The strength of international scientific ties: a novel analysis of inter Country coautorship

Ugo Finardi
Ph.D. - University of Torino – Italy.
CNR-IRCrES, National Research Council of Italy, Research Institute on Sustainable Economic Growth
Official web: https://www.cnr.it/people/ugo.finardi ; http://www.ircres.cnr.it/index.php/it/organizzazione/staff/19-finardi-ugo
https://orcid.org/0000-0002-2568-8185
E-mail: ugo.finardi@ircres.cnr.it

Andrea Buratti
Ph.D. - University of Urbino – Italy. University of Urbino, PhD Research Fellow, Department of Economics, Society, Politics
https://orcid.org/0000-0001-5748-3582
E-mail: andrea.buratti@uniurb.it

Data de submissão: 18/06/2020. Data de aceite: 12/04/2021. Data de publicação: 10/12/2021.

ABSTRACT

The aim of this work is to highlight the strongest inter-country scientific collaborations and the factors that drive them, also in order to offer policy suggestions. The experimental activity performed in this context exploits a less common methodology, preferred to other ones due to its ability to better resolve collaborations. It calculates in fact a powerful indicator, the Probabilistic Affinity Index, starting from the internationally coauthored scientific products of the 100 most scientifically productive Countries. The Probabilistic Affinity Index is able to measure the strength of a collaboration without being influenced by the scientific production of a Country. Once assessed the strength of collaboration, networks of the strongest ones are built using Network Analysis instruments. While results substantially confirm most of the past findings on factors driving scientific collaboration, they also show previously unseen strong collaboration paths existing between Countries. At the end of the paper policy suggestions are drawn.

Keywords: International cooperation. Science policy. Probabilistic affinity index. International network. Cooperation factors.

A força dos laços científicos internacionais: uma nova análise da coautoria entre os países

RESUMO

O objetivo deste trabalho é destacar as colaborações científicas internacionais mais fortes e os fatores que as impulsionam, também a fim de oferecer sugestões de políticas. A atividade experimental realizada neste contexto explora uma metodologia menos comum, preferida a outras devido à sua capacidade de resolver melhor as colaborações. Calcula, de fato, um poderoso indicador, o Índice de Afinidade Probabilística, a partir dos produtos científicos com coautoria internacional dos 100 países mais produtivos cientificamente. O Índice de Afinidade Probabilística é capaz de medir a força de uma colaboração sem ser influenciado pela produção científica de um País. Uma vez avaliada a força da colaboração, as redes dos mais fortes são construídas usando instrumentos de Análise de Rede. Embora os resultados confirmem substancialmente a maioria das descobertas anteriores sobre os fatores que impulsionam a colaboração científica, eles também mostram fortes caminhos de colaboração nunca antes vistos entre os países. No final do artigo, são apresentadas sugestões de políticas.

Palavras-chave: *Cooperação internacional. Política científica. Índice de afinidade probabilística. Rede internacional. Fatores de cooperação.*

La fuerza de los lazos científicos internacionales: un análisis novedoso de la coautoría entre países

RESUMEN

El objetivo de este trabajo es destacar las colaboraciones científicas internacionales más sólidas y los factores que las impulsan, también para ofrecer sugerencias de políticas. La actividad experimental realizada en este contexto explota una metodología menos común, preferida a otras por su capacidad para resolver mejor las colaboraciones. De hecho, calcula un poderoso indicador, el Índice de Afinidad Probabilística, a partir de los productos científicos de coautoría internacional de los 100 países científicamente más productivos. El Índice de Afinidad Probabilística es capaz de medir la fuerza de una colaboración sin estar influenciado por la producción científica de un País. Una vez evaluada la solidez de la colaboración, las redes de las más sólidas se construyen utilizando instrumentos de análisis de redes. Si bien los resultados confirman sustancialmente la mayoría de los hallazgos anteriores sobre los factores que impulsan la colaboración científica, también muestran rutas sólidas de colaboración no vistas que existen entre países. Al final del documento se extraen sugerencias de política.

Palabras clave*: Cooperación internacional. Política científica. Índice probabilístico de afinidad. Red internacional. Factores de cooperación.*

INTRODUCTION

The study of international scientific collaboration is a relevant topic of bibliometrics and of evaluation of research since their inception (BARRIOS *et al*., 2019). The wide number of scientific works dealing with this topic witnesses its importance. The complexity of the paths followed by such collaboration phenomena makes difficult their investigation (ZANOTTO *et al*., 2016). As a result many features of international scientific collaboration are still underexplored or totally unexplored, and thus need further deepening, or methodological improvement (see for instance González-Alcaide *et al*., 2017).

Aim of the present work is to contribute to this research area. To this end it adopts an original and improved methodology in order to define with greater precision the strength of the scientific collaborations measured by coauthorship. This methodology is not influenced by the total internal scientific production of the considered Countries.

This paper offers an analysis that might be able to support policy decisions aiming at increasing international scientific collaboration of a Country. As UNESCO (2015, p. 75) report states, "International scientific collaboration is obviously invaluable for tackling global scientific issues". Thus an analysis of the most relevant and strongest inter-Country scientific collaborations, as well as of the factors that enhance them, may prove relevant for those wishing to build policies aimed at fostering research and improve its results.

In order to perform the above-mentioned analysis this work exploits a dataset built starting from numbers of scientific products (articles, congress contributions, book chapters etc.) coauthored by scientists of different Countries. These data are used to calculate a specific index, the Probabilistic Affinity Index. This index has the relevant advantage of being independent from absolute values of scientific production of the different Countries. In this it is different from other similar indexes often used to describe scientific collaboration. In fact, as also the theoretical framework will show, the use of absolute values of coauthored scientific products is not a truly reliable index of cooperation strength. This because they are, under a general point of view, dependent from the total scientific production of a Country. Many research works show that also some indexes used to this end are, at least in part, dependent from the total scientific production of a Country. On the contrary the index used in this work is able to offer a measure of the strength of the scientific collaborations without being influenced by the total scientific strength production of a Country. The index is able to disentangle also collaborations that present a low absolute value of common scientific products, which are indeed a high fraction of a Country's scientific production. In other words, a slightly productive Country might, in principle, present a very high fraction of products co-authored with another slightly productive Country. This would result in a very strong collaborative connection between the two, while a high number of products co-authored with a highly productive Country (but involved in a very high number of collaborations with other Countries) would result in a weaker connection.

The research questions at the basis of the present work are then: which are the strongest inter-country collaborations and how are they networked? How are they influenced by factors such as, for instance, geography, culture or history? This article extends and generalizes the results of Author (2015) and of Author (2016). Obtained results show the networked nature of the strongest international scientific collaborations, and offers an interpretation of the factors that drive them. Social network analysis instruments presents graphically the networking of the strongest links existing between collaborating Countries. Strongly connected Countries are scientifically less central, while strong actors present only weak connections with a higher number of Countries.

The rest of this article is structured as follows. Section 2 presents a literature overview and a theoretical framework, centred on the topics of the index used in the experimental activity and of the analysis of the factors driving international collaboration. Section 3 presents the methodology of the paper, while section 4 reports the results. Finally, sections 5 and 6 respectively discuss results and present conclusions, as well as policy indications.

LITERATURE OVERVIEW AND THEORETICAL FRAMEWORK

The theoretical framing of the present work involves the analysis of past literature related to two different topics. The first one is methodological, while the second one is related to the studies on international scientific collaboration.

The first topic to consider here is relative to the index used in the experimental activity. In particular this article exploits an index of the family of the probabilistic indexes, the Probability Affinity Index (PAI from now on). This index has been chosen in alternative to other well know indexes, in particular to Jaccard index and to Salton's index (or Salton's cosine), that have often been employed in the analysis of scientific collaboration. This choice is due to the fact that PAI offers advantages with respect to other indexes due to its specific features. In fact all three indexes were used, together with gross count of publications, by Author (2015) in his analysis of international collaborations of the BRICS Countries. This study shows that Jaccard and Salton's indexes present some degree of dependence from absolute values of publications, while PAI is size independent. In a following article Author (2016) did confirm this result. Thus the use of PAI seemed more reliable in order to describe the complex network of international scientific collaboration.

The two cited works haven't been the first ones to use PAI as an indicator to study international collaboration. A past relevant example is the work of Luukkonen *et al*. (1992) who present a case study performed on a wide sample of countries and exploiting a measure of the family of the "probabilistic indexes". This is one of the earliest examples of the use of probabilistic indexes to study international collaboration.

In more recent years Zitt *et al.* (2000) used probabilistic indexes, deeply inspiring the use of PAI in the present study. In fact Zitt *et al.* (2000) affirm that "The PAI removes the effect of size, at the expense of large significance intervals for the index when values of marginal are small. A small Country concentrating its collaboration with a few partners will record very high ratios. The PAI is a convenient means for highlighting small specific relationships" (p. 633).

