DCI-Based Attack Strategies for Optimal Percolation of Duplex Networks

Lucia Pezzetti[a,b]

This manuscript was compiled on January 27, 2024

The study of optimal percolation in complex networks is a critical subject for understanding the resilience and vulnerability of complex systems, ranging from biological networks to infrastructure and social systems. Multiplex networks, characterized by multiple layers of interconnected networks with the same set of nodes, exhibit unique percolation properties that are not fully captured by traditional single-layer network models. This work aims to contribute to the study of targeted attack strategies in duplex networks and offers a more refined lens through which the impact of node removal can be assessed. Specifically, by employing modified versions of the Duplex Collective Influence (DCI) algorithm — DCI_harmonic and DCI_Katz — we aim to capture the effect of node removal not only in its immediate vicinity but also in its farthest-reaching neighborhood. These new strategies are then compared to the standard DCI version, as well as to other state-of-the-art generalizations of single-layer procedures, in the presence of non-zero inter-layer degree correlation and edge overlap. The analysis is performed by presenting the results of large-scale simulations over both synthetic and real-world duplex networks. Our findings offer a new perspective on the understanding of network dynamics, confirming that optimal strategies for network disruption or reinforcement are highly context-specific, varying significantly across different types of duplex networks. However, the promising results on the robustness of the DCI_harmonic procedure to variations in edge overlap probability open avenues for future research, particularly in exploring the applicability of these findings in more diverse and dynamic network configurations.

Optimal Percolation | Duplex Networks | Centrality Measures |

In the realm of network science, the concept of multiplex networks stands out as a key model for understanding complex systems, particularly when viewed through the lens of percolation theory. These networks, with their multi-layered structure, offer an advanced framework for modeling and analyzing a wide range of real-world systems in which it is essential to depict the diversity of interactions between the same nodes. Indeed, each layer of a multiplex network captures a distinct type of connection or relationship, but all layers are unified by a common set of nodes. This intricate setup is especially interesting for the application of percolation theory, which studies the robustness and connectivity of networks under various conditions of stress and disruption.

The modeling potential of multiplex networks in this context extends across numerous fields, with significant implications for comprehending the resilience and vulnerability of a wide range of complex systems. In epidemiology, for instance, these networks can model the spread of diseases through different social interactions offering insights into how a disease percolates through social layers and informing strategies for containment and control [1–3]. In the framework of urban planning and infrastructure, multiplex networks represent different interconnected utility systems, allowing researchers to simulate and understand how breakdowns in one system can propagate to others [4]. Finally, multiplex networks are essential in mitigating cascading failures in cyber systems. For example, an attack on the physical infrastructure layer, like a data center, can have far-reaching impacts on the software and user layers, potentially leading to a systemic collapse of services. Percolation theory is key to building more resilient cyber systems and formulating effective countermeasures against cyber threats [5, 6]. Due to the enhanced versatility and modeling potential of multiplex networks, the study of optimal percolation in these networks has emerged as a crucial area of research. However, even if the concept of percolation has been extensively studied in the framework of single-layer networks [7], its extension to multi-layer networks [8–10] is far from being fully satisfactory.

This work aims to contribute to this evolving framework. Specifically, we focus on duplex networks, a special case of multiplex networks with two layers, to explore how network structure affects the impact of node removal. The core of our investigation lies in the extension of the Duplex Collective Influence (DCI) algorithm [9, 11]. The novel versions, DCI_harmonic and DCI_Katz, are designed to capture the effect of a node removal not only in its immediate vicinity, but also into its farther-away neighborhood, throught the use of centrality measures, the harmonic centrality and the Katz centrality respectively. The new strategies are then compared through extensive simulations to the standard DCI version, as well as to other generalisations of single-layer procedures [8, 12], in the presence of non-zero inter-layer degree correlation and edge overlap [13–15].

Significance Statement

In this study, we delve into the resilience and vulnerability of complex networks, topics of extreme importance in our increasingly interconnected world. Networks, whether digital, biological, or social, are all susceptible to disruptions that could lead to dramatic consequences. We have developed and tested novel methods for identifying critical sets of nodes whose removal could trigger cascade failures in these networks. Identifying and managing these focal nodes can lead to improved designs and strategies, maintaining the integrity of various networked systems. This research is crucial not only for the field of network theory but also has far-reaching implications across a wide range of applications, including urban planning, disease control, and cybersecurity.

