Social network analysis of international scientific collaboration on family farming research

Péter Lengyel[1], János Pancsira[2], Péter Balogh[3], Judit Oláh [4], István Füzesi[5]

A B S T R A C T

In our research, we analysed the co-authorship network of articles on family farming. Based on the definition of social network analysis, we did this on a macro level, as it examines relationships between individuals and interactions. Social Network Analysis (SNA) is a wide strategy to investigate the social structure. SNA is used to obtain patterns of relationships between nodes to ascertain underlying social structure. Node is represented as people and group, meanwhile relation shows the relationship between nodes. The problem is how to find the most influential node in the network. The focus is how to measure centrality in the network. Our sample is based on 254 scientific articles written on this topic between 1976 and 2015. In this paper we modeled a co-authorship network based on this sample. We identified why network analysis is useful and who the most important authors are in the network.

1. Introduction

In this article we present a co-authorship network based on the scientific articles on family farming published over the past 40 years. The sample is made up of 254 scientific articles written between 1976 and 2015 on this topic. Nowadays, an analysis of co-authorship network has become more and more popular in academic life. Since this is a relatively new field of science, especially in domestic terms, there are not many researches dealing with this issue. At the beginning of the research the following hypotheses were formulated:

- The most authoritative author is the most significant in the network sense.
- Authors who publish the topic on family farming also feature joint work with co-authors.

After the theoretical review, the methodology chapter gives a description of the sample used for the research, and then we present the workflows that have been completed. The results of the research are presented with data visualization, tables, brief summary descriptive explanations and analysis.

At the micro level, network analysis examines the behavior of individuals, and macro-level relationships (network structure) between individuals and their interaction (Stokman 2004). The basic principle of network analysis is the socio-psychological analysis, also known as sociometry, which was later learned by experts from other disciplines. One of the first pioneers in sociometry was Jacob Levy

---

[1] Péter Lengyel
University of Debrecen
lengyel.peter@econ.unideb.hu

[2] János Pancsira
University of Debrecen
pancsira.janos@econ.unideb.hu

[3] Péter Balogh
University of Debrecen
balogh.peter@econ.unideb.hu

[4] Judit Oláh
University of Debrecen
olah.judit@econ.unideb.hu

[5] István Füzesi
University of Debrecen
fuzesi.istvan@econ.unideb.hu

Moreno Romanian-born psychiatrist. He began to deal with sociometric with "Who shall survive?" study appeared in 1932. This volume has grown to more than 700 pages in subsequent editions and has been translated into more than 20 languages under the "Basics of Sociometry". In 1937 he published his independent journal, Sociometry (Mérei 2006). Based on these basics, today's advanced social network analysis (SNA) has been developed.

The basic source of SNA methodology was created by Siegfried Frederick Nadel's work (Tóth 2009). The development of the methodology has been well promoted by the appearance of computers, computer access, growth of administration methods and mathematics development. There are many areas that benefit greatly from the development of these methods, especially the merging of businesses, the further spread of innovation, the decision making of political, consumer and market actors, the productivity of teamwork and the spread of diseases (Pál & Vörös 2011).

In our research, we are investigating a network of authors who have published family farming articles. For the definition of family farm we found several definition. Family farm is a farm owned and operated by a family of one or more generations. Most of the land and capital are provided by the family. Most of the workforce is provided by family members living on the farm, but occasionally external work can also be used (Galeski & Wilkening 1987). Contrary to the terms of Takácsné (1995) and Galeski & Wilkening (1987), Williams (1973) excludes the use of external workers, even seasonally. Williams agrees with Buzás (1994), saying that it is not possible to talk about family farming when relying mostly on employees and wage-earners.

In Hungary we can speak about a "more traditional" form of farming, because labor is only provided by family members and can be operated at low cost. In the "more modern" concept, not only family members, but also wage-earners work in the economy, but are typically seasonal.

