EECS4414 Project Proposal: Analysis of Annual Global Trade through Information Networks

|  |  |  |  |
| --- | --- | --- | --- |
|  | Matt MacEachern York University 4700 Keele St. Toronto, ON, M3J 1P3 mcmaceac@my.yorku.ca | Andrew Jaramillo York University 4700 Keele St. Toronto, ON, M3J 1P3 andrew2jaramillo@gmail.com |  |

ABSTRACT

KEYWORDS

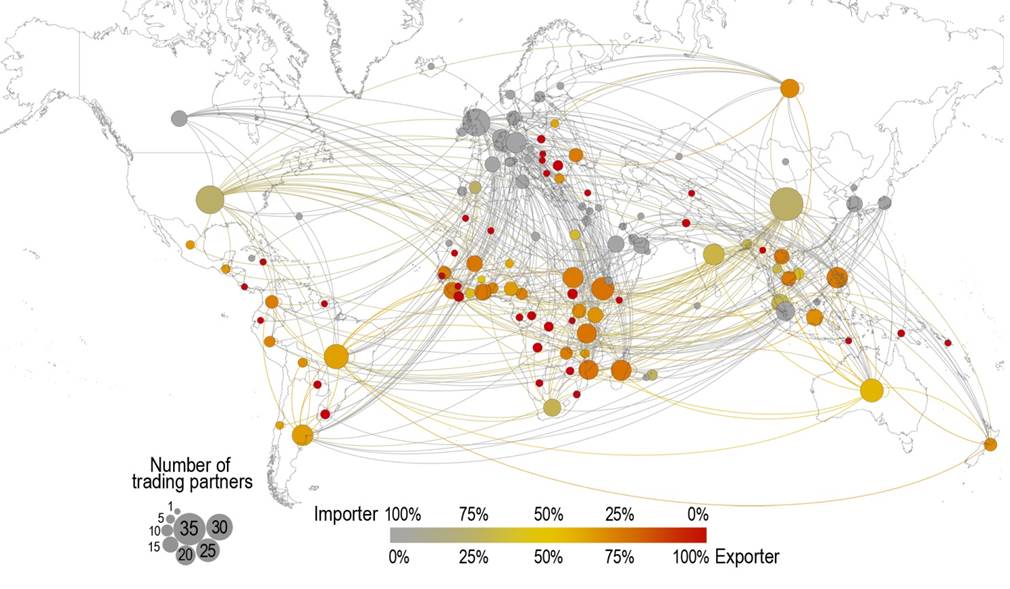
Global Trade, Link Prediction, Network Evolution, Trends, Emerging Markets, Market Community Detection

1 INTRODUCTION AND MOTIVATION

The analysis of global trade has been an essential part of the development and maintenance of economies around the world. Through recessions, war, global catastrophes, sanctions, etc., the global trade network can give interesting and insightful information not only at a historical level, but when looking towards the future of the world economy. This information is valuable and no doubt being analyzed by economists to determine trends in world trade. This trade network from a high level may not seem as complex and interesting as some networks with higher node counts, however the relationships between nodes are incredibly complex, and even small changes to the topology of the graph could have huge implications when it comes to the overall structure of the network.

This development has interesting implications when looking from the perspective of information networks that trade has developed. At a country by country basis, the trade of goods forms a strongly connected, directed graph with edge weights representing the annual amount of goods sold by one country to another. An example of what this graph may look like after it has been developed throughout this project is shown in Figure 1. The graph itself may not seem grandiose or groundbreaking when graphs of millions of nodes are being analyzed in the social networks that have emerged over the past fifteen years. However, the network itself is still an incredibly interesting one in that it has been evolving for centuries (albeit the data is not readily available or accurate for years before the emergence of computers). Also, the number of interesting graph algorithms that can be applied to this network is not limited by the number of nodes in the network.

In this project, there are many interesting analyses that can be performed on the network itself. There are many properties of this graph that can be analyzed. These include (but are not limited to): link prediction, community detection, time-series analysis, as well as topological analysis to see if the global trade network follows a certain already well-known model.



**Figure 1:** What the global trade network could look like after the development of the graph.

2 METHODOLOGY

2.1 Graph Representation

The idea behind representing the data on global trade as an information would be to have each node representing a country. An edge would be between two countries if those countries trade with each other as discovered through the data we collect (see section 3.1 for further information on this data). That is, if node A has an edge connecting to node B, countries A and B are trade partners in some way. The graph ideally will be directed with edge weights applied to each edge. Node A has a directed edge E to node B with weight W if country A has exported a dollar amount of goods equivalent to W in dollars to country B in that particular year. Another idea would be to have the net total amount traded between the two countries A and B and have an undirected edge between the two, but this method would be less descriptive.