Author affiliations: [a]Department of Mathematics, University of Zurich, Rämistrasse 71, 8006 Zürich, Switzerland; [b]Department of Mathemtics, University of Turin, Via Giuseppe Verdi, 8, 10124 Torino, Italy

1. Theoretical Framework

In the framework of multiplex network, we consider a duplex network model $\mathcal{M}$ composed of two independently fabricated Erdős-Rényi layers. Mathematically, the model is defined by two layers $G_A = (V, E_A)$ and $G_B = (V, E_B)$, each of which is an Erdős-Rényi graph. We say that two nodes belong to the same mutually connected component (MCC) if they are connected by at least a path in each layer of the system. Following the approach of [8], the parameter of interest in this analysis is the relative size of the largest mutually connected component (LMCC). The motivating idea behind this approach is that LMCC is the largest maximal subgraph of the network consisting of mutually connected nodes and thus, it is a natural generalization, in the context of multiplex, of the giant component of single-layer networks. As it is common for single-layer networks, in the framework of duplex, we study the effect of percolation on the size of LMCC and, more specifically, the optimal percolation problem consists in finding the minimal set of nodes which, if removed, would reduce the size of the LMCC of $O(N^{1/2})$. This set is called the critical set and we denote its size with q, moreover, we call its members structural nodes (SNs).

2. Measures of Centrality

In order to delve into the core of the work, we need to understand the subtleties and differences between the different measures of centrality used. In general, centrality measures [16] aim at scoring how important a node is inside a network and, therefore, give a way to identify influential nodes. Each of them has been developed to capture a different characteristic of the network and thus has different applications. In the following we will always consider the case of undirected networks. All measures of centrality are also defined for directed networks, paying attention to handle properly in-edges and out-edge.

A. Degree Centrality. The degree centrality of a node is simply given by its degree. Thus, it takes into account information about the amount of closest neighbors a node has. For this reason, degree centrality alone is not enough to tackle all the complexities in a network. Consider, for example, the case in which a node has degree 2 but its two adjacent nodes have a very high degree, Fig.1a. The degree centrality will rank this node as one of the least important in the network, without being able to capture its role of bridge between two important clusters.

B. Betweenness Centrality. Betweenness centrality [17] ranks the importance of a node according to the number of times it acts as a bridge along the shortest paths between pairs of other nodes. We refer to this capability as the brokering power of the node. Mathematically, consider $\sigma(i,j)$ as the number of shortest paths between node i and node j and $\sigma(i,j|k)$ as the number of shortest paths between node i and node j that pass through the brokering node k. Then the betweenness centrality of node k is defined as:

$$C_b = \sum_{i,j \in V \setminus \{k\}} \frac{\sigma(i,j|k)}{\sigma(i,j)} \in [0, N]$$

where V is the set of nodes of the considered network and $N = |V|$. Clearly this new notion of centrality allows to encode the linking role of node A Fig1a, however, in other situations this may not be so desirable. Looking at Fig.1b, node A has the highest betweenness centrality, even if it may not be considered an important crossroad.

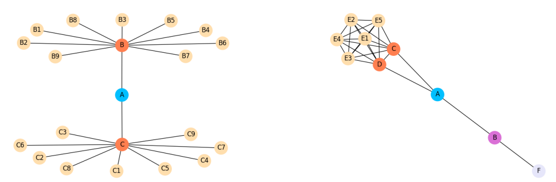
A final remark is that the ability of betweenness centrality to successfully identify nodes that connect different parts of the network is achieved at the expense of intensive computations, especially in case of large networks.

C. Katz-Centrality. Katz-centrality [18] can be considered as an extension of eigenvector centrality, with respect to witch it solves the problem of vanishing scores. Differently than the betweenness centrality, Katz-centrality does not take into account only the shortest paths between two nodes, but all the paths and weights them accordingly to some parameter $\alpha \in (0, \rho(A))^{-1}$ that needs to be carefully tuned. Here, A is the adjacency matrix of the considered network and $\rho(A)$ is its spectrum.

Formally, the Katz-centrality of a node i is defined as:

$$C_K(i) = 1 + \sum_{l=1}^{\infty} \sum_{j \in V} \alpha^l (A^l)_{ij}$$

As a side note, it is worth mentioning that Katz centrality may be more meaningful in tightly knit networks.



(a) Degree centrality and Katz centrality do not to capture the influence of node A in the network.

(b) Node A has the highest betweenness centrality but it may not be the most influential node overall.

Fig. 1. Issues with centrality measures

D. Harmonic Centrality. Harmonic centrality [19] has been proposed to solve the ill-definition of closeness centrality in the context of unconnected networks. Closeness centrality of node i, indeed, is defined as the reciprocal of its total distance, where the total distance is the sum of the lengths of the shortest paths from node i to any other node j. Clearly, such a definition does not provide meaningful values in case of disconnected networks. Conversely, harmonic centrality sums the reciprocal of the shortest path lengths from a node to all other nodes in the network. In this case, unreachable nodes contribute zero to the sum and this allows harmonic centrality to assign to each node a finite value, even when some nodes are unreachable from it, and handle the case of unconnected networks appropriately, without skewing the centrality values.

$$C_h(i) = \sum_{j \in V, j \neq i} \frac{1}{d(i,j)}$$

Due to its nature, harmonic centrality offers a more nuanced understanding of centrality in large networks, especially where path distances vary greatly.

