Multiplex Structures in the World Trade Web

Daryl DeFord
Advisor: Dan Rockmore
Department of Mathematics

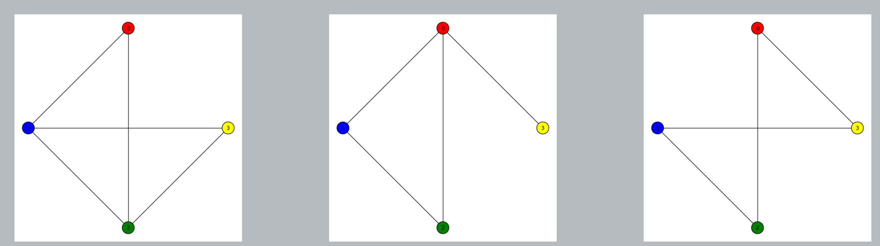
Abstract

Many frequently studied network models admit a multiplex structure that distinguishes the exogenous relationships connecting the nodes from the endogenous processes associated to each object. Unfortunately, it is difficult to understand these models from the structural perspective of graph theory. Here, we present a dynamical approach to this problem that allows us to perform standard techniques from network analysis on multiplex structures in a meaningful way.

As a case study, we analyze the World Trade Web (WTW) from this perspective, showing that many of the traditional analyses that have been performed on the WTW are incomplete with respect to the layer structure. In particular, we present results comparing the connectivity, clustering, assortativity, and centrality of the represented countries and commodities.

Multiplex Networks

A multiplex structure is a collection of networks all defined on the same set of nodes. Each of these "layers" can be used to capture a distinct type of relation between the nodes. For example, a social multiplex might have separate layers for professional relationships and online social interactions. Besides social networks, multiplex methods can be used for economic models, time delay models, and biological networks.



Layer 1 Layer 2 Layer 3

Figure 1 : A toy multiplex network model[1]

Early approaches to studying network problems in this context tried to address this problem from a structural perspective [5] (summing matrices or adding edges between copies). These approaches tend to distort the metrics of interest by conflating the intra and inter relationships.

Multiplex Dynamics

Instead of trying to add new structural components we connect the dynamics using a collection of scaled orthogonal projections. To each node, we associate a projection operator P_n that gathers the information stored at each copy of the node and proportionally redistributes it among the layers. This allows us to respect the independence of the endogenous dynamics.

Given a collection of operators D_i on our layers, this is equivalent to constructing the new operator:

$$M = \begin{bmatrix} \alpha_{1,1}C_1D_1 & \alpha_{1,2}C_1D_2 & \cdots & \alpha_{1,k}C_1D_k \\ \alpha_{2,1}C_2D_1 & \alpha_{2,2}C_2D_2 & \cdots & \alpha_{2,k}C_2D_k \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{k,1}C_kD_1 & \alpha_{k,2}C_kD_2 & \cdots & \alpha_{k,k}C_kD_k \end{bmatrix}$$

where the $C_i = \operatorname{diag}(c_{i,1}, c_{i,2} \ldots, c_{i,\ell})$ represent the coefficients for the node projections with the condition that $\sum_j c_{i,j} = 1$ for all i.

Theoretical Results

This dynamical approach to multiplex modeling allows us to prove general theorems about the spectral structure of the derived operator. Some of these results are summarized in the following theorems:

Theorem 1. *If the original dynamics satisfy any of the following {Irreducible, Primitive, Stochastic, Positive(Negative) (semi–)Definite} then so does the multiplex operator.*

