# Network Analysis of Citations Among the Authors of *The Oxford Handbook of the Economics of Networks*

**By Gulsah Ayhan**

Supervisors: Norma Olaizola & Federico Valenciano

October, 2016 Bilbao

ABSTRACT

Citation analysis is a way of ranking journals or authors based on their mutual citations. Networks are a useful way to visualize the relationships between objects. In our study, the mutual citations of the authors in *The Oxford Handbook of The Economics of Networks* have been been used to evaluate the 45 authors of the *Handbook* by a method which is based on the Invariant Method. Different centrality measures have also been applied in order to rank them from other points of view.

# *Table of Contents*

## 1. INTRODUCTION

A network basically represents an accumulation of objects and the links that make connections among them. In a network, the objects are called *nodes* or *vertices* and the ties that connect them are called *edges*. A network visualization as a collection of nodes and edges can be used by several disciplines to show the interaction between related nodes and edges. A node can be an individual, a disease, a city, an airport, a journal or anything that we want to categorize and explain the relationships among them. These relationships can be friendships between individuals, common molecules that diseases have, railway connections of cities, routes between airports or the mutual citations between the journals. All these links that connect the nodes are represented by the edges.

In Figure 1.1 the images represented as a node, a network with two nodes and an edge, a star network, and a circle network respectively.

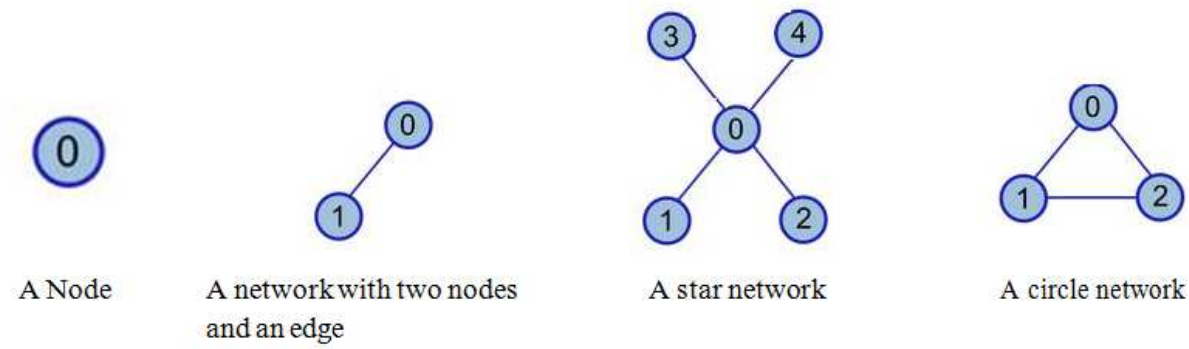
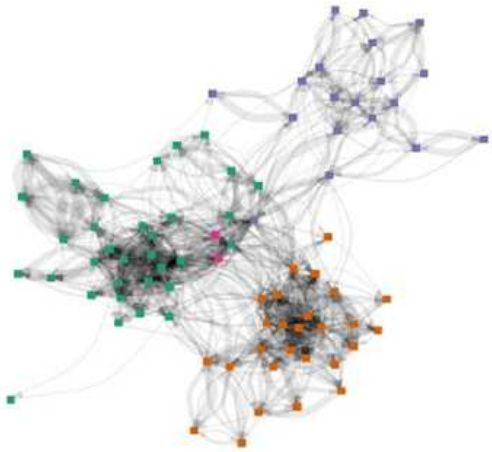


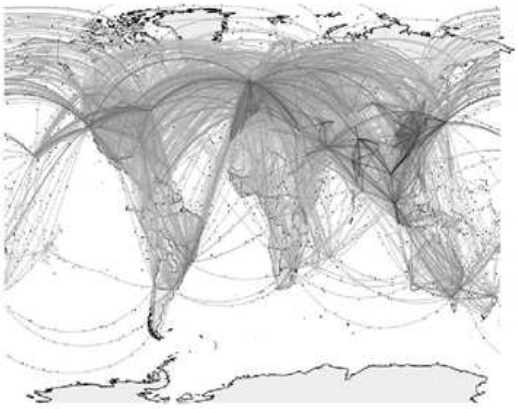

Figure 1.1

Therefore, networks are used for several disciplines to visualize the interactions between related nodes and edges. In Figure 1.2(a), the image shows the personal friendship network of a faculty of a UK university, consisting of 81 vertices (individuals) and 817 connections. The school affiliation of each individual is stored as a vertex attribute (Nepusz T., Petroczi A., Negyessy L., Bazso F., 2008). In Figure 1.2 (b), the image shows Airport network, where the line colour is based on the number of routes. (openflight.com)



UK Faculty Friendship Network
Source:Nepusz T., Petroczi A., Negyessy L., Bazso F.:
Fuzzy communities and the concept of bridgeness in complex networks.

(a)



Airport network
Source:openflight.com

(b)

Figure 1.2

As it can be easily understood, network structures underlie a huge variety of situations in many different fields.

A citation network is a network which considers the journals, papers or authors as the nodes and their murtual citation relationships as the edges. It enables us to do citation analysis which is a way to evaluate the journals or the authors by their citation transfers. In our study, we are going to focus on the citation network of the authors in *The Oxford Handbook of The Economics of Networks* linked by citations among themselves.

## 2. PRELIMINARIES

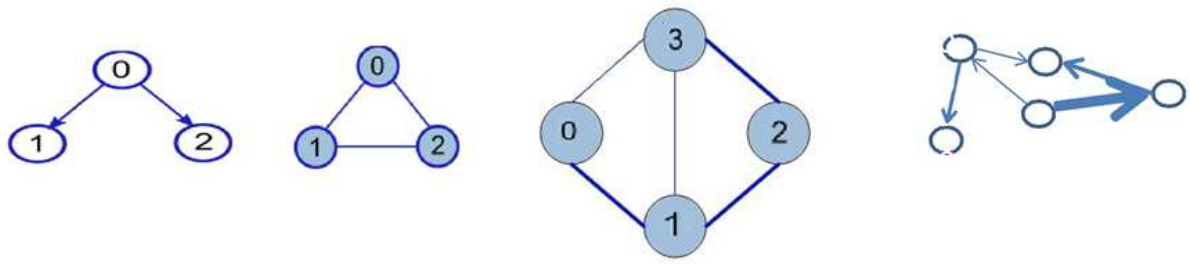
In this section, we are going to explain briefly the concepts that we have used in our study. Different types of network and related examples will be given and also the measurements to understand the importance of the nodes in the network will be investigated in this part.

### 2.1. TYPES OF NETWORKS

In an unweighted network all links have the same strength or importance, while in a weighted network different links may have different strength or importance. This is represented by associating weights with links according to their importance. In our study the weights are

derived from the evaluations of citations among authors. However, directed and both weighted and unweighted network analysis are going to be applied in this study.

A network can be *directed* and *undirected* depending on whether the edges have a direction or not. For instance coauthorship would yield an undirected network. In a directed network the direction of an edge represents the direction of the relationship between the nodes when such direction exists. For example, a twitter account considered as a node in the network by its relationships with the other accounts can follow the other accounts and can be followed by the other accounts. Thus, the links have a direction. In a citation network, the relationship between nodes is represented by two directions: a node can cite the other nodes and a node can be cited by the other nodes in the network. Besides being directed or not, a network can be classified as *weighted* or *unweighted*. In an unweighted network all links have the same strength or importance, while in a weighted network different links may have different strength or importance. The weight of an edge represents how strong or important the tie between two nodes is. For instance, it may reflect the different frequencies of flights in an airport network or the different evaluation of citations among journals. In Figure 2.1 different types of networks are represented.




An unweighted and directed network | An unweighted and undirected network | A weighted and undirected network | A weighted and directed network

Figure 2.1

## 3. CENTRALITY MEASURES

One of the key instruments to interpret the results derived from the network analysis is investigating centrality measures. To carry out the network analysis, calculation of centrality is significantly important. Interpretation of the results and deduction of the importance of the nodes derived from the graph is the main purpose of our study. Thus, focusing on the centrality measures is going to become more of an issue to understand the position of the

nodes in the graph. Centrality measures are considered for the unweighted and weighted networks differently. We will concentrate on mostly used centralities such as degree, betweenness, closeness and eigenvector.

## 3.1. Centrality Measures For Unweighted Networks

### 3.1.1. Degree Centrality

*Degree centrality* indicates the number of edges attached to a node. In other words, it counts the ties which connect a node. Degree centrality can be evaluated under two subtitles for the directed networks, which are *in-degree* and *out-degree* centrality measures. In-degree centrality can be described as the number of edges directed to the node, while out-degree centrality counts the number of edges from the node to the other nodes. The nodes which have high in-degree centrality are referred as prominent and the ones with high out-degree centrality are referred as influential.