Also Mattsson *et al*. (2010) use PAI to analyse the network of European collaboration. To this end they use both co-publication and co-participation to the Framework Programmes for research of the European Union. Their findings show that researchers tend to engage in bilateral collaborations rather than in multilateral ones, preferring extra-European collaborations to European ones, and that geographical proximity performs an influential role of geographical proximity.

Another relevant topic is related to the different paths of international scientific collaboration. Yet Luukkonen *et al.* (1992) discuss the fact that coauthorship is only one (though the most relevant) of the paths followed by (international) scientific collaboration. In fact authors state that "These represent only some of the possible indicators of collaboration. […] Nevertheless, we assume that in most cases coauthorship indicates a fairly active cooperation between the authors" (p. 103, passim). Thus, even if coauthorship may not capture the entire collaboration between two Countries, it can be considered a fair proxy of its strength.

The main topic of this theoretical framework is the discussion of the studies on international scientific collaboration. Besides the review of specific literature we also try to summarize the most studied factors that drive international scientific collaboration.mBeaver and Rosen (1978) in their seminal contribution on the study of scientific collaboration listed the more general reasons that drive researchers to collaborate.

Some early contributions from the 1990s witness both the relevance of Country scientific power in influencing its path of international collaboration and the existence of Country networks. For instance the work of Schubert and Braun (1990), based on data from Science Citation Index database, shows an inverse correlation between the number of internationally coauthored articles and the scientific size (number of published articles) of the Country. Moreover, once discounted the Country size effect, the network structure of international collaboration shows a set of clusters, some of which "probably of historical and/or political origin" (p. 10). Narin *et al.* (1991) obtain similar results in their analysis of intra-EU scientific collaborations. This work is based on a database of articles – retrieved on the Science Citation Index – published between 1977 and 1986. It suggests again that "The magnitude of international coauthorship is only weakly dependent on the scientific size of a country" and "The direction of international coauthorship is heavily dependent on linguistic and historical factors" (p. 323, passim). Finally Kraut *et al.* (1988) highlight the importance of physical proximity at short scale for collaboration.

Also Okubo *et al.* (1992) are among the authors describing the determinants of networking and, more in general, of international collaboration. In fact their work performs an analysis of the collaborations carried on between a wide number of Countries on some scientific fields, using two complementary methods. They end up affirming that "One may add geographical closeness to socio-political factors" (p. 342) as a cause of the links between Countries, and that "It is likely that scientific levels of countries and socio-cultural factors play a major role in constructing similarities in patterns". (p. 343, passim). Also Katz (1994) upholds the relevance of geographical proximity in fostering collaboration in his study of intra-national coauthorship.

Melin and Persson (1996), who dedicate a part of their study to international collaboration, describe similar determinants. Their results are similar to those described in the above revised works. They state in fact that "The dependence on the international scene is proportionately higher for small countries […] also other factors explaining the pattern of country collaboration […] the interactions within the network depend on the geographical distance separating the nodes, cultural, linguistic and political barriers" (p. 373, passim).

Leclerc and Gagné (1994) go even further, as they uphold the trend of continentalization of science, substituting national science. In doing so they affirm that "the historical determining factors […] predominate […] in some cases contrary to economic logic. […] But economic logic continues to gain ground in scientific relations […] despite political conflicts, distance, linguistic barriers, cultural differences and development disparities." (p. 288, passim).

The study of international scientific collaboration and of its drivers continued also in the new century. We underline here the relevant contribution of Okubo and Zitt (2004) who perform an analysis of research collaboration in the European Union starting from a French perspective. While stressing the difference between "top-down" and "bottom-up" drivers, they conclude affirming that, while the main number of publications is still nationally oriented, "geographical proximity plays a conspicuous role in determining the level of exchange" (p. 224). And spatial proximity is directly analysed – though at a Country level – by Ponds *et al.* (2007).

Their study of intra-Netherlands coauthorship shows that "geographical proximity is more important for collaboration between organisations with different institutional backgrounds" suggesting that "geographical proximity is more important in an indirect way by overcoming institutional differences" (p. 441-442, passim).

In more recent years scientific literature dealing with this topic has considered other drivers of cooperation. In fact the authors of another national-based case study (Aksnes *et al.*, 2008) stress the differences between "policies" (top-down) and "personal contacts" (bottom-up) in creating opportunities of international collaboration. Their analysis of the evolution of Norwegian international coauthorship shows an enormous growth, and the fact that "bottom-up" drivers are still the main driving force behind internationalisation. Also Glänzel and Schubert (2005a) present interesting findings in this sense. First of all their work highlights the growth of international collaboration. Then it stresses the role of Country size, of political and economic motivations, of mobility of researchers, as well as of personal affinity.

The process of categorization of factors driving scientific collaboration continues in the thorough literature review performed by Sonnenwald (2007). This work lists a series of factors intervening in the various phases of scientific collaboration (Foundation, Formulation, Sustainment, Conclusion). In specific the factors emerging during foundation are Scientific, Political, Socio-economic, Resource accessibility, Social networks & personal. Nevertheless it must be noted that such factors are relative to collaboration in general and not to the specific analysis of international collaboration.

The last decade witnesses the publication of some case studies on the determinants of international scientific collaboration. For instance Hoekman *et al*. (2010) apply a gravity model to collaboration between European regions. They find the presence of a Country border effect limiting coauthorship, rather than the effects of physical distance. Nevertheless this effect is diminishing with time. Also Acosta *et al*. (2011) perform a region-based case study exploiting a gravity model across all regions of 15 Countries of the EU. The analysis is based on economic differences between regions and finds that the "centre-periphery hypothesis" does not hold, while authors find that "number of publications in the initial year, geographical distance and border contiguity, similarities in scientific specialization between the two regions, and the sharing of similar languages, cultures and policies, also help explain Scientific Collaboration" (p. 72-73).

Also Thijs and Glänzel (2010) study Intra-European networks. In fact they consider, with a peculiar insight, research institutions rather than Countries. Their analysis tries to disentangle collaboration and productivity, and shows that research institutions working on Earth and space sciences present a very high rate of international collaboration.In general collaboration leads to higher visibility, and multidisciplinary institutes are preferred as collaborators.

The works of Hennemann *et al*. (2012) and of Waltman *et al.* (2011) consider again distance as a determinant. The former of the two works finds "a strongly decreasing relation between spatial distance and the probability of co-authoring […]. Moreover, this effect is much more pronounced for collaboration within countries than in cross-country collaboration" (p. 224). The latter of the two instead measures the average distance between co-authors. This distance has grown dramatically from 1980 to 2009. Also Frenken *et al.* (2009) address the effects of distance in collaboration discussing the problem of the "death of distance" from an experimental perspective. Their complex dataset addresses both nation-states and inter-regional collaboration (within and between countries) and is analysed using a gravity model. Findings, in contrast with previous studies, show that geographical distance, as well as national borders, "still hamper research collaboration" (p. 56). Nevertheless this effect is weaker in science than in other fields.

On the contrary Choi (2012) does not find that factors such as affinities in terms of geography, language or economy contribute meaningfully to collaboration. His work is based on the study of coautorship of OECD Countries, assumed as "advanced" ones and performed through the use of several indicators. He shows that "Rather than these extraneous variables to research, scientific expertise or common interest/issues/problems are likely to affect the tie formation" (p. 38).

Some works study cases of scientific collaboration in Africa. This topic, as the rest of the paper will show, is particularly relevant for our work. Mêgnigbêto (2013) addresses collaboration of West Africa Countries. He finds that these Countries present an extremely high rate of international coauthorship, and that they "seemed preferring other African region's local giants or former colonisers or Western countries. […] Overall, language, colonial ties and culture drive collaboration in West Africa science. This study didn't find any effect of geographical close up" (p. 782-782, passim). In a similar way Boshoff (2009) finds relevant levels of "neo-colonial" ties with former colonizing Countries for Central Africa Countries. Finally, Landini *et al*. (2015) perform a study on the collaboration existing between Northern Africa Countries. Their work encompasses both scientific publications and patents from Morocco, Algeria, Tunisia and Egypt. The region has undergone strong internationalisation since the 2000s, while interaction within the region is limited. Egypt is the most active Country and the central node of the Region, as well as international hub. The topic of the work of Hassan *et al*. (2016) on the collaboration among 11 States of the Organization of Islamic Cooperation is partly connected with that of the previous one. Their results show a growth in international collaboration within the network of the 11 Countries, while there is no growth in external collaboration. Our results will confirm some of those ones reported in the above described works.