3. Optimal Percolation in Duplex Networks

The issue when dealing with the optimal percolation problem in duplex networks concerns how to aggregate information in different layers in order to meaningfully rank the nodes on the basis of their percolated power and perform a targeted attack to the network. The idea is that the more influential the node, the more disruptive its removal. A first choice is to simply consider the two different layers separately, rank the node of each layer and then combine the two scores via multiplication (HDA). However, it is easy to see that this trivial approach fails to consider meaningful situations, just think about a node that has zero degree on one layer, but a very high centrality measure on the other network. Different strategies have been proposed to address this problem and try to better capture the structure of the duplex as a whole. We quickly present an overview of some of the state-of-the-art targeted attack procedures and then explain our contribution to the field.

A. Effective Multiplex Degree (EDM). The EMD strategy [12] consists in assigning a weight w^i to each node of the network by trying to encode the effect of the node removal on the LMCC. Mathematically, the EMD of a node in a network is determined by considering the node's connections across different layers of the network and the degree of these connections. Specifically, the EMD of a node takes into account not just the number of connections it has in each layer, but also the importance of these connections based on the interconnectedness and the degree of the nodes at the other ends of these connections.

B. Duplex Collective Influence (DCI). DCI [9] refines and adapts to duplex networks the Collective Influence measure [11] by effectively considering the potential of edge overlap and interlayer degree correlation in triggering a cascade of node removals. DCI is defined as

$$DCI(i) = \frac{k_i^{(1)}k_i^{(2)} - k_i^{(int)}}{k_i^{(aggr)}} \left[\sum_j a_{ij}^{(1)} \left(k_j^{(2)} - 1 \right) + a_{ij}^{(2)} \left(k_j^{(1)} - 1 \right) \right] \quad [1]$$

where $k_i^{(1)}, k_i^{(2)}$ are the degrees of node i in layer 1 and 2, respectively, $k_i^{(int)}$ is its degree in the graph obtained by the intersection of the two layers and $k_i^{(aggr)}$ is its degree in the graph obtained by the union of the two layers. Finally, we denote as $A^{(1)} = (a_{ij}^{(1)})_{i,j}$, $A^{(2)} = (a_{ij}^{(2)})_{i,j}$ the adjacent matrices of the two layers.

The initial factor in the multiplication evaluates the node's score based on the combination of its degrees, and it increases as the degree in the intersection graph increases. On the other hand, the second factor accounts for potential cascades of node removals that can occur when node i is eliminated.

DCI has been shown to outperform EDM in case of non-negligible edge overlap probability, suggesting a new approach to better encode the multiplex structure.

C. Duplex Collective Influence with centrality measures. DCI has achieved remarkable performances in tackling optimal percolation, particularly in the situation of non-negligible edge-overlap probability. Our contribution consists in extending this targeted attack strategy by using different centrality measures aimed at capturing the relevance of a node not just in its immediate vicinity, but also into its farthest away neighborhood.

C.1. Duplex Collective Influence with Harmonic Centrality. As already mentioned in D, harmonic centrality inherently deals with unconnected networks and efficiently captures how influential nodes are in sparse networks. Moreover, it's computational efficiency makes it practical for analyzing large-scale networks, particularly with respect to other centrality measures, like betweenness centrality, that require extensive path-finding across the entire network. In summary, harmonic centrality with its ability to account for indirect influences, handle disconnected components, and its computational efficiency emerges as a particularly suitable choice for addressing optimal percolation problems. We modified DCI as follows:

$$DCI(i) = \frac{(C_h(i)^{(1)} + 1)(C_h(i)^{(2)} + 1) - (C_h(i)^{(int)} + 1)}{C_h(i)^{(aggr)} + 1} \times$$
$$\times \left[\sum_j a_{ij}^{(1)} C_h(j)^{(2)} + a_{ij}^{(2)} C_h(j)^{(1)} \right]$$

where a '+1' has been added to the harmonic centrality of each node to prevent the numerator of the first factor to become negative, due to multiplications of numbers smaller than 1. In this way, we recover the monotonicity of the first factor with respect to the harmonic centrality of the intersection graph.

