A SPATIAL ANALYSIS OF INTERNATIONAL WORLD EXCHANGE USING COMPLEX NETWORKS ANALYSIS

Renata-Graziela BOAR[1], Alexandra FORTIŞ[2], Alexandru IOVANOVICI[1], Horia CIOCARLIE[1]

[1] **Politehnica University Timişoara, Department of Computers and Information Technology**
[2] **Tibiscus University of Timişoara, Department of Computer Science**

Corresponding author: Renata-Graziela BOAR, renata.boar@gmail.com; renata.boar@cs.upt.ro

ABSTRACT: Modern economies are highly interconnected systems dominated by financial exchange. We propose a novel approach of analyzing the structure of the world financial markets by applying methodologies from the field of complex network analysis on a data set comprising he evolution of exchange rates on 23 major currencies over a time span of 14 years. Interesting economical insight is given based only on the topological properties of the correlation network which validate our methodology.
KEY WORDS: econometrics, graphs, scale free networks, currency, exchange rate, community structure, correlation.

1. INTRODUCTION

Economy is a major constant in our everyday lives. Most of our decisions and actions are being influenced by various economical factors, with a great deal of emphasis on economic indicators. Getting started with payments for food and up to running a big business or buying a house our decisions are governed by what is usually called economy. There are various types of "economies" but from our point of view of particular interest for research is the macroeconomy. In the field of macroeconomy, seen as the abstract representation of the functioning of the economy of various state, there is a great deal of interest in the analysis of the relationships between the currency exchange of various entities. There are a few classical approaches to this investigation and we are going to give a short overview, but as a common denominator, all of them rely on statistics.

In the last years we can observe a rapid rising in booth academia and the industry of so-named complex networks or when dealing with aspected related to the human nature and behaviour social networks. A large amount of literature exists on what is usually called the "new network science" [AB02] Barabasi presents in his seminal paper [AB02] a methodology which allows us to model anything that can be expressed as a relationship between some entities and provides us with another insightful means of exploring and understating this newly structured data consisting of new metrics and visualizations of the data. Instead of simply visualizing a bunch of numbers we are now capable of visualizing the dependency between data and by using the specific metrics we can obtain a better understanding of the influence of various nodes/entities to the entire network.

The structure of the paper is as follows: in Section 2 we present a general view of the most important research directions in this field, Section 3 presents the way we analyzed the raw data sets, construction of the associated network and subsequent specific analysis we carried on the relationships between the currency exchange rate of 23 currencies, with novel pertinent visualizations and in Section 4 we carry a discussion on some of the results and draw specific conclusions.

2. STATE OF THE ART

As presented above, complex network lie at the crossroads of many important scientific field and as a consequence it would be expected the economics to benefice from applying similar analysis techniques in order to answer to their specific questions. Econometrics is a major field of study which draws its roots from economics and mathematics, especially statistics. When working with various time series, on of the major desires of the practitioners and academics alike is the possibility of predicting the trend of a time series. Major investigations in this area were done using tools and methods designed for statistical analysis.

Our work is mostly based on the research carried by Mantegna [Man99] which involved building a matrix of correlation indices with data taken form stock trade and transactions. This methodology further allowed building the network of trade interactions. Tuminello et al. present in [T+05] an extension of Mantegna's work which actually can be presented as "generation the planar maximal graph" which in this case was able to carry more information to the analyst.

When working with correlation matrices and associated graphs the key point is to find meaningful data which you have to correlates. It's a classic axiom that “correlation does not mean causality” but when chosen well it can get insightful data. Kennet et al. present in [K+10] a methodology for constructing correlation matrix using only the partial correlation. We use the threshold correlation presented there in order build a high pass filter on the complete graph. Mathematically everything is correlated with everything else, but from a practical point of view most of the time the correlation is so weak we can discard it.

The financial crisis of the last years provided the motor for a plethora of studies in the field of econometrics. Kennet and his team present in [K+11] another interesting research carried also on data provided by stock market which uses the Dow Jones index and they were able to identify strictly algorithmically some episodes of crisis.

Similar initiatives to our one which exploit also the methodology of complex network analysis in order to represent phenomenon of the global financial market are presented in [KR07, L+11 and Car13]. Kali developed this work on data taken from international trade and he tries to explain the roots of the financial crisis, finding what is called a “contagious connection” between some smaller crisis which occurrence in Mexico, Russia and Asia [L+11]. Lee in [L+11] addressed also the problem of global financial crisis through a novel approach, more explicit, the connection between county’s GDP and the amount of economical exchange between countries. The authors find a “strong correlation between the size of the GDP and the force of the blow of that particular country to the crisis”.

In [H+07] the authors also present an algorithm for comparing various countries by the structure of the products and goods.

Table 1: Currency codes and corresponding countries

Currency code	Country	Currency code	Country
AUD	Australian Dollar	LKR	Sri Lankan Rupee
BRL	Brazilian Real	MXN	Mexican Peso
CAD	Canadian Dolar	MYR	Malaysian Ringgit
CHF	Swiss franc	NOK	Norvegian Krona
CNY	Chinese Yuan	NZD	New Zealand Dolar
DKK	Danish Krone	SEK	Swedish Krona
EUR	Euro	SGD	Singapore Dolar
GBP	Great Britain Pound	THB	Thai Baht
HKD	Hong Kong Dolar	TWD	Taiwan New Dolar
INR	Indian Rupee	VEB	Venezualan Bolivar
JPY	Japanese Yen	ZAR	South African Rand
KRW	South Korean Wo		

3. METHODOLOGY AND RESULTS

Our work is based on the dataset retrieved from European Commission's Eurostat Database and consists of the currency exchange rates of 23 major currencies in report to the US Dollar. Data cover the time span from 1st of January 2003 and up to 1stof August 2017 with a resolution of 1/day. There are 3646 data points for each of the 23 currencies. For now on we are going to refer currency arrays by their internationally consecrated 3 letter abbreviation (see Table 1).

Our investigation is going to present a two fold approach: one is the classical one regarding the temporal evolution of the data and the correlogram visualization for the index of correlation and the other one, based on the complex network analysis is going to show us the spatial distribution of the data, emphasizing on the clusters of strongly inter-correlated sets (currencies).

3.1 Pearson correlation visualization and interpretation

Based on statistics a lot of researches found in literature [AB08] uses the Pearson correlation get insight on how two data sets are relate. This mean we want to find if there is a link between the direction of trend of one set in relation to the other.

We applied formula (1) taken from [Car13] for computing the correlation between all the pairs of currencies.

$$R_{i,j} = \frac{<(y_i - <y_i>) - (y_j - <y_j>)>}{\sigma_i \sigma_j} \qquad (1)$$

where $< >$ is used to signify the average of data set y_i and σ_i and σ_j are correspondingly the standard deviation of data sets y_i.

By computing the formula (1) on all distinct datasets we obtained *253* correlation values on which have we charted initially a simple histogram in order to visualize the distribution of data (Figure 1a).

One can observe that the distribution looks similar to a normal distribution half of the samples having a threshold greater than 0.3

A newer approach in displaying and interpreting correlation matrices is the so-called correlogram (Figure 2). Introduced by Friendly in 2002 [Fri02], it is an alternative to the raw matrix, keeping in the same time the spatial characteristics of a matrix but adding the element of color and shade.

A corrgram is read as like: the output is a matrix, with the main diagonal populated with the indices of the datasets, while the SE corner is depicted as rectangles of various shades, the lightest shade having the smaller value of correlation and the most intense one having the highest.

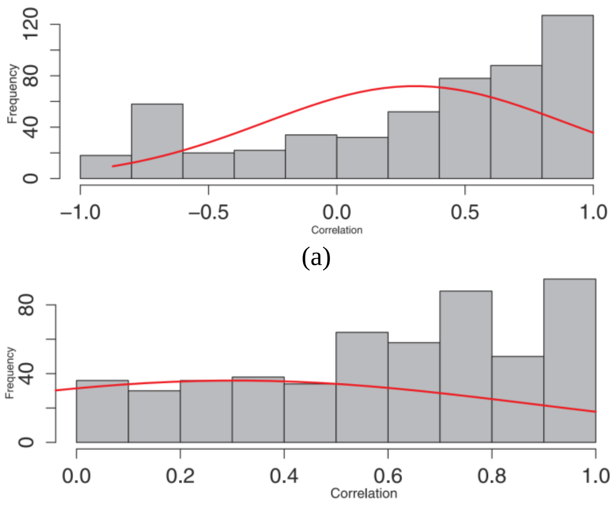


Figure 1. The statistical analysis of the correlation indices was done as a histogram in order to show the composition of the samples. One can perceive the almost normal distribution of the real values (1a). Because we are more interested in the strength of the correlation than of the sign we plotted also the distribution of the absolute values of the correlation (1b). In this case the distribution is much closer to the uniform one, with a bias towards greater values (stronger correlations)

The second, much more interesting approach, is the one based on the theory of complex networks in which we are going to build the corresponding network and offer new insights relying on the topological characteristics.