Two recent works by Cassi *et al*. (2015) and Bergé (2017) offer a specific insight on proximity. The former of the two works presents a specific case study on research activities in the wine industry. Results show the presence in this specific context of relevant effects of both geographical and scientific proximity/distance over collaboration. Besides other specificities of the sector, similarity has a positive impact on trade patterns: economic and knowledge globalisation influence each other. In the latter and more recent paper of the two a gravity model helps disentangling physical proximity from other types of proximity over a European set of regions in the field of Chemistry. In consistency with previous literature, author finds "a significant, negative effect of separation variables, such as geographical distance and national borders. The cognitive distance was also found to have a significant hampering effect on collaboration". Moreover "network proximity alleviates the impeding effects of distance" (p. 22, passim). The work of Jeong *et al*. (2014) offers instead a very peculiar approach. In fact it explores the effects of drivers of collaboration such as input factors (financial and attentional resources, academic excellence), individual and project motivation, passive and active informal communication. The study is based on national (Korean) research project data, and results show that input factors have positive effect on collaboration, that there is a negative effect of the ageing of the researchers and a positive effect of active informal communication.

Finally, Wagner and Leydesdorff (2005) present a systematic review of the drivers of international collaboration in science. They specifically tackle the growth in international collaboration at the end of 20th century. Moreover they discuss, starting from their experimental results, the different theories that shape collaboration patterns. Their results show that most socio-political drivers do not look to be causative, or even correlated, to the rise in international collaboration. In fact the centre-periphery model of international collaboration is not a working model. Instead, a multi-centre model (with many centres collaborating among them) should be considered.

Also, "clustering retains features related to geographical proximity and historical relationships, but these are no longer the strongest features affecting links" (p. 205). The findings of the present work will partially respond to these visions.

Summing up, the second part of this literature overview presents the evolution of the studies on the factors driving international scientific collaboration. Many of these factors, such as the role of history, culture and politics, are highlighted also by less recent literature. Physical or geographical proximity, addressed also as geographical closeness or spatial distance, as well as the effects of Country borders, is considered as a relevant factor of enhancement/detriment of collaboration (not only at international level) though at different degrees. Kraut *et al*. (1988), Melin and Persson (1996) and Okubo and Zitt (2004), as well as Ponds *et al.* (2007) and Frenken *et al.* (2009), all discuss the importance of these effects. On the other side also historical, political, linguistic, cultural, economic features, factors, barriers and motivations are enlisted and discussed as elements that might drive or hinder international collaboration. Schubert and Braun (1990), Narin *et al.* (1991), Melin and Persson (1996), as well as Sonnenwald (2007) and Mêgnigbêto (2013) discuss these factors.

Other factors such as the role of mutual scientific size, level and proximity, as well as that of Country size have been discussed, for instance by Schubert and Braun (1990), Okubo *et al.* (1992), Melin and Persson (1996), Glänzel and Schubert (2005a). Finally other factors are considered by a lesser number of authors, for instance scientific expertise (Choi 2012), neo-colonial ties for Central African Countries (Boshoff 2009), research mobility and personal affinity (GLANZEL; SCHUBERT, 2005a; SONNENWALD, 2007), or even complex mixes of reasons (see for instance Acosta *et al*., 2011 and Jeong *et al*., 2014).

METHODOLOGY

The dataset on coauthored collaborative scientific products exploited in the present work has been built starting from data obtained from the online database Scopus®[1]. The first step in the building of the dataset has been to prepare a list of the 100 Countries that are most scientifically productive in the database. To this end we have considered the total scientific production of each Country in the years going from 1996 to 2014[2]. The alphabetic list of the 100 considered Countries is presented in Appendix for sake of clarity.

Once we have prepared the list of Countries we have obtained from Scopus numbers of scientific articles produced in collaboration. The values have been obtained using the AFFILCOUNTRY (Affiliation Country) search key. For each Country a search has been performed, and numbers of articles presenting also another Country authorship have been recorded. Numbers of articles written in collaboration by scientists of two different Countries have been obtained for each couple of Countries present in our list. Due diligence has been taken in order to obtain reliable values of coauthored scientific products. To this end it must be noted that in some cases values did slightly differ. That is, sometimes searching for Country B in the list of coauthorship for Country A did present slightly different values from those of the search of Country A in the list of Country B. In this case values have been averaged.

The present work considers only bilateral cooperation, as several previous ones did, like for instance Zitt *et al.* (2000) and Choi *et al.* (2015). Thus articles coauthored by scientists from three (or more) Countries are considered in the count in terms of three (or more) bilateral coauthorship, as done for instance by Glänzel and Schubert (2005b).

1 https://www.scopus.com/, data download performed April 2016.

2 1996 has been chosen as Scopus starts collecting a complete set of data starting from this year. 2015 has not been considered due to the fact that data might still have been incomplete at the time of the preparation of the dataset.

Past literature shows that the use of absolute values of coauthored scientific products is problematic and is not a reliable evaluator of the strength of cooperation, also at international level. In fact absolute values of scientific works written in collaboration are directly dependent on the absolute value of the scientific production of a Country. As an example, let us suppose that Country A produces n times the number of scientific products than Country B in the same time period.

Then also the scientific products written by scientists of Country A in collaboration with other Countries will be in principle n times those of Country B. Thus, the use of absolute values of scientific literature is not able *per se* to highlight the strength of international collaboration in terms of intensity with respect to the total scientific production of single Countries. On the other side absolute values of scientific publications can form an apt basis in order to calculate indexes able to properly measure the intensity of collaboration and of networks of collaboration.

As above introduced we have taken due diligence in order to find an index able to overcome the problems deriving from the use of absolute values. A wide number of previous works exploits common indexes such as Salton's index and Jaccard index. Nevertheless such indexes present dependence from absolute values of scientific production, as also Author (2015) shows. For the above introduced reasons this dependence should be avoided in order to obtain reliable indicators of the strength of collaboration. To this end we have chosen to use in this work the Probability Affinity Index PAI. PAI is an index belonging to the family of the probabilistic indexes and, as above cited work shows, is independent from absolute values when calculated using a specific methodology.

PAIs are calculated as follows. All the values of scientific works written in collaboration are arranged in a n×n contingency table (where in our case n=100). Then marginal sums are calculated. Probabilistic affinity indexes are calculated as:

$$PAI_{xy} = \frac{C_{tot}C_{xy}}{C_xC_y}$$

where:

– C_{tot} is the grand total of the contingency table;

– C_{xy} is the value of the collaboration (number of scientific products) between the x^{th} and y^{th} Country;

– C_x and C_y are the marginal sums for each of the two Countries.

Diagonal values of the contingency table are not defined in principle. The most obvious values to insert in the table would be those of the total scientific production of a Country. Nevertheless, as also Author (2015) shows, the values of PAI obtained from such a contingency table are not totally independent from absolute values of scientific production. Thus diagonal values should be calculated in order to obtain PAIs that are independent from the absolute value of the Country's production. This is done via an iterative procedure performed starting from the values present in the contingency table (see de Solla Price, 1981). In the first step of the process all the diagonal values of the contingency table are set as 0. Then marginal values of this contingency table are calculated. The following step is the calculation of diagonal values, calculated as:

$$D_{xy} = \frac{C_xC_y}{C_{tot}}$$

where:

- D_{xy} is the value of diagonal element;

- C_x and C_y are again equal and are the two marginal sums corresponding to the row and column of D_{xy};

- C_{tot} is again the grand total of the table.

After this first step the procedure is iterated inserting in the diagonal of the table the obtained values of D_{xy}. This is repeated n times (n = 18 in the present case) until two reiteration render the same result. The diagonal values obtained in this way are those used to calculate PAIs. All the references, as well as the procedure, are carefully described in Author (2015) and in Author (2016). In particular reference on the use and characteristics of PAI can be found for instance in Zitt *et al*. (2000).

In order to better highlight the structure of collaboration networks, we have exploited the instruments of the Social network analysis, which have a relevant role among the methodologies apt to this scope. This relevance is witnessed by its use in several scientific works (see for instance Garner *et al*., 2012; Eblen *et al*., 2012). In specific, the use of social network analysis is of great help in order to show how the strongest bilateral collaboration links are networked. In this work we use social network analysis in order to better the networks resulting from inter-Country collaboration devised by our study. In order to visualize the networking we have used the Pajek software, v. 4.04. The results are presented in figure 1. The visualization of the network has been improved rescaling in figure 1 all values of PAIs in a 1–10 scale. Thus thickness and grey shade of links are proportional to the values of the PAIs. Obviously, the higher the Index, the darker and thicker the line between two points, and vice versa. Size of nodes is not representative. The graphical representation of the network has been performed using the Kamada and Kawai (1989) layout algorithm[3].

RESULTS

The values of PAI were calculated for each couple of Countries. The symmetric 100 ×100 contingency table contains (obviously) 10,000 values. Out of these ones 100 are diagonal values. Being the table symmetrical the remaining 9,900 are doubled for each couple of Countries (abscissa-ordinate vs. ordinate-abscissa).