Potential Issues with Harmonic Centrality. Sensitivity to Network Size and Structure: harmonic centrality is calculated based on the inverse of the shortest path distances to all other nodes. In large networks, or in networks with a highly modular structure, this can lead to underestimate the importance of locally influential nodes. The reason being that their influence is diluted by the presence of many distant nodes to which they are weakly connected.

C.2. Duplex Collective Influence with Katz Centrality. Katz-centrality C captures both the local and global influences of a node by considering all the weighted paths starting from the considered node. DCI based on Katz centrality is defined as follows:

$$DCI(i) = \frac{C_K(i)^{(1)}C_K(i)^{(2)} - C_K(i)^{(int)}}{C_K(i)^{(aggr)}} \times$$
$$\times \left[\sum_j a_{ij}^{(1)} \left(C_K(j)^{(2)} - 1 \right) + a_{ij}^{(2)} \left(C_K(j)^{(1)} - 1 \right) \right]$$

It is important to note that in the practical implementation using the NetworkX module's katz_centrality, the parameter *normalized* should be set to *False*. This is because normalizing Katz centrality scores confines them to a range between 0 and 1. Therefore, when combining the scores of the two layers via multiplication, the result is inherently smaller than each individual factor. This approach contradicts our objective since we want the multiplication of the Katz centralities of the two layers to act as an upper bound for the Katz centrality of the entire multiplex network. Additionally, in this scenario, $C_K(i)^{(int)}$ is not guaranteed to be less than the product, which could ultimately result in a negative score.

Potential Issues with Katz-Centrality. Parameter Sensitivity: the effectiveness of Katz centrality depends on the choice of the attenuation factor used in its calculation. An inappropriate choice of this parameter can lead to misleading results, either overemphasizing distant connections or failing to capture the influence exerted through these connections. As we tried to incorporate into the DCI information about the influence of a node in far away regions of the network, we choose a alpha that is 'as large as possible':

$$\alpha = min\{\lambda_1^{-1}, \lambda_1^{-1}, \lambda_{aggr}^{-1}, \lambda_{int}^{-1}\} - 0.01$$

where the subtraction helps to avoid issues with approximations and increases stability.

Bias Towards High-Degree Nodes: Katz centrality assigns higher scores to nodes with many direct connections and to those connected to these high-degree nodes. This can create a bias towards nodes that are structurally important but may not be the most critical in terms of network functionality or flow disruption.

C.3. Duplex Collective Influence with Betweenness Centrality. Betweenness centrality seems to be a suitable choice to address the optimal percolation issue. So that, in the context of single-layer networks, percolation centrality [20] has been developed by extending and subsuming the concept of betweenness centrality. It seems therefore natural to rely on this measure even in the context of duplex networks. The main problem is however related to the high computational cost associated with the procedure. The heavy amount of required computations and non promising results obtained in mock tests discouraged us to try this approach. Further explorations however may lead to insightful findings. We leave this interesting research direction for future work.

4. Numerical Analysis:

A. Synthetic Networks. We perform an analysis of the effectiveness of the proposed modified versions of the Duplex Collective Influence by comparing them with other three state-of-the-art procedures: HDA, EMD, DCI (and it's extension DCIz[9]). To do so, we study the optimal percolation problem on synthetic duplex networks, composed of $N = 10^4$ nodes and two Erdős-Rényi layers with mean degree $< k >= 5$, having different edge overlap and interlayer degree correlation. The choice is consistent with [9, 12] and the code is available at [21].

A.1. *Dependence on Edge Overlap.* The edge overlap o [13, 22] of edge (i, j) between the two layers of the duplex is defined as $o_{ij} = a_{ij}^{(1)} + a_{ij}^{(2)}$.

Following the approach of [9], we assign a summary value of edge centrality to the whole network and we normalize it via the Heaviside step function, $\Theta(\cdot)$:

$$o = 2\left(\frac{\sum_{i,j=1}^{N} o_{ij}}{2\sum_{i,j=1}^{N}\Theta(o_{ij})} - \frac{1}{2}\right) \in [0, 1]$$

A.2. *Dependence on Interlayer Degree Correlation.* Interlayer degree correlation (Spearman's rank correlation [23]) ρ has been shown to be a significant factor in studying the robustness of multiplex networks to targeted attacks, as it can heavily influence the position of the random percolation threshold [12, 24, 25].

To isolate the effect of interlayer degree correlation, we consider duplex networks with fixed edge overlap probability. For each of them, we then tune ρ, covering the entire range of values between the maximally disassortative case ($\rho = -1.0$) and the maximally assortative one ($\rho = 1.0$) [13].

A.3. *Simulations.* We start by reproducing the numerical analyses in [9]. More precisely, we created a graph for each combination of three distinct edge overlap probabilities, $o \in \{0.0, 0.2, 0.4\}$ and 11 different interlayer degree correlations, $\rho \in \{-1, -0.9, \ldots, 1\}$ with a step of 0.1. This allows us to visualize effectively the dependence of the proposed strategies on both o and ρ. The results in Fig.3 are averaged over 10 different iterations to reduce variance and increase stability. As evidenced by Fig.3, the size of the critical sets identified by the considered procedures depends significantly on both the interlayer degree correlations and the edge overlap probability.

Analyses suggest that all methods are affected by variations in the edge overlap probability of the duplex, especially for negative values of ρ. The most affected procedure seems to be EDM, whose effectiveness decreases heavily as o increases. In particular, it can be observed that EDM is the best performing strategy in case of no edge overlap and the worst performing method in case of $o = 0.4$. DCI, instead, finds almost consistently the smallest critical set in case of significant edge overlap. On the other hand, the values of the size of the critical set returned by the proposed variant of the Duplex Collective Influence based on harmonic centrality, namely DCI_harmonic, appears as the most stable. Indeed, even if it has the poorest performance in the case of 0.0 edge overlap probability, its outputs remain almost constant for $o = 0.2$ and $o = 0.4$, leading to a constant improvement in its behavior compared to the other strategies. This suggests that the proposed targeted attack strategies may be robust with respect to changes in the edge overlap of the networks.

A deeper intuition of the role of edge overlap probability can be grasped looking at Fig.2. The analysed networks have been constructed starting from two identical Erdős-Rényi layers ($o = 1$ and $\rho = 1$) with $N = 10^4$ nodes and mean degree $< k >= 5$. The edges of one of the two layers have then been iteratively rewired, while maintaining untouched the degree sequence of each layer, to reduce the edge overlap until we get to $o = 0$ [9, 26]. The plot confirms our speculations: EDM is the most affected strategy, followed by DCI, HDA and DCI_Katz, whilst DCI_harmonic, despite not performing optimally in networks with negligible edge probability, is the most robust of all methods to changes in o and have the best performances for high edge overlap probabilities. Specifically, the relative size of the critical set found by DCI_harmonic for $0 = 0.9$ is just of the 9.5% greater than the one obtained for $o = 0.1$, whereas the corresponding increase in the relative size of the critical set returned by EMD is of the 21.9%.

Focusing on the dependence on the interlayer degree correlation, we can observe that for $o = \{0.2, 0.4\}$ all methods tend to find smaller critical sets as ρ increases. This suggests that positive interlayer degree correlations favors targeted attack procedures. Moreover, this pattern appears to be more pronounced as the edge overlap becomes non-negligible.

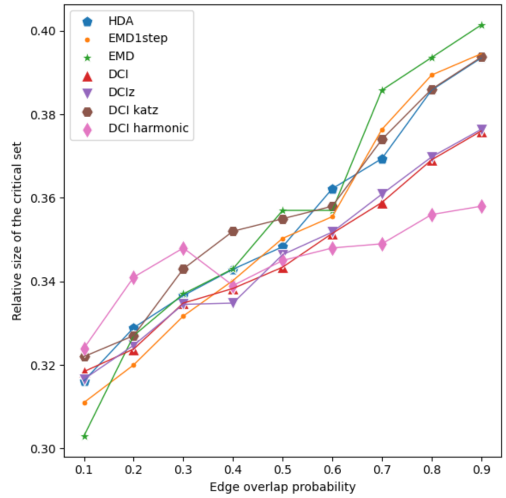


Fig. 2. Size of the critical set q as a function of the edge overlap probability o for duplex networks with $N = 10^4$ nodes, whose layers have been constructed from two identical Erdős-Rényi layers with $< k >= 5$. The edge overlap probability has been gradually reduced by rewiring the edges of one of the two layers, while keeping the other fixed. Results have been averaged over 10 iterations to reduce variance and increase stability.

A.4. *Considerations on DCI_Katz and DCI_harmonic.* Numerical analyses on synthetic networks do not highlight significant improvements related to the newly introduced strategies. EDM, in multiplex networks with no edge overlap, and DCI, in configurations with non-negligible edge overlap probability, are still preferable due to both their performances and small computational cost. This can be seen as an indirect confirmation of the ability of these methods to capture the overall structure of the network and not just the local effects that node removal can have.

However, as highlighted in A.3, the DCI_harmonic targeted attack procedure seems to be almost independent of edge overlap. This could suggest that the method is highly robust, versatile, and capable of generalizing well regardless of this factor. As a consequence, this approach may be applied to a wider range of real-world networks in which the precise overlap is uncertain or varies. Evidence of this speculation will be given in Section B.