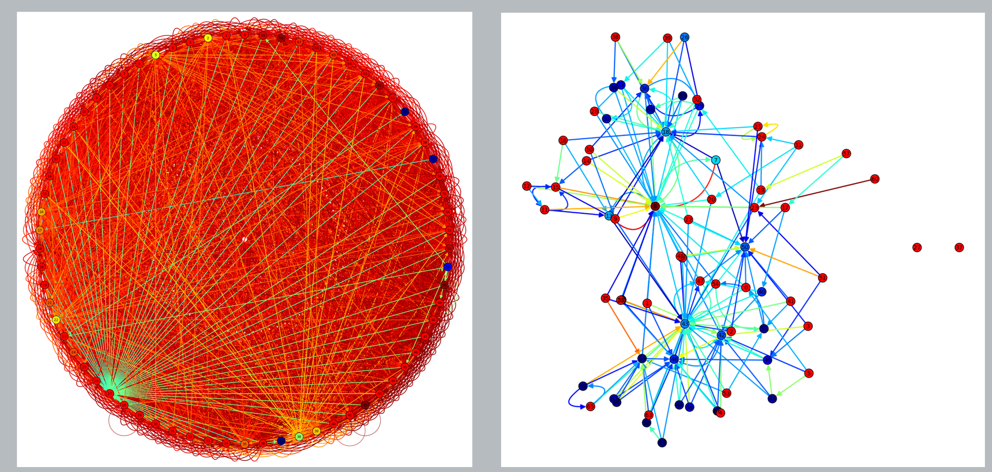
Theorem 2. *If the flows between layers are equidistributed and the individual dynamics are the associated network Laplacians we have the following eigenvalue bounds:*

- ▷ *Fiedler Value:* $\max_i(\lambda_f^i) \le \lambda_f \le \min_i(\lambda_1^i) + \sum_{j\neq\ell} \lambda_f^j$,
- ▷ *Leading Value:* $\min_i(\lambda_1^i) \le \lambda_1 \le \sum_i \lambda_1^i$,
- ▷ *Synchronization Stability:* $\frac{\min_i(\lambda_1^i)}{\min_i(\lambda_1^i)+\sum_{j\neq\ell}\lambda_f^j} \le G_{ss} \le \frac{\sum_i \lambda_1^i}{\max_i(\lambda_f^i)}$.

Theorem 3. *If the original dynamics are stochastic, the steady state vector of M cannot be expressed as a linear combination of the stationary vectors of the $\{D_i\}$.*

World Trade Web

The World Trade Web is a network representation of the trade interactions between countries [4]. In this network, the nodes represent countries and the edges represent the amount of goods sent from country i to country j each year. For simplicity, all of the data presented here is from the year 2000. The topology of the WTW has been quite stable [1, 3] over recent decades, so this analysis is representative of recent experiences.



Entire WTW Digraph Aggregate 3–out WTW

Figure 2 : Visual Realizations of the year 2000 WTW[1]

Commodity Parameters

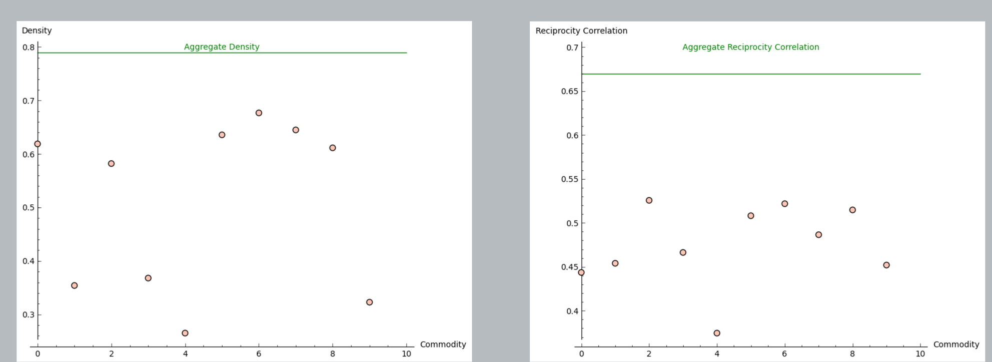
Trade data is usually disaggregated by commodity. Intuitively, we expect that the structure of various markets should vary widely. The table below presents some basic statistics of the layers represented in our data.

Layer	Description	Volume	% Total	Transitivity
0	Food and live animals	291554437	5.1	.82
1	Beverages and tobacco	48046852	0.9	.67
2	Crude materials	188946835	3.3	.79
3	Mineral fuels	565811660	10.0	.62
4	Animal and vegetable oils	14578671	0.3	.64
5	Chemicals	535703156	9.5	.83
6	Manufactured Goods	790582194	13.9	.87
7	Machinery	2387828874	42.1	.85
8	Miscellaneous manufacturing	736642890	13.0	.83
9	Other commodities	107685024	1.9	.56
All	Aggregate Trade	5667380593	100	.93

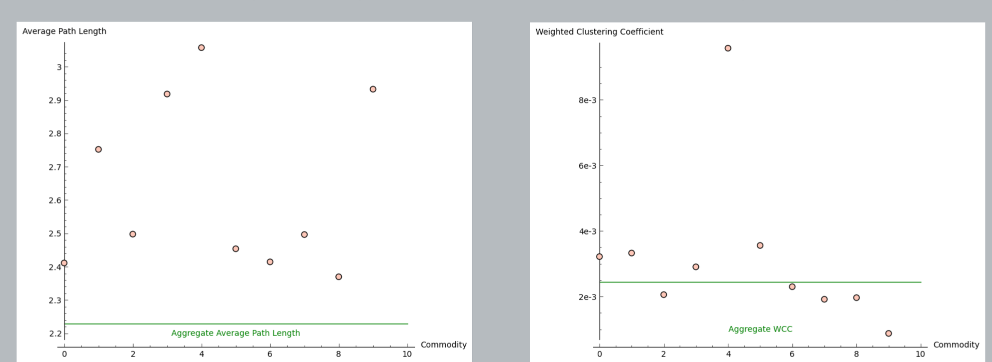
Table 1 : Commodity information for the WTW

Layer Metrics

Most of the prior analysis of the WTW [2, 3, 6, 8] has focused on the aggregate network, but the figures below show that a great deal of information is hidden by this aggregation. The results shown here are similar to those obtained in [1].



Commodity Density Commodity Reciprocity



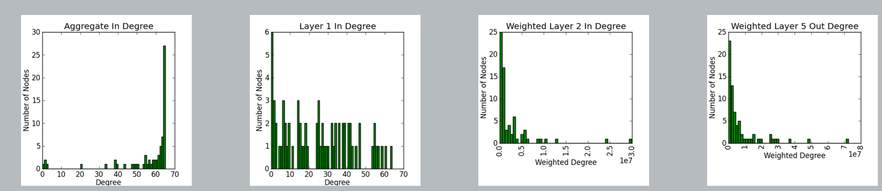
Commodity Path Length Commodity Clustering

Figure 3 : Comparison of standard network metrics between the aggregate WTW and the individual commodity networks[2]

Degrees and Clustering

The degree distribution of a network is a probability distribution measuring the probability that an arbitrary node has degree k. This is one of the most frequently studied aspects of complex networks, since many standard network examples, such as citation networks and transportation networks, have scale–free or power law distributions. We computed four degree distributions for each layer of the WTW reflecting both the directed structure and the weighted structure.

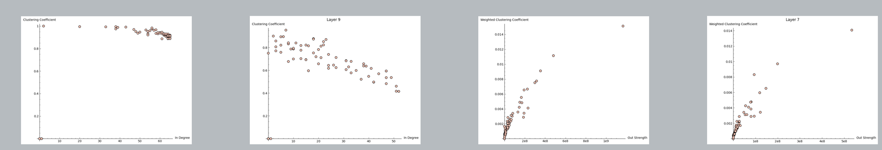
The unweighted degree distributions highlight the differences between the layers and the aggregate network. The aggregate distributions are skewed heavily to the right, while the individual layers are nearly uniform. On the other hand, the weighted distributions at all levels seem to satisfy a power law distribution for both in and out degrees. This is particularly interesting since it implies that a proportionally small number of nodes dominate both the import and export markets for each commodity.