This results in a total of 4,950 values of PAI. Such values range from a minimum of 0 (some cases) to a maximum of 55.743, relative to the collaboration between Cameroon and Senegal.

Table 1 presents the subdivision in classes of the values of PAIs. The table shows that most values are extremely low. In fact 3,123 of them (more than 63 % of the total) are below the value of 1. Another group of 1,015 (20.5 % of the total) are between 1 and 2. Only 812 values of PAI (16.4 %) are above 2. Thus the relative strength of most part of international cooperation links is rather low.

Table 1 − subdivision in classes of values of the PAIs.

Range	PAIs
>55	1
50-55	3
45-50	2
40-45	3
35-40	3
30-35	4
25-30	6
20-25	14
15-20	17
10-15	58
9-10	24
8-9	22
7-8	18
6-7	36
5-6	49
4-5	75
3-4	149
2-3	328
1-2	1015
<1	3123
TOTAL	4950

Out of the 812 PAIs presenting a value above 2, only 111 (2.2 % of the total) are above the value of 10. Our analysis of the strong international cooperation patterns is concentrated on this group of values. The value of 10, besides being chosen for sake of simplicity, is also meaningful as it is only slightly below the 20 % of the highest one in the contingency table.

In order to represent the network analysis pattern of the considered PAIs these values have been rescaled as above described. Besides the graphical representation of the collaboration network, values above 10 are also reported in tables 2-9. Values are arranged in different tables according to the different interconnected sub-networks that are described below.

[3] This algorithm produces regularly spaced results, in particular for not very large connected networks (below five hundred vertices).