2. Material and method

2.1. Sample for research

We queried our sample from Scopus. We filtered it on "family farm" and "family farming" and resulting in 254 hits. This amount was manageable and processable, so we didn't have to narrow down the search criteria. The hit list was exported and inserted into the database of our software that communicated with Scopus API. With the help of the software we have received data tables related to articles (articles, authors, research institutes, keywords, themes).

2.2. Social network analysis

The SNA methodology, which is a relatively young discipline, is appropriate to examine graphical models of publishing models of scientific societies. In Hungary, such analyses are still in the early stages, mostly for journals (Popp, Balogh, Kovács & Jámbor 2015). The association of authors may have several reasons, such as: organizational reasons, availability of laboratory equipment, investments requiring greater financial resources, speeding up procedures, mitigating errors and increasing productivity (Micsinai 2011). In co-authored publications, Yoshikane (2006) states that two different functions can be observed, the leader and the followers. The leader is the person who first appears in writing. Previous research has revealed that the growth of these types of co-operative articles has an impact on individual performance, and thus their own productivity is growing at the same time (Glänzel & Schubert 2004).

It can be said that authoring groups or publishers publishing within a subject can also be considered as a contact network where relationships are the collaborations between authors. The most important method of depicting and analysing contact data is graph theory. This procedure is practical because it has the right vocabulary to demonstrate network formations and provides a mathematical basis for measurable data. Using graphs, we get a clear picture of the pattern of contacts. It is indispensable to note that when displaying the graphs, the position of the points or the length of the lines connecting them do not contain information. In this case, it represents a line of lines, dots, and edges where each edge appears only once.

The graph theory approach is ideal for us because it is excellent for identifying key people. In this case, key people are those writers who have the most co-authorship connections. These determinants occupy the strategic points of the network. In the case of non-directed graphs we are talking about centrality when we consider it important that a participant is involved in relationships, and it is less important he or she has a sending or receiving role (Kürtösi 2004).

One of the known centrality calculation modes is degree centrality. At this point, the number of connections for each point is compared to all connections. We can expect a centrality with closeness centrality, according to which a person is placed in a central position if all authors are easily reached in a short time and thus there is no need to involve other members. The following centrality calculation method is betweenness centrality. This is based on the fact that the actors are the most influential among many others. At this point, we must actually consider the shortest paths that include the intermediate points.

3. Results

Figure 1 shows a world map showing the individual points in which research institutes have been published, and the numbers show how many scientific articles have been generated in a given city. The locations of the research sites were saved in a separate Excel workbook and imported into the MapsData online app. From the chart, it can be seen at first glance that the articles are mostly from Europe and America. This may also be due to the fact that family farming is the most common in these continents.
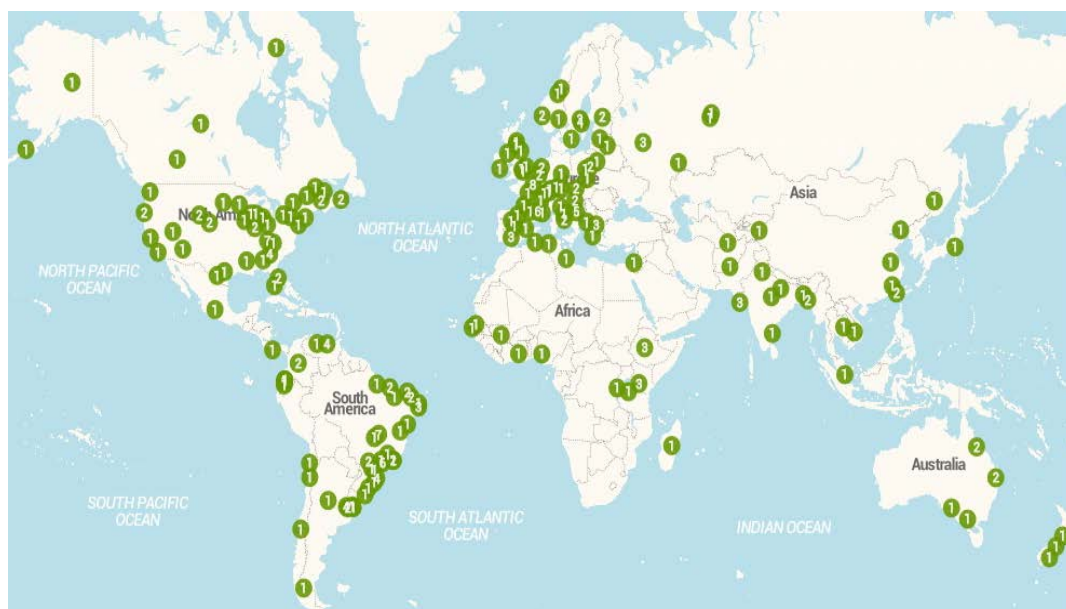