### 3.1.2. Betweenness Centrality

Another centrality measurement we are interested in, is *betweenness* centrality. It is equal to the number of shortest paths from all vertices to all others that pass through that node (Brandes,2001). It measures the extent to which a vertex lies on best paths between other vertices. This can be formulated as below.

For each pair of vertices (s,t), the betwenness centrality of a node $v$ is;

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where $\sigma_{st}$ is the total number of shortest paths from node s to node t, and $\sigma_{st}(v)$ is the number of paths which pass through $v$.

Having high betweenness centrality for a node in a network represents its control over the network, omitting high betweenness nodes from the network ruins the transfers among other nodes as their positions. (Newman, 2010)

### 3.1.3. Closeness Centrality

Another indicator used to measure centrality of a node is closeness. First, we are going to discuss about the farness of a node to derive the closeness centrality. Farness is the total sum of a node's shortest distances from all other nodes. Closeness is defined as the inverse of the farness. Closeness centrality measure indicates the average closeness of each node to all of the other nodes (Freeman, 1978). According to this approach, the less the total sum of a node's shortest distances from all other nodes is, the more central the node is. In other words, closeness centrality measure points out the importance level of a node by the idea of how easily a node can access to the other nodes (Weisstein, E.W.).

### 3.1.4. Eigenvector Centrality

Another measure indicating the importance of a node is *Eigenvector Centrality*. The idea of eigenvector method arises from the invariant method (Pinski and Narin, 1976). The principle behind eigenvector centrality is that "a node is important if it is linked to other important nodes" (Brath and Jonker, 2015). Under this principle, the importance of a node in the network depends on not only the number of connections it has, but also the importance of connections of the node.

While the degree centrality counts the ties attached to a node and determines the centrality respect to those amount of ties; being attached to many edges does not mean that the node's eigenvector centrality is high. For eigenvector centrality, it is crucial being linked by the nodes which also have high eigenvector centrality. Google's PageRank method is based on this idea as well.

## 3.2. Centrality Measures For Weighted Networks

Calculating centrality measures for weighted networks requires some additional considerations compared to centrality measurements for unweighted networks. Calculations take into account *not only quantity but also weights* of edges. Now we are going to discuss

three of the centrality measures that can be considered in weighted network, which are Degree, Closeness and Betweenness.

There have been several studies to develop the idea of centrality measures into the weighted networks. Instead of the number of edges attached to a node, degree centrality has referred to the sum of weights of the edges for the weighted networks (Barrat et al., 2004). And Newman (2001) adjusted Freeman's (1978) approach for closeness to weighted networks by using Dijkstra's (1959) algorithm. Besides, Brandes (2001) did the adjustment using the same algorithm for betweenness centrality. It is important to adjust Freeman's (1978) idea to weighted networks, however these approaches have restrictions such as not focusing on the number of ties. Hence, Opsahl et al. (2010) proposed a more flexibl method. This method considers not only the weights of ties but also the number of them. It uses the support of a tuning parameter (α) to apply closeness centrality on weighted networks. The outcome of the nodes' centralities are evaluated via the tuning parameter.

According to the study of Opshal et al. (2010), the value of the tuning parameter (α) is determined due to the data. If $\alpha = 0$, the method considers the network as a binary network and all the ties as equally weighted, as we have considered for the previous part of our network analysis. In the case that $\alpha = 1$, the evaluation of centrality measures will be the same as the ones calculated by Dijkstra's (1959) algorithm. In this case, the only criteria considered will be the weight of ties. It does not capture the number of ties to determine the importance level of nodes. In the case of $\alpha = 1$, the length of the shortest path is determined solely due to the weight of ties for Closeness and Betweenness centralities. Also, setting the parameter α between 0 and 1 or greater than 1 assigns the balance of portion of those two aspects to calculate the importance level. If α is set smaller than zero, the quantity of edges are more important than the strength of them. In this case, the shortest path determination is based on the distances rather than the weights. Conversely, arranging tuning parameter greater than 1 determines the importance level of nodes based on the weight of ties. And the strength of ties becomes more of an issue rather than the distances to detect the importance (Opshal et al., 2010). Hence, the different values of α parameter enable us to interpret the centralities of the nodes (i.e authors) by combining both the weights and quantity of the ties.

## 4. CASE STUDY: CITATION ANALYSIS IN THE OXFORD HANDBOOK OF THE ECONOMICS OF NETWORKS

### 4.1. DATA

*The Oxford Handbook of The Economics of Networks*, (from now on, the *Handbook*) edited by Yann Bramoulle, Andrea Galeotti and Brian Rogers, who also contribute as authors of chapters in the book, surveys the main developments in the last 20-25 years of Economics of Networks. The *Handbook* consists of seven parts including 31 chapters written by outstanding researchers in their fields. The book has 45 contributing authors, and each chapter has been written by one or several authors. In our study, the database consists of the citations among these authors within the *Handbook*. As might be expected, plenty of papers whose authors have not been involved in the book are cited by the authors in the chapters of the book. In order to detect the influence by using citation analysis of the authors within the book, we concentrate in the citation by chapters of the Handbook of references among whose authors at least one is one of its 45 authors. However, as we presently explain, the references none of whose authors are among the 45 authors have not been completely disregarded in order to weight citation in a 'fair' way.

Among all the 45 authors of the book, Matthew O. Jackson poses a difficulty due to the fact that this author does not cite any articles in his chapter. According to this, his column in the citation matrix would consist of zeros. But Jackson is a very important author in the field, and this would mean a seriously distorting impact in the evaluation process. In order to avoid this, we take his citations in his survey *'Networks in the Understanding of Economic Behaviors'* (Jackson, 2014) as those of him in his chapter of the Handbook. In other words, we replaced his Chapter 4 of the Handbook by this survey and references therein for the citation analysis.

### 4.2. METHODOLOGY

In this part, we explain the difficulties that we have met and the criteria we have used to overcome them, in order to measure citations among authors of the *Handbook*.

This is the first necessary step to rank the authors of the *Handbook* according to their mutual citations in the book. Even the simplest ranking method, direct counting of citations, would require some adjustments (e.g. different chapters have different number of authors, cited

articles may have different number of authors, maybe some not within the 45 of the *Handbook,* etcetera). Therefore it was necessary to find the appropriate method to evaluate or weight the citations in a reasonable way.

Our idea was to adapt to the problem at hand the Invariant method, that was firstly proposed by Pinski and Narin (1976) and later axiomatized by Palacios-Huerta and Volij (2004)*. The invariant method was developed to measure influence of scientific journals based on their mutual citations. Thus, we first explain briefly the original method before explaining the adjustments we have made from our data in order to make evaluations based on the citations of authors in the book.

Let J be the set of all journals and suppose $J \subseteq$ J is a finite subset of journals. For each $i,j \in J$; $c_{ij}$ is the *number of citations to journal* i *by journal* j. For each i∈J, $a_i$ is the number of articles that journal i published. And for a given citation matrix[1] C = $(c_{ij})$; $c_j = \sum_{i \in J} c_{ij}$ is the sum of j's references.

According to the invariant method, vector v = $(v_i)_{i \in J}$, given by the formula below, provides a *cardinal ranking of the journals* in $J$:

$v_i = \sum_{j \in J} \frac{c_{ij}/a_i}{c_j/a_j} v_j$ , i∈ $J$ , where $J \subseteq$J,

where $v_i$ can be interpreted as the value of a representative article in journal i. Note that vector $v = (v_i)_{i \in J}$ can be interpreted as a *fixed point* of the linear function

$f_i(v) = \sum_{j \in J} \frac{c_{ij}/a_i}{c_j/a_j} v_j$ , i∈ $J$.

In other words, if $v \rightarrow f(v)$ is the ranking method based on a weighted sum of $\frac{c_{ij}/a_i}{c_j/a_j}$, vector $v$ weights journals so that the resulting ranking is prescisely $v$.

The formula entails that the value of a journal (i.e. of a representative article in the journal) depends on the values of the other journals as well.

[1] A citation matrix is the fundemal entity which contains the information describing the flow of influence among units. It has the form; C = $\begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}$ (Pinski and and Narin, 1976)

In order to evaluate the citations of authors in our case, the first necessary step is to form the citation matrix based in to our data. We construct a 45x45 square matrix in which their rows and columns correspond to the authors in the *Handbook* ordered alphabetically.