Figure 1 – Network analysis of the international collaborations with PAI > 10

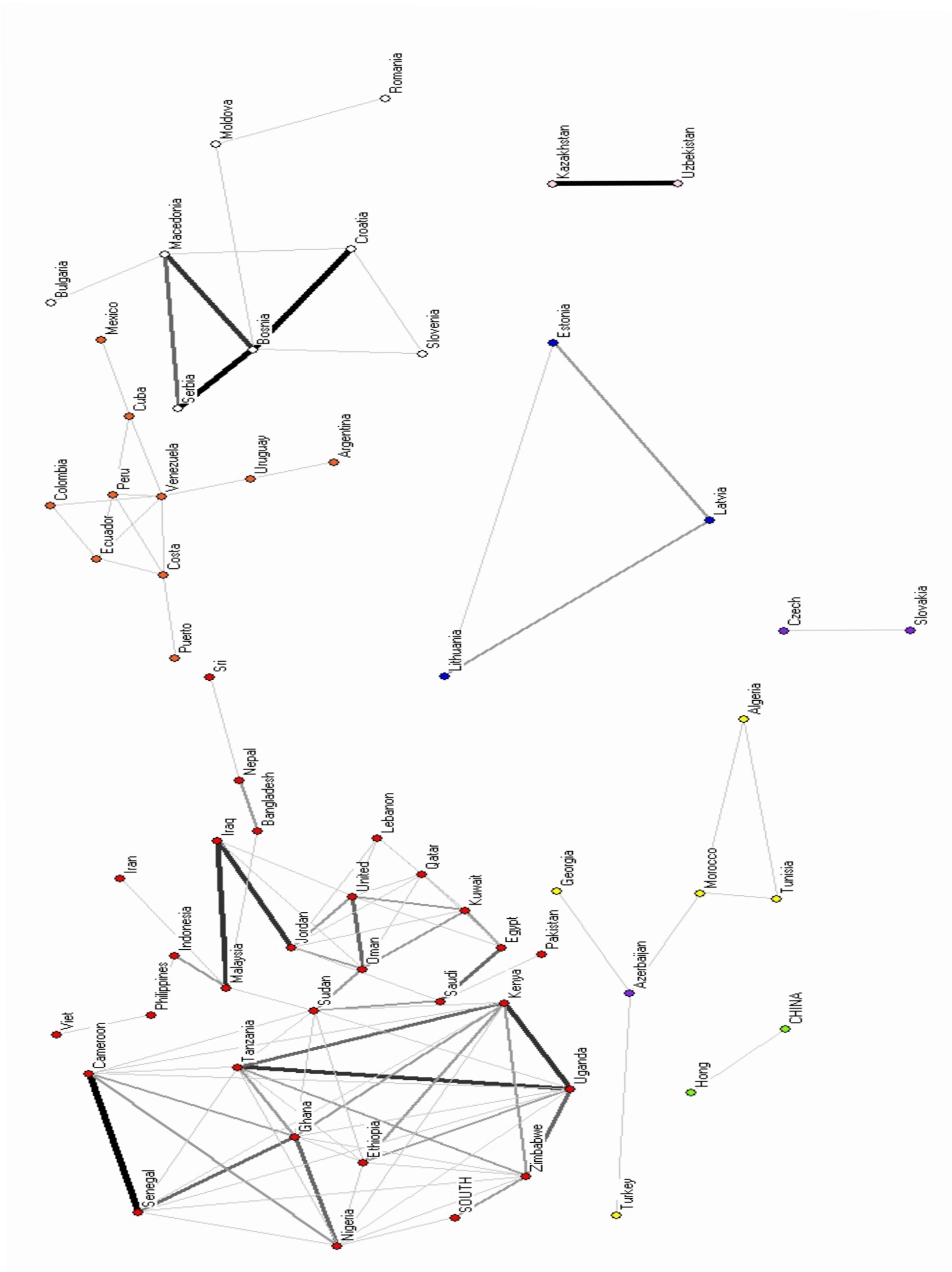

Table 2

	Bangladesh	Cameroon	Egypt	Ethiopia	Ghana	Indonesia	Iran	Iraq	Jordan	Kenya	Kuwait	Lebanon	Malaysia	Nepal	Nigeria	Oman	Pakistan	Philippines	Qatar	S. Arabia	Senegal	S. Africa	Sri Lanka	Sudan	Tanzania	Uganda	U. A. E.	Viet Nam	Zimbabwe
Bangladesh	1.000	1.302	1.440	4.148	4.903	4.637	1.026	3.864	1.022	3.654	4.925	1.307	17.524	20.016	2.862	5.048	5.659	7.908	0.555	2.701	2.769	1.208	6.345	2.091	3.953	2.762	1.789	4.642	3.541
Cameroon	1.302	1.000	1.332	9.630	20.781	4.487	0.369	1.262	0.811	13.996	0.588	1.728	1.038	2.724	27.575	1.483	3.058	2.081	0.513	0.610	55.743	6.909	3.535	10.576	11.610	11.796	0.368	1.990	9.830
Egypt	1.440	1.332	1.000	1.665	1.889	0.998	1.920	7.954	5.546	1.036	22.739	5.540	2.608	0.568	1.564	9.761	3.369	1.114	11.358	39.058	0.789	0.747	2.205	8.801	0.735	1.020	11.612	0.573	0.903
Ethiopia	4.148	9.630	1.665	1.000	14.957	2.702	0.513	1.636	2.189	25.134	0.429	1.260	1.148	8.296	10.892	1.874	2.378	2.764	0.998	1.318	9.134	6.538	6.317	19.556	18.585	21.783	1.013	2.548	14.531
Ghana	4.903	20.781	1.889	14.957	1.000	4.311	0.518	1.860	1.151	20.444	1.733	0.934	1.899	2.141	32.025	2.039	1.402	2.657	0.672	0.633	30.211	6.829	7.505	11.910	22.905	15.836	1.024	3.481	11.869
Indonesia	4.637	4.487	0.998	2.702	4.311	1.000	0.685	2.222	0.952	3.961	1.446	0.913	22.799	5.358	2.078	2.160	1.411	15.159	0.904	0.975	1.945	0.892	4.651	2.443	2.239	3.938	1.295	7.885	2.909
Iran	1.026	0.369	1.920	0.513	0.518	0.685	1.000	5.471	1.787	0.431	3.940	1.498	12.707	0.620	1.517	3.974	3.907	0.784	3.506	2.158	0.253	1.045	2.337	1.525	0.384	0.387	2.530	0.631	0.592
Iraq	3.864	1.262	7.954	1.636	1.860	2.222	5.471	1.000	44.730	0.981	6.005	8.010	46.315	0.285	5.655	16.075	3.563	0.872	6.316	4.793	0.000	0.658	2.150	8.915	1.073	2.625	11.464	0.839	0.000
Jordan	1.022	0.811	5.546	2.189	1.151	0.952	1.787	44.730	1.000	1.541	14.840	10.002	5.577	0.978	1.645	23.152	2.264	2.054	12.549	14.930	0.862	0.564	1.918	8.679	0.985	0.926	26.864	0.490	0.000
Kenya	3.654	13.996	1.036	25.134	20.444	3.961	0.431	0.981	1.541	1.000	0.971	0.705	0.647	3.458	12.220	1.902	1.396	3.531	0.465	0.448	13.279	7.303	3.899	17.241	35.864	42.132	1.144	2.207	25.466
Kuwait	4.925	0.588	22.739	0.429	1.733	1.446	3.940	6.005	14.840	0.971	1.000	12.714	3.164	1.196	4.400	24.424	5.708	1.447	14.183	9.324	0.704	0.920	2.254	9.347	1.393	0.756	24.576	1.332	2.024
Lebanon	1.307	1.728	5.540	1.260	0.934	0.913	1.498	8.010	10.002	0.705	12.714	1.000	0.901	0.352	3.551	6.987	1.628	2.417	10.435	5.163	2.481	1.179	1.840	3.496	0.819	1.459	11.636	0.940	0.925
Malaysia	17.524	1.038	2.608	1.148	1.899	22.799	12.707	46.315	5.577	0.647	3.164	0.901	1.000	2.841	9.325	9.159	8.081	6.742	3.980	8.026	0.542	0.891	5.213	16.967	0.796	0.959	4.091	3.154	1.013
Nepal	20.016	2.724	0.568	8.296	2.141	5.358	0.620	0.285	0.978	3.458	1.196	0.352	2.841	1.000	2.948	0.603	5.012	9.126	0.928	0.828	2.607	1.151	15.775	3.849	3.440	3.600	1.164	5.464	3.055
Nigeria	2.862	27.575	1.564	10.892	32.025	2.078	1.517	5.655	1.645	12.220	4.400	3.551	9.325	2.948	1.000	4.062	2.379	2.956	0.781	1.891	14.734	12.035	2.776	9.814	12.273	16.305	1.555	2.413	12.882
Oman	5.048	1.483	9.761	1.874	2.039	2.160	3.974	16.075	23.152	1.902	24.424	6.987	9.159	0.603	4.062	1.000	9.794	1.844	17.907	8.917	0.710	1.987	4.420	26.570	1.405	2.504	30.563	1.021	0.907
Pakistan	5.659	3.058	3.369	2.378	1.402	1.411	3.907	3.563	2.264	1.396	5.708	1.628	8.081	5.012	2.379	9.794	1.000	1.980	4.445	12.604	0.366	1.284	6.033	3.535	1.003	1.085	3.607	1.016	0.974
Philippines	7.908	2.081	1.114	2.764	2.657	15.159	0.784	0.872	2.054	3.531	1.447	2.417	6.742	9.126	2.956	1.844	1.980	1.000	0.620	0.659	3.762	1.252	7.620	2.005	3.336	2.444	1.144	10.586	3.077
Qatar	0.555	0.513	11.358	0.998	0.672	0.904	3.506	6.316	12.549	0.465	14.183	10.435	3.980	0.928	0.781	17.907	4.445	0.620	1.000	7.937	0.000	0.721	6.775	6.757	0.748	0.628	15.358	0.651	0.262
S. Arabia	2.701	0.610	39.058	1.318	0.633	0.975	2.158	4.793	14.930	0.448	9.324	5.163	8.026	0.828	1.891	8.917	12.604	0.659	7.937	1.000	0.433	1.064	1.949	21.519	0.346	0.436	7.312	0.510	0.441
Senegal	2.769	55.743	0.789	9.134	30.211	1.945	0.253	0.000	0.862	13.279	0.704	2.481	0.542	2.607	14.734	0.710	0.366	3.762	0.000	0.433	1.000	3.514	2.001	9.052	13.850	10.819	0.196	3.716	12.087
S. Africa	1.208	6.909	0.747	6.538	6.829	0.892	1.045	0.658	0.564	7.303	0.920	1.179	0.891	1.151	12.035	1.987	1.284	1.252	0.721	1.064	3.514	1.000	1.638	4.492	6.224	7.439	1.265	0.830	23.338
Sri Lanka	6.345	3.535	2.205	6.317	7.505	4.651	2.337	2.150	1.918	3.899	2.254	1.840	5.213	15.775	2.776	4.420	6.033	7.620	6.775	1.949	2.001	1.638	1.000	4.613	2.908	2.344	1.149	4.175	1.221
Sudan	2.091	10.576	8.801	19.556	11.910	2.443	1.525	8.915	8.679	17.241	9.347	3.496	16.967	3.849	9.814	26.570	3.535	2.005	6.757	21.519	9.052	4.492	4.613	1.000	12.217	12.688	7.679	1.590	4.735
Tanzania	3.953	11.610	0.735	18.585	22.905	2.239	0.384	1.073	0.985	35.864	1.393	0.819	0.796	3.440	12.273	1.405	1.003	3.336	0.748	0.346	13.850	6.224	2.908	12.217	1.000	43.189	0.536	1.698	26.562
Uganda	2.762	11.796	1.020	21.783	15.836	3.938	0.387	2.625	0.926	42.132	0.756	1.459	0.959	3.600	16.305	2.504	1.085	2.444	0.628	0.436	10.819	7.439	2.344	12.688	43.189	1.000	0.225	1.746	36.685
U.A. E.	1.789	0.368	11.612	1.013	1.024	1.295	2.530	11.464	26.864	1.144	24.576	11.636	4.091	1.164	1.555	30.563	3.607	1.144	15.358	7.312	0.196	1.265	1.149	7.679	0.536	0.225	1.000	0.978	1.595
Viet Nam	4.642	1.990	0.573	2.548	3.481	7.885	0.631	0.839	0.490	2.207	1.332	0.940	3.154	5.464	2.413	1.021	1.016	10.586	0.651	0.510	3.716	0.830	4.175	1.590	1.698	1.746	0.978	1.000	1.114
Zimbabwe	3.541	9.830	0.903	14.531	11.869	2.909	0.592	0.000	0.000	25.466	2.024	0.925	1.013	3.055	12.882	0.907	0.974	3.077	0.262	0.441	12.087	23.338	1.221	4.735	26.562	36.685	1.595	1.114	1.000

Table 2a

	Cameroon	Ethiopia	Ghana	Kenya	Nigeria	Senegal	S. Africa	Sudan	Tanzania	Uganda	Zimbabwe
Cameroon	1	9.630	20.781	13.996	27.575	55.743	6.909	10.576	11.610	11.796	9.830
Ethiopia	9.630	1	14.957	25.134	10.892	9.134	6.538	19.556	18.585	21.783	14.531
Ghana	20.781	14.957	1	20.444	32.025	30.211	6.829	11.910	22.905	15.836	11.869
Kenya	13.996	25.134	20.444	1	12.220	13.279	7.303	17.241	35.864	42.132	25.466
Nigeria	27.575	10.892	32.025	12.220	1	14.734	12.035	9.814	12.273	16.305	12.882
Senegal	55.743	9.134	30.211	13.279	14.734	1	3.514	9.052	13.850	10.819	12.087
S. Africa	6.909	6.538	6.829	7.303	12.035	3.514	1	4.492	6.224	7.439	23.338
Sudan	10.576	19.556	11.910	17.241	9.814	9.052	4.492	1	12.217	12.688	4.735
Tanzania	11.61	18.585	22.905	35.864	12.273	13.850	6.224	12.217	1	43.189	26.562
Uganda	11.796	21.783	15.836	42.132	16.305	10.819	7.439	12.688	43.189	1	36.685
Zimbabwe	9.830	14.531	11.869	25.466	12.882	12.087	23.338	4.735	26.562	36.685	1

Table 2b

	Bangladesh	Egypt	Indonesia	Iran	Iraq	Jordan	Kuwait	Lebanon	Malaysia	Nepal	Oman	Pakistan	Philippines	Qatar	S. Arabia	Sri Lanka	Sudan	U. A. E.	Viet Nam
Bangladesh	1	1.44	4.637	1.026	3.864	1.022	4.925	1.307	17.524	20.016	5.048	5.659	7.908	0.555	2.701	6.345	2.091	1.789	4.642
Egypt	1.44	1	0.998	1.92	7.954	5.546	22.739	5.54	2.608	0.568	9.761	3.369	1.114	11.358	39.058	2.205	8.801	11.612	0.573
Indonesia	4.637	0.998	1	0.685	2.222	0.952	1.446	0.913	22.799	5.358	2.16	1.411	15.159	0.904	0.975	4.651	2.443	1.295	7.885
Iran	1.026	1.92	0.685	1	5.471	1.787	3.94	1.498	12.707	0.62	3.974	3.907	0.784	3.506	2.158	2.337	1.525	2.53	0.631
Iraq	3.864	7.954	2.222	5.471	1	44.73	6.005	8.01	46.315	0.285	16.075	3.563	0.872	6.316	4.793	2.15	8.915	11.464	0.839
Jordan	1.022	5.546	0.952	1.787	44.73	1	14.84	10.002	5.577	0.978	23.152	2.264	2.054	12.549	14.93	1.918	8.679	26.864	0.49
Kuwait	4.925	22.739	1.446	3.94	6.005	14.84	1	12.714	3.164	1.196	24.424	5.708	1.447	14.183	9.324	2.254	9.347	24.576	1.332
Lebanon	1.307	5.54	0.913	1.498	8.01	10.002	12.714	1	0.901	0.352	6.987	1.628	2.417	10.435	5.163	1.84	3.496	11.636	0.94
Malaysia	17.524	2.608	22.799	12.707	46.315	5.577	3.164	0.901	1	2.841	9.159	8.081	6.742	3.98	8.026	5.213	16.967	4.091	3.154
Nepal	20.016	0.568	5.358	0.62	0.285	0.978	1.196	0.352	2.841	1	0.603	5.012	9.126	0.928	0.828	15.775	3.849	1.164	5.464
Oman	5.048	9.761	2.16	3.974	16.075	23.152	24.424	6.987	9.159	0.603	1	9.794	1.844	17.907	8.917	4.42	26.57	30.563	1.021
Pakistan	5.659	3.369	1.411	3.907	3.563	2.264	5.708	1.628	8.081	5.012	9.794	1	1.98	4.445	12.604	6.033	3.535	3.607	1.016
Philippines	7.908	1.114	15.159	0.784	0.872	2.054	1.447	2.417	6.742	9.126	1.844	1.98	1	0.62	0.659	7.62	2.005	1.144	10.586
Qatar	0.555	11.358	0.904	3.506	6.316	12.549	14.183	10.435	3.98	0.928	17.907	4.445	0.62	1	7.937	6.775	6.757	15.358	0.651
S. Arabia	2.701	39.058	0.975	2.158	4.793	14.93	9.324	5.163	8.026	0.828	8.917	12.604	0.659	7.937	1	1.949	21.519	7.312	0.51
Sri Lanka	6.345	2.205	4.651	2.337	2.15	1.918	2.254	1.84	5.213	15.775	4.42	6.033	7.62	6.775	1.949	1	4.613	1.149	4.175
Sudan	2.091	8.801	2.443	1.525	8.915	8.679	9.347	3.496	16.967	3.849	26.57	3.535	2.005	6.757	21.519	4.613	1	7.679	1.59
U.A. E.	1.789	11.612	1.295	2.53	11.464	26.864	24.576	11.636	4.091	1.164	30.563	3.607	1.144	15.358	7.312	1.149	7.679	1	0.978
Viet Nam	4.642	0.573	7.885	0.631	0.839	0.49	1.332	0.94	3.154	5.464	1.021	1.016	10.586	0.651	0.51	4.175	1.59	0.978	1

Table 3

	Argentina	Colombia	Costa Rica	Cuba	Ecuador	Mexico	Peru	Puerto Rico	Uruguay	Venezuela
Argentina	1.000	4.682	4.536	4.117	6.432	3.457	4.959	2.509	19.299	5.834
Colombia	4.682	1.000	7.871	5.757	11.271	5.038	6.064	4.898	4.341	10.637
Costa Rica	4.536	7.871	1.000	8.442	12.826	6.284	11.595	10.298	9.922	13.116
Cuba	4.117	5.757	8.442	1.000	9.352	14.831	10.812	4.706	5.855	13.860
Ecuador	6.432	11.271	12.826	9.352	1.000	5.879	18.547	2.186	8.312	16.459
Mexico	3.457	5.038	6.284	14.831	5.879	1.000	4.266	3.362	4.934	6.750
Peru	4.959	6.064	11.595	10.812	18.547	4.266	1.000	4.821	8.790	11.459
Puerto Rico	2.509	4.898	10.298	4.706	2.186	3.362	4.821	1.000	2.249	8.679
Uruguay	19.299	4.341	9.922	5.855	8.312	4.934	8.790	2.249	1.000	12.544
Venezuela	5.834	10.637	13.116	13.860	16.459	6.750	11.459	8.679	12.544	1.000

Table 4

	Algeria	Azerbaijan	Georgia	Morocco	Tunisia	Turkey
Algeria	1.000	0.152	0.190	11.049	13.265	1.821
Azerbaijan	0.152	1.000	11.880	10.586	0.441	11.623
Georgia	0.190	11.880	1.000	5.357	0.381	3.332
Morocco	11.049	10.586	5.357	1.000	10.107	2.152
Tunisia	13.265	0.441	0.381	10.107	1.000	1.399
Turkey	1.821	11.623	3.332	2.152	1.399	1.000

Table 5

	Bosnia and Herzegovina	Bulgaria	Croatia	Macedonia	Moldova	Romania	Serbia	Slovenia
Bosnia and Herzegovina	1.000	3.584	54.631	49.369	11.009	2.598	53.013	16.645
Bulgaria	3.584	1.000	3.976	14.758	2.522	3.555	4.477	1.497
Croatia	54.631	3.976	1.000	12.742	2.105	2.275	7.203	11.693
Macedonia	49.369	14.758	12.742	1.000	8.118	3.231	30.774	9.742
Moldova	11.009	2.522	2.105	8.118	1.000	16.257	1.696	1.114
Romania	2.598	3.555	2.275	3.231	16.257	1.000	3.622	2.664
Serbia	53.013	4.477	7.203	30.774	1.696	3.622	1.000	6.663
Slovenia	16.645	1.497	11.693	9.742	1.114	2.664	6.663	1.000

Table 6

	Czech Republic	Slovakia
Czech Republic	1.000	10.806
Slovakia	10.806	1.000

Table 7

	Estonia	Latvia	Lithuania
Estonia	1.000	20.280	10.704
Latvia	20.280	1.000	24.334
Lithuania	10.704	24.334	1.000

Table 8

	Kazakhstan	Uzbekistan
Kazakhstan	1.000	54.817
Uzbekistan	54.817	1.000

Table 9

	People's Republic of China	Hong Kong
People's Republic of China	1.000	10.515
Hong Kong	10.515	1.000

Table 10

Country	PAI	Connected Country
USA	2.483	Puerto Rico
Germany	2.226	Austria
U.K.	2.156	Ireland
Japan	5.047	Indonesia

Notwithstanding the fact that only 2.2 % of the total values of PAI are above 10, these involve a total of 62 Countries, that is, almost two/thirds of the total. This does not mean, obviously, that all the connections of the 62 Countries are strong. 5 African Countries present the maximum number of connections above PAI 10 (9 connections each): Ghana, Kenya, Sudan, Tanzania and Uganda. These Countries are part of a wide clustered network. This "cluster" contains 29 Countries, interconnected with a variety of links, and is somewhat bipartite. It is in fact composed by a group of African Countries and by one of Asian/ Northern African Countries, joined by Sudan. This Country belongs to both sub-networks. On the other side in some cases small groupings of Countries present strong PAIs with one or two other Countries. These "micro networks are those involving: People's Republic of China – Hong Kong; Czech Republic – Slovakia; Kazakhstan – Uzbekistan; Estonia – Latvia – Lithuania.

Besides these two extremes three other networks are also present. The first one is centred around Danubian Eastern Europe. The second one is centred in Latin America. The third one, interestingly, comprises three Countries from Northern Africa and three in Western Asia.

The pattern of strongly connected Countries is somewhat aligned, at least in part, with the findings of previous literature. Nevertheless it presents some strong peculiarities, which possibly are due to the use of PAI in place of other indicators. The use of a reliable, size-independent, indicator, less influenced than others by gross count of publication, coupled to the analysis of a wide number of Countries, is able to reveal a wide range of strong inter-country connections regardless of their overall scientific strength. In the following section the results will be discussed, also keeping in account the outcome of the theoretical framework.

DISCUSSION

The clustered networks of collaborating Countries present different features. First of all, the "bilateral" cooperation (Country-Country) highlighted by the PAIs are Czech Republic – Slovakia, Kazakhstan – Uzbekistan and People's Republic of China - Hong Kong. In the first case the two Countries were actually the same Country, Czechoslovakia, until January 1st, 1993. In the second case the two Countries are neighbouring, were both part of the USSR until 1991, share the same religious majority (Islam) and partly ethnicity of the population. Finally, regarding People's Republic of China and Hong Kong, it must be noted that formally Hong Kong is a Special Administrative Region of the People's Republic of China. Thus, though considered separately in the databases of scientific publications, actually it is not a separate Country. A similar interpretation can be offered for the three-Country "network" of Estonia, Latvia and Lithuania. The three Baltic Countries are also neighbouring, former USSR Countries until 1990/1991, and share cultural, religious and language similarities. Moreover, the three Countries joined together the European Union on May 1st, 2004, and are in the Euro zone.

