

# Title of submission to PLOS journals

Name1 Surname<sup>1,2❶</sup>, Name2 Surname<sup>2❷</sup>, Name3 Surname<sup>2,3❸</sup>, Name4 Surname<sup>2</sup>, Name5 Surname<sup>2❹</sup>, Name6 Surname<sup>2❺</sup>, Name7 Surname<sup>1,2,3❻</sup>, with the Lorem Ipsum Consortium<sup>❻</sup>

**1** Affiliation Dept/Program/Center, Institution Name, City, State, Country

**2** Affiliation Dept/Program/Center, Institution Name, City, State, Country

**3** Affiliation Dept/Program/Center, Institution Name, City, State, Country

❶These authors contributed equally to this work.

❷These authors also contributed equally to this work.

❸Current Address: Dept/Program/Center, Institution Name, City, State, Country

❹Deceased

❺Membership list can be found in the Acknowledgments section.

\* [correspondingauthor@institute.edu](mailto:correspondingauthor@institute.edu)

## Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

## Author summary

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

## Introduction

Traditionally, for down stream tasks related to protein structure (e.g., protein structure classification), features were directly extracted from the 3d structure of proteins. For instance, Cui and Lui [1] consider the geometrical properties of a protein structure by viewing a protein molecule to be comprising of several cube-like regions and computing probability based statistics of atoms belonging to different regions. However, this kind of feature representation does not capture connectivity properties of atoms/ amino acids. Representing protein structure as graph of atoms/ amino acid residues where edges capture connectivity between amino acids can provide insights into protein folding as network based approaches can be applied to this graph.

## Single layer networks

Protein structure networks are graphs in which nodes represent amino acid residues and edges capture relationships—based on factors such as interaction energy and proximity—between the amino acids [2, 3]. As opposed to directly extracting geometrical properties of a protein structure using the protein’s structural coordinates, a graph-based representation has been found to be beneficial in its mathematical formalism to perform network analysis [4]. Furthermore, much work has been performed using graph based spectral methods to compare performance of such methods with traditional methods that use structural topologies [5–7].

## Graph spectral method for comparing protein structures

Generally, networks can be exactly matched using techniques such as subgraph isomorphism (i.e., given two graphs, determine whether one graph contains a subgraph of the other and is isomorphic or that there is a one-to-one mapping of the vertices between the two graphs such that edge connections are preserved), clique detection [7]. For inexact matching of networks, graph edit distance can be used. Spectral methods that compare networks by approximately matching the clusters capturing local structural properties include k-means based on Fiedler vectors (note that a Fiedler vector of a graph is the eigenvector corresponding to the second smallest eigenvalue) and main components in principle component analysis [7]. Gadiyaram et al. [7] develop a network comparison score that captures differences at both the local edge level as well as at the global cluster level, unlike eigenvalue based methods that cannot differentiate between isomorphic networks and Hamming distance based methods that do not capture differences in clustering of networks (e.g., [8, 9]).

## Network similarity score [7]

Graphs can be represented as matrices taking on various forms such as adjacency, Laplacian, and normalized Laplacian [10–14]. In spectral graph theory, graphs are represented as matrices and algebraic tools are applied to these matrices to capture various structural aspects of the graphs. Let  $G = (V, E)$  be a graph with  $|V| = n$  (where  $n$  is the number of vertices in the  $G$ ),  $A_G$  be  $n \times n$  matrix where  $A_G(i, j) = 1$  iff  $(i, j) \in G$ , and  $D_G$  be the degree matrix where  $D_G(i, i)$  is the degree of vertex  $i$  in  $G$  and is a diagonal matrix. Then, the Laplacian  $L$  of graph  $G$  is defined as  $L_G = D_G - A_G$ .

## Multilayer networks

## Heterogeneous graphs for biological complex scoring

Understanding interactions between molecules of proteins is one of the main aspects in identifying proteins as drug targets [15]. Crampon et al. [16] propose a new scoring function to capture the affinity between a protein and a ligand (i.e., any chemical compound). One of the applications is molecular docking where the affinity score is used to predict the likelihood of a ligand (i.e., a small molecule) will bind to a protein (i.e., a macro molecule or a receptor). This method can be applied to other biological complexes such as protein-protein, protein-sugar, protein-dna, etc. Typically, to predict the affinity score of the complex, either the protein-protein interaction network or the structure of the molecules is used.

## Protein structure classification

One of the ways to classify protein structures is by using network level features (e.g., graphlet features) extracted using the protein structure network [17]. Since structure of a protein hugely influences the protein's functions, the structure classification can be used to further predict protein function. Another approach for structure classification is by using protein sequence features [18]. However, it has been discussed that proteins with low sequence similarity can have similar 3-d structures [19, 20]. Thus, protein sequence and structure data can be used complementary to each other for the task of protein structure classification.

Newaz et al. [17] propose a multilayer sequential network model to classify protein structures. Given a protein domain (i.e., a sequence of amino acids), they construct multiple unweighted graphs (in which nodes are amino acids and edges denote distances between amino acids) using incremental subsequences of the protein domain. Each graph capturing a snapshot of the domain represents a layer in the multilayer network. For protein structure classification, features extracted from the multilayer network include dynamic graphlet degree [21] and graphlet orbit transition [22, 23].