For the citation matrix C = $(C_{ij})$ of us; we construct each term $c_{ij}$ as representing the weights of citations received by unit i from unit j; where i∈N; $1 \leq i,j \leq 45$. Nevertheless, the main issue was to determine the weights to evaluate the citations. Since we were going to rank the authors in the *Handbook* based on their citations inside the book; the starting point was considering the articles included in the references in each chapter in the book one or more of whose authors were authors of some chapter of the *Handbook*. However, there are some difficulties that need to be solved. First of all, the chapters have often more than one author. The same occur with the references cited in the chapters, often with several authors; moreover, often only some of its authors participate as authors of some chapter of the *Handbook.* For these reasons, we must deal with the many to many relationships for the chapters written by several authors and one to many relationships for the ones that owned by a single author. One more issue is that the chapters and the references of these chapters, have often common authors. Although we did not consider the self-citations and we set diagonal entries as 0 in the matrix, the cases where the set of authors of a chapter and of a reference in that chapter intersect, the other authors of the chapter had no limitation to receive the reference from this common author.

In addition to these difficulties, the references include a lot of authors who did not take part in the book as an author, but, even if we do not intend to evaluate them and the citation matrix does not incorporate them, not taking them into account in the evaluating process would be misleading for the sake of preserving the correct weight of citations. We presently explain the way in which this has been done.

In order to handle these problems, we started to examine each chapter seperately and asign one point of each reference of a chapter. Then, to determine the valuation of a citation from author i of a single reference to author j in the chapter, we divide this one point by the number of the authors who share the cited reference and multiply it by the number of authors in the chapter that can receive this citation (i.e the non self-citations). After we apply this method chapter by chapter, we aggregate all values of citations for each author i to author j, and obtain the citation matrix C = $(c_{ij})$. At the end, each entry of the citation matrix, $c_{ij}$ represents the total weight of citations from author i to author j. As the self citations have not

been counted, the diagonal entries of matrix are assigned value 0. And each ith row sums of the matrix, represented as the total weights of citations received by author i.

To clarify the method, let investigate it in examples;

Chapter 19 is written by Golub and Sadler and the first reference of this chapter is *"The network origins of aggragate fluctuations"* by Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi. This reference sets a point and since this reference consists of four authors; we set ¼ point for each author in reference, even Carvalho is not an author of the book. In this case, since the reference article and the chapter has no common author, we multiply ¼ by 2, the number of authors of Chapter 19, to evaluate citations from each author in the cited article to each author in the chapter. In conclusion, the values of citations from Acemoglu to Golub; Acemoglu to Sadler; Ozdaglar to Golub, Ozdaglar to Sadler; Tahbaz-Salehi to Golub, Tahbaz-Salehi to Sadler will be equal to ½ for this particular reference and chapter. It is worth noting that, we did not evaluate the citation of Carvalho to the authors of chapter, by the reason that he is not an author in the *Handbook*. However, we take him into consideration to calculate the others' valuations.

Another example for Chapter 17 which is written by Choi, Gallo and Kariv, and 26th reference article *"Network Architecture, salience and coordination"* of the chapter that is written by Choi, Gale, Kariv and Palfrey. Because we set every references as 1 point and this article has 4 authors,again we start putting ¼ point for each author in the article. This time we have 2 common authors in the chapter and the cited article. As we did not take into account the self citations, Choi can be cited by only Gallo and Kariv in the chapter. Therefore, we multiply ¼ by 2 to calculate the valuation of Choi's reference to Gallo and Kariv seperately. Similarly, Kariv can be cited by only Choi and Gallo in the chapter, thus his valuation will be (¼ x2) and (¼ x2) for Choi and Gallo respectively. On the other hand, Gale can be cited by all of the authors of the chapter. Hence we multiply ¼ by 3 in order to evaluate his citation to Choi, Gallo and Kariv. Finally, the value of citations from Choi to Gallo, Choi to Kariv; Gale to Choi, Gale to Gallo; Kariv to Choi, Kariv to Gallo will be ½; ½; ¾; ¾; ¾; ½; ½, respectively.

## 4.3. Analysis & Results

In this part, we are going to present and interpret the results based on our data. Interpretation of the results and deduction of the importance of the nodes derived from the graphs is the main purpose of our study. Thus, focusing on the centrality measures is going to become more of an issue to understand the position of the nodes in the graph. In our study, at first we will consider our network as directed and unweighted. After, we will concentrate on the mostly used centrality measures such as degree, betweenness, closeness and eigenvector centralities. Then, we will investigate the results obtained from the weighted network of our data

The structure of the study is as follows; firstly, the most important descriptives will be given in a table, then there will be interpretations of general graphs and the graphs of some specific important nodes with respect to some centrality values.

### 4.3.1. Descriptives For Unweighted Network

| Nodes | In degree | Out degree | |
|---|---|---|---|
| Goyal | 25 | 28 | Best |
| Pin | 23 | 11 | Second Best |
| Rogers | 22 | 17 | Third Best |
| Jackson | 21 | 36 | Fourth Best |
| Choi | 18 | 6 | Fifth Best |
| Manea | 4 | 5 | Fifth Worst |
| Dessein | 4 | 1 | Fourth Worst |
| Chaney | 4 | 3 | Third Worst |
| Mayzlin | 1 | 0 | Second Worst |
| Economides | 1 | 0 | Worst |

Table 4.3.1.1: First best 5 and last 5 nodes in the network in terms of their in-degree centralities.

| Nodes | Out degree | In-degree | |
|---|---|---|---|
| Jackson | 36 | 21 | Best |
| Goyal | 28 | 25 | Second Best |
| Galeotti | 25 | 14 | Third Best |
| Vega Redondo | 25 | 14 | Fourth Best |
| Bramoulle | 25 | 9 | Fifth Best |
| Manea | 1 | 17 | Fifth Worst |
| Dessein | 1 | 1 | Fourth Worst |
| Chaney | 0 | 1 | Third Worst |
| Mayzlin | 0 | 4 | Second Worst |
| Economides | 0 | 4 | Worst |

Table 4.3.1.2: First 5 and last 5 nodes in the network in terms of their in-degree centralities for the directed and unweighted network.

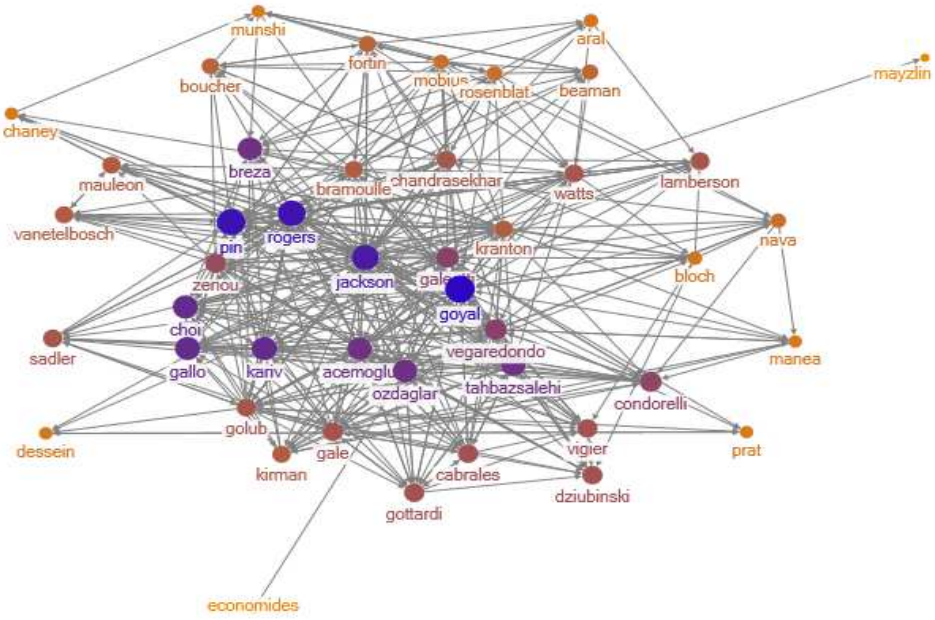



Figure 4.3.1.1: The directed and unweighted network graph of the authors in *The Oxford Handbook of The Economics of Networks* based on their mutual citations, constructed and ordered by the in-degree centrality.

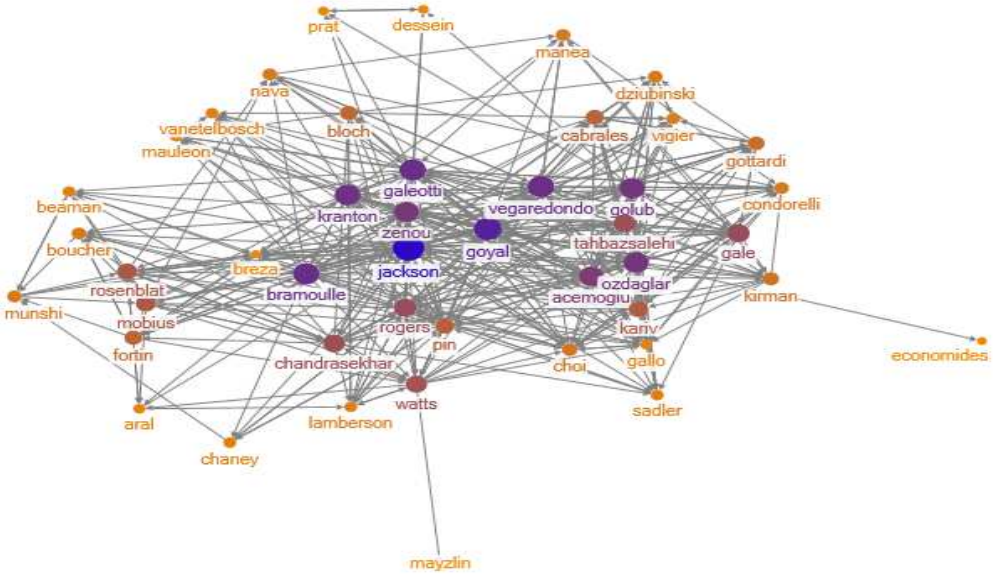



Figure 4.3.1.2: The directed and unweighted network graph of the authors in *The Oxford Handbook of The Economics of Networks* based on their mutual citations, constructed and ordered by the out-degree centrality.

The degree centrality of each author shows the number of authors he is connected to in terms of references. In-degree centrality of the chosen node, (the node represents here a specific author in *The Oxford Handbook of The Economics of Networks*) indicates that the number of authors who have been cited by the chosen author. On the other hand, the out-degree centrality points out the number of authors who have cited to the chosen author according to the derived graph. In other words, one can say that if we choose an author in the graph and count the ties which are towards him, we simply observe the number of authors whose references have been used by the chosen one in his article in the *Handbook.* Above, in Table 4.3.1.1 and Table 4.3.1.2, the highest and the lowest 5 authors are given in terms of in-degree and out-degree values with respect to the results obtained from directed and unweighted version of the citation network among the authors of the *Handbook* by NodeXL. Goyal has the highest in-degree centrality with 25. Therefore, he is the one who cited the highest number of authors. However the out-degree centrality of him, that is 27, is higher than his in degree centrality. This means that he is cited by more authors than he cited. Pin follows Goyal as the second highest in-degree centrality owner with 23, and his out-degree centrality which is 11, is significantly lower than the in-degree centrality he has. This comparison leads us to interpret that, Pin cited more than twice number of authors he is cited by. Conversely, Economides and Mayzlin share the lowest in degree centrality with only value of 1, referring that they have cited only one author. On the other hand, both of them has no out degree value.

It demonstrates that they do not have only weak relationships with authors in terms of references taken but also no author cited them.

According to results, Jackson is the one with the highest out degree centrality with 36, which means he has been cited by 36 authors of *The Oxford Handbook of The Economics of Networks*. Also he has cited a considerable amount of authors. His in degree centrality is 20, therefore we can roughly predict his place in the network graph as a central one, even without investigating it. The second highest out degree centrality belongs to Goyal which we have examined above. The lowest out degree centrality owners are Breza, Economidies, Mayzlin, Dessein and Prat with the values 0,0,0,1,1 respectively. One can expect that although they have the same out degree values as 0, Breza will have much more central position than Economidies and Mayzlin in the network graph, due to the fact that in degree centrality of him is 17 while Economidies' and Mayzlin's are only 1.

Since out-degree is referred as the number of citations an author gives, in our study out-degree centrality is more crucial than in-degree centrality to detect importance of an author in the network. Pin and Rogers can be seen in the first 5 list for the in-degree centrality, however they have not been involved the first 5 list for out-degree.

| Nodes | Betweenness | |
|---|---|---|
| Jackson | 237,587638 | Best |
| Goyal | 136,617686 | Second Best |
| Galeotti | 119,099855 | Third Best |
| Vega Redondo | 115,381839 | Fourth Best |
| Bramoulle | 18 | Fifth Best |
| Sadler | 0,905 | Fifth Worst |
| Prat | 0,118 | Fourth Worst |
| Dessein | 0,118 | Third Worst |
| Economidies | 0 | Second Worst |
| Mayzlin | 0 | Worst |

Table 4.3.1.3: Best and worst five nodes in the network in terms of betweenness centrality.

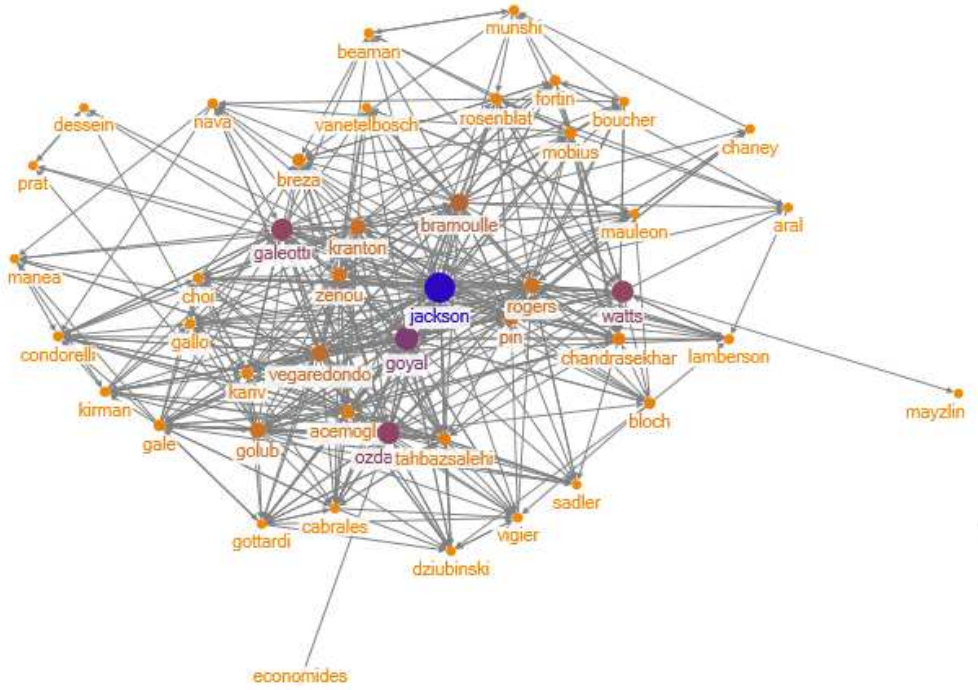



Figure 4.3.1.3: The directed and unweighted network graph of the authors in *The Oxford Handbook of The Economics of Networks* based on their mutual citations ordered by the betweenness centrality.

Above, it is given the 5 highest and lowest betweenness centrality with the authors in our data. Jackson with 237,588 value; is the forefront of high betweenness centrality. It is more likely for Jackson to take place in a shortest path of a randomly chosen pair of authors. Goyal follows him with the value of 136,618. This can be interpreted that Jackson and Goyal have the large amount of control over the relationships of other authors. In contrast, Economides and Mayzlin has the 0 value of betweenness centrality which exhibits that there is no occurrence of them in any shortest path of any pair of nodes. Thus Economides and Mayzlin have no probability to occur in a geodesic path between randomly chosen two authors. Since Having high betweenness centrality for a node in our network represents their controls over mutual references. Omitting these high betweenness nodes may destroy the citation transfers among other nodes as their positions.

| Nodes | Closeness | |
|---|---|---|
| Jackson | 0,02 | Best |
| Goyal | 0,019 | Second Best |
| Ozdaglar | 0,017 | Third Best |
| Galeotti | 0,017 | Fourth Best |
| Vega Redondo | 0,017 | Fifth Best |
| Prat | 0,011 | Fifth Worst |
| Dessein | 0,011 | Fourth Worst |
| Chaney | 0,011 | Third Worst |
| Economidies | 0,01 | Second Worst |
| Mayzlin | 0,009 | Worst |

Table 4.3.1.4: Best and worst five nodes in the network in terms of closeness centralities.

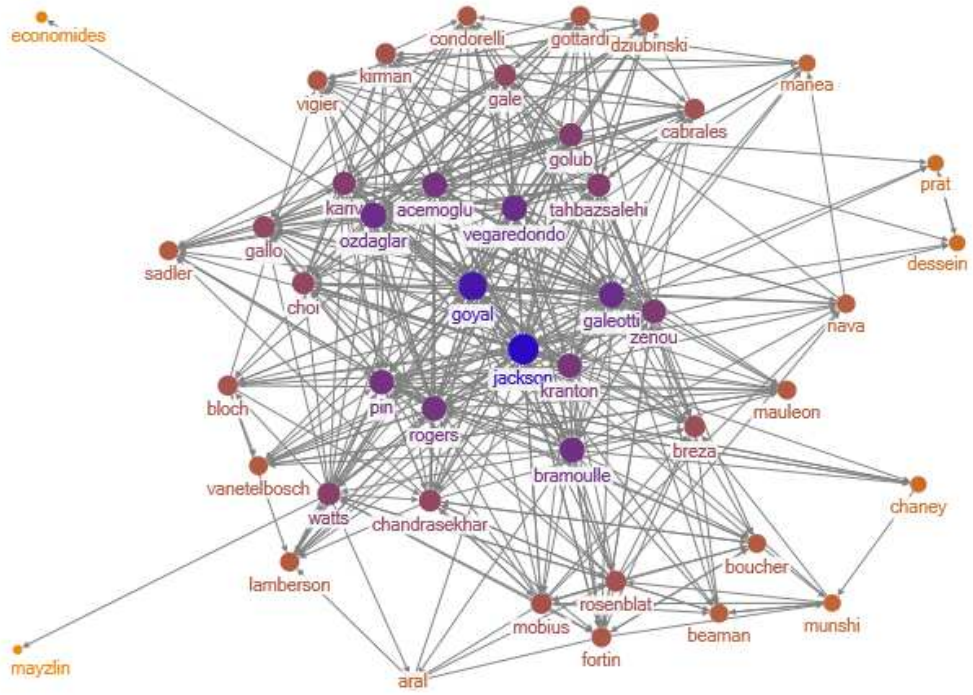



Figure 4.3.1.4: The unweighted network graph of the authors in *The Oxford Handbook of The Economics of Networks* based on their mutual citations ordered by the closeness centrality.

Considering our data, Jackson has the highest closeness degree with 0,02, whereas Mayzlin has the lowest degree with 0,009. Nevertheless, as the Table 4.3.1.4 depicts, they do not differ remarkably from each other. Therefore, for Jackson it easier to access the other authors than any one in the network. Controversely, since the total sum of shortest distances of Mayzlin from the other authors is higher than any one in network, for him it is not that easy to access to other authors in the network.

| Nodes | Eigenvector | |
|---|---|---|
| Jackson | 0,044 | Best |
| Goyal | 0,042 | Second Best |
| Ozdaglar | 0,037 | Third Best |
| Vegaredondo | 0,037 | Fourth Best |
| Acemoglu | 0,036 | Fifth Best |
| Munshi | 0,007 | Fifth Worst |
| Prat | 0,006 | Fourth Worst |
| Dessein | 0,006 | Third Worst |
| Economides | 0,002 | Second Worst |
| Mayzlin | 0,001 | Worst |

Table 4.3.1.5: Best and worst five nodes in the network in terms of eigenvector centralities.

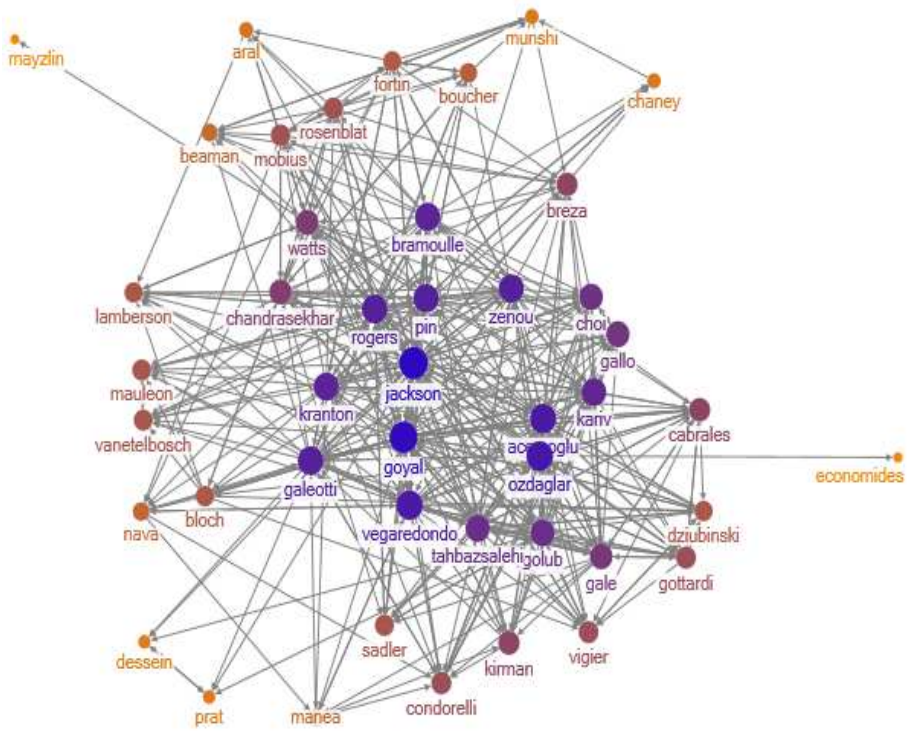



Figure 4.3.1.5: The directed and unweighted network graph of the authors in *The Oxford Handbook of The Economics of Networks* based on their mutual citations ordered by the eigenvector centrality.

Results obtained by the data deduce that Jackson is 44 times more important than Mayzlin in terms of Eigenvector centrality. According to the eigenvector centrality measurement results, Jackson receives 0.044 points, the highest amount of Eigenvector centrality, while Mayzlin receives 0.001 points. Another important point is, even Mayzlin and Economidies have the same in degree, out degree and betweenness centrality measures as 1,0,0 respectively, they differ for the measurement of Eigenvector centrality. Mayzlin has 0.001 point of Eigenvector centrality measure whereas Economidies has 0.002. Also, while the only connection of Economidies is Ozdaglar, Mayzlin's only connection is Watts. (See Figure 4.3.1.7 and 4.3.1.8) This indicates that Ozdaglar is a more important author than Watts. Indeed when the data is examined, one can see that the Eigenvector centrality of Ozdaglar is 0.037 whereas Watts' is 0.027. On the contrary of Eigenvector centrality measures, Mayzlin is 9 times more important than Economidies in terms of Closeness centrality, 0.009 and 0.001 respectively. In conclusion, the importance level of authors in terms of their citations might change depending on the method of measurement used.

| | Mayzlin | Economides |
|---|---|---|
| In degree | 1 | 1 |
| Outdegree | 0 | 0 |
| Betweenness | 0 | 0 |
| Closeness | 0,001 | 0,002 |
| Eigenvector | 0,009 | 0,001 |

Table 4.3.1.6: Unweighted centralities for Mayzlin and Economides.

| | Ozdaglar | Watts |
|---|---|---|
| Eigenvector | 0,037 | 0,027 |

Table 4.3.1.7: Eigenvector centrality for Ozdaglar and Watts.

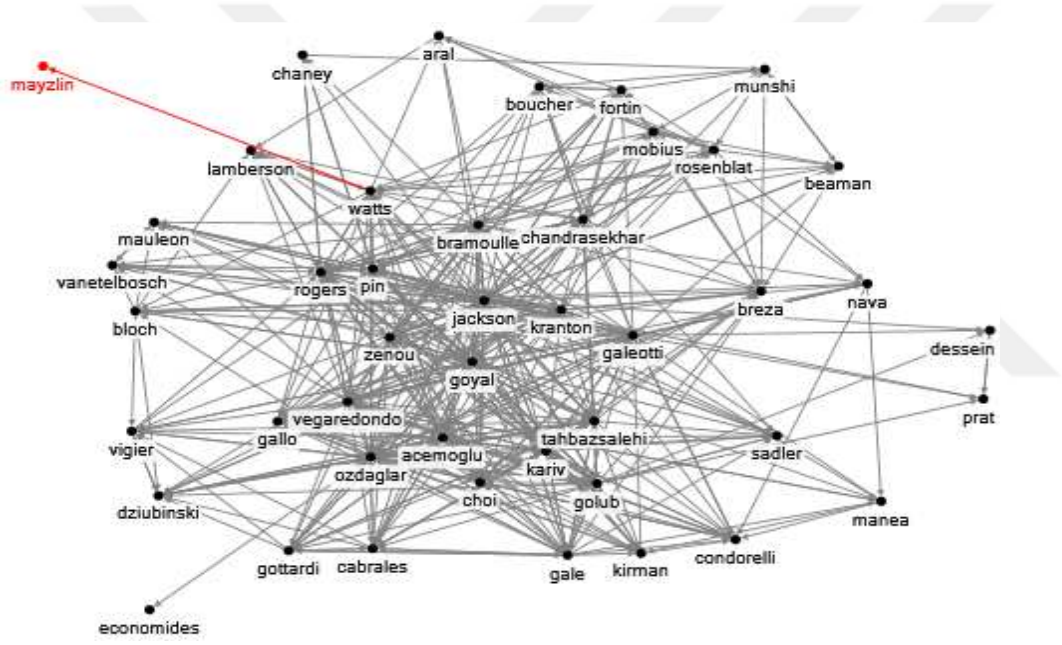



Figure 4.3.1.7

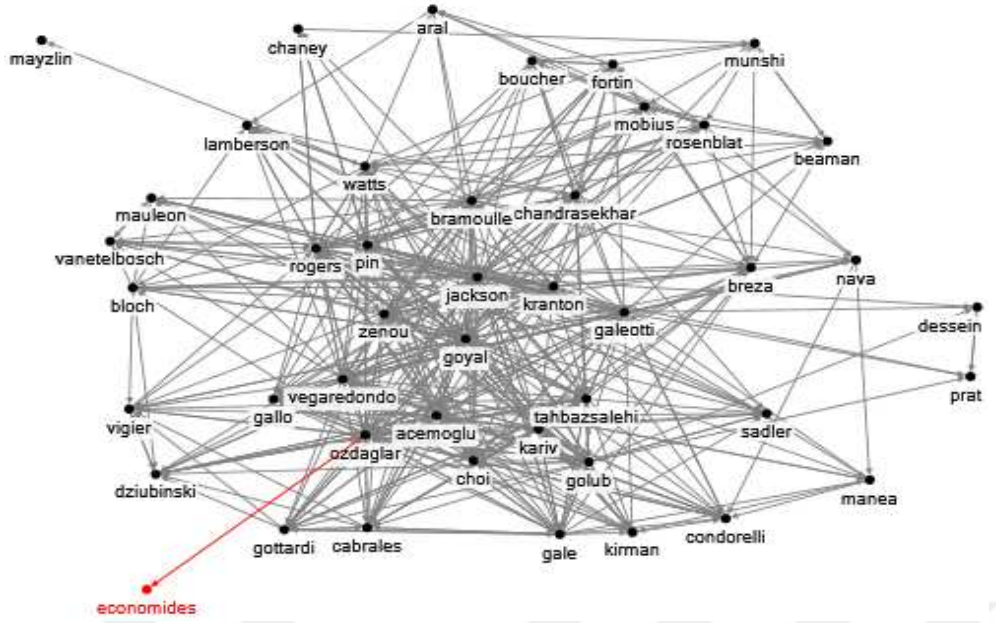



Figure 4.3.1.8

### 4.3.2. DESCRIPTIVES FOR WEIGHTED NETWORK

So far the centrality measures listed above belong to the unweighted forms of the network. They all refer to the relationship among authors with respect to the citations. Nevertheless, the interpretation of the results derived from the unweighted graph covers solely whether an author cites the other authors and is cited by other authors or not. These measurements do not take into account the importance of the citations (i.e the weights of the citations). The approach used above determines the importance or centrality of an author in the network, assuming all citations have the same weight. And also, it does not take into consideration if an author is cited more than once by the same author. Basically, it only focuses on the existence of relationship between authors based on citing or being cited. This is an obstacle to interpret, regarding the weight of citations which is essential for citation analysis. Therefore, in the second part we are going to discuss the directed and weighted form of the network. Because these forms allow us to analyse the citation network more precisely. From now on, citations will occur with respect to their weights which we have evaluated in our study with the method that mentioned in methodology part.

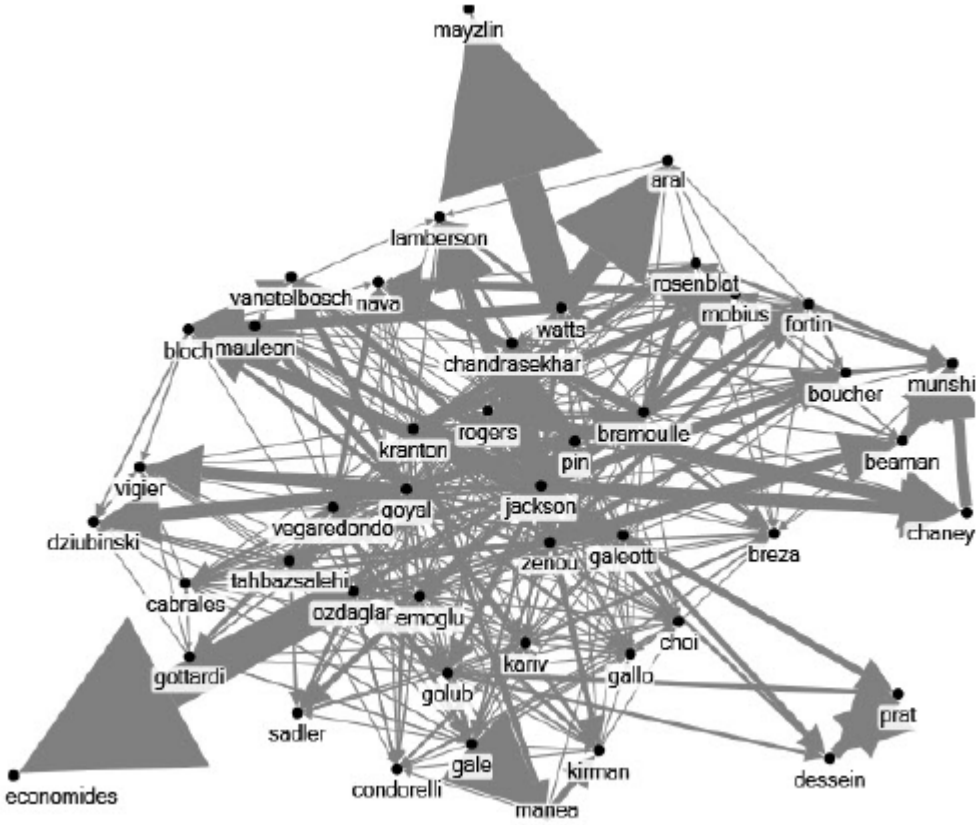



Figure 4.3.2.1: General weighted graph of the network in terms of citations weights.

From the Figure 4.3.2.1, it can be seen that not all the edges have the same thickness. They can be distinguished from each other by the weight of citations. The thickness of the edges depicts how important the citations are that the method for importance of citations' determination given in methodology part. For instance, Nava has cited Bloch, Bramoulle, Kranton, Mobius, Rosenblat and Vega Redondo since the ties are directed to Nava. One can see from the Figure 4.3.2.1 that the citations taken from Bramoulle and Rosenblat do not have the same weights. Since the weights have remarkably different sizes, it is easy to observe without numerically examining that Bramoulle's references to Nava is more weighted (i.e important) than Rosenblat's. (See Figure 4.3.2.2). When we quantitatively compare the derived results of citation weights (0,311 and 0,052), they noticeably differ from each other.

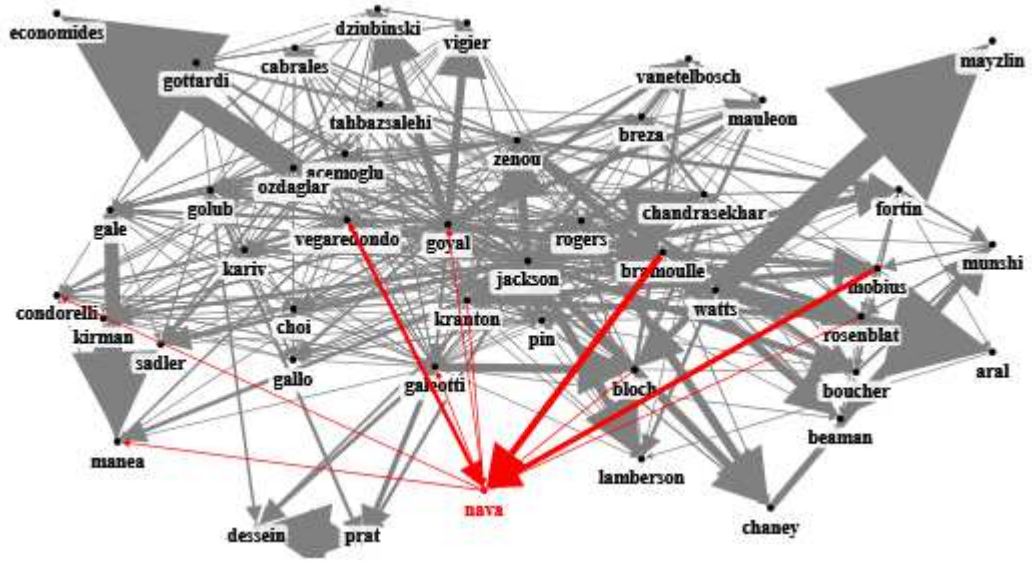



Figure 4.3.2.2

| Node | Out alpha = 0.5 | In alpha = 0.5 | Out alpha = 1 | In alpha = 1 | Out alpha = 1.5 | In alpha = 1.5 |
|---|---|---|---|---|---|---|
| Acemoglu | 5,2990132 | 4,107215 | 1,337121 | 0,9923068 | 0,33740104 | 0,2397422 |
| Aral | 0,3464102 | 2,246691 | 0,06 | 1,0095238 | 0,0103923 | 0,4536175 |
| Beaman | 1,2764777 | 2,645751 | 0,5431318 | 1 | 0,23109853 | 0,3779645 |
| Bloch | 2,9155984 | 2,236068 | 0,9445238 | 1 | 0,30598357 | 0,4472136 |
| Boucher | 1,4142821 | 2,828427 | 0,4000388 | 1 | 0,11315353 | 0,3535534 |
| Bramoulle | 7,6872279 | 3,013861 | 2,3637389 | 1,0092621 | 0,726824 | 0,3379751 |
| Breza | 0 | 4,132776 | 0 | 1,0046962 | 0 | 0,2442461 |
| Cabrales | 1,9759662 | 3,331734 | 0,3904442 | 1,0091321 | 0,07715046 | 0,3056509 |
| Chandrasekhar | 4,0908184 | 3,159196 | 1,0459247 | 0,998052 | 0,26741799 | 0,3153042 |
| Chaney | 1,1629256 | 2 | 0,4507987 | 1 | 0,17474844 | 0,5 |
| Choi | 1,3831159 | 4,241207 | 0,3188349 | 0,9993245 | 0,07349761 | 0,2354635 |
| Condorelli | 0,8260862 | 3,605551 | 0,1364837 | 1 | 0,02254945 | 0,2773501 |
| Dessein | 0,6164414 | 1,999474 | 0,38 | 0,9994737 | 0,23424773 | 0,4996053 |
| Dzibuinski | 1,0647087 | 3,316625 | 0,2267209 | 1 | 0,04827835 | 0,3015113 |
| Economidies | 0 | 1 | 0 | 1 | 0 | 1 |
| Fortin | 2,2845401 | 2,828427 | 0,5799026 | 1 | 0,14720119 | 0,3535534 |
| Gale | 5,4227836 | 3,323189 | 1,7297989 | 1,0039623 | 0,55178385 | 0,3033051 |
| Galeotti | 7,1542984 | 3,73935 | 2,0473594 | 0,9987671 | 0,58589681 | 0,2667671 |
| Gallo | 0,6440031 | 4,242641 | 0,20737 | 1 | 0,06677345 | 0,2357023 |
| Golub | 5,8524694 | 3,162278 | 1,4891912 | 1 | 0,37893243 | 0,3162278 |
| Gottardi | 1,4221694 | 3,316625 | 0,2528207 | 1 | 0,04494424 | 0,3015113 |
| Goyal | 10,6237582 | 4,971196 | 4,0308657 | 0,9885117 | 1,52939079 | 0,1965634 |
| Jackson | 16,074413 | 4,548642 | 7,1774098 | 0,9852448 | 3,20479584 | 0,213406 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Kariv | 3,1087261 | 4,242641 | 0,8053482 | 1 | 0,20863391 | 0,2357023 |
| Kirman | 1,6009266 | 3,00039 | 0,366138 | 1,0002597 | 0,08373715 | 0,3334632 |
| Kranton | 6,8293538 | 2,998926 | 1,9433364 | 0,9992843 | 0,55298883 | 0,3329756 |
| Lamberson | 0,7776113 | 3,161516 | 0,2015598 | 0,9995181 | 0,05224506 | 0,3159992 |
| Manea | 1,679405 | 2 | 0,5640802 | 1 | 0,18946382 | 0,5 |
| Mauleon | 0,5994459 | 3,000297 | 0,1197785 | 1,000198 | 0,02393357 | 0,3334323 |
| Mayzlin | 0 | 1 | 0 | 1 | 0 | 1 |
| Mobius | 3,2460567 | 2,45158 | 0,8105295 | 1,0017073 | 0,20238652 | 0,4092943 |
| Munshi | 1,8331327 | 2 | 0,6720751 | 1 | 0,24640057 | 0,5 |
| Nava | 1,4162513 | 2,448704 | 0,4011535 | 0,9993582 | 0,11362685 | 0,4078553 |
| Ozdaglar | 7,3360582 | 4,12408 | 2,3399022 | 1,0004725 | 0,74633299 | 0,2427076 |
| Pin | 2,2484008 | 4,833012 | 0,4595733 | 1,0155656 | 0,09393682 | 0,2134018 |
| Prat | 0,6882472 | 2,005228 | 0,4736842 | 1,0052345 | 0,32601183 | 0,503931 |
| Rogers | 4,4817716 | 4,686407 | 1,1815457 | 0,9982912 | 0,31149518 | 0,2126545 |
| Rosenblat | 2,4918411 | 2,454175 | 0,4776363 | 1,003829 | 0,09155336 | 0,4105953 |
| Sadler | 0,6519289 | 3,162278 | 0,1416704 | 1 | 0,03078635 | 0,3162278 |
| Tahbaz Salehi | 3,626627 | 4,114564 | 0,8220264 | 0,995861 | 0,18632396 | 0,2410314 |
| Vannatelbosch | 0,6205874 | 2,999849 | 0,1283762 | 0,9998995 | 0,02655622 | 0,3332831 |
| Vegaredondo | 6,9372322 | 3,738402 | 1,9250076 | 0,9982609 | 0,534169 | 0,2665643 |
| Vigier | 0,6331187 | 3,316625 | 0,1002098 | 1 | 0,01586118 | 0,3015113 |
| Watts | 6,2262716 | 3,173551 | 2,5844305 | 1,0071429 | 1,07275774 | 0,319622 |
| Zenou | 7,2506052 | 3,458911 | 2,3896035 | 0,9970056 | 0,7875487 | 0,2873795 |

Table 4.3.2.1: The nodes in weighted network interms of their in-degree and out-degree centralities with different α values, ordered by the colours.

From the results given above, it is observed that, Jackson and Goyal have the highest out-degree for all α values. It means that, when the weights is considered, no matter it is consideration is higher or lower than number of edges, Jackson and Goyal is the most important authors in the *Handbook* according to citations they give. However, we see the inverse order for the in-degree centrality when α is 1 and 1,5. When α = 1 (i.e the only consideration is weights of citations), Jackson and Goyal is in the worst 5 regarding their in-degree centralities. In other words, the citations from other authors to Jackson and Goyal has less weighted comparing to the other mutual citations. This make sense when we think that the citations from Jackson and Goyal to the other authors have the highest weights. Similarly, when α = 1,5 (i.e weights of citations is more important than the number of them); Jackson and Goyal is in the worst 5 list according to their in-degree centralities. On the other hand, Aral has one of the lowest centralities according to out-degree centralities of him when α =

0,5; 1; 1,5, whereas he has one of the highest in-degree level when α = 1. One should not forget that, in our study out-degree is a more reliable measure to order the authors by their importance level regarding their citations.

| Node | Closeness, α = 1 | Closeness α = 0,5 | Closeness α = 1,5 |
|---|---|---|---|
| Acemoglu | 0,015 | 0,0171 | 0,010596509 |
| Aral | 0,0056 | 0,0099 | 0,003423659 |
| Beaman | 0,0123 | 0,0122 | 0,010782704 |
| Bloch | 0,0156 | 0,0157 | 0,01280165 |
| Boucher | 0,01 | 0,0108 | 0,008951983 |
| Bramoulle | 0,0171 | 0,0196 | 0,013384499 |
| Cabrales | 0,0088 | 0,0124 | 0,006036312 |
| Chandrasekhar | 0,0167 | 0,017 | 0,013086093 |
| Chaney | 0,0114 | 0,0114 | 0,009981829 |
| Choi | 0,0075 | 0,0115 | 0,005141981 |
| Condorelli | 0,0072 | 0,0105 | 0,003940575 |
| Dzibuinski | 0,0105 | 0,0126 | 0,006825004 |
| Fortin | 0,0102 | 0,0128 | 0,008969608 |
| Gale | 0,0152 | 0,0153 | 0,012627236 |
| Galeotti | 0,0173 | 0,0192 | 0,01314883 |
| Gallo | 0,0079 | 0,0094 | 0,005523808 |
| Golub | 0,0141 | 0,0173 | 0,009420382 |
| Gottardi | 0,007 | 0,0111 | 0,003259258 |
| Goyal | 0,0313 | 0,0274 | 0,02861829 |
| Jackson | 0,0465 | 0,0355 | 0,049657837 |
| Kariv | 0,0106 | 0,0133 | 0,007153013 |
| Kirman | 0,0106 | 0,0131 | 0,007366471 |
| Kranton | 0,0171 | 0,0188 | 0,013101222 |
| Lamberson | 0,011 | 0,0114 | 0,009445208 |
| Manea | 0,0143 | 0,014 | 0,011849466 |
| Mauleon | 0,0039 | 0,0073 | 0,001564735 |
| Mobius | 0,0128 | 0,015 | 0,009933713 |
| Munshi | 0,0118 | 0,0119 | 0,010383672 |
| Nava | 0,0131 | 0,0134 | 0,010176377 |
| Ozdaglar | 0,0144 | 0,0166 | 0,010341121 |
| Pin | 0,0118 | 0,0147 | 0,007846943 |
| Rogers | 0,0191 | 0,0178 | 0,016409331 |
| Rosenblat | 0,0096 | 0,0122 | 0,006831054 |

