

A new world trade ranking method using link analysis

JIANG Chun-heng¹, LIN Wen-bin²

(1.School of Mathematics, Southwest Jiaotong University, Chengdu 610031, P.R.C.;

2.School of Physical Science and Technology, Southwest Jiaotong University, Chengdu 610031, P.R.C.)

Abstract: World trade requires multifaceted tools for measurement and analysis. This work proposes one new method for world trade ranking of countries, where both the import and export are brought into considered. With the concept of world trade network (WTN), link analysis is utilized to evaluate the influence of countries within the global multilateral trade network. With the proposed method, each country in the network will receives one non-negative score for ranking. Empirical experiments on real trade data indicates that the proposed method provides an effective tool for world trade measurement and analysis.

Key words: world trade network; ranking algorithm; link analysis; PageRank; HITS

中图分类号: TP393.092

文献标识码: A

文章编号: 1003-4271(2014)03-0451-05

1 Introduction

Trading plays a vital role for international interactions among countries. The analysis and understanding of world trade is of primary importance for modern international economics [1]. Usually, the world trade ranking of countries is done in terms of the export and/or import measured in US dollar. However, this approach biases towards the developed nations and is unfavorable to the under-developed countries. To avoid this problem, a natural idea is to present the trade flow of goods and services between two countries, using a directed line that linked two vertices representing the two trading countries. Moreover, the value of the flow can also be attached to the line. Thus, a directed graph called World Trade Network (WTN) can be constructed with all nations in global included. Although the notion of world trade network is not new, and economists have conceived the international trade as a network since long ago [2], there are still numerous efforts are devoted to explore the topology properties of trade network in the past decades [3, 4, 5].

This paper focuses on the ranking of countries involving in the global trading. The ranking problem is also the core of information retrieval, and especially for web search. To efficiently find the most relevant information in the sea of information, various ranking methods have been developed. These ranking methods can be categorized into two classes, one for content analysis, and the other one for link analysis. In the latter category, the methods such as PageRank [6], HITS [7], TrustRank [8], are all built upon the assumption that, web pages are connected via the hyper-links and forms a huge web. The link analysis idea has been applied in various directed networks, such as the citation network of scientific papers and journals [9, 10, 11], the semantic network for machine translation [12, 13]. The World Wide Web is similar to the trade network. For this reason, we can introduce those methods on web search into the world trade ranking of countries on the world trade network.

PageRank [6] is one of the most important algorithms used in the ranking system of Google, the most popular search engine in the world. In PageRank algorithm, each web page v_i has one score π_i and can be modeled mathematically in terms of the random surfer model [14]

$$\pi_i = \alpha \sum_{j: e_{ji} \in E} \pi_j / o_j + (1 - \alpha)1/N . \quad (1)$$

for a network with N web pages. Here, $o_j = |\{k : e_{jk} \in E\}|$ presents the out-degree of the web page v_j . α is a free parameter between 0 and 1, that controls the proportion of time the random surfer follows the hyperlinks. On the other side, $1 - \alpha$ implies the proportion of time to jump at random with a uniform distribution.

收稿日期: 2014-03-31

作者简介: 蒋春恒(1987-), 男, 汉族, 安徽人, 硕士研究生, 研究方向: 数据挖掘, 电子邮箱: chiangchunheng@my.swjtu.edu.cn; 林文斌(1970-), 男, 教授, 博士生导师; 研究方向: 高性能并行计算、数据挖掘、搜索引擎。

基金项目: New Century Excellent Talents in University (No. NCET-10-0702).

Another link analysis method called HITS [7] is invented at the same time. In HITS, each web page v_i has both an authority score x_i and a hub score y_i . The intuition is that a good authority is pointed to by many good hubs, and a good hub points to many good authorities. This mutually reinforcing relationship can be represented in the following form

$$x'_i = \sum_{j: E_{ji} \in E} y_j, \quad y'_i = \sum_{j: E_{ij} \in E} x_j, \quad x_i = x'_i / \|x'_i\|_2, \quad y_i = y'_i / \|y'_i\|_2. \quad (2)$$

The goal of this paper is to customize the two link analysis based methods PageRank and HITS for the world trade ranking of countries. The rest of this paper is organized as follows. In Section 2, we use link analysis to explore the properties of world trade network and develop one new ranking method for countries involved in the network. Then, we conduct empirical experiments to evaluate the proposed method and compare it with other ranking methods in Section 3. Finally, Section 4 arrives at a conclusion of the paper and makes suggestions for further research.