## Why multilayer networks to model protein complexes

Say we have two sets of disjoint nodes,  $m$  and  $n$ , in a single-layer graph, the total number of possible edges between all nodes is  $\binom{m+n}{2}$ . There are no constraints on how these edges are distributed, as in any node can connect to any other node. In a multilayer graph, the graph is split into 2 distinct layers, each with its own set of nodes  $m$  and  $n$ . The edges are divided into: connections within each layer or intralayer edges and connections between nodes in different layers. Intralayer edges are constrained to a maximum of  $\binom{m}{2}$  edges in the first layer and  $\binom{n}{2}$  edges in the second layer. Maximum number of interlayer edges is  $m \times n$ , as each node in the first layer can connect to every node in the second layer.

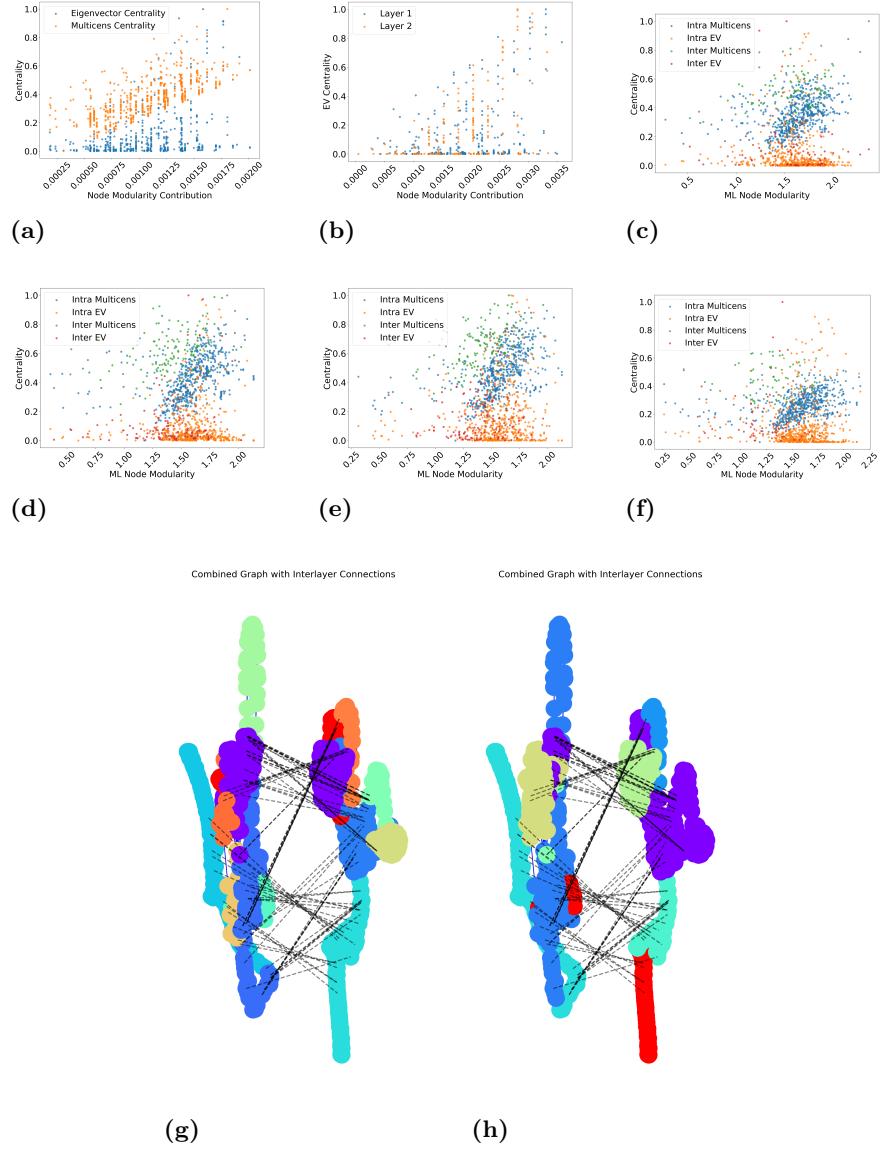
However, presence or absence of interlayer edges often depends on relationships within the individual layers. Nodes in one layer may only connect to nodes in the other layer if there is a certain property (e.g., closeness, similarity, or shared features) established by their intralayer edges. This introduces dependencies between interlayer and intralayer edge configurations, imposing constraints.

In a single-layer graph, all  $\binom{m+n}{2}$  edges are independent, and no specific relationships govern their presence. In a multilayer graph, the structure of the layers directly influences the presence and pattern of interlayer edges. The total edge distribution is shaped by both layer-specific constraints and cross-layer relationships. The resulting graph has far fewer possible configurations than a single-layer graph with the same number of nodes.

Intuitively, in a single layer setting, all individual residues in a protein complex can connect to any residues without restrictions. In a multilayer setting, connections between residues across two chains exist only if they are already part of some underlying structures in their individual chains. That is, in a multilayer setting, a significantly more number of interlayer configurations (out of the possible  $m \times n$ ) are ruled out versus in a single layer setting. For instance, in Clock-Bmal2 complex, there are 354 residues in Clock and 374 residues in Bmal2. While there are  $\sim 2.5k$  edges in individual proteins, there are only  $\sim 100$  edges across them.

