Visualizing the European Trade Graph

Carlos Castillo[1] and Ricardo Baeza-Yates[2]

[1] Università di Roma "La Sapienza"
castillo@dis.uniroma1.it
[2] ICREA Professor, Universitat Pompeu Fabra
ricardo.baeza@upf.edu

Abstract. We show that the Kendall's τ metric over the ranked lists of commercial partners between European countries, induces a distance function that is very useful for visualizing the European trade graph, and that recovers geographical relationships between countries.

1 Commercial Trade Data

In this paper we consider a graph representing commercial trade. In the *trade graph*, each node is a country and each (weighted) arc represents the amount of commercial trade between two countries. We consider an undirected version, in which we sum both exports and imports. The data we analyze was obtained from the United Nations Statistics Division (Commodity Trade Database COMTRADE, http://unstats.un.org/unsd/comtrade/).

We use a subset of the global trade graph which corresponds to the trade between European countries. In Table 1 we show the list of countries and top-level domain codes; we have partitioned the countries into four geographical zones which we represent with shapes and colors in the following pages.

Table 1. List of countries.

Zone	Country	Domain	Zone	Country	Domain
North	Denmark	dk	South	Andorra	ad
	Finland	fi		France	fr
	Iceland	is		Greece	gr
	Ireland	ie		Italy	it
	Norway	no		Malta	mt
	Sweden	se		Portugal	pt
	U.K.	uk		Spain	es
Center/West	Austria	at	Center/East	Bulgaria	bg
	Belgium	be		Czech Rep.	cz
	Germany	de		Estonia	ee
	Luxembourg	lu		Hungary	hu
	Netherlands	nl		Latvia	lv
	Switzerland	ch		Lithuania	lt
				Poland	pl
				Romania	ro
				Slovakia	sk
				Slovenia	si

2 Graphs of Trade Volume

We used the program `neato` of the Graphviz (http://www.graphviz.org/) package, which uses a spring model to obtain a visual representation of a graph. In the next figures, a high level of trade is represented by a shorter, darker edge, while a low level of trade is represented by a longer, lighter edge. The result of using the total amount traded as edge weight is shown in Figure 1.

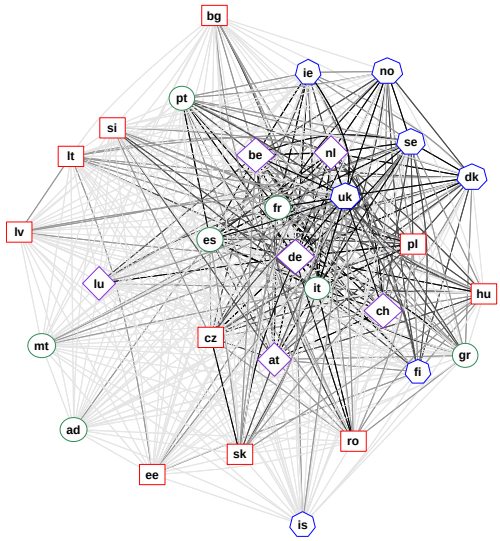

Fig. 1. Graph using the total amount of trade.

As expected, the bigger economies play a more central role in the layout of the graph. Chen [1] proposed to use spanning trees to visualize complex graphs. We obtained the maximal spanning tree (Figure 2), and now it is clear that a star-like structure around Germany dominates the graph, and that the U.K. and France act as "bridges" in this graph. The usage of spanning trees, while losing information, is a good method for better visualizing this graph.

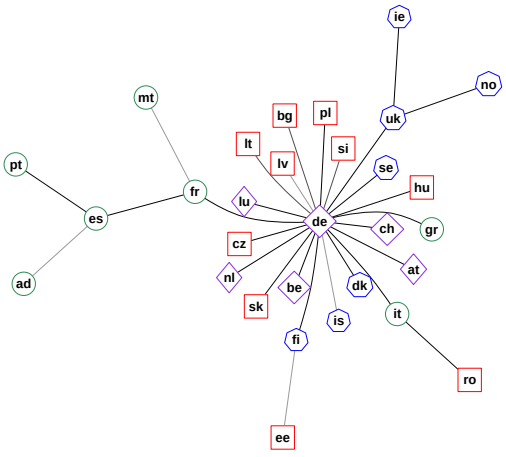

Fig. 2. Maximal spanning tree, using the total trade as cost for the edges.