2 A New Ranking Method: TradeRank

Let $G = (V, E)$ be a graph for the world trade network, where V is the set of vertices or nodes. In our case, the nodes are the countries involving in the world trading. What is more, we assume that the number of nations be $N = |V|$. E presents the set of directed edges connecting countries. Suppose there is a connected path from the country $v_i \in V$ to the country $v_j \in V$, then $e_{ij} \in E$. That is to say, these two countries have trading interactions. $X = (x_{ij})_{N \times N}$ be one non-negative matrix representing the export in terms of US dollar. The diagonal elements of X are all zero, and each off-diagonal element x_{ij} is also zero if $e_{ij} \notin E$. Given two countries A and B , since the export from A to B equals to the import from B to A . Therefore, the total import I_i , as well as the total export O_i of v_i can be derived from the export matrix X , and written as

$$I_i = \sum_{j=1}^N x_{ji}, \quad O_i = \sum_{j=1}^N x_{ij}. \quad (3)$$

The two link analysis based methods PageRank and HITS, are suitable to measure the citation relationships between web pages, sites or scientific articles. However, we cannot directly apply them to measure the world trade network for the following reasons. (1) The citation network for both PageRank and HITS are binary, on the other side, the world trade network provides additional information on the links with trade flow. (2) In general, the size of the citation network is much bigger, while the size of the trade network is bound by the number of countries in the world. (3) Furthermore, no country would like to trade around in random, and the random walk theory does not suitable to deal with the bilateral trading. As for HITS, it is one proper tool to model the economic interdependence among countries. We hope to develop one composite measure with both the authority and the hub properties.

There are two kinds of countries in global trade community, the consumer and the producer countries. As their names suggest, the consumer countries such as the United States have thriving demands on products; the producer countries (e.g. China) are expected to supply considerable products. They are the two sides of one coin and most countries import and export products simultaneously. Viewed from their structures in the network, it is easy to distinguish the two classes of countries, and the primary consumers present as authorities, while the producers paly as hubs.

Given a country $v_i \in V$, there are various factors that can be utilized to measure its strength s_i in trade. Among them, only four primary factors: the outflows, the inflows, the destination countries of the outflows, and the source countries of the inflows are selected to give the definition of the strength score as follows:

$$s_i = \alpha \sum_{j=1}^N \frac{x_{ij} s_j}{I_j} + (1 - \alpha) \sum_{j=1}^N \frac{x_{ji} s_j}{O_j}. \quad (4)$$

where the first term presents the contribution of the outflow to v_j , and the second term describes the influence of the inflow from v_j . We note that, the percentages of the flows are considerably more important for the strength measure. Hence, both terms are given in the form of percentage. For the country v_i , its export to v_j is one part of the import of v_j , then the outflow x_{ij} is divided by the total import I_j of v_j . On the other side, the inflow x_{ji} is divided by the total export O_j of v_j . The free parameter $\alpha \in [0, 1]$ controls the weight of the export. When $\alpha = 1$, the model is completely built upon the export without consideration of the import factor. When $\alpha = 0$, the import factor are ignored. The proposed method brings both the import and export factors into consideration, also the method

attempts to makes full use of the link structure of the trade network, so that to improve the current trade ranking methods.

With the definitions of two diagonal matrices $\Lambda_I = \Lambda(I_1, I_2, \dots, I_N)$ and $\Lambda_O = \Lambda(O_1, O_2, \dots, O_N)$, the model then can be written in matrix form

$$S = [\alpha X \Lambda_I^{-1} + (1 - \alpha) X^T \Lambda_O^{-1}] S. \tag{5}$$

Actually, there are few countries do not have any export flow or import flow, producing diagonal zero elements in Λ_I and Λ_O . To make all diagonal matrices invertible, all the diagonal zero elements are replaced with one. It's easy to prove that the matrix $\alpha X \Lambda_I^{-1} + (1 - \alpha) X^T \Lambda_O^{-1}$ is one stochastic matrix. The Perron-Frobenius theorem [15] ensures that the matrix has one unique eigenvector with the corresponding eigenvalue 1 be the largest one. As a result, the efficient power method will be adopted for the computation of the strength scores.

3 Experimental Evaluations

Unlike other kinds of ranking problems, there is yet no benchmark standard to enable comparison of different methods. Various ranking methods have been proposed to evaluate trading countries from different aspects of trading. We would like to conduct empirical experiments using real data set to investigate the reason behind our experimental results.