Interlayer edges often reflect binding characteristics of residues across chains based on properties within the individual chains such as density or connectivity of intra-chain edges, structural similarity of intralayer properties of residues (e.g. Jaccard similarity, residues are similar if their neighborhoods in their respective chains are similar), interchain edges exist only if residues meet some kind of a threshold criteria derived from

intrachain connectively, interchain edges are formed preferentially between residues that belong to similar or overlapping communities within their chains (e.g., Louvain or Infomap to identify residue communities within each layer; interchain edge presence depends on the overlap of node community memberships), or interchain edges reflect the flow of influence or information across layers, based on some underlying intrachain dynamics.



**Fig 1.** (a) Node modularity vs. Centrality; (b) Node modularity vs. Eigen-vector centrality; (c) Inter vs intra layer modularity of single and multilayer centralities; Inter vs intra layer modularity of single and multilayer centralities using (d) random interface 1, (e) random interface 2 (all pairs shortest path distribution chosen from the original, (f) random interface 3 (all pairs shortest path distribution as close as possible with the original; (g) Node modularity vs. Centrality. (h) Inter-layer modularity depends on intra-layer structures

## Motivation

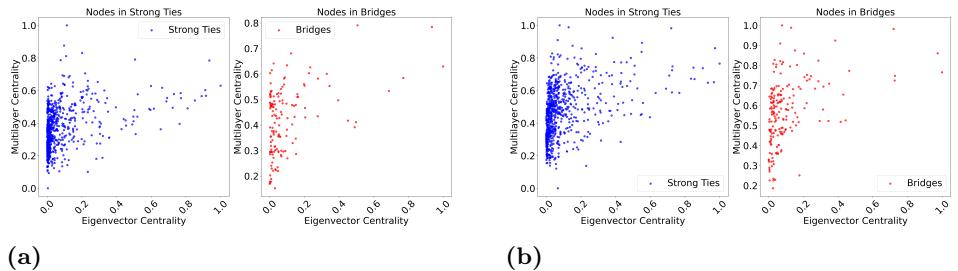
To demonstrate differences in how single-layer and multi-layer centrality measures capture network characteristics, we begin by using modularity to compare intrachain and interchain dynamics.

Modularity captures the underlying structure of the network. It measures how strong a network is with respect to being broken down into communities [24] or subsets of nodes. A high value typically indicates dense graph structure within communities and sparse connections between communities. For an individual node, modularity is not usually computed in isolation but in the context of its community. To understand the presence of dependence of interchain on intrachain structure, we first define node level *greedy* modularity  $Q_i$  for node  $i$  as follows [24–29]:

$$Q_i = \frac{1}{2m} \sum_{i,j} \left[ A_{ij}^{Intra} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (1)$$

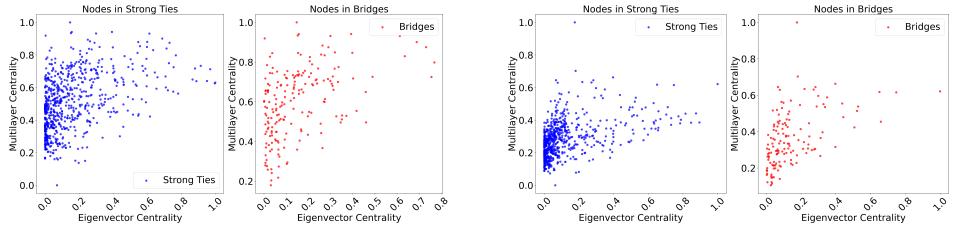
where  $m$  is total number of edges in the graph,  $A_{i,j}^{Intra}$  is the intra-layer connection between  $i, j$ ,  $k_i$  is degree of node  $i$ ,  $k_j$  is degree of node  $j$ ,  $\delta(c_i, c_j)$  is 1 if nodes  $i$  and  $j$  are in the same community and 0 otherwise.

Based on the above definition, nodes with low modularity should have low centrality because centrality captures node's local and global properties while modularity reflects a node's involvement in communities. If a node is not involved in strongly knit sub-networks, such a node should not be central (per degree, betweenness, eigen vector, etc.). Additionally, if indeed interchain structures are dependent on intrachain networks, then node level modularity might have some discernable correlation with multilayer centralities as both capture a node's participation in communities. Single layer centrality measures such as Eigen vector centrality, while capturing a node's influence in local and global network properties, do not particularly measure participation in communities—which may be necessary to quantify its role in interlayer structures, as interlayer structures are much sparser than intralayer structures. Figure 1 (a) shows that node modularity contribution is correlated with MultiCens (a multilayer centrality measure) values, and low centrality values (regardless of whether it is single layer or multi-layer) means low modularity values.

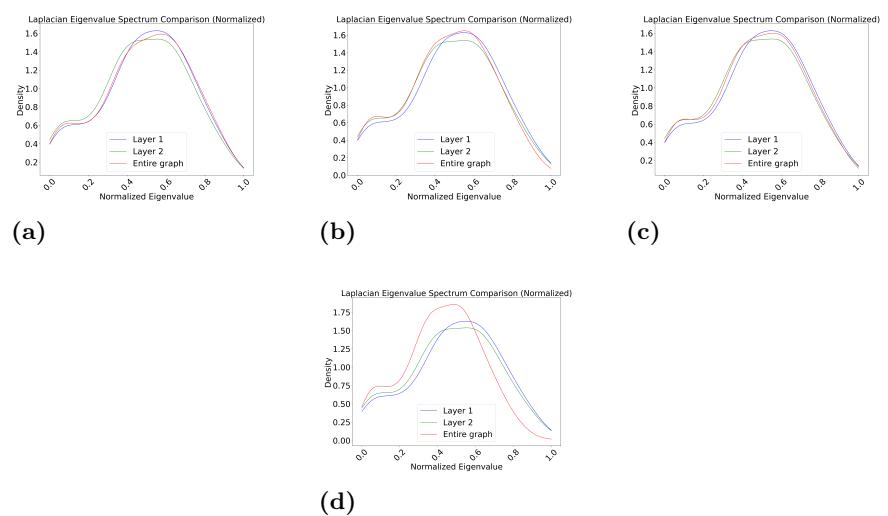


**Fig 2.** (a) Eigen vector vs multicens centrality of strong ties (on the left) and bridges (on the right). Eigen vector vs multicens centrality of strong ties (on the left) and bridges (on the right) for random interlayer configurations; (b) Random interface 1.

