z)$ is the predicted edge strength between x and z, which is obtained by regressing on features of user x and user z.

Similarly, TEM makes prediction by

$$\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=\Lambda^{(v)}_{N_x}(x,N_x)\prod_i \hat{r}^{(v)}_{N_x}(x;z_i),$$
$$\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)-\sum_i \hat{\lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,z_i).$$

We compare the prediction results obtained by these two models with the following baselines:

- **Uniform Strength Model**: Assume uniform and stationary edge strength for all connections of one user. This case aggregated edge strength is proportional to the number of connections:
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=\Lambda^{(v)}_{N_x}(x,N_x)$,
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=\frac{|N_x\cup\tilde{N}^T_x|}{|N_x|}\cdot\Lambda^{(v)}_{N_x}(x,N_x)$.
- **Interference-Free Model**: Assume no interference between existing connections and new connections. The edge strength on existing connections, does not change due to the introduction of new connections:
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=\Lambda^{(v)}_{N_x}(x,N_x)$,
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=\Lambda^{(v)}_{N_x}(x,N_x)+\sum_i \hat{\lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,z_i)$.
- **Constant Evolutionary Model**: For each new connection, assume constant calibration rate α of aggregated edge strength on existing connections, *i.e.*, constant restricted evolutionary rate (w), and constant growth rate β of total aggregated edge strength, *i.e.*, constant total evolutionary rate (r):
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=\alpha^{|\tilde{N}^T_x|}\cdot\Lambda^{(v)}_{N_x}(x,N_x)$,
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=\beta^{|\tilde{N}^T_x|}\cdot\Lambda^{(v)}_{N_x}(x,N_x)$.
- **Proportional Evolutionary Model**: Assume calibration of aggregated edge strength over existing connections – *i.e.*, restricted evolutionary rate (w) – to be proportional to the rate of change in connections with an exponential of γ representing resistance, and growth of aggregated edge strength over total connections – *i.e.*, total evolutionary rate (r) – to be proportional to the rate of change in connections with an exponential of δ modeling resistance:
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)=(\frac{|N_x|}{|N_x\cup\tilde{N}^T_x|})^{\gamma}\cdot\Lambda^{(v)}_{N_x}(x,N_x)$,
  $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)=(\frac{|N_x\cup\tilde{N}^T_x|}{|N_x|})^{\delta}\cdot\Lambda^{(v)}_{N_x}(x,N_x)$.

The parameters α, β, γ and δ in the above baselines are inferred by minimizing squared residual of log-likelihood. For notational convenience, we henceforth use $\Lambda(N_x)$ and $\Lambda(N_x\cup\tilde{N}^T_x)$ as shorten forms for $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x)$ and $\hat{\Lambda}^{(v)}_{N_x\cup\tilde{N}^T_x}(x,N_x\cup\tilde{N}^T_x)$, respectively.

Root mean square relative error (RMSRE) is used as evaluation metric for the prediction task: $\text{RMSRE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{Y}_i-Y_i}{Y_i}\right)^2}$, where $(\hat{Y}_1,\hat{Y}_2,...,\hat{Y}_n)$ represents the vector of n predictions, and $(Y_1,Y_2,...,Y_n)$ is the vector of ground truth values.

We believe that the relative error normalized by the existing quantities is more important than the absolute error. For instance, suppose that there are two users where one user has 100 interactions per month while the other has only 2. If we predict the former user's interaction by 95 but the latter user's interaction by 7, then we may think that the former prediction is acceptable while the latter prediction looks like over estimation. Hence, here we measure the relative error metric and use it to compare each model's predictive performance.

Prediction method	$\Lambda(N_x)$		$\Lambda(N_x\cup\tilde{N}^T_x)$	
	profile_view	*feed*	*profile_view*	*feed*
Uniform Strength	3.38	1.83	1.33	1.25
Interference-Free	3.38	1.83	1.44	1.16
Constant Evolutionary	3.18	1.90	1.28	1.29
Proportional Evolutionary	3.19	1.89	1.39	1.27
Total Evolutionary	2.58	1.37	0.75	**0.81**
Restricted Evolutionary	**1.22**	**1.26**	**0.67**	0.86

Table 3: Root-mean-square relative error (RMSRE) of prediction methods in predicting aggregated edge strength in *profile_view* and *feed*.