Katz-centrality, on the other hand, does not give rise to any new significant enhancement, dissuading its further employment.

To conclude, we also mention that other targeted attack strategies have been tested, however, with discouraging results. Namely, we tried to average the centrality measure of nodes over different layers, simply compute their product and to substitute the Katz centrality with the betweenness centrality.

B. Real World Datasets. We conclude our numerical analysis by providing the results of simulations over real-world multiplex networks. Datasets are available here. Specifically, since several of the multiplex comprehend more than two layers, we select the duplex sub-networks corresponding to the pairs of layers yielding the largest LMCC. This approach is standard and has already been applied by [9, 25].

The study of targeted attack strategies in network systems is critical in a variety of real-world contexts, playing a crucial role in improving the resilience and security of infrastructures. In the realm of cybersecurity, it helps to fortify networks against cyber threats, ensuring the protection of vital information and systems. In public health, it gives insight into understanding the dynamics of disease spread through social networks, thereby aiding in the development of effective containment strategies. In financial networks, these studies are instrumental in safeguarding against systemic risks and potential economic crises. Finally, in ecological and environmental management, understanding targeted

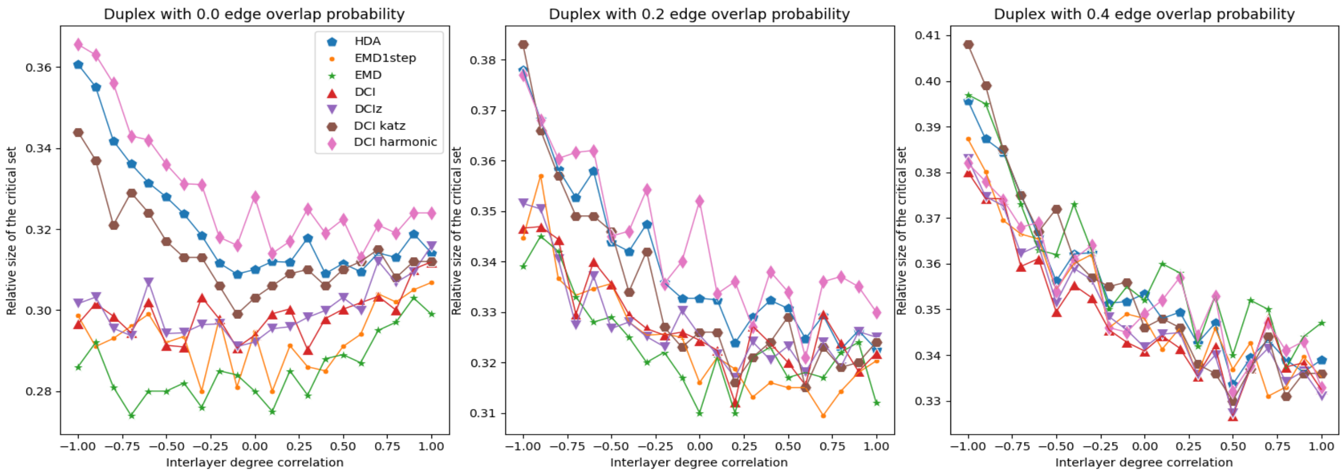


Fig. 3. Size of the critical set q as a function of the interlayer degree correlations coefficient ρ [13] for duplex networks with $N = 10^4$ nodes, whose layers are Erdős-Rényi graphs with $< k >= 5$. We report three different overlap conditions: no edge overlap, low edge overlap, and and moderate edge overlap. The concurrent presence of interlayer degree correlations and edge overlap strongly affects the robustness of a system against targeted attacks. The results are averaged over 10 realizations.