Figure 1. Location of the authors' research institutes

3.1. Descriptive statistics

Figure 2 shows the yearly development of the number of publications on family farming. It can be noticed that most of the articles on family farming have been written over the last few years. This is due to the fact that, today the family economy has a much greater role in law than the delimitation of the concept. Especially since 2009, the number of articles dealing with this topic has been increasing. In 2015, 37 articles were written, representing 15% of all published articles.
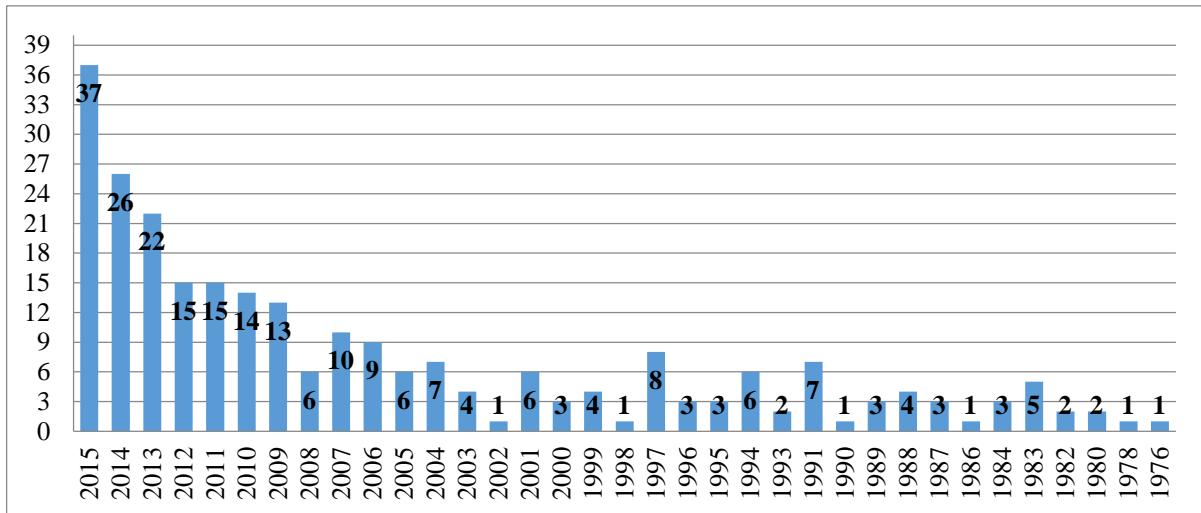
Figure 2. Number of publications between 1976 and 2015

In our opinion, some of the ones written in 2014 are due to the fact this year was the International Year of Family Farming declared by United Nations. "The goal of the International Year 2014 is to put family farming at the center of national agri-environmental and social policies" (FAO 2014). From the 1990s, domestic literature has also dealt with more and more new forms of economics (Magda 2010).