3 Similarity of Commercial Partners

To factor the effect of big economies out, we propose to use a different type of distance in the trade web: *the similarity between the ranked lists of partners of two countries*. For this similarity calculation, the actual amount traded does not matter, only the order of the commercial partners. We used the Kendall's τ metric for similarity, and both the lists of imports and exports separately. A comparison of this metric with using simply the total amount traded is shown in Figure 3.

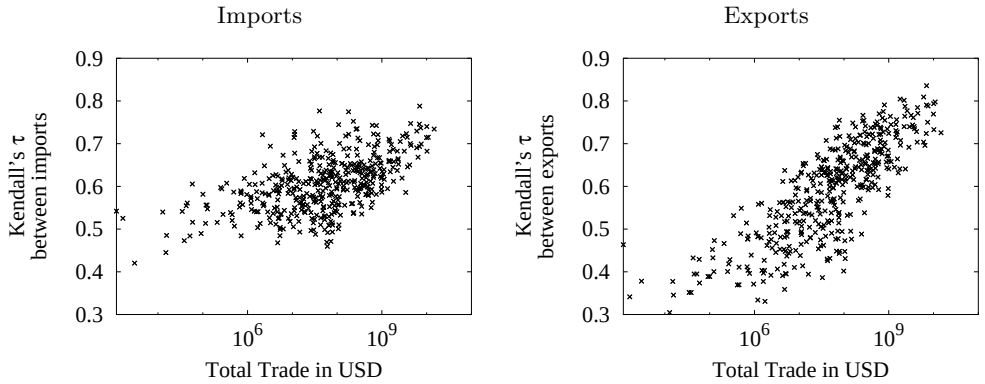

Fig. 3. Kendall's τ of the partners of two countries, versus the total amount traded by those countries; each dot represents a pair of countries.

The ordering of the destinations of exports is more diverse in this sample than for the imports. Also, both the correlations of imports and exports are a different measure than the total amount traded. Using the τ correlation of trade partners as edge weights, we generated Figures 4 and 5; this measure tends to cluster countries in the same geographical zone together.

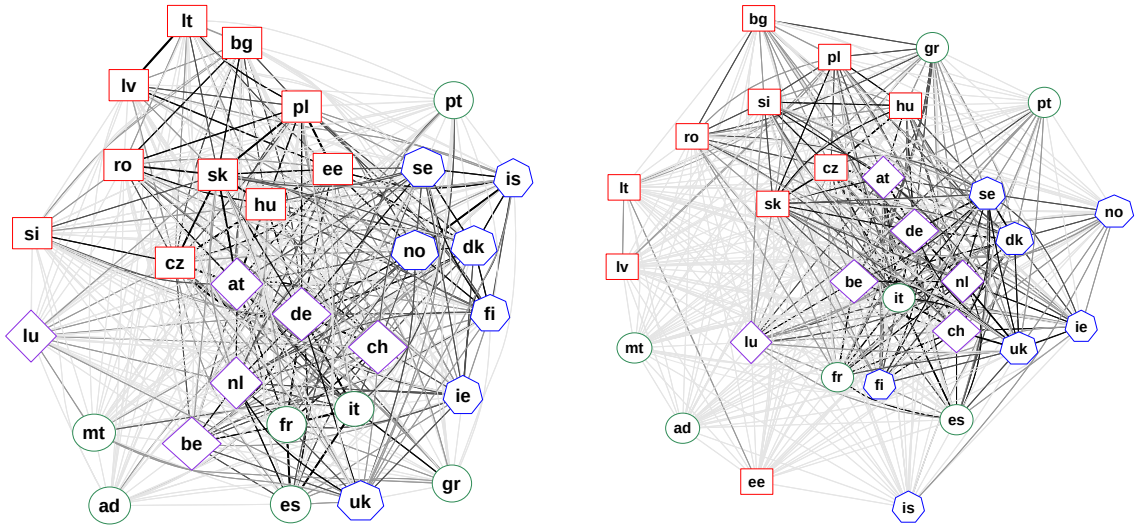

Fig. 4. Graphs using τ correlation of imports (left) and exports (rights).

