Multi-layer Network Composition under a Unified Dynamical Process

Xiaoran Yan[1], Shang-Hua Teng[2], Kristina Lerman[3]

[1] Indiana University Network Science Institute, Indiana University, USA
[2] Computer Science Department, University of Southern California, USA
[3] Information Sciences Institute, University of Southern California, USA

Abstract. In this paper, we take a step towards a principled method of network composition from multi-layer data. We argue that inter-layer dynamics is a essential component of understanding the structure as a whole. Mathematically, we consider the following abstract problem: given multiple layers of network data over a shared vertex set, and additional parameters for inter-layer transitions, construct a (single) weighted network that best integrates the multi-layer dynamics. In this context, we will also study an empirical use case of the composition framework.

1 Introduction

Studies of network structures have lead to fundamental insights into the organization and function of social, biological and technological systems [13]. On top of these network structures, different dynamical processes unfold [4,8], leading to applications ranging from ranking web pages to maximizing social influence and controlling epidemics [14,9]. Traditionally, most research has focused on the simple graph representation where all verticies and edges are of a single type. More recently, there has been a great interest in network models that are capable of capturing multiple types of connections [15,1]. In this paper, we adopt the general notion of *multi-layer networks*, with *multiplex network* being a special case when inter-layer structures are absent [10].

Structure and dynamics of multi-layer networks have been explored in both theoretical graphs and real world data [2,5,12]. However, it remains an open question as how to build a multi-layer network in the first place. They are often constructed simply by stacking or projecting layers into a single network. When inter-layer edges are explicitly modeled, they are usually captured by a simple parameter called the *coupling strength*. One challenge for modeling inter-layer structures is that they are empirically difficult to measure in most cases [7].

In this paper, we propose a two-stage framework for multi-layer network composition based on a unified dynamical process. In Section 2, we will first briefly discuss how to transform the layers into homogeneous Markov processes, followed by the main theorem which infer inter-layer edge weights based on inter-layer dynamics. A real world example will be investigated in Section 3.

2 The multi-layer composition framework

We first introduce some basic notations.

Single-layer data: A standard network is represented by weighted directed graph $G = (V, E, \boldsymbol{A})$, where $V = \{1, ..., n\}$ and for $u, v \in V$, $a_{uv} \geq 0$ assigns a weight to edge $(u, v) \in E$. We follow the convention that $a_{uv} = 0$ if and only if $(u, v) \notin E$. G may have self-loops. For $u \in V$, let $d_u^{out} = \sum_{v=1}^{n} a_{u,v}$ denote the *out-degree* of vertex u. In this paper, we use $\boldsymbol{D_A}$ (or $\boldsymbol{D}$ when the context is clear) to denote the diagonal matrix whose entries are out-degrees.
Multi-layer data: We consider *vertex-aligned* multi-layer networks [10]. We use l to denote number of layers, and use $G^i = (V, E^i, \boldsymbol{A}^i)$ to denote the network at i^{th} layer. For clarity, we will use superscripts i, j, r for the layers and subscripts u, v, w for vertices. Note that the vertex set V is the same across the layers.

The simplest dynamical process on graphs G is the discrete time *unbiased random walk* (URW), represented by the transition matrix $\boldsymbol{M}$.

Lemma 1. *For every directed network $G = (V, E, \boldsymbol{A})$, there is a unique transition matrix, $\boldsymbol{M_A} = \boldsymbol{A}\boldsymbol{D_A}^{-1}$, that captures the URW Markov process on G. Conversely, given a transition matrix $\boldsymbol{M}$, there is in fact multi-layer an infinite family of adjacency matrices whose random walk Markov process is consistent with $\boldsymbol{M}$: $\boldsymbol{\mathcal{A}_M} = \{\boldsymbol{M\Gamma} : \boldsymbol{\Gamma}$ is a positive diagonal matrix.$\}$*

In other words, every directed network uniquely defines a random walk process. However, given a transition matrix $\boldsymbol{M}$, there remains n degrees of freedom to specify the underlying network. They are vertex scaling factors. For undirected graphs, there is only one global scaling factor.

In [8], we argued that perceived network structure is a result of the interplay between the network topology and the dynamical process on top of it. We believe this interplay is even more pronounced in multi-layer networks, with each layer represents a different type of connection. It is thus essential to account for the different intra-layer dynamics before we put them together.

