Information transfer in community structured multiplex networks

Albert Solé-Ribalta, Clara Granell, Sergio Gómez and Alex Arenas
Departament d’Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona (Spain)

Abstract. The study of complex networks that account for different types of interactions has become a subject of interest in the last few years, specially because its representational power in the description of users interactions in diverse online social platforms (Facebook, Twitter, Instagram, etc.). The mathematical description of these interacting networks has been coined under the name of multilayer networks, where each layer accounts for a type of interaction. It has been shown that diffusive processes on top of these networks present a phenomenology that cannot be explained by the naive superposition of single layer diffusive phenomena but require the whole structure of interconnected layers. Nevertheless, the description of diffusive phenomena on multilayer networks has obviated the fact that social networks have strong mesoscopic structure represented by different communities of individuals driven by common interests, or any other social aspect. In this work, we study the transfer of information in multilayer networks with community structure. The final goal is to understand and quantify, if the existence of well-defined community structure at the level of individual layers, together with the multilayer structure of the whole network, enhances or deteriorates the diffusion of packets of information.
1. Introduction

The study of transport properties of networks is becoming increasingly important due to the constantly growing amount of information and commodities being transferred through them. A particular focus of these studies is how to make the capacity of the diffusion of information in the network maximal while minimizing the delivery time. In the basic approach information is formed by units, the “packets”, and the handling of information for processing and distribution takes finite time. Both network packet routing strategies and network topology play an essential role in networks’ traffic flow.

In realistic settings, like online social networks, the knowledge that any one has about the topology of the network is limited to its local area of influence. Consequently, much of the focus in recent studies has been on “searchability”, the process of sending information to a target when the trajectory to reach the target is unknown. Moreover, given the limited capability of nodes to handle information packets and redistribute them, the problem of congestion arises [1, 2, 3]. It has been observed, both in real world networks and in model communication networks, that the network flow collapses when the load (number of packets to be processed) is above a certain threshold [3].

In general, most real and engineered systems include multiple subsystems and layers of connectivity, and it is important to take such features into account when trying to obtain a complete understanding of them. It is thus necessary to generalize the “traditional” network theory to multilayer systems in a comprehensive fashion [4, 5]. Generally speaking, up to now, the description of networks so far has been developed using a single and combined snapshot of the connectivity, which is a reflection of instantaneous interactions or accumulated interactions in a certain time window. This description is limited when trying to understand the intricate variability of real complex systems, which contain many different time scales [6] and coexisting structural patterns forming the real network of interactions [7]. This is the case of e-social networks that are constantly changing [8], having some connections with very short lifetime and others that are persistent. Interest groups [9] are constantly being developed and growing, and individual nodes participate through different interests at the same time. An accurate description of such complexity should take into account these differences of interactions and their evolution through time. In the last couple of years, the scientific community on networks theory has focused on this issue and proposed a solution that has been commonly referred to as the multiplex network structure [4].

General flows on multiplex networks have also been in the focus of network scientists [10, 11, 12, 13, 14, 15, 16, 17], and the consequences of having such topologies have been shown to be far from trivial. For example, in [18] the authors found that a general diffusive process on top of the multiplex structure is able to speed up the less diffusive of the layers. It could also give rise to a super-diffusion process thus enhancing the diffusion of both layers. This striking result appears when the diffusion between the layers of the multiplex is faster than that occurring within each of the layers. These consequences are also observed in the discrete representation of diffusion by random walkers [17], and
have explicit consequences on the navigability of the multiplex structure.

Here we fix our attention in the process of information transfer on top of multiplex networks. Specifically, we aim at determining the structural effects of a multiplex network endowing community structure, i.e. modular at each layer, on the dynamics of information transfer. To this end we have investigated a particular set up in which multiplex networks are built connecting different modular networks, and determining analytically the onset of congestion in the information flow. Our results reveal that when the community structure of the different layers is equivalent and the communities overlap, the multiplex offers higher resilience to congestion and consequently the system may improve information transfer compared to the individual layers. On the other side when the community structure is considerably different and communities still overlap the multiplex structure offers a balancing environment where the efficiency of the system is averaged. On the intermediate situation, that is community structure is similar in both layers and communities overlap, the effect of the multiplex structure is devastating and hinders information transfer by reducing the onset of congestion in the system.