When coming to larger networks of Countries presenting interconnections due to high PAIs the story becomes more complex. The Countries present in the "Danubian-Eastern Europe network" are connected by borders, and partly share common languages. More specifically most part of the Countries that did belong to former-Yugoslavia are present in the cluster and are connected with very high coefficients. These five Countries (Serbia, Bosnia and Herzegovina, Croatia, Republic of North Macedonia and Slovenia) form the core of this cluster. The relations between these Countries, notwithstanding the war period of the 1990s, are obviously strong under several points of view (mainly cultural, linguistic and historical reasons). The other three Countries have a rather accessory role in the network, being connected to only one of the above described Countries (Bulgaria to Macedonia, Romania to Moldova which is then connected to Bosnia).

It is possible to tell a similar story about the "Latin-American" network. The core of the network is built around Venezuela, Peru, Ecuador and Costa Rica, while Colombia and Cuba are connected with two of the above indicated Countries each, and Puerto Rico, Mexico, Cuba, Uruguay and Argentina are out of the core. Nevertheless here the cultural, historical, linguistic and geographical connections existing between the Countries are actually very strong, as most of these Countries share common borders, a common language (Spanish) and a past as former Spanish colonies. This clustering of Countries can be interpreted according to the continentalization of science (LECLERC; GAGNÉ, 1994).

The other two networks (both composed by two smaller sub-networks) are more difficult to interpret at a first glance. The first one, encompassing six Countries, is somewhat bipartite. Nevertheless a deeper analysis of the Morocco-Azerbaijan link shows the presence of a single, very strong collaboration, sharing a large number of scientific works in the considered timespan, thus creating a link (see Choi, 2012 for a discussion on the relevance of common research interests on the formation of links). The two triangular networks resulting in this way (Morocco-Algeria-Tunisia and Azerbaijan-Turkey-Georgia) present again strong cultural, geographical, historical and religious ties (see for instance Landini *et al.*, 2015).

The most evident feature of the last and largest network is that Sudan connects two smaller networks: an "African" one and an "Afro-Asian" one (all the Countries are in Asia except Egypt). Each of the two "sub-networks" presents several of the above described factors. The "African sub-network" (detailed in Table 2a) is composed of Countries that are either contiguous or English speaking, or both. The only exception is Senegal, which nevertheless presents the strongest coefficient of the whole set (that of the Senegal-Cameroon connection). As table 2a shows the network is strongly interconnected. Interestingly the least connected Country is South Africa, which is also the most productive African Country. Our results partly confirm those of Mêgnigbêto (2013) and of Boshoff (2009).

The "Afro-Asian subnetwork" (detailed in table 2b) is instead less interconnected, but several factors can be invoked to explain its structure. All the Countries but Egypt are in Asia, and many are contiguous or geographically close: for instance Pakistan, Nepal, Bangladesh and Sri Lanka surround the "scientific power" India. Then a cultural/religious factor is present. In fact several Countries in the group either speak Arabic as main language[4] or witness a relevant presence, or a majority, of Islamic population[5]. Some Countries are connected to another one or two Countries only (Sri Lanka-Nepal-Bangladesh, Iran, Viet Nam-Philippines-Indonesia are all connected to Malaysia) and Pakistan (linked to Saudi Arabia only). Nevertheless these connections always show the influence of the above described factors.

CONCLUSIONS

Aim of the present work is to measure the strongest scientific inter-Country collaborations and the structure of their network, as well as the main factors that may influence them. This has been done building a database of co-authored scientific products and exploiting it in order to build an instrument, the Probability Affinity Index, less influenced by absolute values of scientific products than the same absolute values of scientific products, as well as than other indicators. This article generalizes the findings of Author (2015) and, more closely, of Author (2016). The interpretative framework sketched in the latter of the two articles did consider three main determinants as drivers of collaboration: "geographical proximity", "cultural proximity" and "historical proximity". "Geographical" meant in that article that the involved Countries are neighbouring, or are at least located at close distance in the same Continent.

4 Egypt, Iraq, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Sudan, United Arab Emirates.

5 Bangladesh, Egypt, Indonesia, Iran, Iraq, Jordan, Kuwait, Lebanon, Malaysia, Oman, Pakistan, Qatar, Saudi Arabia, Sudan, United Arab Emirates.

"Cultural" involved characteristics such as main language or main religion. "Historical" involved former relationship of colonizing/colonized Countries, or having been part of the same Country before it did split up. This interpretative framework still partly holds considering the experimental results of the present work, though it needs to be further integrated.

The network structure of the "clusters" of Countries discussed above shows some interesting peculiarities. First of all, none of the most productive and scientifically "mainstream" Countries (such as, for instance, the U.S.A., Germany or Japan) is present in the picture. The only exception is the easily explainable presence of People's Republic of China in the connection with Hong Kong. This absence could be explained by the fact that such Countries tend to be involved in several, weaker international cooperation rather than in stronger cooperation with fewer Countries. On the other side it is true that many links or cluster are easily explainable: most two- or three-Country link, as well as the group of former-Yugoslavia Countries are strictly linked either from the geographic, cultural or historical point of view. This makes a strong collaboration easily understandable. Nevertheless it is relevant that this study has been able to highlight their strength with respect to other more frequently described research partnerships.

On the other side our data show clearly that most of the Countries connected by strong values of PAI belong to the set of the less scientifically central ones. These are Countries that, due to several reasons – e.g. population, lower funding, short historical path of research activities – often produce less research products and/or products of minor impact. Thus we can imagine that networking depends at least in part by a process of self-selection based on the scientific strength of Countries: "strength is in numbers" and small (in terms of scientific production) and close (in terms of geography, culture or history) Countries may tend to form small networks with peers, or to have a strong bilateral relationship (like some Countries in the "Asian cluster" do). Thus policy makers might decide to incentivize existing collaboration of this type, or to foster the creation of new ones. Conversely, in case further information may indicate that strong collaboration with peer Countries is detrimental, might instead foster collaboration with "big actors".

These facts do not necessarily mean that a process of exclusion of scientifically weaker Countries from elite clusters may be in act. Our results show that scientifically stronger Countries – such as for instance the United States, the United Kingdom, Germany or Japan – present always low or very low values of PAI. Table 10 presents as an example the highest PAIs of these four Countries. Thus scientifically stronger Countries, rather than establishing strong collaboration with few "neighbours" (either physical, cultural, historical, scientific etc.) have enough force to sustain a wide number of collaborations. These are (relatively) weaker than those of Countries with a minor number of collaborations. Thus, rather than "excluding" smaller Countries, the collaboration paths with these ones basically present a strength similar to that of any other connection with scientifically stronger Country. The idea of a multi-centre model, instead of a centre-periphery model, is surreptitiously fostered by these results.

The present work, though still limited in its effort, presents novel results. Further research will deepen the analysis, addressing items such as differences existing between scientific fields among Countries involved in the collaboration, sub-Country geographical subdivisions, time evolution as well as modelling spatial interactions exploiting, e.g., gravity models (see for instance Lata *et al.* 2015; Scherngell and Hu (2011). With respect to previous literature we have been able to design the network of the strongest inter-Country scientific collaboration with the aid of a more reliable instrument. Once the network has been designed we have interpreted the relations between Countries at the light of the various factors that drive scientific collaboration as devised by on-topic literature.

Finally, we have offered some policy suggestion deriving from our results. The use of PAIs and the high number of Countries considered in the analysis allows us to highlight international connections that were not considered in previous works. Our results can be important in order to design policies intended to support international scientific collaboration, in particular for those countries experiencing a lower rate of scientific production with respect to other ones.

REFERENCES

ACOSTA, M. *et al.* Factors affecting inter-regional academic scientific collaboration within Europe: the role of economic distance. *Scientometrics,* v. 87, n. 1, p. 63-74, 2011.

AKSNES, D. W.; FROLICH, N.; SLIPERSAETER, S. Science policy and the driving forces behind the internationalisation of science: The case of Norway. *Science and Public Policy*, v. 35, n. 6, p. 445-457, 2008.

AUTHOR (2015).

AUTHOR (2016).

BARRIOS, C. *et al.* Is there convergence in international research collaboration? An exploration at the country level in the basic and applied science fields. *Scientometrics*, v. 120, n. 2, p. 631-659, 2019.

BEAVER, D.; ROSEN, R. Studies in scientific collaboration. *Scientometrics*, v. 1, n. 1, p. 65-84, 1978.

BERGÉ, L. R. Network proximity in the geography of research collaboration. *Papers in Regional Science*, v. 96, n. 4, p. 785-815, 2017.

BOSHOFF, N. Neo-colonialism and research collaboration in Central Africa. *Scientometrics,* v. 81, n. 2, 2009.