Aggregate Unweighted Layer 1 Unweighted Layer 2 Weighted Layer 5 Weighted

Figure 4 : Representative Degree Distributions[2]

The clustering coefficient of a network is a measure of transitivity, the propensity for triangle formation in the network. Like the degree distribution, clustering coefficients can be computed as weighted or unweighted quantities. We found that for the layers, the unweighted clustering coefficient is negatively correlated with degree, while the weighted clustering coefficient is positively correlated with weighted degree.

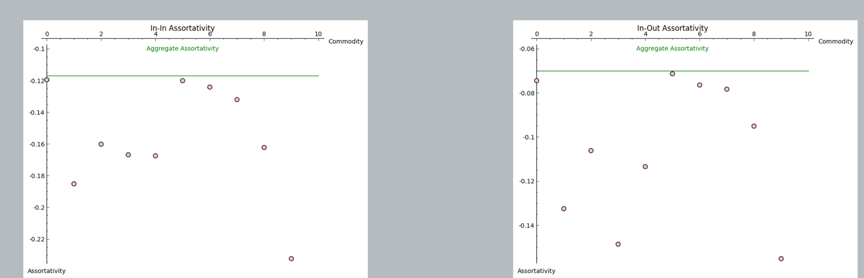


Aggregate Unweighted Layer 9 Unweighted Aggregate Weighted Layer 7 Weighted

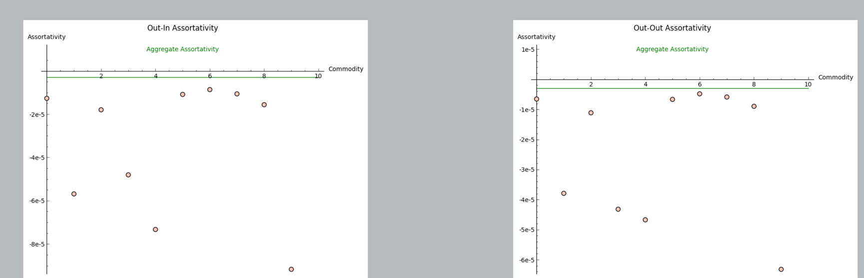
Figure 5 : Representative Clustering Plots[2]

Assortativity

Assortativity is a measure of the correlation between the degrees of connected nodes. For directed networks there are four separate assortativity values {in–in, in–out, out–in, out–out} that capture different information about the connectivity structure. Due to the highly reciprocal structure of the aggregate network, previous studies have only computed symmetrized versions of this statistic for the WTW.



In–In Assortativity In–Out Assortativity



Out–In Assortativity Out–Out Assortativity

Figure 6 : Assortativity values for the commodity networks[2]

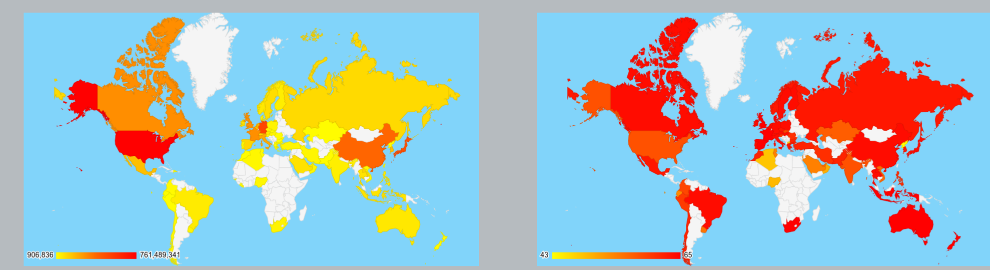
Although the assortativity values for the out–out correlation are small in magnitude, they differ markedly from the expectation as determined by a null model (Table 2). Thus, this data suggests that there are economic constraints determining this behavior, especially when considered with respect to the degree distribution analysis above.