Data Set. The United Nations Conference on Trade and Development (UNCTADStat) is a free online retrieval system based on the United Nations Comtrade Database [16]. The aggregated trade data for all commodities over the time period 1995 -- 2012 are retrieved to build a reduced world trade matrix of size 210. The network is made up of 210 vertices and more than twenty thousands of trade links. To investigate the connectivity of the trading countries, we consider the density of the network, namely the ratio between the number of connected edges and the number of maximum links possible, a measure can be viewed as the probability of taking two countries at random with trade interaction. As for our reduced network, the density has increased from 0.474 in 1995 to the highest 0.621 in 2010, then falls to 0.600 in 2012 as shown in Table 1.

Table 1. Density of the world trade network over the time period 1995 – 2012.

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003
Density	0.474	0.491	0.510	0.517	0.526	0.550	0.564	0.573	0.580
Year	2004	2005	2006	2007	2008	2009	2010	2011	2012
Density	0.593	0.598	0.608	0.613	0.619	0.621	0.620	0.618	0.600

Experiments Results. We carry out experiments to produce various ranking lists of countries using factors such import, export, and the proposed composite measure in TradeRank algorithm. Moreover, we also observe the effect of the free parameter α on the change of the ranking results via two values 0.5 and 1.0 for α . Finally, these ranking results are further analyzed in pairwise fashion.

Table 2. The ranking lists with respect to the import (IR), the export (ER), the TradeRank with $\alpha = 0.5$ and $\alpha = 1.0$ over three representative years 2003, 2008, 2012. Only the top 10 countries (areas) are presented in the ranking lists. Countries are presented with the ISO3 country codes.

Position	1	2	3	4	5	6	7	8	9	10	
2003	IR	USA	DEU	GBR	FRA	CHN	JPN	ITA	NLD	CAN	HKG
	ER	DEU	USA	JPN	CHN	FRA	GBR	ITA	CAN	NLD	BEL
	TR(0.5)	NZL	DEU	EST	USA	IDN	AUS	GBR	POL	IND	ITA
	TR(1.0)	DEU	USA	JPN	CHN	FRA	GBR	ITA	NLD	BEL	HKG
2008	IR	USA	DEU	CHN	FRA	JPN	GBR	NLD	ITA	BEL	HKG
	ER	DEU	CHN	USA	JPN	NLD	FRA	ITA	BEL	RUS	GBR
	TR(0.5)	USA	DEU	CHN	FRA	JPN	GBR	NLD	ITA	BEL	HKG
	TR(1.0)	USA	DEU	CHN	FRA	JPN	GBR	NLD	ITA	BEL	HKG
2012	IR	USA	CHN	DEU	JPN	FRA	GBR	NLD	HKG	KOR	ITA
	ER	CHN	USA	DEU	JPN	NLD	FRA	KOR	ITA	HKG	GBR
	TR(0.5)	USA	CHN	DEU	JPN	FRA	GBR	NLD	HKG	KOR	ITA
	TR(1.0)	USA	CHN	DEU	JPN	FRA	GBR	NLD	HKG	KOR	ITA

The ranking results in terms of import, export and TradeRank for three representative years 2003, 2008, and 2012 are summarized in Table 2. Only the top 10 countries or areas in the ranking lists are shown for ease of visualization. We observe that the ranking results with respect to the import and the export are different from each other at majority of the top positions. The United States is always the largest importer country in the world. Germany is the largest exporter in 2003 and 2008, and then the position is replaced by China since 2009. Meanwhile, China also grows to be the second largest importer, the position once occupied by Germany. The ranking lists produced by the proposed

TradeRank indicates that the free parameter α seems be insensitive to the network with higher density, both TR(0.5) and TR(1.0) achieve the same results as that obtained by the import-based ranker (IR). In 2003, TR(0.5) provides an interesting ranking results in 2003, while TR(1.0) ranks consistent with the export-based ranker (ER). For example, two top-ranked countries New Zealand, Estonia are exceptions. Although some countries do not have significant total import or export, they may be ranked top of the list for their balanced trading with other countries. The proposed method provides insights about potential gains of a balanced trading relationship.