Simulations on Real Duplex Networks											
Dataset Name	N	LMCC	ρ	o_{norm}	HDA	EDM 1step	EDM	DCI	DCIz	DCI Katz	DCI harmonic
CS-Aarhus_Duplex_L_1_5	62	58	0.32	0.63	14	15	15	15	**<u>12</u>**	15	**<u>12</u>**
Arabidopsis_Duplex_L_1_2	6981	442	0.65	0.40	42	**<u>36</u>**	38	48	48	57	38
arXiv_Duplex_L_2_6	14490	916	0.85	0.77	89	106	117	90	89	91	**<u>78</u>**
arXiv_Duplex_L_3_6	14490	790	0.92	0.81	80	83	82	98	93	92	**<u>65</u>**
Homo_Duplex_L_1_2	18223	9312	0.57	0.32	1110	**<u>1082</u>**	1098	1123	1141	1109	1087
Homo_Duplex_L_1_5	18223	3886	0.31	0.16	312	319	**<u>306</u>**	334	324	344	327
Homo_Duplex_L_2_5	18223	4944	0.44	0.17	428	394	**<u>372</u>**	409	425	420	404
SacchCere_Duplex_L_1_2	6571	4531	0.36	0.10	816	809	**<u>765</u>**	815	808	850	868
SacchCere_Duplex_L_1_7	6571	4720	0.28	0.07	1016	1011	964	**<u>961</u>**	980	1079	1068
SacchPomb_Duplex_L_3_4	4093	1112	0.20	0.14	60	54	**<u>45</u>**	64	63	58	**<u>45</u>**
SacchPomb_Duplex_L_3_6	4093	956	0.14	0.06	43	36	32	38	38	40	**<u>26</u>**
SacchPomb_Duplex_L_4_6	4093	2292	0.61	0.01	371	368	368	370	**<u>357</u>**	378	375
HUM_HIV_Duplex_L_1_2	1006	144	0.54	0.41	3	3	3	**<u>2</u>**	4	6	3
FAO_Duplex_L_3_24	215	193	0.94	0.70	87	86	89	89	87	89	**<u>85</u>**
Mus_Duplex_L_1_3	7748	1059	0.56	0.37	61	**<u>59</u>**	61	67	66	57	62
Drosophila_Duplex_L_1_2	8216	299	0.18	0.07	13	9	**<u>7</u>**	11	11	11	11
Drosophila_Duplex_L_1_3	8216	202	0.27	0.06	**<u>3</u>**	**<u>3</u>**	4	**<u>3</u>**	**<u>3</u>**	4	**<u>3</u>**
Drosophila_Duplex_L_1_4	8216	1024	0.13	0.09	29	**<u>26</u>**	29	30	33	28	**<u>26</u>**
Drosophila_Duplex_L_2_3	8216	449	0.68	0.35	46	**<u>44</u>**	48	50	50	49	**<u>44</u>**

Table 1. Size of the critical set for the seven targeted attack strategies considered applied to real-world duplex networks. For each network, useful statistics are reported: the number of nodes N, the interlayer degree correlation within the LMCC ρ, the normalized edge overlap probability [$o_{norm} = o/o_{max}$ where o_{max} is the maximum overlap probability in the set of corresponding configuration models [26, 27]]. The size of the minimal critical set is underlined and highlighted in bold. In general, DCI_harmonic appears to be quite effective on this real-world datasets. Finally, we mention that for the computation of the Katz centrality a parameter $\alpha = 0.01$ has been used, with few exception that required a more attentive tuning and for which we end up using $\alpha = 0.005$.

disruptions can lead to more effective conservation strategies. In all these situations, the network structure is often complex with non-independent layers and non-negligible edge overlap.

We evaluated the performances of the considered methods over 19 different duplex networks constructed from real-world data. The sizes of the critical set returned by the different targeted attack strategies and useful information for each considered dataset are reported in Table1. The examined networks have sizes that vary from a few to thousands of nodes, different edge overlap probabilities, and interlayer degree correlations. As expected, there is not a unique strategy that proves to be the best for all datesets. However, the robustness of the DCI_harmonic method underlined in A.4 is confirmed here. DCI_harmonic proves to be the best targeted attack procedure for 9 datasets and to be among the two best strategies in 12 duplex networks. This confirms that it is promising for dealing with real-world scenarios. Nevertheless, computational cost may be a serious drawback for this method in considering very large sets of nodes. Further improvement in its computation may be necessary before proposing it as a general solution.

Finally, it is worth noting a consideration on the role of the parameter α of the Katz centrality. Tuning the parameter in case of real-world dataset presents serious challenges and could be heavily expensive, especially for large networks. On the other hand, it is undeniable that the parameter plays a crucial role in enhancing the effectiveness of the procedure. A possible solution that would be interesting to address is to incorporate in the strategy a "depht-complexity trade-off", i.e. a balance between the depth of analysis in terms of gathering extensive network information (Depth) and the need to manage computational complexity (Complexity). For example, one way could be to incorporate information related to only a circumscribed neighborhood of each node.

5. Conclusions

This study extensively analyzed targeted attack strategies in duplex networks using variations of the Duplex Collective Influence (DCI) algorithm. Significant findings emerged regarding the comparison of these strategies with state-of-the-art targeted attack methods under different conditions of edge overlap probability and interlayer degree correlations. Existing procedures appears to take into account a significant portion of the network structure, so to not be overshadowed by the newly introduced methods. However, extensive simulations highlighted the significant dependence of the performance of this methods on the specific structure of the duplex, aspect that may not be desirable when targeting a real-world scenario. Conversely, DCI_harmonic stands out for displaying remarkable stability across different network conditions, suggesting its potential applicability in various real-world networks.

