On the Relation Between Tourism and Trade: A Network Experiment

Jigar Brahmbhatt  
BioComplex Laboratory  
Department of Computer Sciences  
Florida Institute of Technology, USA  
jbrahmbhatt2010@my.fit.edu

Ronaldo Menezes  
BioComplex Laboratory  
Department of Computer Sciences  
Florida Institute of Technology, USA  
rmenezes@cs.fit.edu

Abstract—Current literature on world global trade and tourism indicates these two areas are related and may influence each other. This study investigates the relationship between international tourism and international trade flows using an approach based on network sciences. In this paper, bilateral trade-exports data and inbound tourist arrival data were modeled as weighted directed networks and then analyzed for possible common properties. The empirical results show strong weighted degree correlation and strong PageRank correlation between trade and tourism networks. This paper shows the relationship using metrics derived from network sciences.

I. INTRODUCTION

Most systems studied today are non-trivial in nature—they are generally composed of many parts with independent roles but contributing to the whole. Moreover, these systems’ behaviors cannot be explained by the sum of its parts. The literature tends to call systems with these characteristics complex systems. The study of complex systems began with the effort to identify their structure and develop models that can reproduce their characteristics. Graphs are fundamental concepts in complex systems research. By representing complex systems as graphs, we can apply graph-theoretical approaches that are mature and well understood.

The National Research Council defines network science as “the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena [1]”. The term “complex network” has been used widely to represent a variety of systems such as social networks [2], economic networks [3], transportation systems [4], epidemic spreading [5], metabolic networks [6], the World Wide Web [7], and many more.

Due to the interest in complex networks and the tremendous improvements in computer technology, many new network models have been proposed and formalized in the past few years. These models showed that many real networks are not random. Watts and Strogatz [8] described an important measurement in networks called “small-world”. A network with the small-world property is one in which each node can reach any other node via a small number of hops, and also demonstrates the property of large clustering coefficient. Another interesting property was observed by Barabasi et al. [9] and called “scale-free”. It relates to the networks having a power-law degree distribution. They modeled the topology of the World Wide Web in their experiment and showed that such networks are scale-free.

This paper focuses on analyzing the relationship between international tourism and international trade using complex network concepts. Though international trade is not a recent concept, it has been shown to be important economically [10], socially [11], and politically [12] during recent decades. International trade has been studied to find out any causality to macroeconomics [13], exchange rate [10], and financial crises contagion [14], to name but a few.

Travelling around the world has become a lot easier, faster and cheaper in recent years. Over the past six decades, the tourism industry has expanded and diversified, creating a powerful impact on the economy of the world. According to United Nations World Tourism Organization (UNWTO) international tourist arrivals reached 983 million worldwide in the year 2011, up from 940 million in 2010. For many developing countries, tourism is the principal source of foreign currency income, tops most export categories, and creates essential employment and opportunities for future development. Tourism’s direct contribution to the global GDP in 2011 was US$2 trillion and the industry generated 98 million jobs.

Countries which opened their doors to free-trade agreements gained growth in international trade flows and international travel. In one sense, international trade and tourism perform a similar role in the financial sector of a country’s balance of payments accounts. Still, the question arises: does any link exist between international trade and tourism flows? There are studies which deal with world trade networks [17], [18] and trade-tourism analysis [16], [19] based on statistical and economic models. This paper attempts to uncover any relation between tourism and trade using complex network concepts.

The study of the potential complementary relationship between flows of goods and international tourism is of major interest, as it can promote economic growth of a nation because trade and tourism are two of the most important economic activities for any nation. The contribution of this study is twofold: to help (1) policy decision making and

1According to “The Travel & Tourism Competitiveness Report 2011” by World Economic Forum (http://www.weforum.org/)
2According to ‘Economic Impact of Travel & Tourism 2012 Summary’ (World Travel & Tourism Council)
The core-periphery approach was also used by Nemeth and Smith [23], who represented countries in structural positions based on flows separated over five distinct types of internationally traded commodities. Whereas Snyder and Kick [22] placed countries into positions based on trade, military interventions, treaty memberships, and diplomatic exchanges, Nemeth and Smith [23] solely focused on classifying nations based on trade networks and unequal exchange. In their analysis, 89 countries were separated into four different structural positions that validated the dependency theories.

Smith and White [24] wrote a seminal article related to world trade network (WTN) analysis. They apply the core-periphery approach to study the evolution of the WTN using dynamic analysis for three different years (1965, 1970 and 1980). Previous studies analyzed only one point in time. By using a dynamic modeling approach, they were able to confirm the expansion of core countries over the period of time, track the change of position of specific countries, and observe the continuous marginalization of circumferential countries. In a similar study, using a weighted network approach, Fagiolo et al. [25] examined the presence of homophily between countries in terms of richness, structure, and other variables. They also confirm that nations with high positions in the WTN lead to higher growth than nations with low positions.

Apart from studies centered on the core-periphery framework, there are studies in which network analysis has also been employed to investigate the debate over whether regionalization is a stepping stone or a stumbling block to globalization. Some pundits consider regionalization as an ephemeral maneuver that some countries follow to become more competitive on the globalized market and believe that in the end it will promote globalization. Others propose that regionalization hinders the process of globalization by damaging the welfare of individual non-member countries and contributing to inefficient production strategies that may work at the regional level but not at the international level. Kim and Shin [26] also find that the density of the network of countries and node-degree distributions both have increased with time—an effect associated with the globalization process. Furthermore, newly added connections do not seem to be equally distributed. Countries in the core tend to make outbound connections while peripheral countries tend to make inbound connections. The consequences of globalization are further researched by Kastelle et al. [27] who used complex network analysis on International Monetary Fund data from the period 1938–2003 to examine standard hypotheses on the dynamics of the topological properties of the world trade network. They determined that development of the WTN has not reached any stabilized state that would imply one huge integrated global market. Instead, it has been slowly modifying and appears to continue doing so in the future.

Studies related to trade networks have focused on either understanding the network structure and properties or its economic impact. They have used the unweighted undirected network approach to demonstrate the connection between nations which may not be a good idea because consideration
of edge weights can provide interesting insights. Similarly, directed network analysis is important to understand inflow and outflow patterns. In this paper, the world trade network is examined together with the world tourism network to discover any similarity in their structure and flow patterns. In our methodology, we have used edge weights to generate directed trade and tourism networks considering the bilateral global data of 204 and 229 countries respectively.

III. DATA AND IMPLEMENTATION

The source of the trade data is the United Nations Commodity Trade Statistics Database (UN Comtrade)—a United Nations Statistics Division. The merchandise trade data used in this thesis was downloaded from the United Nations Conference on Trade and Development (UNCTADstat) [28]. UNCTADstat provided a compiled bilateral matrix of merchandise trade by trading partner and product based on the November 4, 2011 version of Statistics of International Trade in Services (SITC). It presents the total of all product exports of merchandise trade from reporting country to its trading partner from the year 2006 to 2010. The total number of countries in the dataset is 233, and the trade values are expressed in thousands of dollars.

To get the geographical layout view of global trade flow on the network analysis software package, we downloaded the CIA World Factbook 2011 [29] and obtained latitude and longitude information of nations’ centroids (expressed in degrees and minutes). To support the requirement of the software package, we converted this information into a decimal degree format.

The tourism data was obtained from Compendium of Tourism Statistics 2012 edition [30]—the most comprehensive statistical database available on the tourism sector provided by the World Tourism Organization (a Specialized Agency of United Nations). It provides statistical data and indicators on inbound, outbound and domestic tourism, as well as on the number and types of tourism industries, the number of employees by tourism industries, and macroeconomic indicators related to international tourism. The Compendium of Tourism Statistics for year 2006 to 2010 was kindly provided by the World Tourism Organization in exchange for assurances to maintain privacy of data and commitment to provide the results of this study.

To create a bilateral matrix of tourism data, we used the data from The Yearbook of Tourism Statistics which focuses on data related to inbound tourism (total arrivals and overnight stays), broken down by country of origin. The final dataset contains a 229×229 matrix that represents the origin country and destination country, and the intersecting cell contains the total number of arrivals of non-resident tourists in the destination country from origin country. We used the same latitude-longitude data for tourism analysis.

A. Consistency Checks in the Datasets

In the trade dataset we deleted some nations from the matrix because no trade information was available in the source data (e.g. British Virgin Islands, Vatican City). We also verified that some countries and territories listed in global GDP data (provided by the World Bank Development indicator) did not have any GDP data or had negligible trade data, therefore we eliminated them from the dataset (e.g. Turks and Caicos, Cook Islands). After the elimination process, the Trade dataset consists of 204 countries’ bilateral trade data.

In the tourism dataset provided by the WTO, there is inbound tourism data for 218 countries but there are some countries and territories for which information was not provided. Therefore to make proper bilateral tourism matrix, we added new rows for those countries and territories (e.g. Afghanistan, Gibraltar). After compiling the data, there were two countries (Saudi Arabia and Saba) that had an inbound file but was not listed anywhere in other countries’ inbound file, therefore we did not use them in the matrix. There are also some countries listed in just one or two individual country files, so we also eliminated them from bilateral matrix (e.g. Falkland Islands, Isle of Man). After the elimination and addition process, the tourism dataset consists of 229 countries’ bilateral inbound tourism data.

The total values in our travel dataset may not reflect the same total values and ranks as listed in World Tourism Statistics [30]. The main factor that causes the slight differences is the inbound tourism statistics provided by individual countries. The tourism data files contain tourism arrival statistics in six categories: (1) arrivals of non-resident overnight visitors (tourists) at national borders, (2) arrivals of non-resident visitors (overnight and same-day visitors) at national borders, (3) arrivals of non-resident overnight visitors (tourists) in hotels and similar establishments, (4) arrivals of non-resident overnight visitors (tourists) in all types of establishments providing accommodation services for visitors, (5) overnight stays of non-resident overnight visitors (tourists) in hotels and similar establishments, (6) overnight stays of non-resident overnight visitors (tourists) in all types of establishments providing accommodation services for visitors. The problem is not all the countries have provided data using all categories, so we decided to use the values of category 2 when it is available as it gives us total number of arrivals. If category 2 is not available in the file then we used category 1, category 4, or category 3 according to availability. Apart from that, some countries have provided total arrival data using regions (e.g. Eastern Europe, South Asia etc.) that we could not consider while generating the matrix. For these reasons, minor data inconsistency can be observed. We are confident these issues do not affect our results.

In this paper we also perform Pearson correlation analysis to measure the correlation between trade and tourism variables. For this analysis, we needed the same number of variables (countries) for both trade and tourism, hence we eliminated countries from the tourism dataset which were not also in trade dataset.

http://data.worldbank.org/indicator/NY.GDP.MKTP.CD/countries
IV. EXPERIMENTAL RESULTS

We have worked with networks from five distinct years, although the number of nodes is fixed for all network, the number of edges change (slightly) because of the changing level of trade and tourism between countries and because of the inconsistencies listed in the previous section. Table I lists the number of nodes and edges in all networks. Table I also shows that the trade network is more densely connected than the tourism network.

We first analyzed the full network of both datasets using community detection based on the agglomerative Blondel et al. [31] algorithm, which discerns community structure by grouping nodes in the configuration which maximizes Girvan-Newman modularity [32]. The existence of communities could be used to empirically show whether we can observe regionalism in the trade and tourism networks.

In the tourism network, we found seven different communities between countries. Among them, one community is mostly formed from Europe, Caribbean islands, and British overseas territories; another from northeast Europe and The Commonwealth of Independent States; two from Africa; one from the Americas, one from the Middle East; and one from Asia and the Pacific countries. This reconfirms that people tend to travel more in the neighboring country. Inter-governmental organizations, cost of travel and transport, similar currency, ease of environment etc. can be factors behind the regional flow. On the other hand, community structure in the trade network does not look as regional as in tourism. We found four communities which have countries together from more than one region. Figures 1 and 2 show the community structure in the tourism and trade network for year 2010 respectively.

A. Trade and Tourism Network Measures

Scale-free networks are characterized by having a distribution that follows a power-law. When dealing with weighted networks though it is more appropriate for one to look at the weighted degree distribution as it better represents the distribution of the relationships in the network [33]. The weighted in-degree and out-degree distributions of the entire tourism networks for all years (2006–2010), and the weighted in-degree and out-degree distribution of the trade networks for all years (2006–2010) show a long-tailed distribution that is typical of a scale-free network and distribution exponent that approaches the range of 2 to 3 seen in many scale-free networks. The small-world property is identified by having low average path length and high clustering coefficient; all the networks studied here are classified as small world. We also analyzed other statistical properties of these networks. Table II summarizes our network properties and adds a comparison to the ACM networks.

B. Correlation Between Trade and Tourism Networks

The experimental results in this section demonstrate the correlation between the trade and tourism networks. The correlation analysis focuses on the weighted in-degree, weighted out-degree, and PageRank measures. The PageRank algorithm helps find the most important nodes of the networks using a metric that considers a node important if the nodes pointing to it are themselves important. We ranked the countries based on their PageRank values for all ten networks. We wanted to see if the important countries are similar in both global networks. We found that, based on their PageRank values, top-ranked countries remained the same from the years 2006 to 2010 in both networks. United States and China ranked the top two in the tourism network for all five years.

The Pearson correlation coefficient is used to measure the correlation between all four degree measurements and PageRank. The PageRank algorithm helps find the most important nodes of the networks using a metric that considers a node important if the nodes pointing to it are themselves important. We ranked the countries based on their PageRank values for all ten networks. We wanted to see if the important countries are similar in both global networks. We found that, based on their PageRank values, top-ranked countries remained the same from the years 2006 to 2010 in both networks. United States and China ranked the top two in the tourism network for all five years.

The Pearson correlation coefficient is used to measure the correlation between all four degree measurements and PageRank of both networks. The Pearson correlation coefficient is utilized here because it is widely used in both the physical and social sciences to measure linear dependency between two variables. The correlation coefficient value ranges from -1 (highly uncorrelated) to 1 (highly correlated).

The interpretation of the correlation coefficient is dependent on the application. For example, a high positive value will have different meanings when interpreted in the context of a physical law using high-quality instruments versus relationships...
observed in social sciences. This variance in interpretation exists because physical measurements are performed very precisely in tightly controlled environments. On the other hand, measurements in the area of social sciences are taken in an environment where many complicating and interfering factors exist [36]. In the case of trade and tourism network analysis, the interpretation of the Pearson correlation coefficient (given by $r$) is taken in the context of the social sciences because we are measuring human flows and socio-economic interaction in a global context.

In our analysis we found high positive correlation between degree weights and PageRank for several network pairs. Weighted degree correlation implies that countries have similar importance in terms of their position in the inbound and outbound tourism networks, and the import and export trade network. It may also support the four-way relationship between tourism arrival, import, outbound tourism, and export. Results of this analysis are summarized in Table III.

Given that we have such high correlation coefficients, we double checked it against other datasets that we know should not be highly correlated. In this way we become more confident that the high correlation of Table III is significant. We performed a correlation analysis between weighted out-degree trade data and six global datasets [37] which include: (1) refugee population (by origin of country), (2) time required to start a business (days), (3) average precipitation, (4) percent of the population affected by droughts, floods, and extreme temperatures, (5) adolescent fertility rate, and (6) percent of the population in rural areas. We found that the correlation varies from no correlation to medium-negative correlation. We did not include other trade and tourism dataset for this analysis because they would have given similar values as they all are highly correlated to each other. Table IV shows the results of these correlation analyses.
Fig. 2. Community structure in Trade Network of 2010. The color nodes represent their community; countries in the same community are grouped together in a line. In the Trade Network the division is also based on colors and position in the visualization, the communities here are not as well defined as in the case of tourism (Figure 1) and we see some mixture of regions such as in the case of the largest community that includes African, middle eastern and asian countries.

TABLE III

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Countries</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade$<em>{in}$$\rightarrow$Trade$</em>{out}$</td>
<td>198</td>
<td>0.740</td>
<td>0.734</td>
<td>0.734</td>
<td>0.727</td>
<td>0.725</td>
</tr>
<tr>
<td>Trade$<em>{out}$$\rightarrow$Trade$</em>{out}$</td>
<td>204</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>Trade$<em>{in}$$\rightarrow$Tourism$</em>{out}$</td>
<td>0.87</td>
<td>0.87</td>
<td>0.86</td>
<td>0.86</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Trade$<em>{out}$$\rightarrow$Tourism$</em>{in}$</td>
<td>178</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>PageRank</td>
<td>203</td>
<td>0.67</td>
<td>0.66</td>
<td>0.66</td>
<td>0.65</td>
<td>0.65</td>
</tr>
</tbody>
</table>

V. Conclusion

In this paper we have run several experiments on the world trade network and the world tourism network to examine possible relationships between them using the network analysis approach. The UNWTO tourism dataset and UNCTAD trade dataset were used mainly because (1) the data values are directly reported from an individual country, and (2) the datasets are nearly complete and well structured.

We argue that network science analysis can be useful to find out the correlation, if any, between global trade and tourism flows. As some methods described herein identify some structural similarity between these two global networks, they can be used as an indicator that there are hidden similarities between both networks which can be revealed using deeper network analysis. Clearly, the experimental results show a high correlation between the world trade and tourism networks. More importantly, the high correlation observed in the four-way weighted in-degree and out-degree relationship between both the networks supports our assumption of a similar pattern.
of trade and tourism flows.

The analysis performed in this paper leads to several interesting areas for future work. Ego-network similarity should be tested on groups of countries on the basis of different factors including weighted degree. We expect that the work done in this thesis will raise interest in finding more useful insights about the relationship between global trade and tourism using the network science approach. Further, it would be interesting to analyze the same networks using normalized weights on the basis of GDP per capita, tourist arrivals per capita or any other important economic indicator. In that way, one may find strength of relationship between trade and tourism among nations having similar or dissimilar forms of government. Other interesting work can be done to find out the regionalism patterns among trade and tourism of countries over a long span of years. Apart from regionalism, the trade and tourism networks’ dynamic growths can also be examined over a long span of years. It might help in understanding when the global trade and tourism networks started showing similar structures and make us think about causation between trade and tourism.

ACKNOWLEDGEMENTS

We would like to thank Evan Stoner from the Florida Institute of Technology for the help in the preparation of this manuscript. His suggestions were very useful and improved the readability of this document.

REFERENCES


TABLE IV

<table>
<thead>
<tr>
<th>Miscellaneous Datasets</th>
<th># Countries</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refugee population (by origin of country)</td>
<td>185</td>
<td>0.02</td>
</tr>
<tr>
<td>Time required to start a business (days)</td>
<td>179</td>
<td>-0.20</td>
</tr>
<tr>
<td>Average precipitation in depth (mm per year)</td>
<td>176</td>
<td>-0.23</td>
</tr>
<tr>
<td>Droughts, floods, extreme temperatures affected (% of population)</td>
<td>157</td>
<td>-0.33</td>
</tr>
<tr>
<td>Adolescent fertility rate</td>
<td>186</td>
<td>-0.44</td>
</tr>
<tr>
<td>Rural population (% of total population)</td>
<td>201</td>
<td>-0.44</td>
</tr>
</tbody>
</table>