Layer	0	1	2	3	4	5	6	7	8	9	Aggregate
Z–Score	-.41	-.43	-.52	-.47	-.53	-.40	-.40	-.45	-.39	-1.06	-.28

Table 2 : Z–Scores for Commodity Out–Out Assortativity

WTW Centrality

One of the main techniques of network science is using topological properties to determine the most central nodes in a network. There are many different ways to measure centrality [7] and each metric captures a different aspect of the importance of central nodes. For example, Figure 7 demonstrates the difference between the simplest two centrality values on the WTW. Although the USA has the highest aggregate strength, or trade volume, of any nation, it has a relatively small number of trading partners compared to the rest of the WTW.



Aggregate In Strength Aggregate In Degree

Figure 7 : Simple WTW Centralities[3]

Many measures of centrality incorporate dynamical aspects, such as random walks or geodesics. These metrics are particularly well suited for analysis under our multiplex operator, as well as having natural relations to the WTW data. Our main tools for studying centrality on multiplex structures come from these statistics.

Layer Centralities

We studied the centrality scores of the individual layers using several different metrics, focusing on those that are related to flows across the network. The table below reports the most central node under each metric. Interestingly, although the USA has the highest aggregate trade volume of any nation (Figure 7), it does not dominate all of these measures.

Layer	Out Degree	In Degree	Closeness	Eigenvector	RWBC
0	Japan	Denmark	Japan	USA	USA
1	Germany	UK	USA	France	France
2	Netherlands	Germany	Netherlands	Canada	USA
3	USA	Germany	USA	Canada	Oman
4	Germany	Netherlands	France	Indonesia	Laos
5	Germany	Denmark	France	USA	USA
6	Belgium	France	Belgium	USA	Germany
7	Germany	Denmark	France	USA	USA
8	UK	France	UK	China	China
9	USA	UK	USA	Canada	USA
All	China	8 Countries	China	USA	USA

Table 3 : Highest Layer Centrality Scores

Multiplex Centrality

We applied this same concept in the multiplex setting. Using our derived operator M we were able to compute analogues of these centrality measures for the entire multiplex, comparing the interactions between individual components of each countries trading profile. To distribute the flow between the commodities, we used four separate weighting schemes: an equal distribution, proportional to node in strength, proportional to node out strength, and the layer weight proportions from Table 1.

	Equal		In Strength		Out Strength		Layer Strength	
Rank	Layer	Country	Layer	Country	Layer	Country	Layer	Country
1	All	USA	7	Japan	7	USA	7	USA
2	All	Canada	7	USA	7	Canada	7	Japan
3	All	Japan	7	Mexico	7	Mexico	7	Canada
4	All	China	7	Canada	7	Japan	7	Mexico
5	All	Mexico	7	Germany	7	China	7	China
6	All	Germany	8	China	3	Japan	7	Germany
7	All	UK	7	S. Korea	7	Germany	6	USA
8	All	France	7	China	8	USA	8	USA
9	All	S. Korea	7	Laos	8	Japan	6	Japan
10	All	Italy	8	USA	7	Laos	7	S. Korea

Table 4 : Multiplex Centrality Leaders

Note that 42% of the total trade volume in our data occurs in layer 7, so it is reasonable that most of the important multiplex nodes lie in that layer.

Community Detection

The random walk associated to a digraph can be used to construct a metric known as the commute time on the network. The values of this metric can be computed from the pseudoinverse of a modified Laplacian [9]. After computing the multiplex version of this operator with our method, we used a complete linkage algorithm to cluster the countries according to their distances under the commute time. The eight most significant clusters are displayed in the figure below.