2. Material & Methods

The proposed dynamical model considers that information flows through networks in atomic and discrete packets that are sent from an origin node to a destination node. Each node is an independent agent that can store as many packets as necessary. However, to have a realistic picture of communication we must assume that the nodes have a finite capacity to process and deliver packets. That is, a node will take longer to deliver two packets than just one. This physical constraint of the agents on delivering information can derive in network congestion. When the amount of information a particular agent receives is too large, it is not able to handle all the packets and some of them remain undelivered for extremely long periods of time. In this study, the interest is focused on when congestion occurs depending on the topology of the multiplex network, in particular, in relation to its community structure.

2.1. Dynamics of information transfer

The dynamics of the model is as follows. At each time step \( t \), information packets are created at every node with rate \( \rho \) (injection rate). Therefore, \( \rho \) is the control parameter: small values of \( \rho \) correspond to low density of packets and high values of \( \rho \) correspond to the generation of a large amount of packets. When a new packet is created, a destination node, different from the origin, is chosen (uniformly) at randomly in the network. Thus, during the following time steps \( t + 1, t + 2, \ldots, t + T \), the packet travels toward its destination. Once the packet reaches the destination node, it is delivered and disappears from the network.

The time that a packet remains in the network is related not only to the length of the path between the source and the target nodes, but also to the volume of packets.
that share its path. Nodes with high loads, i.e. high volume of accumulated packets, will take longer to deliver packets or, in other words, it will take more time steps for packets to cross regions of the network that are highly congested. We assume, without loss of generality, that nodes can handle only one packet per time step (i.e. the delivery rate is $\tau = 1$), and undelivered packets are stored in a first-in-first-out queue attached to each node. The paths followed by packets between source and destination nodes are decided using a routing strategy, being shortest paths and random walks the most prominent strategies. It is important to note, however, that the model is not deterministic. For example, there may be several shortest paths between two nodes, one of them chosen randomly in the delivery of the corresponding packet. Moreover, the order in which packets are stored in the queues when several of them arrive in the same time step is undefined.

Previous work on single layer networks [3] shows that for low values of the injection rate of packets $\rho$ there is no accumulation of packets at any node in the network. Moreover, it is stated that the number of packets that arrive at node $i$ is, on average, $\rho B_i/(N-1)$, where $B_i$ is the effective betweenness of node $i$ and $N$ the number of nodes in the network. The effective betweenness is defined as the ratio between the number of paths that pass through node $i$, and the total number of paths traversing the network between any pair of nodes [19].

The onset of congestion is reached when a node receives more packets than it can deliver per time step, i.e. $\rho B_i/(N-1) > 1$. Therefore, the first node that collapses ($i^*$) is the one with largest effective betweenness ($B_{i^*} = \max_i(B_i)$), and the maximum injection rate for which the network is congestion free, the critical injection rate $\rho_c$, is given by

$$\rho_c = \frac{N - 1}{B_{i^*}}.$$  

The rest of the nodes will collapse with larger injection rates. However, up to now, it is not known how to analytically compute their critical injection rates since they not only depend on the topological betweenness but also on the overall network congestion.

In the generalization of the routing dynamics exposed to multiplex networks, the average number of packets arriving to node $i$ in layer $\alpha$ is $\rho L B_{i\alpha}/(N-1)$, where $L$ is the number of layers of the multiplex network and $B_{i\alpha}$ is the effective betweenness of node $i$ in layer $\alpha$. Thus, the critical injection rate also depends on the effective betweenness, which encapsulates the routing strategy and the topology of the network:

$$\rho_c = \frac{(N - 1)/L}{\max_{i,\alpha}(B_{i\alpha})}.$$  

Next, we extend the concept of betweenness to multiplex networks allowing the computation of the onset of congestion.
2.2. Computation of betweenness in the multiplex

The extension of any centrality measures to multiplex networks is not straightforward and requires special care. In many situations several extensions are possible and the choice of it may depend on the problem at hand. Many attempts have been done to extend single layer centrality measures to the multiplex framework [20, 21, 22]. Here, we follow the line described in [23], which is mathematically grounded on the tensorial formalism for multilayer networks [24].