When single layer centrality such as Eigen vector centrality values of individual interlayer networks are compared with node modularity contribution to underlying communities (as seen in Figure 1 (b)), we clearly see that as the intralayer centrality and modularity values increase, there is no correlation between them. This shows that there must be some interdependence between the interlayer networks and intralayer structures.



**Fig 3.** Eigen vector vs multicens centrality of strong ties (on the left) and bridges (on the right) for random interlayer configurations; (a) Random interface 2; (a) Random interface 3.



**Fig 4.** (a) Spectral distribution of intra layer graphs and the entire graph. Spectral distribution of intra layer graphs and the entire graph using random interface configurations. (b) Random interface 1. (c) Random Interface 2. (d) Random Interface 3.

## Intra versus inter layer dynamics

Sparsity of edges in interlayer connections is strictly more than that in intralayer connections. Exploring this further, modularity of individual CLOCK and Bmal2 networks (0.71 and 0.71) is slightly less than that of the entire complex network (0.77). In addition to restrictive interlayer sparsity, this shows that inter-layer modularity depends on the intra-layer communities. From the view of multilayer formalism, let inter-layer modularity be defined as:

$$Q_{inter} = \omega \sum_{i,j} A_{ij}^{Inter} \delta(c_i, c_j) \quad (2)$$

where  $\omega$  is inter-layer coupling strength and  $A_{i,j}^{Inter}$  is the inter-layer connection between  $i, j$ , and as before,  $\delta(c_i, c_j)$  is 1 if nodes  $i$  and  $j$  are in the same community and 0 otherwise. Quantifying the inter-layer contribution in the overall modularity based on Equation 2 gives 0.41. The total modularity thus is 0.71+0.41, which is 1.12. Figure 1 (a) shows that inter-layer communities depend on the intra-layer structure. Figure 1 (h) shows the community structure of the entire protein complex graph. In Figure 1 (g), communities are detected for each intra-layers separately. The inter-layer connections are plotted using dashed lines. The figure (h) shows that inter-layer edges connect to

similar communities between layers.

Figure 1 (c) shows that Multicens captures inter layer modularity better than eigen-vector centrality. Not all inter layer edges contribute to overall modularity directly as the intra layer structures influence inter layer connections. Multicens, by modeling inter layer transitions based on intra layer rankings, models hierarchical relationship between inter and intra layer edges. It distinguishes modularity of intra and inter layer nodes better than eigen vector centrality. Eigen vector centrality has low values for most intra layer nodes and lacks the granularitly of information capturing dependence between intra and inter layer modularity. Intra and inter layer centrality of multicens are better correlated with multi-layer modularity than that of eigen vector centrality. In the case of eigen vector centrality, most of the nodes have low centrality regardless of their modularity (or contribution to community formation).

Figures 1 (d), (e), and (f) show intra and inter layer centrality vs modularity for random inter layer configurations. On the left, inter layer configuration is random. As we move to the right, the topology of the interface gets closer to the original. Relationship between Inter/Intra layer modularity captured by Multicens helps it distinguish random inter layers better than eigen vector centrality. In Multicens, intra layer modularity changes across different random interlayers, whereas in eigen vector centrality, intra layer modularity does not change much across different random inter layer configurations. Furthermore, Multicens places emphasis on the inter layer structures and thus intra layer patterns change across the various random inter layer configuration even though intra layer structures remain unchanged.

Multicens captures relationship between strong ties and bridges better than eigen vector centrality. A strong tie is an edge whose vertices belong to the same community. A bridge is an edge whose vertices belong to different communities. No strong tie is a bridge; some inter layer edges are strong ties and rest are bridges. By modeling interdepence between intralyer and interlayer structures, Multicens distinguishes strong ties and bridges (as shown in Figure 2 (a)) by scoring bridges in a broader range (i.e. strong ties are central while bridges may or may not be) than strong ties' centrality values.

Figures 2 (b), 3 (a), (b) show that Multicens has clear grouping patterns for different random inter layers for bridges and strong Ties; eigen vector centrality does not (indicating existence of dependence between intra/inter layer structures). On the other hand, eigen vector centrality captures spectral properties that model local and global properties of the graph structure. Spectral distribution (as shown in Figure 4 (a) for original interface and Figures 4 (b), (c), (d) for random interlayer configurations) clearly distinguishes random from original interfaces. Local and global structural properties captured by the spectral distribution ensures that the overall network's distribution must lie within those of the individual intra layer networks. Thus, a random interface that is topologically similar to the original interface is detected to modify global structure of the network.

Thus, the hypothesis is that devising a centrality measure that captures both: (a) modularity as a factor of the interdependence of interlayer configuration on intralayer structures, and, (b) properties capturing local and global properties together, can help understand the interlayer and intralayer node properties better in protein complexes.