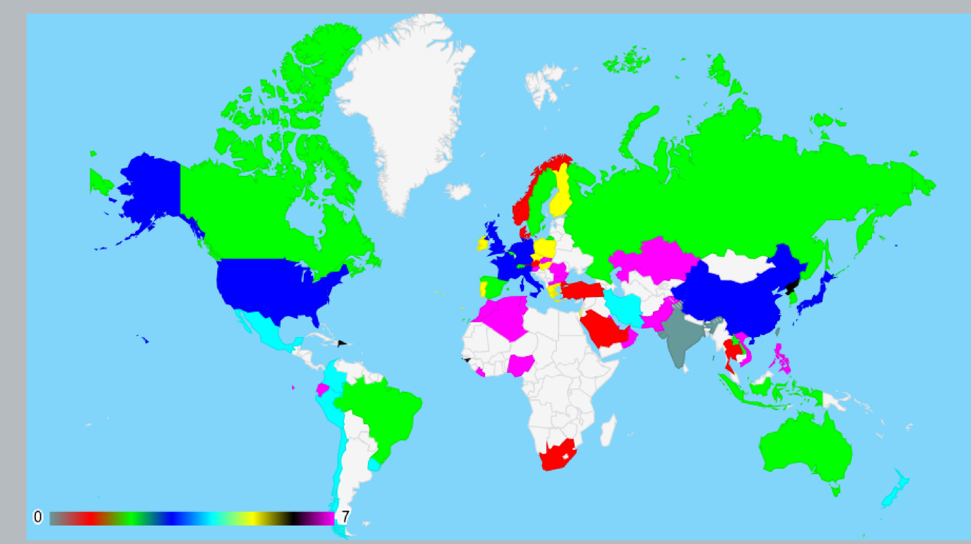


Figure 8 : Communities in the WTW[3]

Conclusions and Extensions

Our dynamically motivated approach allows us to analyze multiplex structures using many of the standard tools and techniques of network theory. This is a significant improvement over prior structural models that did not respect the dynamical interpretations of derived network operators. With regard to the WTW dataset we showed that the commodity layers display distinctly different topological structure than the aggregate network, suggesting that the aggregation process obscures a great deal of economic information. This is particularly important when studying measures of centrality or robustness from a practical perspective.

We intend to extend this research in several ways:

- ▷ Proving stronger eigenvalue bounds for commonly studied operators
- ▷ Developing community detection methods that respect the multiplex structure such as generalized modularity
- ▷ Studying the correlation of spectral data from multiplex networks with traditional layer dynamics
- ▷ Incorporating external control variables into the dynamic analysis

References

[1] M. Barigozzi, G. Fagiolo, and D. Garlaschelli: *Multinetwork of international trade: A commodity–specific analysis*, Physical Review E, **81**, 1–23, (2010).

[2] G. Fagiolo, J. Reyes, and S. Schiavo: *On the topological properties of the world trade web: A weighted network analysis*, Physica A, **387**, (2008), 1868–1873.

[3] G. Fagiolo, J. Reyes, and S. Schiavo: *World–Trade web: Topological properties, dynamics, and evolution*, Physical Review E, **79**, (2009), 1–19.

[4] R. Feenstra, R. Lipsey, H. Deng, A. Ma, H. Mo: *World Trade Flows: 1962–2000*, Working Paper 11040, NBER.

[5] S. Gomez, A. Diaz-Guilera, J. Gomez-Gardenes, C.J. Perez-Vicente, Y. Moreno, and A. Arenas: *Diffusion Dynamics on Multiplex Networks*, Physical Review Letters, **110**, (2013), 1–5.

[6] N. Foti, S. Pauls, and D. Rockmore: *Stability of the World Trade Web over time –An extinction analysis*, Journal of Economic Dynamics & Control, **37**, (2013), 1889–1910.

[7] M. Newman: *Networks: An Introduction*, New York, Oxford University Press, (2010).

[8] M. Serrano and M. Boguna: *Topology of the World Trade Web*, Physical Review E, **68**, (2003), 1–4.

[9] L. Yen, D. Vanvyve, F. Wouters, F. Fouss, M. Verleysen, and M. Saerens: *Clustering using a random walk based distance measure*, Proc. ESANN '05, (2005), 317–324.