We start by defining a walk between two individuals $s$ and $t$ on a multiplex network as a sequence of nodes, following intralayer and/or interlayer edges, which starts at node $s$ “in any layer” and finishes in node $t$ “in any layer”. Note that in this definition we do not care about the layer, just the node. The reasoning behind this lack of discrimination is that, in the multiplex structure, the different node replicas in the different layers correspond to the same individual (social networks) or location (transportation networks), thus it is only important to know if the packet has arrived, but not in which layer. Fig. 1 shows an example of a walk between two nodes in a multilayer network where non-trivial effects can be observed because of the presence of interlayer connections that affect the navigation through the networked system [17].

Given the definition of a walk in the multiplex topology, the effective betweenness of a node $i$ in layer $\alpha$, $B_{i\alpha}$, can be directly obtained as the fraction of walks that pass through node $i$ in layer $\alpha$ for every possible origin-destination pair $(s,t)$. In some cases it might be convenient to obtain the betweenness of node $i$ irrespective of the layer. In this case, the betweenness can be obtained just by accumulation of the individual
contributions of each layer where \( i \) is represented, \( B_i = \sum_\alpha B_{i\alpha} \).

For the specific case of the \textit{shortest path betweenness}, every walk is restricted to be the path with minimum distance that starts from the source node \( s \) in any layer, and reaches the destination node \( t \) in any layer. The distance function may take into account the weights of the edges the path traverses. In this work, without loss of generality, we assume the edges’ weights are unitary and define the distance as the number of traversed edges in the path. The shortest path between two locations may be degenerated and consequently the set of shortest paths may contain paths using a single layer (classical shortest paths) and paths which change layer (pure multiplex paths). For an accurate computation of the shortest path betweenness special care must be taken with the path degeneration. A good and efficient algorithm can be found in [25].

Equivalently to the shortest path betweenness, the \textit{random walk betweenness} depends on the particular definition of the network traversal procedure. In this case, a random walk is defined as a walk in which, at each time step, the next visited node is chosen with uniform random probability among the neighbors of the last visited node. The random walk betweenness is usually computed considering a transition matrix obtained from the adjacency matrix of the network. For a detailed description of random walks in a multiplex network, see [17]. In this document we will use the classic random walk definition. For the random walk betweenness the walk degeneracy is enormous and consequently is impractical to compute the betweenness accounting for all the possible individual random walks. Fortunately, the random walk betweenness can be efficiently computed using matrix inversion and absorbing random walks [26].

2.3. Community structure in multiplex

Networks representative of complex systems are characterized by having community structure, meaning the presence of dense groups of nodes with sparse connections between them [27]. It is known that dynamical processes running on top of such networks have a big dependency on community structure, which affects the process either by fostering or hampering it [28, 29, 30]. As evidenced in several works [31, 7], when the different layers of a multiplex network exhibit community structure, the influence on the overall system is not trivial to determine.

Here, to uncover the basic effects of communities in information flow process, we propose a simple setting with imposed community structure where communities fully overlap between layers and the degree of each node of the network is kept constant. Each multiplex network consists of two layers, and each layer has 256 nodes distributed in four communities (64 nodes per community) [32]. The links are generated in such a way that the density of links inside the communities is always higher than the density between them. The networks are generated independently for each layer, resulting in a two-layer multiplex network with different community structure in each layer.