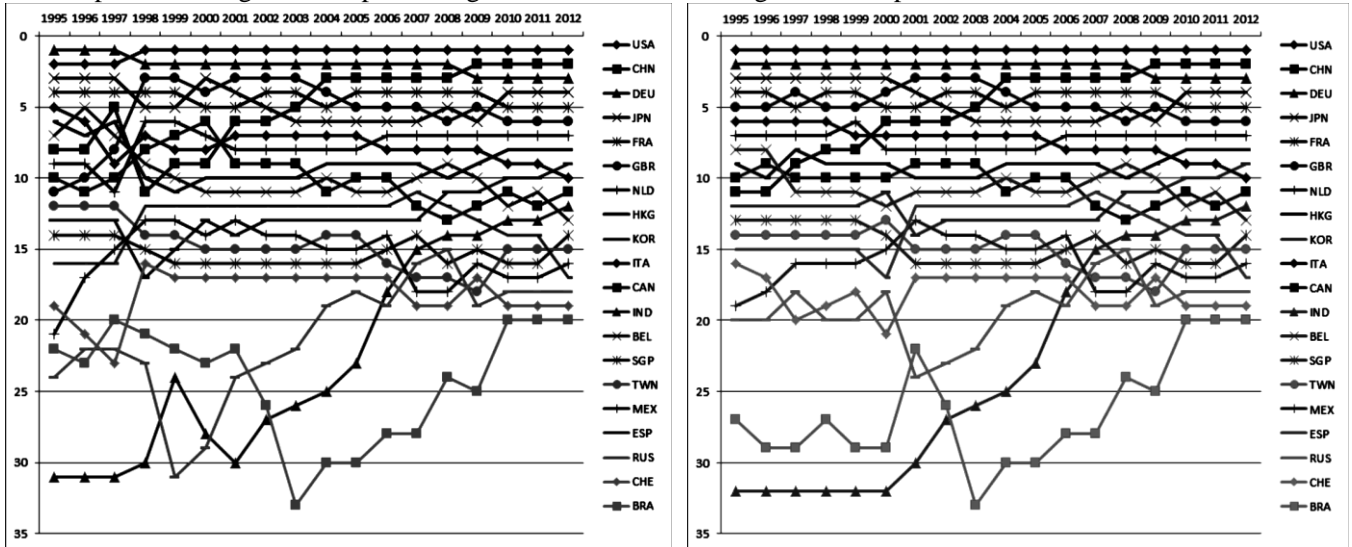


Fig 1. The Trade Ranking Lists of Twenty Countries (Areas) over the Time Period 1995 – 2012 (left $\alpha = 0.5$, right $\alpha = 1.0$).

Furthermore, we report the advancement of trade for all countries that are ranked top 20 in terms of trade strength. As depicted in Fig. 1, the most important countries in the current global trading are U.S., China, Germany, Japan and France. During the past decades, two Asian emerging economies India and China show apparent increasing in trade. For example, China ranks around tenth in 1990s, and then jumps to second position in 2009. Similar patterns could be observed for India, which climbs from thirtieth in 1995 to tenth in 2012. On the other side, some developed countries, such as Japan and Korea are decreasing in terms of trade strength. Also, we explore the relationship of the number of trade partner countries and the trade strength, and discover one interesting pattern that the trade strength of one country is positively correlated with the number of its trade partners as shown in Fig. 2. On the average, those best ranked 50 countries have 190 trade partners. Conversely, the average number falls below 90 for those lowest ranked 50 countries. It implies that the countries that actively develop trade partnership with other countries are likely to gain most from the strongly connected global trade network.

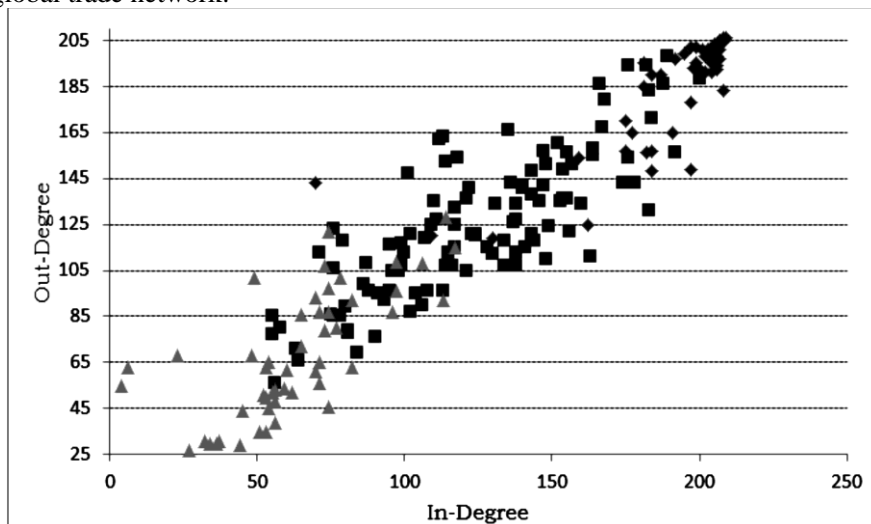


Fig 2. The relationship between the number of trade partners and the trade strength scores of the best ranked 50 countries (blue), the lowest ranked 50 countries (green) and the other middle-ranked countries (red).