2.2 Link Prediction

Building recommendation engines has been studied extensively in different fields. Common methods involved link prediction on heterogeneous graphs. Also, [7] attempts to improve this method, include supervised random walk, where the algorithm assumes the network is homogeneous, and hence the random walk has no constraints. Additionally, link prediction can also be looked through information diffusion. This principle could be applied in trade relationships, where the propagation of the information is the trade itself. Meaning, does the trading patterns of country A affect country B and the countries spanning from B in the trade network.

2.3 Community Detection

Finding communities in any graph is not a simple task, and the WTW network is no different. Although detecting communities using algorithms such as the Girvan–Newman algorithm is very helpful, trying to infer meaning from those findings and knowing when one has sufficiently subdivided a network into communities is not an exact science. Nonetheless, studies have been done specifically on the WTW to find the communities that inevitably exist due to trade deals, geographical location, sanctions, etc. [6]. Barigozzi et al. have performed an interesting analysis on 14 commodities, and the communities that arise due to these commodities in the WTW [6].

2.4 Time-Series Analysis

Trade relationships and their evolution over time is an important aspect of the WTW in understanding how the network came to be in its current state. Fagiolo et al. do an incredibly in depth analysis of the evolution of the WTW from 1981 to 2000 which uncovered numerous interesting facts about network itself [2]. They uncovered that certain many countries have weak trade links, while there seems to be a core structure of rich countries that are more highly connected to other countries in the network. This goes back to the idea of community detection presented in 2.3. For the purpose of this project, a more recent analysis of the WTW would be interesting in light of recent economic events (specifically the financial crisis of 2008), and how the WTW adapted and evolved in response to these events.

An important model relating to the evolution of the WTW is the fitness network model discussed by Garlaschelli and Loffredo [5]. This model states that each node in the network has an inherent competitive factor called the nodes’ fitness. This measure is related to the idea of “the rich get richer” in that nodes that have a higher fitness tend to attract stronger links at the expense of other nodes. The exact math behind this method will be left for when the data is actually collected, however this could potentially be a very interesting factor in looking at the annual network evolution of the WTW.

2.5 Topological Analysis

Topology analysis is a quintessential part of any graph analysis, and the topology of the WTW is no different. The structure of the WTW will be verified through the analysis of the data set discussed in 3.1, however many researchers have discovered that some well known properties of the WTW are that it seems to follow the power-law distribution, has a high clustering coefficient, and follows the small-world network model [1, 3]. Another interesting phenomenon found in these networks in the past has been the correlation between GDP per capita, and the centrality of these nodes in the WTW network [1]. It has been revealed that countries with higher GDP per capita tend to have a more central position in the network, and have more trade relationships (edges) in the network than lower GDP countries.

3 EVALUATION

3.1 Data Set

The data set uncovered for the purpose of this project comes from the World Bank and is managed by the World Integrated Trade Solution or WITS [4]. WITS allows users to retrieve data on a country by country basis, and filtering on a number of aspects. The data for most countries dates back to 1989, which will suffice for the purpose of our study considering we will be more interested in a recent analysis of the WTW. WITS also allows for filtering on certain product categories such as fuels, chemicals, plastics, etc. to allow for a more in depth analysis if necessary.

REFERENCES

|  |  |
| --- | --- |
| [1] | Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. 2009. World-trade web: Topological properties, dynamics, and evolution. Physical Review E 79, 3 (2009). DOI:http://dx.doi.org/10.1103/physreve.79.036115 |
| [2] | Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. 2007. The Evolution of the World Trade Web. Physical Review E 79, 3 (2007). |
| [3] | Ma Ángeles Serrano and Marián Boguñá. 2003. Topology of the world trade web. Physical Review E 68, 1 (November 2003). DOI:http://dx.doi.org/10.1103/physreve.68.015101 |
| [4] | "World Integrated Trade Solution (WITS) | Data on Export, Import, Tariff, NTM", Wits.worldbank.org, 2018. [Online]. Available: https://wits.worldbank.org. [Accessed: 07- Feb- 2018]. |
| [5] | Diego Garlaschelli and Loffredo Maria. 2004. Structure and evolution of the world trade network. Physica A (2004). |
| [6] | Matteo Barigozzi, Giorgio Fagiolo, and Giuseppe Mangioni. 2011. Identifying the community structure of the international-trade multi-network. Physica A: Statistical Mechanics and its Applications 390, 11 (2011), 2051–2066. DOI:http://dx.doi.org/10.1016/j.physa.2011.02.004 |
| [7] | Lars Backstrom and Jure Leskovec. 2011. Supervised random walks. Proceedings of the fourth ACM international conference on Web search and data mining - WSDM 11(2011). DOI:http://dx.doi.org/10.1145/1935826.1935914 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |