Interdependence Between Network Layers

Cristopher Moore, Santa Fe Institute

joint work with

Caterina De Bacco, SFI/Columbia/Max Planck
Daniel Larremore, SFI/Boulder
Eleanor Power, SFI/London School of Economics
Multilayer networks

Different types of links: social relationships, modes of transportation, loci of similarity…

Hypothesis: explained by common community structure…

…even if the layers seem very different

Don’t ask whether two layers are correlated: ask whether knowing one helps predict the other
Multilayer networks

- assortative
- disassortative
- aggregated
Modeling the adjacency tensor

\[ A_{ij}^{(\alpha)} = \# \text{ of links from } i \text{ to } j \text{ in layer } \alpha \]

\[ A_{ij}^{(\alpha)} \sim \text{Poisson with mean } M_{ij}^{(\alpha)} \]

# of groups

\[ M_{ij}^{(\alpha)} = \sum_{k,\ell=1}^{K} u_{ik} v_{j\ell} w_{k\ell}^{(\alpha)} \]

Membership of vertex i in group k

Membership of vertex j in group l

Density of edges of type \( \alpha \)

from group k to group l

Mixed-membership

Allows inhomogeneous degrees

Different roles for incoming and outgoing links

Arbitrary structure in each layer: assortative, core-periphery, etc.
Aside: tensor factorization

rank-one matrix: \( A_{ij} = u_i v_j \)

rank-\( K \) matrix: \( A_{ij} = \sum_{\ell=1}^{K} u_{i}^{(\ell)} v_{j}^{(\ell)} \)

to compute matrix rank, just use linear algebra

can approximate (minimize \( L_2 \)) using singular values

rank-one tensor: \( A_{ijk} = u_i v_j w_k \)

rank-\( K \) tensor: \( A_{ijk} = \sum_{\ell=1}^{K} u_{i}^{(\ell)} v_{j}^{(\ell)} w_{k}^{(\ell)} \)

tensor rank is NP-complete!

rank of matrix multiplication tensor is unknown

even minimizing \( L_2 \) requires an iterative algorithm
The EM algorithm

\( \rho_{ijkl}^{(\alpha)} \) is our estimate of the probability that an edge \( i \rightarrow j \) in layer \( \alpha \) is due to their being in groups \( k \) and \( l \)

maximizing the likelihood gives update equations:

\[
\begin{align*}
    u_{ik} &= \frac{\sum_{j,\alpha} A_{ij}^{(\alpha)} \sum_{\ell} \rho_{ijkl}^{(\alpha)}}{\sum_{\ell} \left( \sum_j v_{j\ell} \right) \left( \sum_{\alpha} w_{k\ell}^{(\alpha)} \right)} \\
    v_{j\ell} &= \frac{\sum_{i,\alpha} A_{ij}^{(\alpha)} \sum_k \rho_{ijkl}^{(\alpha)}}{\sum_k \left( \sum_i u_{ik} \right) \left( \sum_{\alpha} w_{k\ell}^{(\alpha)} \right)} \\
    w_{k\ell}^{(\alpha)} &= \frac{\sum_{ij} A_{ij}^{(\alpha)} \rho_{ijkl}^{(\alpha)}}{\left( \sum_i u_{ik} \right) \left( \sum_j v_{j\ell} \right)} \\
    \rho_{ijkl}^{(\alpha)} &= \frac{u_{ik} v_{j\ell} w_{k\ell}^{(\alpha)}}{\sum_{k',\ell'} u_{ik'} v_{j\ell'} w_{k'\ell'}^{(\alpha)}}
\end{align*}
\]

iterate until fixed point = local maximum of the likelihood

(gives a point estimate: there are also Bayesian versions)
Cross-validation and link prediction

Can the model fill in missing data?

Hide 20% of the links, use 80% as training data

Generative models assign probabilities to missing links: use AUC to measure accuracy

Avoid overfitting, and select # of communities

But! “Best” model depends on what kind of data is hidden, and what we are trying to predict: application-dependent (as it should be)
Using link prediction to measure interdependence between layers

Does knowing one layer help predict links in another one?

“Similar” (not correlated!) layers have common community structure, and help predict each other.
Using link prediction to measure interdependence between layers

Does knowing one layer help predict links in another one?

“Similar” (not correlated!) layers have common community structure, and help predict each other.

Layers are redundant if they reveal same latent features of the nodes: don’t need to ask about both.

But knowing one layer may make it harder to predict another, if their structures are inconsistent.
Experiments on social networks

Two Indian villages
~400 nodes, ~7000 edges
One village consists of two separate hamlets
12 layers: looking for work, babysitting, borrowing, discussing important issues
A common community structure

Correlated with caste, gender, and geography
Knowing more layers helps

FIG. 1. Teṇpatṭi Village community partition. On the left we show the division by caste membership. To the right we show the membership in each of the 4 communities for each node (each figure represents one community), with color ranging from white if the normalized out-going membership \( u_{ik} = 0 \) to black if \( u_{ik} = 1 \). Values in between denote overlapping membership (grey).

The fact that caste membership partially overlaps with the communities identified by our algorithm suggests a relationship between topological structure and caste, a topic that will be investigated in a future paper.

Work Position Babysit Loan Advice Talk Impiss Close Errand Borrow Defend Cash

AUC

L=1

L=2

L=3

L=9

FIG. 2. Layer interdependence in the Indian social support networks. On the x-axis are the layers' labels used in the test dataset, and the y-axis shows the AUC obtained through the cross-validation schemes for measuring layer interdependence. Bold lines are for Teṇpatṭi Village, dashed for Aḻakāpuram Village. \( L=1 \) refers to single-layer AUC, where the algorithm is only given access to that layer. \( L=2, 3, 12 \) show the increase in the AUC for that layer when the algorithm is given access to 1 additional layers; for \( L=2 \) and \( L=3 \) we choose the best set of 1 additional layers using the greedy procedure described in the previous section.

First, we analyze social support networks from two villages in the Indian state of Tamil Nadu, which we call by the pseudonyms "Teṇpatṭi" and "Aḻakāpuram". As part of a survey questionnaire, village residents were asked to name those individuals who provided them with 12 different types of support, ranging from lending them household items to helping them navigate government bureaucracy. The resulting directed networks have \( N=362 \) and \( N=420 \) nodes, respectively. Each type of support corresponds to a layer in these networks, giving each of them \( L=12 \) layers, with average degrees ranging from 2.0 to 4.4.

Second, we analyze the patterns of shared genetic substrings among a set of malaria parasite virulence genes. Each of the \( N=307 \) nodes represents a single gene, and an edge connects two genes if they share a substring of significant length. Due to the fact that the same set of genes was analyzed at nine different genetic loci (i.e., locations on the genes themselves) which are
Interdependence and similarity

“Who do you discuss important issues with?” helps predict many layers; looking for work, babysitting less so.

Can also cluster layers by affinity matrices $w^{(\alpha)}$: 

![Diagram showing clusters of layers in principal component analysis](image)
In contrast: malaria genetics

~300 variants of a gene in the malaria parasite

9 layers, similarity at different loci

highly variable to avoid immune response

more layers, less accurate: no consistent communities

the joy of negative results…
Conclusions and questions

Can use tensor factorization to express models of multilayer networks

Instead of correlation, ask whether one layer helps predict another: do they reveal similar latent structures?

This is model-dependent… what are we missing?

Local rules, e.g. if (a,b) and (b,c) have relation #1, then (a,c) have relation #2
To put it bluntly: this book rocks! It somehow manages to combine the fun of a popular book with the intellectual heft of a textbook.

Scott Aaronson, MIT

This is, simply put, the best-written book on the theory of computation I have ever read; one of the best-written mathematical books I have ever read, period.

Cosma Shalizi, Carnegie Mellon