The distribution of the number of publications per country can be observed in Figure 3. Articles on the subject have come out from more than 50 countries. Most of them were published in Brazil, out of 254 articles, 50 were published in this country. This is followed by the United States of America 25, followed by France 22. Most of the writings in Brazil are due to the fact that they are dominant in agriculture, world-leading in coffee, sugar cane, orange and banana production, as well as in soy, cocoa, beef and tobacco production (Balogh, Békési, Gorton, Popp & Lengyel 2016). In 2008, the Brazilian government proposed a Small Farmers' Program to strengthen family farming (Marti 2008).
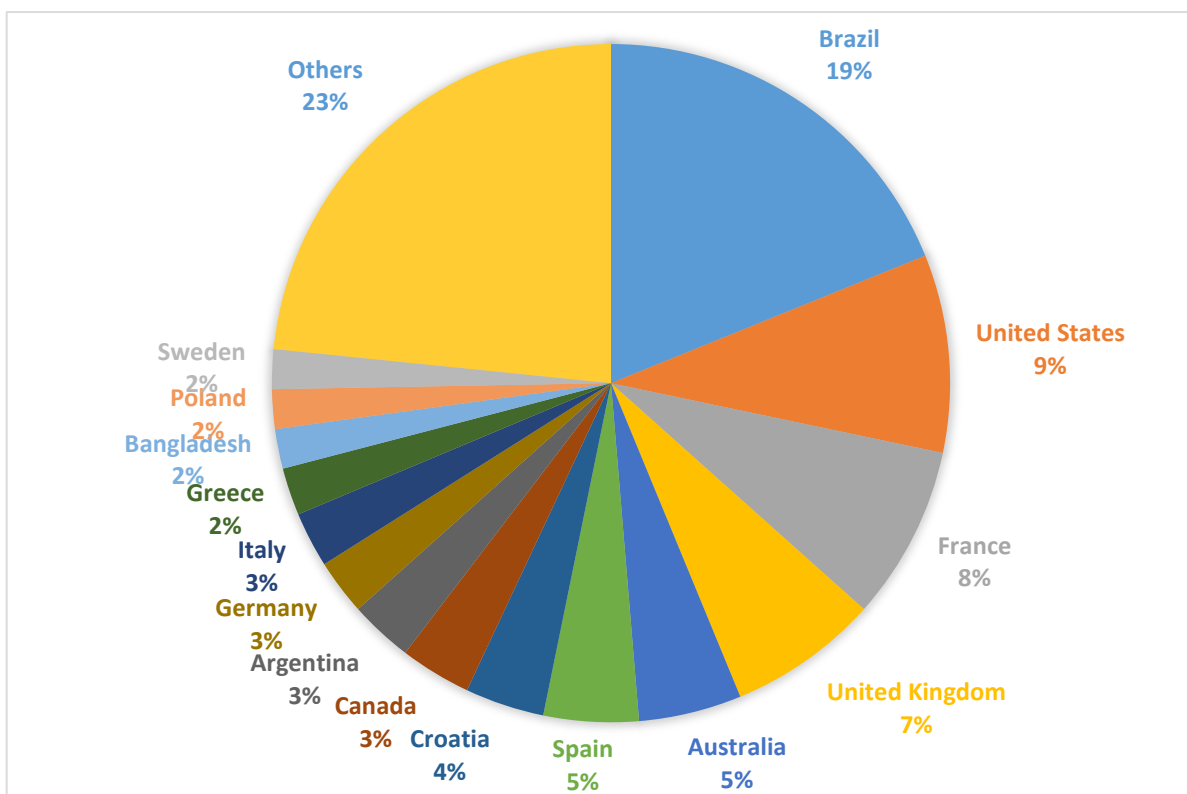

Figure 3. Distribution of publications by country

Péter Lengyel, János Pancsira, Péter Balogh, Judit Oláh, István Füzesi: Social network analysis of international scientific collaboration on family farming research

More than 20 different European countries have been published. Most of these countries came from France (22), UK (19), Spain (12) and Croatia (10).

Table 1 clearly shows that the most widely published writer Linda Price wrote 6 articles, of which three were written alone and three were co-authors. Three of the six publications deal with the situation of women in family farms. In these writings, she describes how the role of women in farms has been evaluated.