4 Conclusion

In this paper we present a link analysis based ranking algorithm called TradeRank to bring orders to countries based on the import, export factors. The proposed algorithm is also empirically studied on the real international trade data set from 1995 to 2012. The experimental results demonstrate that the proposed approach generates reasonable ranking results within the world trade network. We will further investigate the critical factors influencing the trade strength for ranking, as well as the effect of the economic interdependence among countries [18] with some special links in the network removed.

References

- [1] PAUL R KRUGMAN. International economics: Theory and policy[M]. Pearson Education India, 2009.
- [2] FOLKE HILGERDT. The case for multilateral trade [J]. The American Economic Review[J]. 1943, 33(2): 393-407.
- [3] GIORGIO FAGIOLO, JAVIER REYES, STEFANO SCHIAVO. World-trade web: Topological properties, dynamics, and evolution [J]. Physical Review E, 2009, 79(3): 036115.
- [4] JIANKUI HE, MICHAEL W DEEM. Structure and response in the world trade network [J]. Physical review letters, 2010, 105(19): 198701.
- [5] LUCA DE BENEDICTIS, LUCIA TAJOLI. The world trade network [J]. The World Economy, 2011, 34(8):1417-1454.
- [6] SERGEY BRIN, LAWRENCE PAGE. The anatomy of a large-scale hypertextual web search engine [J]. Computer networks and ISDN systems, 1998, 30(1):107-117.
- [7] JON M KLEINBERG. Authoritative sources in a hyperlinked environment [J]. Journal of the ACM (JACM), 1999, 46(5):604-632.
- [8] ZOLTÁN GYÖNGYI, HECTOR GARCIA-MOLINA, JAN PEDERSEN. Combating web spam with trustrank [C]. Proceedings of the Thirtieth international conference on Very large data bases-Volume 30. VLDB Endowment, 2004: 576-587.
- [9] P Yu and H Van de Sompel. Networks of scientific papers [J]. Science, 1965, 169:510-515.
- [10] SIDNEY REDNER. Citation statistics from more than a century of physical review [EB/OL]. (2004-08-06) [2014-03-10]. <http://arxiv.org/abs/physics/0407137>.
- [11] NAN MA, JIANCHENG GUAN, YI ZHAO. Bringing pagerank to the citation analysis [J]. Information Processing & Management, 2008, 44(2):800-810.
- [12] RADA MIHALCEA, PAUL TARAU, ELIZABETH FIGA. Pagerank on semantic networks, with application to word sense disambiguation [C]. Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, 2004: 1126-1129.
- [13] ENEKO AGIRRE, AITOR SOROA. Personalizing pagerank for word sense disambiguation [C]. Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2009: 33-41.
- [14] AMY N LANGVILLE, CARL D MEYER. Google's PageRank and beyond: The science of search engine rankings [M]. Princeton University Press, 2011.
- [15] CARL D MEYER. Matrix analysis, applied linear algebra [M], volume 2. SIAM, 2000.
- [16] UN COMTRADE. United nations commodity trade statistics database [EB/OL]. (2013-10-02) [2014-3-10]. <http://unctadstat.unctad.org/ReportFolders/reportFolders.aspx>.
- [17] KENNETH N WALTZ. Structural realism after the cold war [J]. International security, 2000, 25(1):5-41.

一种基于链接分析的国际贸易排名新方法

蒋春恒¹, 林文斌²

(1.西南交通大学数学学院, 四川 成都 610031; 2.西南交通大学物理学院、数学学院, 四川 成都 610031)

摘要: 国际贸易需要多元化的度量与分析工具. 提出一种新的国际贸易排名方法, 兼顾贸易进口与出口的影响, 对参与国际贸易的国家进行排名. 在国际贸易网络的框架下, 使用链接分析评估各个国家在全球多边贸易下的影响力, 为每个国家赋予权值并进行排名. 真实贸易数据上的实证分析表明本文提出的方法提供了一种度量与分析国际贸易的有效工具.

关键词: 世界贸易网络; 排名算法; 链接分析; PageRank, HITS