For this purpose, we reintroduce the parametrized Laplacian [8], $\boldsymbol{\mathcal{L}} = (\boldsymbol{D'} - \boldsymbol{BA})(\boldsymbol{D'T})^{-1}$, where $\boldsymbol{A}$ is the adjacency and $\boldsymbol{D'}$ is the reweighed degree matrix now defined as: $d'_u = \sum_v [\boldsymbol{BA}]_{uv}$. The diagonal matrix $\boldsymbol{T}$ controls the time delay factors at each vertex. The bias factors form the other diagonal matrix $\boldsymbol{B}$. Under the framework, we can transform each input layer to equivalent graphs underlying continuous time URWs as the unifying dynamical process (Please refer to for the details and proofs).

Theorem 1. *For a directed network $G = (V, E, \boldsymbol{A})$, the dynamics $\boldsymbol{\mathcal{L}} = (\boldsymbol{D'} - \boldsymbol{BA})(\boldsymbol{D'T})^{-1}$ is equivalent to a continuous time URW with uniform delay factors on another transformed graph.*

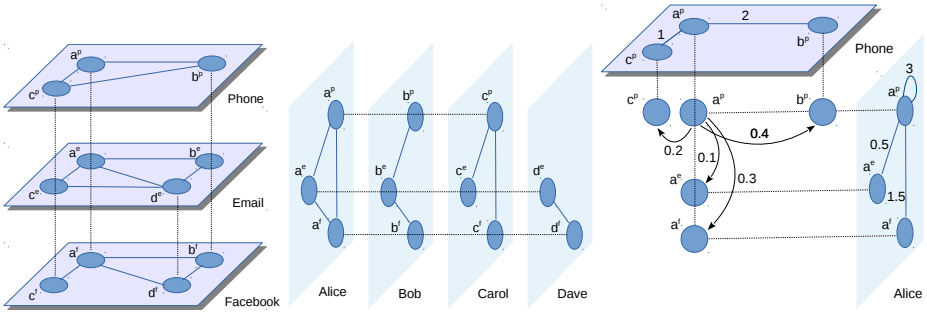
Corollary 1. *For an undirected network $G = (V, E, \boldsymbol{A})$, and the dynamical parameters $B, \boldsymbol{T}$, the interaction matrix $\boldsymbol{W} = \alpha(B\boldsymbol{A}B + (\boldsymbol{T} - \boldsymbol{I})\boldsymbol{D'})$ is unique up to a global scaling factor.*

We are now ready to discuss the second stage of the framework. For multiplex composition, simple matrix addition does the trick. However, we need a more general framework when inter-layer structures matter. Consider the following:

Formulation 1 *Given l transformed layers $G^1 = (V, E^1, \boldsymbol{W}^1), ..., G^l = (V, E^l, \boldsymbol{W}^l)$, and egocentric inter-layer dynamics $(\boldsymbol{M}_v : v \in V)$, compose a $(ln \times ln)$ weighted super-adjacency matrix, $\mathbb{W} = \begin{bmatrix} \boldsymbol{W}^1 & \boldsymbol{W}^{12} & ... & \boldsymbol{W}^{1l} \\ ... & & & \\ \boldsymbol{W}^{l1} & \boldsymbol{W}^{l2} & ... & \boldsymbol{W}^{l} \end{bmatrix}$ to integrate the multi-layer network data. In addition, $\mathbb{W}$ represent a diagonal multi-layer networks, as defined in [10], which means that all inter-layer edges are between the same vertex at different layers. Here, in $\mathbb{W}$, the l diagonal $(n \times n)$-blocks are directly fed from the first stage $\boldsymbol{W}^1, \boldsymbol{W}^2, ..., \boldsymbol{W}^l$.*

We have used the model of egocentric inter-layer dynamics for each vertex $(\boldsymbol{M}_v : v \in V)$, with $\boldsymbol{M}_v$ being the stochastic transition matrix for the inter-layer instances of the same vertex v. Such egocentric models are considered to be fundamental in the formation of social structures[6].

Figure 1a 1b is a toy example with three horizontal layers, consisting of (hypothetical) phone contacts, email exchanges and Facebook friendships. At the same time, egocentric inter-layer dynamics form a vertical perspective of the same system. Figure 1c represents the Markov transitions of this joint system when Alice receives a phone call. She might pass on the message directly by calling others with probability $0.6 = 0.4 + 0.2$, or relay the message through emails with probability 0.3, or post it on a Facebook wall with probability 0.1.