We perform the prediction experiments for both *profile_view* and *feed*, and report the results in Table 3. Overall, both Total Evolutionary Model and Restricted Evolutionary Model outperform the baseline models in every network view on each of two aggregated edge strength quantities.

Between the baseline Interference-Free Model and the Restricted Evolutionary Model (REM), the only difference is whether or not we model the interference between existing connections and new connections. Modeling such interference leads to huge improvement in the predictive capability for the proposed REM, as seen in Table 3. This results bolster that synergy and cannibalization exist between existing connections and new connections, as previously observed.

The Constant Evolutionary Model and the Proportional Evolutionary Model consider interference between existing connections and new connections by either directly modeling calibration on existing connections or controlling overall aggregated edge strength. However, proposed models still outperform these baselines by a large margin. This makes sense because, as we have observed, the aggregated edge strength (Λ) over a group of users in one view interrelates with other user group as well as other views. The proposed models have taken such interrelation into account.

6.4 Cross-View Interplay in Dynamic Networks per New Connection Formation

In this section, we aim at studying how the dynamics of one network view would correlate with not only itself but also other network views at earlier timestamps. The proposed model, which was verified in the previous prediction task, enables us to study how the dynamics of one view is affected by other views whenever a new connection is established.

By solving Optimization Problem (8) and (9), such cross-view interplay can be depicted by piece-wise linear functions mapping each covariate to concerned evolutionary descriptor, such as $r^{(conn)}$, $r^{(feed)}$, $w^{(p\_v)}$, *etc.* We use the Restricted Evolutionary Model (8) to examine the relationship between each covariate and restricted evolutionary rate (w), while using the Total Evolutionary Model (9) to see the correlation between each covariate and total evolutionary rate (r).

$w^{(p\_v)}$: restricted evolutionary rate for *profile_view*								$w^{(feed)}$: restricted evolutionary rate for *feed*								$r^{(conn)}$: total evolutionary rate for *connection*							
covariate	Rank	Trend	Significance					covariate	Rank	Trend	Significance					covariate	Rank	Trend	Significance				
			Seg 1	Seg 2	Seg 3	Seg 4	Seg 5				Seg 1	Seg 2	Seg 3	Seg 4	Seg 5				Seg 1	Seg 2	Seg 3	Seg 4	Seg 5
x-Σp_v	1	↘	$***$	$***$	$***$	$***$	$***$	x-$\Sigma feed$	1	↘	$***$	$***$	$***$	$***$	$***$	x-$\Sigma conn$	1	↘	$***$	$***$	$***$	$***$	$***$
x-Σact_conn	2	↘↗	$***$	$***$	$***$	$***$	$***$	x-Σact_conn	2	↗	$***$	$***$	$***$	$***$	$***$	x-$\Delta conn$	2	↗	$***$	$***$	$***$	$***$	$***$
x-$\Delta conn$	3	↘	$***$	$***$	$***$	$***$	$***$	x-Δact_conn	3	↘	$***$	$***$	$***$	$***$	$***$	x-Σact_conn	3	↗↘	$***$	$***$	$***$	$***$	$***$
x-$\Sigma conn$	4	↗	$***$	$***$	$***$	$***$	$***$	x-Δp_v	4	↗	$***$	$***$	$***$	$***$	$***$	x-Σp_v	4	↗	$***$	$***$	$***$	$***$	$***$
z-Σp_v	9	↗	$***$	$**$	$***$		$***$	x-$\Delta feed$	5	↗↘	$***$	$***$	$***$	$***$	$***$	z-$\Sigma conn$	7	↗	$***$	$***$	$***$		$***$