## Multilayer Graph Centrality

Centrality measures such as degree, betweenness [30], and closeness centrality [30] are based on static, single-layer networks. However, in complex systems such as protein complexes, connectivity is often multilayered. Here, we introduce a new node scoring metric derived from k-hop Jaccard similarity and matrix factorization. The goal of this

metric is to quantify node importance by leveraging localized graph overlap patterns while incorporating a learned latent structure. The metric integrates multilayered structural information using k-hop Jaccard similarity across different layers, followed by a matrix factorization approach to refine the node importance score.

## Methodology

For a protein complex network composed of two chains or layers,  $G_{l1}$  and  $G_{l2}$  (for simplicity, we consider complexes of two chains, however, the formulation can easily be extended to model complexes with many chains), and the entire graph  $G$ , we compute the  $k$ -hop Jaccard similarity for every pair of nodes  $A$  and  $B$  separately for each layer ( $L_1$  and  $L_2$ ) and for the entire graph:

$$J(A, B) = \frac{|N_k(A) \cap N_k(B)|}{|N_k(A) \cup N_k(B)|}$$

where  $N_k(A)$  represents the  $k$ -hop neighborhood of node  $A$  and  $N_k(B)$  represents the  $k$ -hop neighborhoood of node  $B$ . The  $k$ -hop Jaccard similarity matrix captures local node connectivity by measuring the fraction of shared neighbors within a  $k$ -hop neighborhood. The self-similarity matrices for  $l1$  and  $l2$  are scaled and placed into respective blocks of a larger matrix  $W$ , capturing intra-layer relationships. We construct the following overlap matrices:

$$W = \begin{bmatrix} scale(W_{L1}) & 0 \\ 0 & scale(W_{L2}) \end{bmatrix}$$

where  $W_{L1}$  is similarity matrix for Layer 1,  $W_{L2}$  is similarity matrix for Layer 2, and scaling ensures comparable magnitudes across layers.

An initial random matrix  $A$  is generated, which serves as the latent structure to be learned. The optimization objective minimizes the difference between  $WA$  and  $X$ , incorporating an  $L_2$ -regularization term:

$$\min_A ||WA - X||_F^2 + \lambda ||A||_F^2.$$

To refine node scores, we perform an iterative learning process using gradient-based optimization [31]:

$$A^{(t+1)} = A^{(t)} - \eta [2W^T(WA - X) + 2\lambda A]$$

where  $\eta$  is the learning rate,  $\lambda$  is a regularization term,  $A$  is the learned weight matrix, randomly initialized, and  $X$  is the similarity matrix for the entire network ( $G$ ). Finally, the multilayer centrality (MLC) score for each node is obtained by summing the row-wise product of  $W$  and  $A$ :

$$NodeScore = \sum (WA)_{row}.$$

This approach integrates local and global structural properties while leveraging the multilayer structure. This effectively learns a transformation matrix that aligns local structures (as captured by  $W$ ) with global node relationships (encoded in  $X$ ). This aggregation captures how much each node contributes to the learned latent structure, effectively ranking nodes based on their structural significance in the graph.

Unlike degree centrality, which considers only local edges, our metric integrates layer-wise connectivity and global network structure. The use of  $k$ -hop Jaccard similarity ensures that node importance reflects not just immediate neighbors but also higher-order neighborhood overlap. Instead of treating all edges equally, the matrix factorization learns node importance, improving robustness over static measures.

## Multicens [32]

Multicens addresses the challenge of modeling interlayer influence in complex networks by separating intralayer influence of nodes with their interlayer influences. Local centrality isolates a node's influence within its layer using a modified PageRank algorithm applied to intralayer adjacency matrices ( $A[i]$ s). By nullifying interlayer connections, it identifies nodes that coordinate localized influences for a particular layer,  $i$ . The iterative equation:

$$l_{layer_i}^{(k+1)} = pA^{[i]}l_{layer_i}^{(k)} + \frac{(1-p)}{n}1^i$$

prioritizes the node influence within each layer. Here,  $l_{layer_i}^{(k)}$  represents layer  $i$ 'th node scores in iteration  $k$ ,  $1^i$  is a vector with entries for nodes in layer  $i$  set to 1 and 0 otherwise,  $n$  is the number of nodes, and  $1 - p$  being the probability that a random walk restarts from any node in the network.

The global centrality captures the remaining effect of a random walk continuing to any node regardless of layer and is given by:

$$g^{(k+1)} = p[(A + C)g^k + Cl] + \frac{(1-p)}{N}1$$

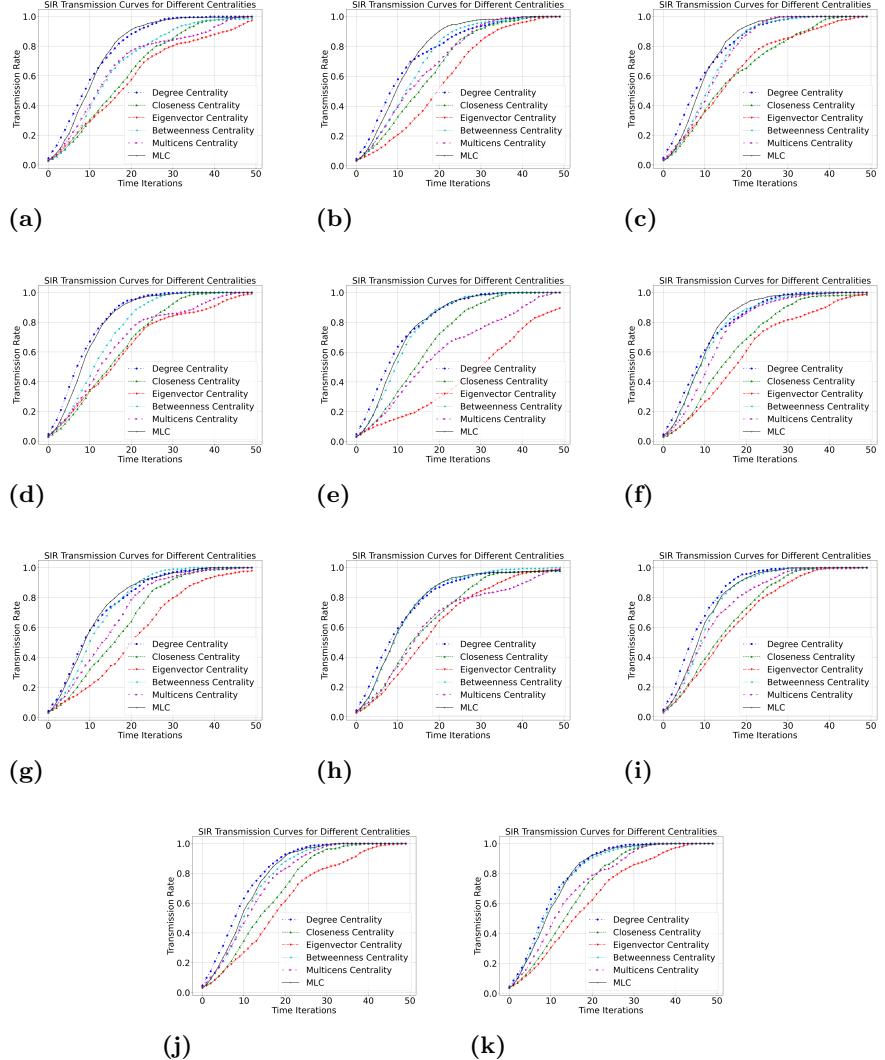
Here  $g^k$  represents the overall node scores in iteration  $k$ ,  $1$ , as previously, is a vector with entries for nodes in the network set to 1 and 0 otherwise,  $N$  is the total number of nodes in the network, and  $p$  being the probability that the random walk continues either within the same layer or jumps to another layer,  $A$  represents the adjacency matrix with intralayer connections, and  $C$  represents the adjacency matrix with interlayer connections.

## Discussion

To evaluate the effectiveness of different centrality measures in determining the spread of an infectious disease, we utilize the Susceptible-Infected-Recovered (SIR) model [33–35]. The model operates on a network where nodes interact via edges, and disease transmission occurs probabilistically based on the network structure. Each node in the network can exist in one of three states: susceptible (S), not yet infected but vulnerable to infection; infected (I), currently infected and capable of spreading the disease; recovered (R), no longer infectious and immune to reinfection. The probability of infection is governed by the transmission rate  $\beta$ , while the probability of recovery is determined by  $\gamma$ . The infection propagates through the network over discrete time steps.

To assess the role of centrality in transmission dynamics, we compare the impact of initializing infections in nodes ranked highest by various centrality measures: degree centrality, nodes with the highest number of direct connections; closeness centrality, nodes that have the shortest average path length to all others; eigenvector centrality, nodes that are well-connected to other influential nodes; betweenness centrality, nodes that frequently serve as intermediaries in shortest paths; multicens centrality, a composite measure incorporating multilayer centrality scores; multilayer centrality, centrality derived from interactions across multiple network layers.

For each centrality measure, the simulation follows these steps: *selecting initial infected nodes*, the top 2% of nodes based on each centrality measure are chosen as the initial infected group; *disease transmission*, at each time step, infected nodes attempt to transmit the disease to their susceptible neighbors with probability  $\beta = 0.1$ ; *recovery*, infected nodes recover with probability  $\gamma = 0.1$  and can no longer transmit the disease; *tracking epidemic spread*, the fraction of infected nodes over time is recorded. The



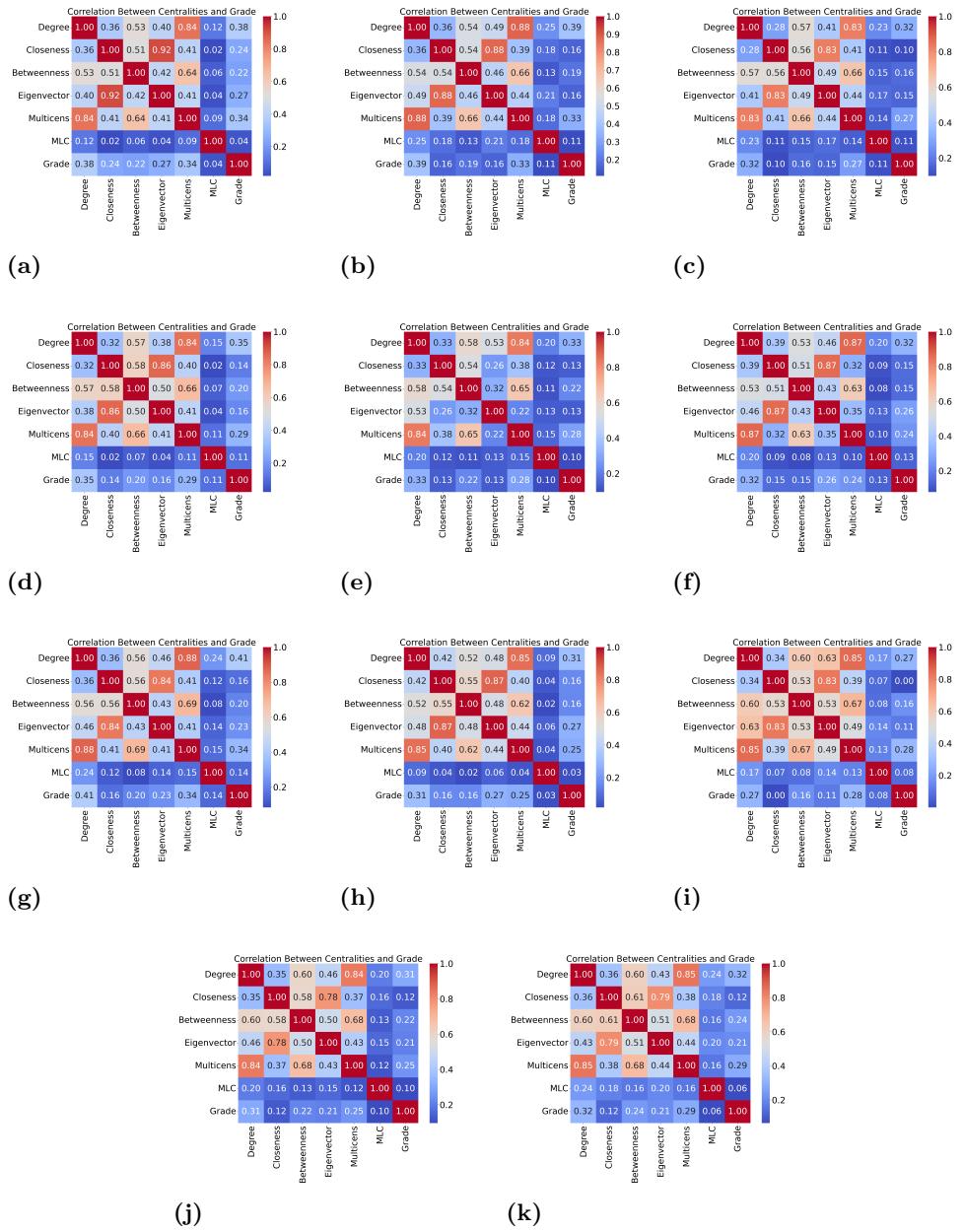
**Fig 5.** SIR transmission curves for degree, closeness, eigenvector, betweenness, multicens, and MLC centralities for proteins: (a) AHR-ARNT; (b) CLOCK-BMAL1; (c) HIF1A-ARNT; (d) HIF2A-ARNT; (e) HIF3A-ARNT; (f) NPAS1-ARNT; (g) NPAS2-BMAL1; (h) NPAS3-ARNT; (i) NPAS4-ARNT; (j) SIM1-ARNT; (k) SIM2-ARNT

simulation is run for 50 time steps for each centrality measure to analyze how the infection propagates across the network.

Figure 5 presents the infection curves corresponding to different centrality measures for proteins: (a) AHR-ARNT; (b) CLOCK-BMAL1; (c) HIF1A-ARNT; (d) HIF2A-ARNT; (e) HIF3A-ARNT; (f) NPAS1-ARNT; (g) NPAS2-BMAL1; (h) NPAS3-ARNT; (i) NPAS4-ARNT; (j) SIM1-ARNT; (k) SIM2-ARNT. The key observations include: degree centrality results in rapid early-stage infection due to high connectivity of initially infected nodes resulting in high overall transmission rate; closeness centrality facilitates rapid disease spread across the entire network resulting in average transmission rate; eigenvector centrality causes sustained transmission as high-scoring nodes reinforce long-term infection spread which in the protein networks leads to poor transmission rate; betweenness centrality delays the initial infection peak

but leads to widespread outbreaks due to its control over network pathways, leading to relatively high transmission rate; multicens centrality exhibits a balanced spread pattern, combining aspects of individual centrality measures leading to average transmission rate; multilayer centrality (MLC) captures complex inter-layer interactions, altering the disease spread trajectory compared to single-layer measures leading to competitive transmission rates, higher than that achieved by most of the other centrality measures.

Figure 6



**Fig 6.** Correlation between degree, closeness, betweenness, eigenvector, multicens, and MLC centralities and conserved grade scores for proteins: (a) AHR-ARNT; (b) CLOCK-BMAL1; (c) HIF1A-ARNT; (d) HIF2A-ARNT; (e) HIF3A-ARNT; (f) NPAS1-ARNT; (g) NPAS2-BMAL1; (h) NPAS3-ARNT; (i) NPAS4-ARNT; (j) SIM1-ARNT; (k) SIM2-ARNT

## Conclusion

307

## Acknowledgments

308

Cras egestas velit mauris, eu mollis turpis pellentesque sit amet. Interdum et malesuada fames ac ante ipsum primis in faucibus. Nam id pretium nisi. Sed ac quam id nisi malesuada congue. Sed interdum aliquet augue, at pellentesque quam rhoncus vitae.