CASSI, L.; MORRISON, A.; RABELOTTI, R. Proximity and Scientific Collaboration: Evidence from the Global Wine Industry. *Tijdschrift voor Economische en Sociale Geografie,* v. 106, n. 2, p. 205-219, 2015.

CHOI, S. Core-periphery, new clusters, or rising stars?: international scientific collaboration among 'advanced' countries in the era of globalization. *Scientometrics,* v. 90, n. 1, p. 25-41, 2012.

CHOI, S.; YANG, J. S.; PARK, H. W. The triple helix and international collaboration in science. *Journal of the Association for Information Science and Technology,* v. 66, n. 1, p. 201-212, 2015.

EBLEN, M. *et al. Social network analysis comparing researcher collaborations in two cardiovascular cohort studies. Research Evaluation,* v. 21, n. 5, p. 392-405, 2012.

FRENKEN, K. *et al.* Death of Distance in Science? A Gravity Approach to Research Collaboration. In: PYKA, A.; SCHARNHORST, A. (Eds.)., *Innovation Networks:* new approaches in modelling and analyzing. Berlin: Heidelberg, 2009. p. 43-57.

GARNER, J. G. *et al. Assessing research network and disciplinary engagement changes induced by an NSF program. Research Evaluation,* v. 21, n. 2, p. 89-104, 2012.

GLANZEL, W.; SCHUBERT, A. Analysing Scientific Networks Through Co-Authorship. In: MOED, H. F.; GLANZEL, W.; SCHMOCH, U. (Eds.)., *Handbook of Quantitative Science and Technology Research:* The Use of Publication and Patent Statistics in Studies of S&T Systems. Dordrecht: Springer Netherlands, 2005a. p. 257-276.

GLANZEL, W.; SCHUBERT, A. Domesticity and internationality in co-authorship, references and citations. *Scientometrics,* v. 65, n. 3, p. 323-342, 2005b.

GONZÁLEZ-ALCAIDE, G. *et al.* Dominance and leadership in research activities: Collaboration between countries of differing human development is reflected through authorship order and designation as corresponding authors in scientific publications. *PLOS ONE*, v. 12, n. 8, 2017.

HASSAN, S. U.; SARWAR, R.; MUAZZAM, A. Tapping into intra- and international collaborations of the Organization of Islamic Cooperation states across science and technology disciplines. *Science and Public Policy,* v. 43, n. 5, p. 690-701, 2016.

HENNEMANN, S.; RYBSKI, D.; LIEFNER, I. The myth of global science collaboration—Collaboration patterns in epistemic communities. *Journal of Informetrics,* v. 6, n. 2, p. 217-225, 2012.

HOEKMAN, J.; FRENKEN, K.; TIJSSEN, R. J. W. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, v. 39, n. 5, p. 662-673, 2010.

JEONG, S.; CHOI, J. Y.; KIM, J. Y. On the drivers of international collaboration: The impact of informal communication, motivation, and research resources. *Science and Public Policy,* v. 41, n. 4, p. 520-531, 2014.

KAMADA, T.; KAWAI, S. An algorithm for drawing general undirected graphs. *Information Processing Letters,* v. 31, n. 1, p. 7-15, 1989.

KATS, J. S. Geographical proximity and scientific collaboration. *Scientometrics,* v. 31, n. 1, p. 31-43, 1994.

KRAUT, R.; EGIDO, C.; GALEGHER, J. Patterns of Contact and Communication in Scientific Research Collaboration. In *Proceedings of the 1988 ACM Conference on Computer-supported Cooperative Work.* New York, ACM, 1988. p. 1-12.

LANDINI, F.; MALERBA, F.; MAVILIA, R. The structure and dynamics of networks of scientific collaborations in Northern Africa. *Scientometrics,* v. 105, n. 3, p. 1787-1807, 2015.

LATA, R.; SCHERNGELL, T.; BRENNER, T. Integration Processes in European Research and Development: A Comparative Spatial Interaction Approach Using Project Based Research and Development Networks, Co-Patent Networks and Co-Publication Networks. *Geographical Analysis,* v. 47, n. 4, p. 349-375, 2015.

LECLERC, M. ; GAGNÉ, J. International scientific cooperation: The continentalization of science. *Scientometrics,* v. 31, n. 3, p. 261-292, 1994.

LUUKKONEN, T. ; PERSSON, O. ; SIVERTSEN, G. Understanding Patterns of International Scientific Collaboration. *Science, Technology, & Human Values,* v. 17, n. 1, p. 101-126, 1992.

MATTSSON, P. *et al.* What do European research collaboration networks in life sciences look like? *Research Evaluation,* v. 19, n. 5, p. 373-384, 2010.

MÊGNIGBÊTO, E. International collaboration in scientific publishing: the case of West Africa (2001–2010). *Scientometrics,* v. 96, n. 3, p. 761-783, 2013.

MELIN, G.; PERSSON, O. Studying research collaboration using co-authorships. *Scientometrics,* v. 36, n. 3, p. 363-377, 1996.

NARIN, F.; STEVENS, K.; WHITLOW, E. S. Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics,* v. 21, n. 3, p. 313-323, 1991.

OKUBO, Y. *et al*. Structure of international collaboration in science: Typology of countries through multivariate techniques using a link indicator. *Scientometrics,* v. 25, n. 2, p. 321-351, 1992.

OKUBO, Y.; ZITT, M. Searching for research integration across Europe: a closer look at international and inter-regional collaboration in France. *Science and Public Policy,* v. 31, n. 3, p. 213-226, 2004.

PONDS, R. VA OORT, F.; FRENKEN, K. The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, v. 86, n. 3, p. 423-443, 2007.

PRICE, D. The analysis of square matrices of scientometric transactions. *Scientometrics*, v. 3, n. 1, p. 55-63, 1981.

SCHERNGELL, T.; HU, Y. Collaborative Knowledge Production in China: Regional Evidence from a Gravity Model Approach. *Regional Studies,* v. 45, n. 6, p. 755-772, 2011.

SCHUBERT, A.; BRAUN, T. International collaboration in the sciences 1981–1985. *Scientometrics,* v. 19, n. 1, p. 3-10, 1990.

SONNENWALD, D. H. Scientific collaboration. *Annual Review of Information Science and Technology,* v. 41, n. 1, p. 643-681, 2007.

THIS, B.; GLANZEL, W. A structural analysis of collaboration between European research institutes. *Research Evaluation*, v. 19, n. 1, p. 55-65, 2010.

UNESCO. *Science Report*. [Europa]: Unesco, 2015. Disponível em: https://en.unesco.org/unesco_science_report. Acesso em: 02 maio 2021.

WAGNER, C. S.; LEYDESDORFF, L. Mapping the network of global science: comparing international co-authorships from 1990 to 2000. *International Journal of Technology and Globalisation,* v. 1, n. 2, p. 185-208, 2005.

WALTMAN, L.; TIJISSEN, R. J.; ECK, N. J. Globalisation of science in kilometres. *Journal of Informetrics,* v. 5, n. 4, p. 574-582, 2011.

ZANOTTO, S. R.; HAEFFNER, C.; GUIMARÃES, J. A. Unbalanced international collaboration affects adversely the usefulness of countries' scientific output as well as their technological and social impact. *Scientometrics,* v. 109, n. 3, p. 1789-1814, 2016.

ZITT, M.; BASSECOULARD, E.; OKUBO, Y. Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science. *Scientometrics,* v. 47, n. 3, p. 627-657, 2000.

Appendix – List of the Countries considered in the present study

Algeria
Argentina
Armenia
Australia
Austria
Azerbaijan
Bangladesh
Belarus
Belgium
Bosnia and Herzegovina
Brazil
Bulgaria
Cameroon
Canada
Chile
Colombia
Costa Rica
Croatia
Cuba
Cyprus
Czech Republic
Denmark
Ecuador
Egypt
Estonia
Ethiopia
Finland
France
Georgia
Germany
Ghana
Greece
Hong Kong
Hungary
Iceland
India
Indonesia
Iran
Iraq
Ireland
Israel
Italy
Japan
Jordan
Kazakhstan
Kenya
Kuwait
Latvia
Lebanon
Lithuania
Luxembourg
Macedonia
Malaysia
Mexico
Moldova
Morocco
Nepal
Netherlands
New Zealand
Nigeria
Norway
Oman
Pakistan
People's Republic of China
Peru
Philippines
Poland
Portugal
Puerto Rico
Qatar
ROC
Romania
Russian Federation
Saudi Arabia
Senegal
Serbia
Singapore
Slovakia
Slovenia
South Africa
South Korea
Spain
Sri Lanka
Sudan
Sweden
Switzerland
Tanzania
Thailand
Tunisia
Turkey
Uganda
Ukraine
United Arab Emirates
United Kingdom
United States
Uruguay
Uzbekistan
Venezuela
Viet Nam
Zimbabwe