$r^{(p\_v)}$: total evolutionary rate for *profile_view*								$r^{(feed)}$: total evolutionary rate for *feed*								$r^{(act\_conn)}$: total evolutionary rate for *active_connection*							
covariate	Rank	Trend	Significance					covariate	Rank	Trend	Significance					covariate	Rank	Trend	Significance				
			Seg 1	Seg 2	Seg 3	Seg 4	Seg 5				Seg 1	Seg 2	Seg 3	Seg 4	Seg 5				Seg 1	Seg 2	Seg 3	Seg 4	Seg 5
x-Σp_v	1	↘	$***$	$***$	$***$	$***$	$***$	x-$\Sigma feed$	1	↘	$***$	$***$	$***$	$***$	$***$	x-Σact_conn	1	↘	$***$	$***$	$***$	$***$	$***$
x-Σact_conn	2	↘↗	$***$	$***$	$***$	$***$	$***$	x-Σact_conn	2	↗	$***$	$***$	$***$	$***$		x-$\Sigma conn$	2	↘↗	$***$	$***$	$***$	$***$	$***$
x-$\Sigma conn$	3	↘↗	$***$	$***$	$***$	$***$	$***$	x-Δact_conn	3	↘	$***$	$***$	$***$	$***$	$***$	x Δact_conn	3	↘	$***$	$***$	$***$	$***$	$***$
x-$\Delta conn$	5	↘	$***$	$***$	$***$	$***$	$***$	x-Δp_v	4	↗	$***$	$***$	$***$	$***$	$***$	x-Σp_v	4	↗	$***$	$***$	$***$	$***$	$***$
z-Σp_v	6	↗	$***$	$***$	$***$		$***$	x-$\Delta feed$	5	↗↘	$***$	$***$	$***$	$***$	$***$	x-$\Sigma feed$	5	↗	$***$		$***$	$***$	$***$

Significance codes: $***$ – 0.001, $**$ – 0.01, $*$ – 0.05, $\cdot$ – 0.1.

Table 4: Rank determined by total variance and significance codes derived from p-value reflecting how significant each covariate is correlated with each descriptor.

With definition in Formula (7), the piece-wise linear relationship would be given by $h^{(\tilde{v},\phi)}(\phi^{(\tilde{v})}) = \sum_{i=1}^{5} f_i^{(\tilde{v},\phi)}(\phi^{(\tilde{v})})$.

Given a concerned evolutionary descriptor, we would like to know how each covariate is correlated to the given descriptor and how crucial the correlation is. For the former problem, we plot out the learned piece-wise linear function and observe the trend. For the latter, we use total variation of the piece-wise linear function to quantify the significance of the covariate to this descriptor, and the total variation of a function h defined on an interval $[a, b] \subseteq \mathbb{R}$ is the quantity

$$V_a^b(h) \coloneqq \sup_{P \in \mathcal{P}} \sum_{i=0}^{n_P - 1} |h(\phi_{i+1}) - h(\phi_i)|,$$

where $P = \{\phi_0, ..., \phi_{n_P}\}$ is any partition of $[a, b]$, and $\mathcal{P}$ is the set of all such partitions.

Out of consideration for company confidentiality, we cannot explicitly present the relationships with plots and exact values. However, we can alternatively show how each covariate is ranked by corresponding total variation in a descending order as well as the general trends of these covariates (*e.g.*, ↗: monotonically increasing; ↘: monotonically decreasing; ↗↘: first increasing then decreasing, *etc.*). Additionally, we can also present the p-value which reflects the confidence of the inferred piece-wise linear relationship. Besides, in order to prevent calculated total variation from affected to much by outliers, we compute total variation on the interval where 95% of records are populated. In practice, we exclude the upper fifth percentile for covariates with only positive value, and exclude both lower and upper 2.5% for other covariates. We report the importance of covariates by the rank determined by total variation and the significance derived from p-value as well as their trend in Table 4. For each descriptors, we always report covariates ranked top 3 with additional 2 of interests.

Based on the trained piece-wise linear functions, we make the following observations.

Non-monotonic relationship exists. Thanks to the flexibility of piece-wise linear covariate functions, we can discover non-monotonic relationship between descriptor and covariates.

The most prominent non-monotonic relationship is that between total evolutionary rate for *profile_view* ($r^{(p\_v)}$), and source user's node strength in *connection* (x-$\Sigma conn$) as well as in *active_connection* (x-Σact_conn). This relationship is descending at first when x-$\Sigma conn$ or x-Σact_conn is small, and then turns upward.

This is actually expected because x-$\Sigma conn$ is just the number of connections the source user has, and at the time users just sign up on LinkedIn, they could fervidly view the profiles of the newly added connections who are also their friends in life. Note that the more connections one has, the bigger pool of profiles one can view. However, after these users have got connected with most of their friends in life who use LinkedIn, this new connection formation boom will stop. As a result, profile views would no longer be fast growing. However, if a user falls into the upper segments in terms of connection numbers, he or she might be a recruiter, who would be willing to connect a wide range of people without confining to people he or she knows in life. For these users, they may keep actively viewing more profiles. The same scenario can also be applied to explaining this non-monotonicity in x-Σact_conn.

Different correlation to total evolutionary rate and restricted evolutionary rate. For the same network view *profile_view*, two descriptors, total evolutionary rate (r) and restricted evolutionary rate (w), can respond differently to one covariate.

The first case is about the covariate on the number of new connections the source user makes between Phase 0 and 1 (x-$\Delta conn$), which is very critical – monotonically decreasing, and ranked 3rd – to restricted evolutionary rate in *profile_view* ($w^{(p\_v)}$), but less crucial to total evolutionary rate in *profile_view* ($r^{(p\_v)}$). As a collateral evidence, for $r^{(p\_v)}$, the total variation of x-$\Sigma conn$ (ranked 3rd) is nearly three times as great as that of x-$\Delta conn$, while for $w^{(p\_v)}$, x-$\Delta conn$ exceeds x-$\Sigma conn$ and stands among the top 3.

This makes sense because x-$\Delta conn$ is identical to the number of new connection established over the previous observation period. The more connections a user makes, the more energy viewing profile might be spared to new connections. According to observation we have made about cannibalization in *profile_view* in section 6.2, this means the time viewing profiles of existing connections would be limited. As a result, the higher x-$\Delta conn$ is, the smaller $w^{(p\_v)}$ would likely to be.

The second case is about the covariate on destination user's profile view frequency (z-Σp_v), which is more crucial to $r^{(p\_v)}$ (monotonically increasing) than to $w^{(p\_v)}$. This is exactly in line with the definition of restricted evolutionary rate (w). When connecting to user with high node strength in *profile_view*, such as a recruiter, who is like to contribute more profile view interactions, the overall profile view interactions would tend to increase, leading to a bigger $r^{(p\_v)}$. However, once restricted on existing users, the interaction involving the new connection will no longer contribute to the aggregated edge strength (Λ), meaning less significance to $w^{(p\_v)}$.

Expanding your network by connecting to social hubs. The relationship between destination user's node strength in *connection* (z-$\Sigma conn$) and total evolutionary rate for *connection* ($r^{(conn)}$) is monotonically increasing.

Users with high node strength in *connection*, *i.e.*, number of connections, are essentially social hubs. Recall that LinkedIn is a professional networking site. Hence, connecting to social hubs could expose users to more professionally relevant potential connections, their own connection network may grow as well due to the increasing visibility to more relevant yet unconnected users.

The diminishing returns are observed in all views. For all six descriptors, their relationship with source user's node strength from the same view (x-Σv) is always monotonically decreasing. This basically implies the more interactions or relationship a user has, the harder it is to further increase the same kind of interaction or relationship, *i.e.*, the law of diminishing returns holds for all views.

7. CONCLUSION AND FUTURE WORK

In this paper, we proposed to study the dynamics of large multi-view social networks. We developed scalable model that can capture synergy, cannibalization and cross-view interplay in fine temporal granularity, which, as application, also sets the stage for a new strategy in guiding network expansion. On a large social network dataset, the proposed model uncovered cannibalization within profile view interactions and synergy within feed update interactions. Synergy was also observed across network views, specifically between connection and profile views. Aggregated edge strength prediction experiments verified the superior performance of the proposed model relative to various baselines. The proposed model effectively revealed a wide variety of relationships correlated to network growth at a fine temporal scale. Such relationships could potentially be non-monotonic.

Directions for future work include the use of triadic and pairwise covariates to derive additional features. It would also be of interest to extend the experiments to other datasets and explore other forms of interactions. On the dynamics of multi-view social network, another direction could be to model and uncover non-additive joint effects from multiple covariates.

Acknowledgments. We thank our colleagues and friends for the enlightening discussions: Mario Rodriguez, Bee-Chung Chen, Aastha Jain, Jieying Chen, Yang Yang, Guanfeng Liang, Rupesh Gupta, Hsiao-Ping Tseng, Xi Chen, Mo Yu, Xin Fu, Honglei Zhuang, Huan Gui, and many members of the Growth Relevance Team and Communication Relevance Team at LinkedIn.

8. REFERENCES

[1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47, 2002.

[2] A. Ansari, O. Koenigsberg, and F. Stahl. Modeling multiple relationships in social networks. *Journal of Marketing Research*, 48(4):713–728, 2011.

[3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4):175–308, 2006.

[5] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.

[6] R. J. Carroll and D. Ruppert. *Transformation and weighting in regression*, volume 30. CRC Press, 1988.