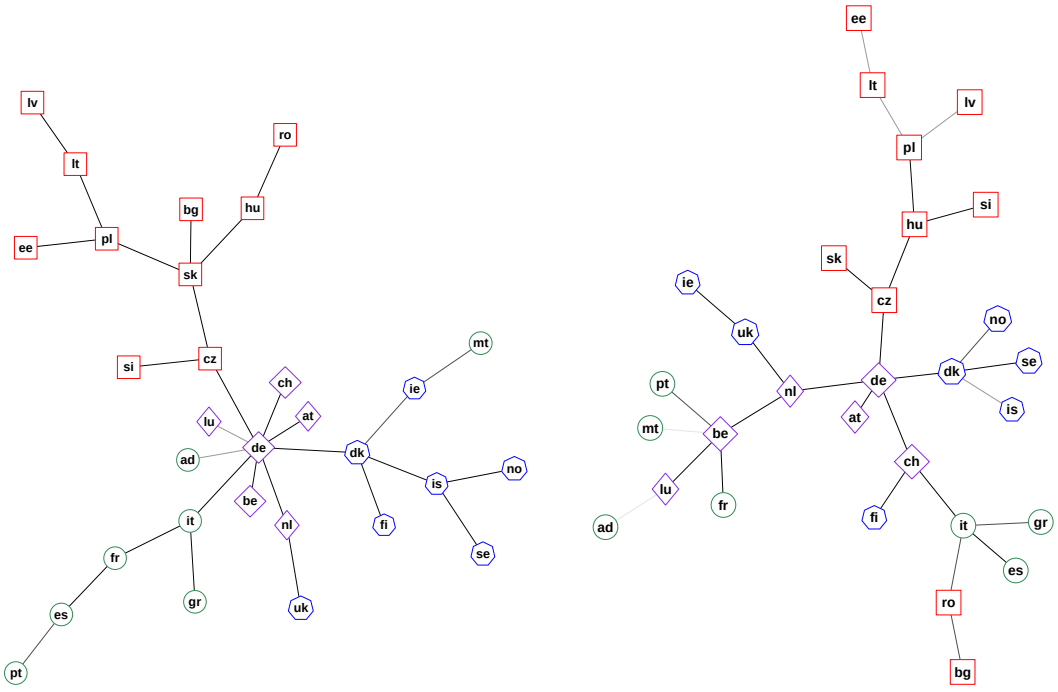

Fig. 5. Maximal spanning trees of correlation of imports (left) and exports (rights). The correlation of imports recover most of the geographical relationships between countries.

In the maximal spanning tree, we can see that the similarity of import partners is better to partition the graph into geographical zones; this can be explained because this measure is also more independent from the total amount traded.

4 Conclusions

Previously, [2] used geographical distance for the layout of the graph before adding trade data; other attempts of visualization have used only one country [3] or report a failure to draw an intelligible graph for more than 10 countries [4]. We have shown how by using the similarity between the ranked lists of trade partners, both with spring models and spanning trees, we can recover several geographical relationships between 30 European countries using only trade data.

References

1. Chen, C.: Visualising semantic spaces and author co-citation networks in digital libraries. Inf. Process. Manage. **35**(3) (1999) 401–420
2. Krempel, L., Plümper, T.: Exploring the dynamics of international trade by combining the comparative advantages of multivariate statistics and network vizualizations. Journal Of Social Structure **4**(1) (2003)
3. Hagman, J.: Visualizing quantitative aspects of communicative and commercial interaction. In: Proceedings of the Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), Minas Gerais, Brazil (1996) 79–86
4. Batagelj, V., Mrvar, A., Zaversnik, M.: Partitioning approach to visualization of large graphs. In: GD '99: Proceedings of the 7th International Symposium on Graph Drawing, London, UK, Springer-Verlag (1999) 90–97