Table 1. TOP 5 most authoritative authors

No.	Author	No. of publications
1.	Price L.	6
2.	Guanziroli	3
3.	Sourisseau J.-M.	3
4.	Radinovic	3
5.	Kasimis C. - Papadopoulos, A.G.	3

3.2. Co-authorship network

Figure 4 shows the publication link graph. It can be seen that authors who have reported an article alone or with one and two co-autors are occupying the periphery. This diagram is based on a modularity index, so larger co-authorship cliques are located in the center. There are five relatively large cliques that are made up of ten or more authors.
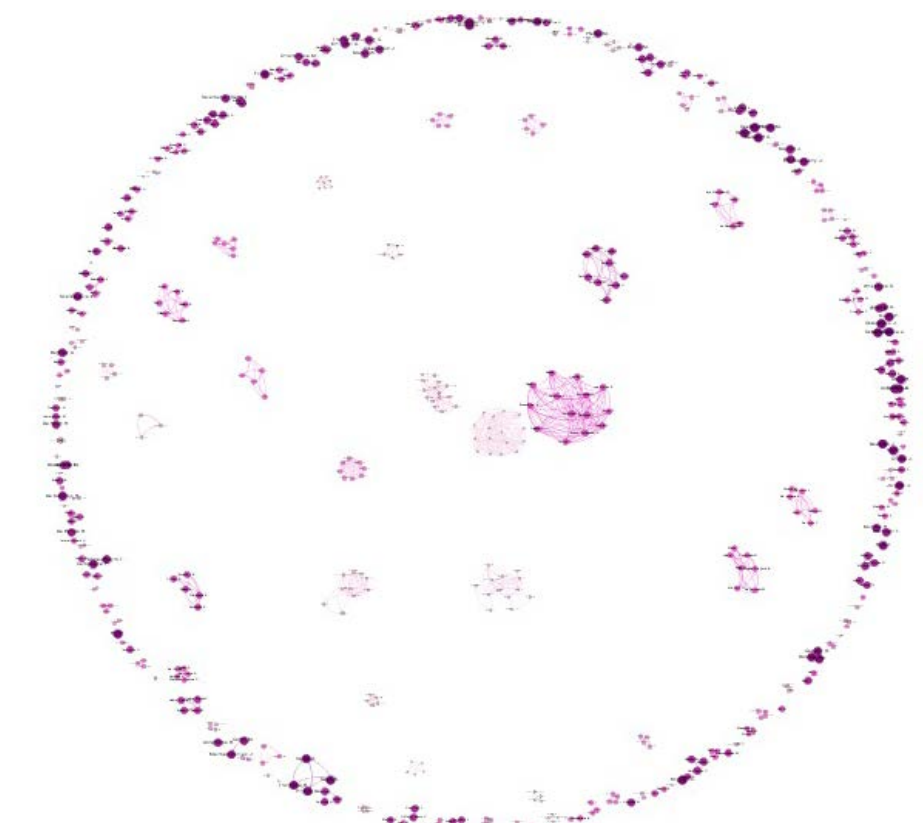

Figure 4. Co-authorship network

It can be deduced that close authoring groups have been created and it is difficult to enter among them and there are no links between the larger groups.

3.3. Authors' cliques

Table 2 shows the top 5 collections of authors, affecting a total of 53 people. This represents 9.8% compared to the total data plate with 539 people. It is important to note that besides the 5 largest groups, there are associations of 4, 5, 6 and 7 people.

Table 2. Members of the 5 largest cliques

Clique 1	Clique 2	Clique 3	Clique 4	Clique 5
Bodekær	Gehring	Bosc	Rozon	Van Vliet
Harrison	Spithoven	Marzin	Lucotte	Schut
Philipsen	Schmid	Bélières	Davidson	Reidsma
Petersen	Bitter	Sourisseau	Sousa Passos	Descheemaeker
Rogowski-Tylman	Braun-Fahrländer	Pédelahore	Oestreicher	Van de Ven
Schmalwieser	Dalphin	Losch	Mertens	Slingerland
Triguero-Mas	Hyvärinen	Bonnal	Paquet	Giller
Dadvand	Pekkanen	Parrot	Romana	
Lesiak	Riedler			
Narbutt	Weiland			
Eriksen	Büchele			
Heydenreich	Mutius			
Nieuwenhuijsen	Vuitton			
Thieden	Brunekreef			
Young				
Wulf				

This also shows that most authors have worked in smaller groups or alone. The number of publications in these groups is negligible compared to all publications. It is characteristic for these groups that one article has been published jointly. Figure 5 shows the graphs of these groups depicted by modularity.