The study opens avenues for future research. Possible directions may involve analyzing the impact of centrality measures in other network structures, such as duplex constructed using scale free layers. Additionally, it would be interesting and meaningful to study not only the size of the critical set returned by targeted attack strategies, but also their composition. One could address the question of whether the critical sets identified through different strategies are consistent within each other or present several variations. Critically, indeed, in almost all real-world scenarios in which we are interested in identifying the smallest set of nodes whose removal would maximally disrupt the connectivity of the network, knowledge of the specific nodes that belong to it is crucial. Consider, for example, fields ranging from epidemiology and infrastructure resilience to cybersecurity and financial stability, where targeted interventions can significantly improve effectiveness and preparedness. In fact, by focusing on these key nodes, resources can be used more efficiently, risks can be managed more effectively, and overall system resilience will be significantly improved.

1. P Clusella, P Grassberger, FJ Pérez-Reche, A Politi, Immunization and targeted destruction of networks using explosive percolation. *Phys. Rev. Lett.* **117**, 208301 (2016).
2. Y Liu, Y Deng, B Wei, Local immunization strategy based on the scores of nodes. *Chaos: An Interdiscip. J. Nonlinear Sci.* **26**, 013106 (2016).
3. G Strona, C Castellano, Rapid decay in the relative efficiency of quarantine to halt epidemics in networks. *PHYSICAL REVIEW E* **97** (2018).
4. J Raimbault, Multi-dimensional urban network percolation (2019).
5. A Zhaikhan, MA Kishk, H ElSawy, MS Alouini, Safeguarding the iot from malware epidemics: A percolation theory approach. *IEEE Internet Things J.* **8**, 6039–6052 (2021).
6. JH Cho, T Moore, *Software diversity for cyber resilience: percolation theoretic approach.* pp. 313–342 (2019).
7. S Pei, J Wang, F Morone, HA Makse, Influencer identification in dynamical complex systems (2019).
8. S Osat, A Faqeeh, F Radicchi, Optimal percolation on multiplex networks. *Nat. Commun.* **8** (2017).
9. A Santoro, V Nicosia, Optimal percolation in correlated multilayer networks with overlap. *Phys. Rev. Res.* **2**, 033122 (2020).
10. M Kivela, et al., Multilayer networks. *J. Complex Networks* **2**, 203–271 (2014).
11. F Morone, HA Makse, Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68 (2015).
12. GJ Baxter, G Timár, JFF Mendes, Targeted damage to interdependent networks. *Phys. Rev. E* **98** (2018).
13. V Nicosia, V Latora, Measuring and modeling correlations in multiplex networks. *Phys. Rev. E* **92** (2015).
14. G Bianconi, F Radicchi, Percolation in real multiplex networks. *Phys. Rev. E* **94** (2016).
15. B Min, SD Yi, KM Lee, KI Goh, Network robustness of multiplex networks with interlayer degree correlations. *Phys. Rev. E* **89** (2014).
16. A Saxena, S Iyengar, Centrality measures in complex networks: A survey (2020).
17. LC Freeman, A set of measures of centrality based on betweenness. *Sociometry* **40**, 35–41 (1977).
18. L Katz, A new status index derived from sociometric analysis. *Psychometrika* **18**, 39–43 (1953).
19. MA Beauchamp, An improved index of centrality. *Behav. science* **10**, 161–3 (1965).
20. M Piraveenan, M Prokopenko, L Hossain, Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks. *PloS one* **8** (2013).
21. L Pezzetti, Complex networks/project (https://github.com/lucia-pezzetti/Complex_Networks/tree/main/Project) (2024).
22. F Battiston, V Nicosia, V Latora, Structural measures for multiplex networks. *Phys. Rev. E* **89** (2014).
23. C Spearman, The proof and measurement of association between two things. *The Am. J. Psychol.* **100**, 441–471 (1987).
24. SV Buldyrev, NW Shere, GA Cwilich, Interdependent networks with identical degrees of mutually dependent nodes. *Phys. Rev. E* **83**, 016112 (2011).
25. KK Kleineberg, L Buzna, F Papadopoulos, M Boguñá, MA Serrano, Geometric correlations mitigate the extreme vulnerability of multiplex networks against targeted attacks. *Phys. Rev. Lett.* **118**, 218301 (2017).
26. M Diakonova, V Nicosia, V Latora, MS Miguel, Irreducibility of multilayer network dynamics: the case of the voter model. *New J. Phys.* **18**, 023010 (2016).
27. A Santoro, V Nicosia, Algorithmic complexity of multiplex networks. *Phys. Rev. X* **10**, 021069 (2020).