[7] R. K. Chandy and G. J. Tellis. Organizing for radical product innovation: The overlooked role of willingness to cannibalize. *Journal of Marketing Research*, pages 474–487, 1998.

[8] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[9] G. Facchetti, G. Iacono, and C. Altafini. Computing global structural balance in large-scale signed social networks. *PNAS*, 108(52):20953–20958, 2011.

[10] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, 1999.

[11] O. Frank and K. Nowicki. Exploratory statistical anlaysis of networks. *Annals of Discrete Mathematics*, 55:349–365, 1993.

[12] I. Gollini and T. B. Murphy. Joint modelling of multiple network views. *Journal of Computational and Graphical Statistics*, pages 00–00, 2014.

[13] D. Greene and P. Cunningham. Producing a unified graph representation from multiple social network views. In *Web Science Conference*, pages 118–121. ACM, 2013.

[14] J. W. Harris and H. Stöcker. *Handbook of mathematics and computational science*. Springer Science & Business Media, 1998.

[15] H. Hu, X. Yan, Y. Huang, J. Han, and X. J. Zhou. Mining coherent dense subgraphs across massive biological networks for functional discovery. *Bioinformatics*, 21(suppl 1):i213–i221, 2005.

[16] W. Kim, O.-R. Jeong, and S.-W. Lee. On social web sites. *Information Systems*, 35(2):215–236, 2010.

[17] T. Kollmann, A. Kuckertz, and I. Kayser. Cannibalization or synergy? consumers' channel selection in online–offline multichannel systems. *Journal of Retailing and Consumer Services*, 19(2):186–194, 2012.

[18] G. Kossinets and D. J. Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, 2006.

[19] A. Kumar and H. Daumé. A co-training approach for multi-view spectral clustering. In *ICML*, pages 393–400, 2011.

[20] A. Kumar, P. Rai, and H. Daume. Co-regularized multi-view spectral clustering. In *NIPS*, pages 1413–1421, 2011.

[21] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *WWW*, 2010.

[22] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007.

[23] S. Lipovetsky and M. Conklin. Finding items cannibalization and synergy by bws data. *Journal of Choice Modelling*, 12:1–9, 2014.

[24] J. Liu, C. Wang, J. Gao, and J. Han. Multi-view clustering via joint nonnegative matrix factorization. In *SDM*, volume 13, pages 252–260. SIAM, 2013.

[25] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz. Continuous-time model of structural balance. *PNAS*, 108(5):1771–1776, 2011.

[26] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, pages 415–444, 2001.

[27] P. Pattison and S. Wasserman. Logit models and logistic regressions for social networks: II. multivariate relations. *British Journal of Mathematical and Statistical Psychology*, 52(2):169–194, 1999.

[28] J. Pei, D. Jiang, and A. Zhang. On mining cross-graph quasi-cliques. In *KDD*, pages 228–238. ACM, 2005.

[29] D. M. Romero, B. Meeder, V. Barash, and J. M. Kleinberg. Maintaining ties on social media sites: The competing effects of balance, exchange, and betweenness. In *ICWSM*, 2011.

[30] M. Salter-Townshend and T. H. McCormick. Latent space models for multiview network data. Technical Report 622, Department of Statistics, University of Washington, 2013.

[31] F. Schneider, A. Feldmann, B. Krishnamurthy, and W. Willinger. Understanding online social network usage from a network perspective. In *SIGCOMM*, pages 35–48. ACM, 2009.

[32] N. Shi, M. K. Lee, C. M. Cheung, and H. Chen. The continuance of online social networks: how to keep people using facebook? In *System Sciences (HICSS)*, pages 1–10. IEEE, 2010.

[33] V. Sindhwani and P. Niyogi. A co-regularized approach to semi-supervised learning with multiple views. In *ICML Workshop on Learning with Multiple Views*, 2005.

[34] Z. Zeng, J. Wang, L. Zhou, and G. Karypis. Coherent closed quasi-clique discovery from large dense graph databases. In *KDD*, pages 797–802. ACM, 2006.

[35] D. Zhang, F. Wang, C. Zhang, and T. Li. Multi-view local learning. In *AAAI*, pages 752–757, 2008.

[36] D. Zhou and C. J. Burges. Spectral clustering and transductive learning with multiple views. In *ICML*, pages 1159–1166. ACM, 2007.