309

310

311

## References

1. Cui C, Liu Z. Classification of 3d Protein based on Structure Information Feature. In: 2008 International Conference on BioMedical Engineering and Informatics (ICBEI). IEEE; 2008. p. 27–30.
2. Vijayabaskar MS, Vishveshwara S. Interaction energy based protein structure networks. *Biophys J*. 2010;99(11):3704–15.
3. Bhattacharyya M, Ghosh S, Vishveshwara S. Protein Structure and Function: Looking through the Network of Side-Chain Interactions. *Curr Protein Pept Sci*. 2016;17(1):4–25.
4. Schieber TA, Carpi L, Díaz-Guilera A, Pardalos PM, Masoller C, Ravetti MG. Quantification of network structural dissimilarities. *Nat Commun*. 2017;9(8):13928.
5. Bhattacharyya M, Bhat CR, Vishveshwara S. An automated approach to network features of protein structure ensembles. *Protein Sci*. 2013;22(10):1399–416.
6. Chakrabarty B, Parekh N. PRIGSA: protein repeat identification by graph spectral analysis. *J Bioinform Comput Biol*. 2014;12(6):1442009.
7. Gadiyaram V, Ghosh S, Vishveshwara S. A graph spectral-based scoring scheme for network comparison. *Journal of Complex Networks*. 2017;5(2):219–244.
8. Jurman G, Visintainer R, Filosi M, Riccadonna S, Furlanello C. The HIM glocal metric and kernel for network comparison and classification. In: IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE; 2015. p. 1–10.
9. Jurman G, Visintainer R, Furlanello C. An introduction to spectral distances in networks. *Frontiers in Artificial Intelligence and Applications*. 2011;226:227–234.
10. Butler S, Chung F. Spectral Graph Theory. Florida: O'Reilly; 2006.
11. Butler S. Algebraic aspects of the normalized Laplacian. Switzerland: Springer International Publishing; 2015.
12. Chung F. Spectral Graph Theory. USA: American Mathematical Society; 1997.
13. Luxburg Uv. A tutorial on spectral clustering. *Statistics and Computing*. 2007;17:395–416.
14. Zumstein P. Comparison of spectral methods through the adjacency matrix and the Laplacian of a graph. ETH Zürich; 2005.

15. Xie X, Yu TY, et al. Recent advances in targeting the “undruggable” proteins: from drug discovery to clinical trials. *Signal Transduct Target Ther.* 2023;8(1):335.
16. Crampon K, Giorkallos A, Vigouroux X, Baud S, Steffenel LA. Heterogeneous graph convolutional neural network for protein-ligand scoring. *Exploration of Drug Science.* 2023;1:126–39.
17. Newaz K, Piland J, Clark P, Emrich S, Li J, Milenković T. Multi-layer sequential network analysis improves protein 3D structural classification. *Proteins.* 2022;90(9):1721–1731.
18. Xia J, Peng Z, Qi D, Mu H, Yang J. An ensemble approach to protein fold classification by integration of template-based assignment and support vector machine classifier. *Bioinformatics.* 2016;22(6):863–870.
19. Sousounis K, Haney J Carland Cao, Sunchu B, Tsonis P. Conservation of the three-dimensional structure in non-homologous or unrelated proteins. *Hum Genomics.* 2012;6(1):10.
20. Kosloff M, Kolodny R. Sequence-similar, structure-dissimilar protein pairs in the PDB. *Proteins.* 2008;71(2):891–902.
21. Hulovatyy Y, Chen H, Milenković T. Exploring the structure and function of temporal networks with dynamic graphlets. *Bioinformatics.* 2015;31(12):i171–80.
22. Aparício D, Ribeiro P, Silva F. Graphlet-orbit Transitions (GoT): A fingerprint for temporal network comparison. *PLoS One.* 2018;13(10):e0205497.
23. Aparício D, Ribeiro P, Milenković T, Silva F. Temporal network alignment via GoT-WAVE. *Bioinformatics.* 2019;35(18):3527–3529.
24. Newman M. Modularity and community structure in networks. *Proceedings of the national academy of sciences.* 2006;103(23):8577–8582.
25. Newman M. Analysis of weighted networks. *Physical Review E.* 2004;70(5):056131.
26. Reichardt, Bornholdt. Statistical Mechanics of Community Detection. *Phys Rev.* 2006; p. E74.
27. Clauset A, Newman ME, Moore C. Finding community structure in very large networks. *Physical Review E.* 2004;70(6).
28. Newman M. Networks: An Introduction. Oxford University Press. 2011; p. 224.
29. Newman ME. Fast algorithm for detecting community structure in networks. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Top.* 2004;69(5):0309508.
30. Newman M. Networks : An introduction. New Delhi: Oxford University Press; 2012.
31. Ruder S. An overview of gradient descent optimization algorithms. *CoRR.* 2016;abs/1609.04747.
32. Kumar T, Sethuraman R, Mitra S, Ravindran B, Narayanan M. MultiCens: Multilayer network centrality measures to uncover molecular mediators of tissue-tissue communication. *PLOS Computational Biology.* 2023;19(4):1–33.

33. Bailey NT. The mathematical theory of infectious diseases and its applications. *Nat Immunol.* 1977;28(2):479–480.
34. Wang Y, Xiao G. Effects of interconnections on epidemics in network of networks. In: Proc. 7th Int. Conf. Wireless Commun., Netw. Mobile Comput.; 2011. p. 1–4.
35. Hai-Feng Z, Small M, Xin-Chu F. Different epidemic models on complex networks. *Commun Theor Phys.* 2009;52(1):180–184.