For the experiments, we consider 12 different multiplex community structures and 300 different realizations for each. For all of them, we fix the bottom layer (L1) to
Information transfer in community structured multiplex networks

\[ \mu = 0.03 \]

\[ \mu = 0.23 \]

\[ \mu = 0.37 \]

Figure 2. Samples of the multiplex networks generated, represented by means of their superposed adjacency matrices. From left to right, the top layer diffuses its community structure (increasing mixing parameter), while the bottom layer remains fixed.

\[ k_{in} = 31 \text{ and } k_{out} = 1 \] (i.e., 31 edges inside the community and 1 link outside, per node), which displays strong and clear communities, and we vary the community structure of the top layer (L2), which ranges from the previous strong block structure to a more diluted one \((k_{in} = 20 \text{ and } k_{out} = 12)\) where the communities are almost imperceptible. We quantify the strength of the community structure of the L2 layer using a mixing parameter defined as \( \mu = k_{out}/\langle k \rangle \). Figure 2 depicts three examples of such generated networks.

3. Results

To evaluate the influence of the multiplex networks with community structure in information transfer we assess several aspects of the information transfer dynamics, namely the shortest paths distribution, the packets ingoing rate of each node and the critical injection rate of the network.

Figure 3 shows the obtained distribution of shortest paths in the different layers of the multiplex. In the case of having equivalent community structure in both layers (leftmost points in the plot), the multiplex structure provides a very good load balance where the same fraction of paths traverse using layer 1 and 2. In general, we can conclude that the effect of the multiplex is negligible for the overall system behavior since only a very small fraction of paths \((0.5\%)\) makes use of the full multiplex structure. In fact, paths using both layers of the multiplex are only used in the case where the origin and destination are in different communities. As we increase the mixing parameter of the second layer, its community structure dilutes, enhancing the communication between communities but slightly hindering the transfer of information internally. This effect is evident in Fig. 3, which shows a large increase of intercommunity trips in the second layer and a small increase of intracommunity paths in the first layer. At the same time, the improvement of intercommunity paths in the more diffuse layer yields a disappearance of the (already small number of) shortest paths using both layers.

To assess the microscopic behavior of the system we show how the ingoing rate of
packets to each node varies with respect to the mixing parameter. We compute the ingoing rate of each node of the multiplex structure as

$$\hat{\sigma}_{i\alpha} = \frac{L B_{i\alpha}}{N - 1}.$$  (3)

Results are shown in Fig. 4. For the shortest path routing strategy (subplot A) we observe a clear distinction between the behavior of nodes in layers 1 and 2. As can be seen in Fig. 3, the main effect on the increasing of the mixing parameter is clearly a migration of the shortest paths from layer 1 to layer 2, i.e. paths that traversed layer 1 now find a more efficient route through layer 2, which has a more diluted community structure. This migration of paths should increase the ingoing rate of nodes in layer 2 similarly to the observed decrease of ingoing packets of layer 1. This is the situation for small mixing parameters, but increasing the mixing parameter means also an increase in the efficiency of layer 2 routing packets between nodes in different communities, which in turn substantially reduces the overall node betweenness provoking an interesting tradeoff that will prescribe the final efficiency of the full system. These two opposed effects (migration of shortest paths and reduction of node betweenness) have a huge impact on the ingoing rate of nodes in layer 2, which experience a constant decrease after the maximum ingoing rate is reached. For the random routing strategy the scenario is completely different. The increase of the mixing parameter has an equivalent impact in both layers, which experience an important decrease of the ingoing rate.

Figures 5 and 6 show the effect of the community structure on the critical injection rate $\rho_c$. For the shortest path routing strategy (Fig 5) the critical injection rate of the
A multiplex network reaches its minimum value around $\mu = 0.1$. This minimum indicates that there is a worst-case scenario for which the multiplex topology is less efficient than the individual layers. On the other side, the behavior of the critical injection rate of layer 2 is monotonically increasing. This situation is expected since a less clear community structure leads to a reduction of the average shortest path, which in turn is positively correlated with a decrease of the node betweenness.

In general, if we compare the values of $\rho_c$ for the multiplex network and the separate layers L1 and L2, we clearly observe three possible situations: (i) the multiplex is more resilient to congestion (efficient) than the individual layers. This situation arises when both layers have a similar community structure. (ii) The multiplex is less efficient than any of the layers. This setup corresponds to the minimum resilience of the multiplex structure. And (iii), the multiplex efficiency achieves a medium value which is a trade-off between the resilience of both layers. In a real setup, this situation would mean that joining those two layers in a multiplex improves the resilience observed in one layer at the cost of deteriorating the resilience observed in the second layer. However, as we can observe in the plot, the reduction of the efficiency is larger than the average of the efficiency of both layers and consequently the coupling of layers is inefficient.

With respect to the random walk routing strategy (Fig 6), the situation is completely different. For similar community structures the multiplex worsens the efficiency, because the paths get trapped within the communities. For different community structures, in general, we obtain a efficiency that corresponds to the average efficiency of both layers.

Figure 4. Ingoing rate of each node in the network for different mixing parameter values, and two routing strategies: shortest paths (A) and random walks (B). The different colours indicate the layer to which the node belongs to.
4. Discussion

We have shown that packet information flow can be compromised when community structure is considered in some layer of the multiplex network structure. Since
Community structure implies the presence of topological bottlenecks, the information flows migrate to those layers where these constraints are relaxed (diluted communities). We have shown that community structures produce a non-trivial effect in the transfer of information and in the resilience to information flow congestion, that here defines the efficiency of the structure. Essentially, the better defined the communities, the more affected the packet transportation. Information tries to avoid bottlenecks and packets migrate towards the layer where the community structure is diluted, because it is more efficient, but as a direct consequence of this migration the most efficient layer becomes overloaded. This nonlinear relation makes the problem of assessing the performance of the multiplex structure particularly challenging.

Using the analytical approach presented, we are able to determine for any multiplex topology what is the onset of congestion in the information flow and how it compares with the onset of the individual layers. We have also provided results, for very specific scenarios, of shortest path and random walk routing strategies respectively. The results show that the shortest path approach heavily depend on the particular sharpness of the prescribed communities. This work provides the starting point for the discrete flow analysis of more complicated scenarios of community structure in multiplex networks.

Acknowledgments

This work has been supported by Ministerio de Economía y Competitividad (Grant FIS2012-38266) and European Comission FET-Proactive Projects PLEXMATH (Grant 317614). A.A. also acknowledges partial financial support from the ICREA Academia and the James S. McDonnell Foundation.

References

[1] Roger Guimerà, Alex Arenas, and Albert Díaz-Guilera. Communication and optimal hierarchical networks. *Physica A: Statistical Mechanics and its Applications*, 299(1):247–252, 2001.

[2] Alex Arenas, Leon Danon, Albert Díaz-Guilera, and Roger Guimerà. Local search with congestion in complex communication networks. In Computational Science-ICCS 2004, pages 1078–1085. Springer, 2004.

[3] Roger Guimerà, Albert Díaz-Guilera, Fernando Vega-Redondo, Antonio Cabrales, and Alex Arenas. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.*, 89(24):248701, 2002.

[4] Mikko Kivelä, Alexandre Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014.

[5] S. Boccaletti, G. Bianconi, R. Criado, C.I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1 – 122, 2014.

[6] Laetitia Gauvin, André Panisson, and Ciro Cattuto. Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach. *PloS one*, 9(1), 2014.

[7] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *science*, 328(5980):876–878, 2010.
Information transfer in community structured multiplex networks

[8] Antonia Godoy-Lorite, Roger Guimera, and Marta Sales-Pardo. Long-term evolution of technosocial networks: Statistical regularities, predictability and stability of social behaviors. arXiv preprint arXiv:1506.01516, 2015.

[9] Barry Wellman and Stephen D Berkowitz. Social structures: A network approach, volume 2. CUP Archive, 1988.

[10] Charles D Brummitt, Kyu-Min Lee, and K-I Goh. Multiplexity-facilitated cascades in networks. Physical Review E, 85(4):045102, 2012.

[11] Osman Yağan and Virgil Gligor. Analysis of complex contagions in random multiplex networks. Physical Review E, 86(3):036103, 2012.

[12] Renaud Lambiotte, Jean-Charles Delvenne, and Mauricio Barahona. Random walks, markov processes and the multiscale modular organization of complex networks. Network Science and Engineering, IEEE Transactions on, 1(2):76–90, 2014.

[13] Renaud Lambiotte, Vsevolod Salnikov, and Martin Rosvall. Effect of memory on the dynamics of random walks on networks. Journal of Complex Networks, 3(2):177–188, 2015.

[14] Franco Bagnoli and Emanuele Massaro. Risk perception and epidemic spreading in multiplex networks. In ISCS 2014: Interdisciplinary Symposium on Complex Systems, pages 319–332. Springer, 2015.

[15] Dawei Zhao, Lixiang Li, Haipeng Peng, Qun Luo, and Yixian Yang. Multiple routes transmitted epidemics on multiplex networks. Physics Letters A, 378(10):770–776, 2014.

[16] Sergio Gómez, Albert Díaz-Guilera, Jesús Gómez-Gardeñes, Conrad J. Pérez-Vicente, Yamir Moreno, and Alex Arenas. Diffusion dynamics on multiplex networks. Phys. Rev. Lett., 110:028701, Jan 2013.

[17] Manlio De Domenico, Albert Solé-Ribalta, Sergio Gómez, and Alex Arenas. Navigability of interconnected networks under random failures. Proceedings of the National Academy of Sciences USA, 111(23):8351–8356, 2014.

[18] Albert Solé-Ribalta, Manlio De Domenico, Nikos E Kouvaris, Albert Díaz-Guilera, Sergio Gómez, and Alex Arenas. Spectral properties of the laplacian of multiplex networks. Physical Review E, 88(3):032807, 2013.

[19] Mark Newman. Networks: An Introduction. Oxford University Press, Inc., New York, NY, USA, 2010.

[20] Arda Halu, Raúl J. Mondragón, Pietro Panzarasa, and Ginestra Bianconi. Multiplex PageRank. PLoS ONE, 8, 2013.

[21] Federico Battiston, Vincenzo Nicosia, and Vito Latora. Structural measures for multiplex networks. Physical Review E, 89(3):032804, 2014.

[22] Luis Sola, Miguel Romance, Regino Criado, Julio Flores, Alejandro Garcia del Amo, and Stefano Boccaletti. Eigenvector centrality of nodes in multiplex networks. Chaos, 3:033131, 2013.

[23] Manlio De Domenico, Albert Solé-Ribalta, Elisa Omodei, Sergio Gómez, and Alex Arenas. Ranking in interconnected multilayer networks reveals versatile nodes. Nature communications, 6, 2015.

[24] Manlio De Domenico, Albert Solé-Ribalta, Emanuele Cozzo, Mikko Kivelä, Yamir Moreno, Mason A. Porter, Sergio Gómez, and Alex Arenas. Mathematical formulation of multilayer networks. Phys. Rev. X, 3:041022, Dec 2013.

[25] Albert Solé-Ribalta, Manlio De Domenico, Sergio Gómez, and Alex Arenas. Centralities in multiplex networks. In ACM Web Science Conference, WebSci ’14, Bloomington, IN, USA, June 23-26, 2014, pages 149–155, 2014.

[26] Albert Solé-Ribalta, Manlio De Domenico, Sergio Gómez, and Alex Arenas. Random walk centrality in interconnected multilayer networks. arXiv:1506.07167, 2015.

[27] Santo Fortunato. Community detection in graphs. Physics Reports, 486:75 – 174, 2010.

[28] Zonghua Liu and Bambi Hu. Epidemic spreading in community networks. EPL (Europhysics Letters), 72(2):315, 2005.

[29] Alex Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. Synchronization reveals
topological scales in complex networks. *Phys. Rev. Lett.*, 96:114102, Mar 2006.

[30] Alex Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. Synchronization processes in complex networks. *Physica D: Nonlinear Phenomena*, 224:27 – 34, 2006. Dynamics on Complex Networks and Applications.

[31] Laura Bennett, Aristotelis Kittas, Gareth Muirhead, Lazaros G Papageorgiou, and Sophia Tsoka. Detection of composite communities in multiplex biological networks. *Scientific reports*, 5, 2015.

[32] Mark E. J. Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, Feb 2004.