(a) Traditional view (b) The vertical perspective (c) Markov model around a^p

Fig. 1: A hypothetical toy example

By Lemma 1 each layer $\boldsymbol{W}^i$ uniquely defines a Markov model, $\boldsymbol{M}_{\boldsymbol{W}^i}$. Our task is to combine them with the n egocentric Markov models $(\boldsymbol{M}_v : v \in V)$. Therefore, we aim to identify a weighted $(ln \times ln)$-adjacency matrix $\mathbb{W}$, whose random-walk Markov model, $\boldsymbol{M}_{\mathbb{W}}$, satisfies the following two basic conditions:

1. **Layer Consistency**: The random-walk Markov model of each layer, $\boldsymbol{M}_{\boldsymbol{A}^i}$, $i \in [1, 2, .., l]$, is the *projection* of $\boldsymbol{M}_{\mathbb{W}}$ to that layer, and
2. **Ego Consistency**: The egocentric inter-layer dynamics, $\boldsymbol{M}_v$, of vertex $v \in V$, is the *layer marginals* of $\boldsymbol{M}_{\mathbb{W}}$ at vertex v.

The projection of $\boldsymbol{M}$ onto a subset is simply the stochastic normalization of corresponding principal submatrix of $\boldsymbol{M}$. Thus, Condition 1 is automatically achieved by setting diagonal blocks of $\mathbb{W}$ as $\boldsymbol{W}^1, ..., \boldsymbol{W}^l$ in Formulation 1.

For Condition 2, notice that $\mathbb{W}$ also defines an $l \times l$ interlayer adjacency matrix $\boldsymbol{W}_v$ at each $v \in V$. The corresponding random-walk process, $\boldsymbol{M}_{\boldsymbol{W}_v}$, is the projection of $\boldsymbol{M}_{\mathbb{W}}$ to these vertical slices. Let $q_{v,i}$ denote the transition probability of going from vertex v in the i^{th} layer to some u in the same layer, according to $\boldsymbol{M}_{\mathbb{W}}$. Let $\boldsymbol{Q}_v$ be the $l \times l$ diagonal matrix of $[q_{v,i} : i \in [l]]$. Then, $\boldsymbol{Q}_v + \boldsymbol{M}_{\boldsymbol{W}_v} \cdot (\boldsymbol{I} - \boldsymbol{Q}_v)$ denote the layer marginals of the joint Markov model $\boldsymbol{M}_{\mathbb{W}}$ at vertex v. Consequently, Ego Consistency requires that $\boldsymbol{M}_v = \boldsymbol{Q}_v + \boldsymbol{M}_{\boldsymbol{W}_v} \cdot (\boldsymbol{I} - \boldsymbol{Q}_v)$. Intuitively, $\boldsymbol{M}_v$ bridges between the orthogonal projections by including both $\boldsymbol{Q}_v$ and $\boldsymbol{M}_{\boldsymbol{W}_v}$. Now we present the main theorem of this paper:

Theorem 2. *For any multi-layer data $(\boldsymbol{A}_i : i \in [l], \boldsymbol{M}_v : v \in V)$, there exists a unique and feasible super-adjacency $\mathbb{W}$ that satisfies both Layer Consistency and Ego Consistency.*

Proof. Because Formulation 1 requires that all off-diagonal blocks of $\mathbb{W}$ are diagonal matrices, we have $(l^2 - l)n$ degrees of freedom after meeting Layer Consistency. Notice that $(\boldsymbol{M}_v : v \in V)$ are n stochastic $l \times l$ matrices. Thus, Ego Consistency represents $(l^2 - l)n$ dimensional constraints, which matches perfectly with the remaining degrees of freedom. Uniqueness proven.

To prove the feasibility of the unique solution, we introduce Algorithm 1,

Algorithm 1 Multilayer network composition

Input: weighted network layers: $G^1 = (V, E^1, \boldsymbol{A}^1), G^2 = (V, E^2, \boldsymbol{A}^2), ..., G^l = (V, E^l, \boldsymbol{A}^l)$, parameters of the dynamics: $\boldsymbol{T}^1, B^1, \boldsymbol{T}^2, B^2, ..., \boldsymbol{T}^l, B^l$, and n $l \times l$ egocentric inter-layer dynamics M_u for each vertex $u \in V$
Algorithm: For each layer i,

- Apply the bias transformation $\boldsymbol{A}'^i = B^i \boldsymbol{A}^i$ ($\boldsymbol{A}'^i = B^i \boldsymbol{A}^i B^i$ for undirected graphs)
- Apply the delay transformation with a global scaling $\boldsymbol{W}^i = \alpha^i(\boldsymbol{A}'^i + (\boldsymbol{T}^i - \boldsymbol{I})\boldsymbol{D}'^i)$
- Create a $ln \times ln$ empty matrix $\mathbb{W}$
- Fill the l diagonal blocks (each of size $n \times n$) with $\boldsymbol{W}^1, \boldsymbol{W}^2, ..., \boldsymbol{W}^l$
- Construct the off diagonal blocks $\boldsymbol{W}^{ij}$ (each of size $n \times n$) for all layer pairs i and j based on Algorithm 2 with $\boldsymbol{W}^1, \boldsymbol{W}^2, ..., \boldsymbol{W}^l$ as inputs

Output The super adjacency matrix $\mathbb{W}$

We need a subroutine Algorithm 2 to satisfy the Ego Consistency. Rearrange $\mathbb{W}$ so that the counterparts of the same vertex are grouped together. $\bar{\mathbb{W}} = \begin{bmatrix} \boldsymbol{W}_1 & \boldsymbol{W}_{12} & ... & \boldsymbol{W}_{1n} \\ ... & & & \\ \boldsymbol{W}_{n1} & \boldsymbol{W}_{n2} & ... & \boldsymbol{W}_{n,} \end{bmatrix}$ where $\boldsymbol{W}_{u,v}$ are $l \times l$ matrices that have already been fixed by Layer Consistency. The diagonal blocks $\boldsymbol{W}_v : v \in V$ contains all entries set by Ego Consistency. The diagonal entries of $\boldsymbol{W}_v$, $v \in V$, are also set by Layer Consistency, because they are unaffected by the rearrangement of $\mathbb{W}$. The rest $n(l^2 - l)$ entries lead to the same degrees of freedom we discussed earlier.

The reordered $\boldsymbol{W}_u$ blocks are closely related to the egocentric adjacencies $\boldsymbol{X}_u$ underlying the egocentric inter-layer dynamics $\boldsymbol{M}_u$. The vertical slice in Figure 1c demonstrates such a $\boldsymbol{X}_a$, where intra-layer transitions are captured using self-loops. Subroutine Algorithm 2 can now be specified as

Algorithm 2 Building inter-layer blocks

Input: transformed layers: $G_1 = (V, E^1, \boldsymbol{W}^1), G_2 = (V, E^2, \boldsymbol{W}^2), ..., G_l = (V, E^l, \boldsymbol{W}^l)$, and a $l \times l$ egocentric inter-layer transition matrix $\boldsymbol{M}_u$ for vertex $u \in V$.
Algorithm:

- Create a $l \times l$ empty matrix $\boldsymbol{X}_u$
- Fill the diagonal elements with $\boldsymbol{X}_u^{ii} = d_u^i(out)$
- Construct the off diagonal elements $\boldsymbol{X}_u^{ij} = \frac{\boldsymbol{M}_u^{ij}}{\boldsymbol{M}_u^{ii}} d_u^i(out)$

Output Block $\boldsymbol{X}_u$ and repeat for each $u \in V$

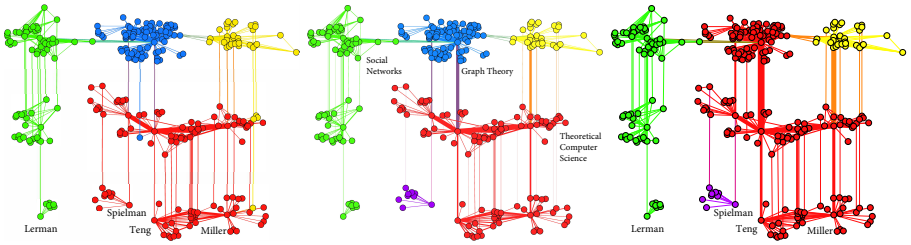
Using Lemma 1, we can rewrite the steps in Algorithm 2 as $\boldsymbol{X}_u = \boldsymbol{M}_u \boldsymbol{\Gamma}$, by setting the i^{th} entry of $\boldsymbol{\Gamma}$ uniquely as $d_u^i(out)/\boldsymbol{M}_u^{ii}$, where $d_u^i(out)$ is the total out edge weights of vertex u in layer i. Intuitively, we are simply using the intra-layer dynamics to determine the vertex scaling factor.