The graph of the 1st clique shows clearly that each author is in contact with each author. The same applies to members of the 2nd clique. Clique 3, 4, and 5 show that not all authors have direct contact with everyone. Clearly visible, for example, in the case of clique 4 Jacquez Marzin has only an indirect relationship with Even, Koné and Wampfler. Also in Figure 5 we can observe that Novo and Jansen are in direct contact with Slingerland alone.
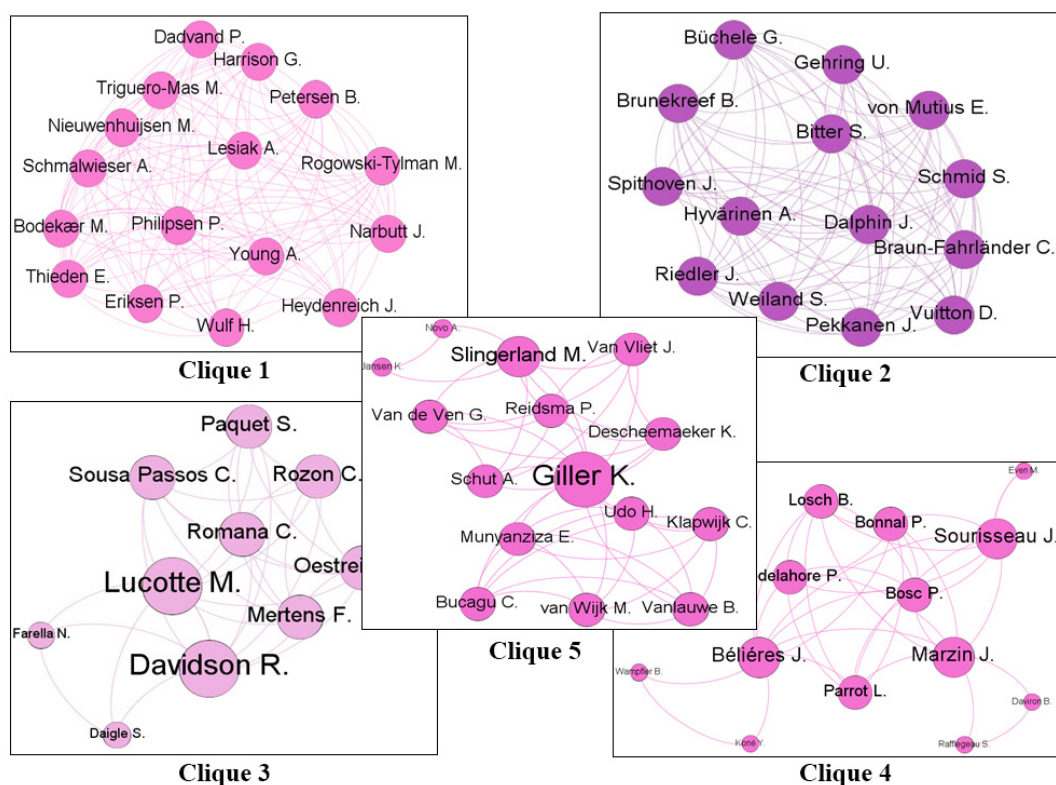
Figure 5. Co-authorship cliques

3.4. Author rankings

The most useful index for us is the betweenness centrality, as the values differed significantly in this case, so a spectacular, accurate order emerged. We have examined whether there is a match between the order of the most published authors based on degrees and betweenness. With this we got the answer to our first hypothesis, in which we assumed that most authoritative publishers are the most prominent in the network.