| | | | |
|---|---|---|---|
| Sadler | 0,0081 | 0,0108 | 0,004856159 |
| Tahbaz Salehi | 0,0122 | 0,0154 | 0,007846215 |
| Vannatelbosh | 0,0035 | 0,0072 | 0,001458647 |
| Vega Redondo | 0,0205 | 0,02 | 0,016913367 |
| Vigier | 0,0069 | 0,0106 | 0,003313008 |
| Watts | 0,0151 | 0,0155 | 0,012711667 |
| Zenou | 0,0248 | 0,0221 | 0,022481502 |

Table 4.3.2.2: The nodes in weighted network interms of their closeness centralities with different α values, ordered by the colours.

In Table 4.3.2.2, the authors of *The Oxford Handbook of The Economics of Networks* ordered with respect to their closeness centrality values for α = 1, α = 0,5, α = 1,5. From the results, Jackson is the most central author in the network no matter α he has. His closeness degree increase when we increase α from 0,5 to 1,5. His closeness degree is higher if we consider longer path with high weighted ties is better than shorter path with less weighted ties. On the other hand, Vannatelbosh has the lowest closeness degree for all α's. His closeness is greater if we decrease α = 1,5 to α = 0,5 . Than means if one suggests that high weighted and longer edges are less preferable than less weighted and shorter edges between two nodes Vannetelbosh will be better of in centrality.

For Gallo; he has the third lowest closeness degree centrality when α is 0,5. However, for the other α's he is not in the worst five. Deducting from this result that if a high weighted edges supposed no less important than a short distance edge to measure closeness, than Gallo is not going to participate the *worst* list. And we can understand from that, he is better at connecting with high weighted citations to the other authors than occuring in a closer distance position to them.

| Node | Betweenness α = 1 | Betweenness α = 0,5 | Betweenness α = 1,5 |
|---|---|---|---|
| Acemoglu | 12 | 31 | 14 |
| Aral | 0 | 0 | 0 |
| Beaman | 90 | 37 | 125 |
| Bloch | 219 | 57 | 386 |
| Boucher | 1 | 2 | 2 |
| Bramoulle | 75 | 63,5 | 105 |
| Breza | 0 | 0 | 0 |

| | | | |
|---|---|---|---|
| Cabrales | 0 | 0 | 0 |
| Chandrasekhar | 14 | 39 | 5 |
| Chaney | 52 | 34 | 92 |
| Choi | 0 | 17 | 0 |
| Condorelli | 0 | 0 | 0 |
| Dessein | 0 | 0 | 0 |
| Dzibuinski | 0 | 0 | 0 |
| Economidies | 0 | 0 | 0 |
| Fortin | 2 | 7 | 0 |
| Gale | 95 | 41 | 143 |
| Galeotti | 149 | 94 | 195 |
| Gallo | 0 | 0 | 0 |
| Golub | 131 | 67 | 174 |
| Gottardi | 0 | 0 | 0 |
| Goyal | 308 | 220 | 470 |
| Jackson | 383 | 420 | 450 |
| Kariv | 95 | 44 | 100 |
| Kirman | 0 | 4 | 0 |
| Kranton | 43 | 42,5 | 25 |
| Lamberson | 33 | 3 | 43 |
| Manea | 86 | 12 | 130 |
| Mauleon | 18 | 0 | 43 |
| Mayzlin | 0 | 0 | 0 |
| Mobius | 100 | 41 | 123 |
| Munshi | 132 | 56 | 199 |
| Nava | 91 | 30 | 145 |
| Ozdaglar | 40 | 39 | 39 |
| Pin | 58 | 20 | 84 |
| Prat | 0 | 0 | 14 |
| Rogers | 92 | 143 | 74 |
| Rosenblat | 0 | 0 | 0 |
| Sadler | 0 | 0 | 0 |
| Tahbaz Salehi | 0 | 0 | 0 |
| Vannatelbosch | | | 0 |
| Vegaredondo | | | 195 |
| Vigier | | | 0 |
| Watts | | | 337 |
| Zenou | | | 17 |

Table 4.3.2.3: The nodes in weighted network interms of their betweenness centralities with different α values, ordered by the colours.

In Table 4.3.2.3, we ranked the authors according to their betweenness centrality measures for $\alpha = 1$, $\alpha = 0{,}5$ and $\alpha = 1{,}5$. Jackson has the highest betweenness for $\alpha = 1$ and $\alpha = 0{,}5$; but this time Goyal has the highest betweenness when $\alpha = 1{,}5$. Therefore, we can say that Goyal is more likely to be in the closest path between two authors than Jackson if a path between two authors with a high weight and long distance is considered *shorter* than a path with low weight and high distance. As we have seen there are lots of zeros for all α's, for that reason the comparison of the last five node is not meaningful. We only say that the ones with 0 betwenness do not occur in the shortest paths of any pair of author. However, from the results Mauleon has 0 betwenness for $\alpha = 0{,}5$ while he has 18 and 43 for $\alpha = 1$ and $\alpha = 1{,}5$ respectively. If we increased α from 0,5 to 1; in other words, if the only criteria to detect the *shortest distance* will be weights of edges (i.e weights of mutually citations.) instead of a criteatia combines both distances and weights but gives more importance to closer distances than high weighted edges between two nodes, Mauleon gains positive betwenness centrality. And if we consider a criteria that combines both weights and distances but give more account on weights to detect the *shortest path* between two nodes; his possibility to occur in a *shortest path* of a randomly chosen two nodes has increased.

From the results we can observe that Choi has occured in some authors' *shortest distance* to the other authors, if only in the criteria which combines distances and weights but gives more value to short distances than high weights between two nodes.

## 5. REFERENCES

Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A. (2004) The architecture of complex weighted networks. Proceedings of the National Academy of Sciences 101 (11), 3747-3752

Bramoullé, Y., Galeotti, A., Rogers, B.W., Eds. (2016) The Oxford Handbook of The Economics of Networks, Oxford University Press.

Brandes, U. (2001) A Faster Algorithm for Betweenness Centrality. Journal of Mathematical Sociology 25 (2): 163–177.

Brath, R. and Jonker, D. (2015) Graph analysis and visualization: Discovering business opportunity in linked data. Hoboken, NJ, United States: John Wiley & Sons.

Dijkstra, E. W. (1959) A note on two problems in connexion with graphs. Numerische Mathematik 1, 269-271.

Freeman, L. C. (1978) Centrality in social networks: Conceptual clarification. Social Networks 1, 215-239

Igraphdata: Friendship network of a UK university faculty (2010) Available at: https://rdrr.io/cran/igraphdata/man/UKfaculty.html (Accessed: 19 October 2016).

Jackson, M. O. (2008) Social and Econmic Networks, Princeton University Press..

Jackson, M. O. (2014) Networks in the understanding of economic behaviors, Journal of Economic Perspectives 28, 3-22.

Nepusz T., Petroczi A., Negyessy L., Bazso F.: Fuzzy communities and the concept of bridgeness in complex networks. Physical Review E 77:016107, 2008.

Newman, M.E.J. (2010) Networks: An introduction. New York: Oxford University Press.

Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Physical Review E 64, 016132.

Nykamp, D.Q. An introduction to networks. http://mathinsight.org/network_introduction

Openflight.org Available at: http://Openflight.org

Opsahl, T., Agneessens, F., Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32, 245-251.

Palacios-Huerta, I. and Volij, O. (2004) 'The measurement of intellectual influence', Econometrica, 72(3), pp. 963–977. doi: 10.1111/j.1468-0262.2004.00519.x.

Pinski, G. and Narin, F. (1976) 'Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics', Information Processing & Management, 12(5), pp. 297–312. doi: 10.1016/0306-4573(76)90048-0.

Weisstein, E. W.: Least Squares Fitting - Power Law. MathWorld - A Wolfram Web Resource (http://mathworld.wolfram.com/LeastSquaresFittingPowerLaw.html).