From Figure 1c, it is clear that the off-diagonal parts of $\boldsymbol{X}_u$ is exactly what we are looking for in $\boldsymbol{W}_u$ blocks. Or $\boldsymbol{W}_u = \boldsymbol{X}_u - D_u(out)$, where the diagonal matrix $D_u(out)$ is composed of $d_u^i(out)$ entries. With the uniquely solvable $\boldsymbol{X}_u$ blocks, we can now complete the output $\mathbb{W}$ by filling its off diagonal blocks $\boldsymbol{W}^{ij}$ with reordered $\boldsymbol{W}_u$ blocks. On top of that, Algorithm 2 will always lead to feasible solutions with the constrains $\boldsymbol{X}_u^{ij} \geq 0$, provided that $\boldsymbol{M}_u$ entries are well defined. Uniqueness and feasibility proven.

3 A Real World Example

Figure 2 presents collaboration networks (undirected) centered around four authors: Shang-hua Teng, Daniel Spielman, Gary Miller and Kristina Lerman, as well as their coauthors on papers appearing in the ACM Digital Library. Each layer represents a separate time period: from bottom to top, 1985-1994, 1995-2004, and 2005-2014. The weight of an intra-layer edge represents the number of times two authors collaborated during that decade. The traditional approach, as shown by Figure 2(a), simply connect the same author between neighboring decades with a constant coupling strength of 2. To use our framework, we assume that an author has a 10%/20% transition probability to connect with former self in the previous decade. Specifying inter-layer edges using Algorithm 1 and 2, first between the top tow layers then the bottom two, we have Figure 2(b)/(c).

For comparison, we have visualized the community structures with different multi-layer compositions in Figure 2, using the Louvain algorithm [3], with the resolution parameter set to 5 [11]. The traditional approach produced some counter-intuitive cross-layer communities, because the constant coupling strength is too strong for peripheral vertices. Our framework, on the other hand, leads to much more sensible results with different inter-layer strength for vertices with different degrees. With a 10% transition probability, we can see that authors surrounding Teng and Spielman later separated from the red community (which became yellow) of theoretical computer scientists like Miller. As Teng started



(a) Constant couplings (b) 10% transition (c) 20% transition

Fig. 2: Community structures of coauthor networks using different compositions

to collaborate with the social network community surrounding Lerman (green), the newly formed blue community now represents the filed of graph theory.

References

1. Acar, E., Yener, B.: Unsupervised multiway data analysis: A literature survey. Knowledge and Data Engineering, IEEE Transactions on 21(1), 6–20 (2009)
2. Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J.J., Vespignani, A.: Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences 106(51), 21484–21489 (Dec 2009)
3. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 10, 10008 (Oct 2008)
4. Borgatti, S.: Centrality and network flow. Social Networks 27(1), 55–71 (Jan 2005)
5. De Domenico, M., Sole, A., Gomez, S., Arenas, A.: Random Walks on Multiplex Networks. ArXiv e-prints (Jun 2013)
6. Dunning, D., Cohen, G.L.: Egocentric definitions of traits and abilities in social judgment. Journal of Personality and Social Psychology 63(3), 341–355 (1992)
7. Gallotti, R., Barthelemy, M.: The multilayer temporal network of public transport in Great Britain. Scientific Data 2, 140056 (Jan 2015)
8. Ghosh, R., Teng, S.h., Lerman, K., Yan, X.: The Interplay Between Dynamics and Networks: Centrality, Communities, and Cheeger Inequality. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1406–1415. KDD '14, ACM, New York, NY, USA (2014)
9. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: KDD '03. pp. 137–146. ACM (2003)
10. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A.: Multilayer Networks. ArXiv e-prints (Sep 2013)
11. Lambiotte, R., Delvenne, J.C., Barahona, M.: Laplacian dynamics and multiscale modular structure in networks. arXiv preprint arXiv:0812.1770 (2008)
12. Mucha, P.J., Richardson, T., Macon, K., Porter, M.A., Onnela, J.P.: Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. Science 328, 876– (May 2010)
13. Newman, M.: Networks: An Introduction. OUP Oxford (2010)
14. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web (1999)
15. Wasserman, S., Faust, K.: Social network analysis: Methods and applications, vol. 8. Cambridge university press (1994)