In case the most authoritative author's name is found in the order column set by centrality metrics, our hypothesis is correct. Otherwise, if most authors have not been listed in any of the other columns, our first assumption is false. To make them easier to review, we have prepared a summary table (Table 3), where these sequences can be seen together.

Table 3. Author rankings

No. of publications		Betweenness centrality		Degree centrality	
Author	**Value**	**Author**	**Value**	**Author**	**Value**
Price	6	Marzin	10	Bodekær	15
Guanziroli	3	Slingerland	7	Harrison	15
Sourisseau	3	Giller	7	Philipsen	15
Radinovic	3	Žutinić	6	Petersen	15
Kasimis	3	Radinovic	4	Triguero-Mas	15
Papadopoulos	3	Sourisseau	3	Schmalwieser	15
Lucotte	2	Bélières	2	Rogowski-Tylman	15

Slingerland	2	Pritchard	2	Dadvand	15
Marzin	2	Lucotte	1	Lesiak	15
Giller	2	Davidson	1	Narbutt	15

From Table 3, it can be stated at first glance that the most authoritative authors are not necessarily the most prominent ones in the network sense. With respect to the degrees, all 10 authors have equally published with co-authors, each of them having a value of 15. There is no match for the authors selected for the degrees and the most authoritative publishers. This suggests that most authoritative authors are not characterized by cooperation with co-authors.

The betweeness centrality index of Jacques Marzin is in the first place, suggesting a prominent "mediator" role. That is, in this network he is the author who is the easiest to achieve. The number of Marzin's publications is 2, which is relatively low but can not be considered a bad one, considering that the database is very scattered and most authors have published only one publication on this subject. Both documents were made by more people, so they were not published alone. In terms of betweenness centrality index, the following two authors, Maja Slingerland and Ken Giller, also published two publications. Apart from Graham Harrison and Paul Eriksen, the same names appear in the two columns. Jacques Marzin is not included in the betweenness index, among the top 10 authors, which suggests a smaller but more stable co-author network.

Overall, these results demonstrate that, unlike the first findings, there is no correlation between the number of authors and their role in the network. It seems that some authors have published a lot with a few co-authors relative to the divisions of the database, while others have a broader co-authoring network, but they have relatively few articles.

6. Conclusion

The articles in the samples come in roughly half the way between America and Europe. This is not surprising, as this mode of farming plays a significant role on these two continents. In America, most of all, we can talk about farm management, which is slightly different from the family farming we interpret but is essentially the same.

It is apparent from the year of publication of articles that this topic does not look back on a long past, it is a relatively young area. It was surprising that the published articles came from such a large proportion (44%) of alone publishing authors. We can only find a few major co-authorship groups that are clearly distinct from the graph depicted on the basis of the modularity index. This suggests that these groups are closed so it is difficult to get in. Within these groups, we can also see that there are typically cliques in which each author is in contact with each author. Typically, such a clique has just written one article together.

Our first hypothesis that the most authoritative author is the most significant in the network sense has not been proven. The most authoritative author was Linda Price, however it is not decisive in terms of its roles as an intermediary, or in terms of easy access or rank. Our second hypothesis that authors of these articles are characterized by co-authorship, proved to be true. This is evidenced by the fact that 56% of the published articles were made by co-authors, while the articles were made in a smaller proportion (44%) due to individual work. It is important to note, however, that not only writers are more likely to be grouped in smaller groups.

References

Balogh, P, Békési, D, Gorton, M, Popp, J & Lengyel, P 2016 'Consumer willingness to pay for traditional food products', Food Policy 61: pp. 176-184. https://doi.org/10.1016/j.foodpol.2016.03.005

Buzás, Gy 1994 'A családi gazdaságok jövedelmezősége és helyük a vállalati struktúrában' (The profitability of family farms and their place in a corporate structure) Gazdálkodás, XXXVIII. évfolyam 6. sz. pp. 25-34.

FAO 2014 '2014 a családi gazdálkodás nemzetközi éve' (2014 is the International Year of Family Farming) http://www.fao.org/docrep/019/as281hu/as281hu.pdf letöltés dátuma: 2016.08.04.

Galeski, B & Wilkening, E 1987 'Family Farming in Europe and America.', Boulder, Colorado Westviwe Press, pp. 1-4.

Glänzel, W & Shubert, A 2004 'Analyzing scientific networks through coauthorship' In Moed HF (ed) Handbook of Quantitative Science and Technology Report. Kluwer Academic Publishers, pp. 257-276.

Kürtösi, Zs 2004 'A társadalmi kapcsolatháló-elemzés módszertani alapjai' (Methodological basics of social networking analysis) In Letényi, L (ed): Településkutatás. Budapest, L'Harmattan, pp. 663-684.

Magda, S 2000 'A régiók mezőgazdaság- és vidékfejlesztése' (Agricultural and rural development of the regions) Gazdálkodás XLIV. 2, pp. 60-67.

Marti, MA 2008 'Brazília mezőgazdasága' (Brazil's agriculture) Brüsszel, pp. 17-18. http://www.europarl.europa.eu/RegData/etudes/note/join/2008/397242/IPOL-AGRI_NT(2008)397242_HU.pdf [2016.09.03]

Mérei, F 2006 'Közösségek rejtett hálózata' (Hidden Network of Communities) Osiris Kiadó, Budapest, pp. 44-49. ISBN 9789633894583

Micsinai, I 2011 'Társszerzői hálózatok topográfiája és dinamikája. - A Közgazdasági Szemle, a Magyar Pszichológiai Szemle, a Pszichológia, valamint a Szociológiai Szemle társszerzői hálózatainak elemzése' (Topography and Dynamics of Co-authorship Networks) Szociológiai szemle. pp. 21(1): 4–28.

Pál, J & Vörös, A 2011 'Bevezetés a társadalmi kapcsolatháló-elemzés elméletébe és módszertanába.' (Introduction to the theory and methodology of social networking analysis) ELTE TÁTK Survey Szakest, Budapest http://surveyszakest.tatk.elte.hu/Pal_Voros.PDF [2016.08.11]

Popp, J, Balogh, P, Kovács, S & Jámbor, A 2015 'Hálózatosodás az agrárgazdaságtanban - Szerzői és hivatkozási kapcsolatok a Kelet-Közép-Európáról szóló szakirodalomban' (Networking in Agricultural Economics – Co-authorship and referral connections in the literature on Central and Eastern Europe) Közgazdasági Szemle. LXII. évf. pp. 525-543.

Stokman, FN 2005, 'What Binds Us When with Whom? Content and Structure in Social Network Analysis' English version of Stokman, FN & Vieth, M 2004, 'Was verbindet uns wann mit wem? Inhalt und Struktur in der Analyse sozialer Netzwerke', Kölner Zeitschrift für Soziologie. Sonderheft 44, pp. 274-302.

Takácsné, Gy 1995, 'A családi gazdaságok méretére ható tényezők vizsgálata, különös tekintettel a növénytermesztés gépesítésére' (Investigating factors affecting the size of family farms, in particular the mechanization of crop production) Kandidátusi értekezés, Gödöllő pp. 11-23.

Tóth, Zs 2009 'A hálózatelmélet alapfogalmai és rövid története' (Basic concepts and brief history of network theory) in: Tudáshálók a gazdasági felsőoktatásban PhD értékezés NYME-Közgazdaságtudomyányi Kar, Sopron. pp. 11–26.

Williams, WM 1973 'The social Study of family farming' In Mills, DR. (ed), English Rusal Communities, London: Macmillan, pp. 116-133. https://doi.org/10.1007/978-1-349-15516-3_5

Yoshikane, F, Nozawa, T & Tsuji, K 2006 'Comparative analysis of co-authorship networks considering authors' roles in collaboration: Differences between the theoretical and application areas' Scientometrics, 81 (2) pp. 499–511. https://doi.org/10.1007/s11192-006-0113-1