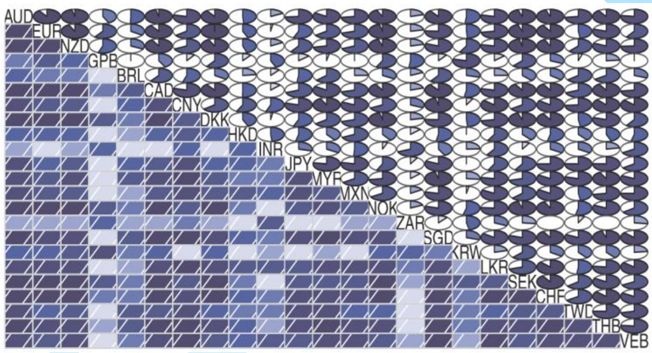


Figure 2. The correlogram of the 23 currency exchange data sets shows a 2D spatial visualization of the correlation between any two sets as shades of color or proportionally filled discs. If we want to see from a qualitative point of view the strength of the correlation, between two sets we have to identify the corresponding two "cells" on the main diagonal and look at the intersection between them in the lower or upper side triangle of the matrix

3.2 Computing the correlation based network

Of them one of the most widely used methods of building networks based on financial time series is the representation of the correlation coefficient in a graphical way. We have to define our graph as a set nodes represented by currency exchange rates *Si = {AUD, BRL, ..., ZAR}* $i = [1,n_s]$, (where n_s is the number of countries, in our case *23*) and the population of undirected links with the weight given by by the correlation index $r_{i,j}$ between sets S_i and S_j. This way a $n_s \times n_s$ matrix, R, is built.

One can observe that we get by this approach a complete graph, (Figure 3a) not yielding significant spatial interpretations. In order to obtain a non-complete graph we have to build a partial graph of the initial one and we do this by imposing a cut-off threshold on the correlation index.

Using a literature referenced threshold of 0.3 [Car13] we filter the edges and obtain

Table 2: Correlation clusters and their color coding

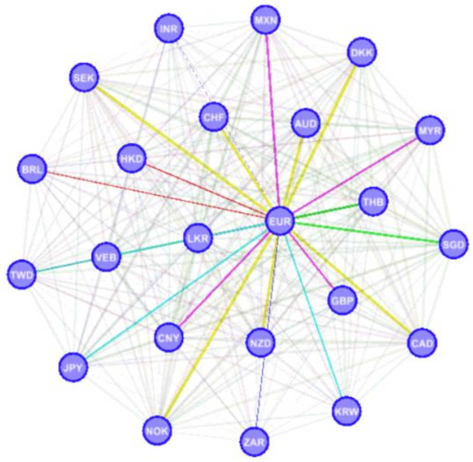
Correlation index	Color
[0.3,0.45)	Yellow
[0.45,0.6)	Red
[0.6,,0.75)	Green
[0.75,0.9)	Blue
>0.9	Black

a non complete graph. We further discretize the range of correlation indices in five disjoint sets, as depicted in table 2. The ranges and colours are taken from [Car13] for a uniform view across all the studies.

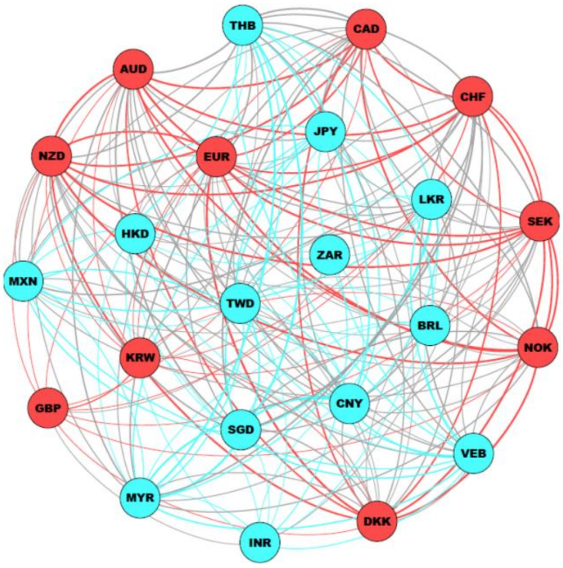
Figure 3b shows the graph after deleting the links with less than *0.3* correlation. On the obtained graph we can perceive a few interesting observations.

The newer network has only *202* edges, which hold for *79.84%* of the initial one and is consistent with the correlation distribution from figure 3a. The average degree is exactly *17.56* which is equivalent to the number of nodes (currencies) which have some kind of connection in their trend between them.

Our investigation does not stop here and in order to get an even more interesting insight into our data we are going to apply some complex network analysis techniques on this graph. Of particular importance and with good results we found to be the algorithm of *Modularity based Community Detection*. This is using the clustering coefficient metric running Newman's modularity metric [New03] in order to detect subgraphs with similar characteristics, in this case having a high average degree in the cluster and a low connectivity with other clusters.



(a) Complete correlation graph with emphasize on the node for the EUR currency



(b) Subgraph of the correlation network showing only edges with a correlation index $c_{i,j} > 0.3$

Figure 3. The correlation network of the currency exchange rates is a complete graph as shown in figure 3a. Each of the nodes is linked to all the others trough edges having the correlation index in place of edge weight. We emphasize on the connections of the EUR currency. One can perceive the varying thickness and color of the edge, dependant of the weight. In figure 3 we show the partial graph obtained after filtering out the edges with the imposed threshold of 0.3. After running the community detection algorithm on the new graph there can be seen two distinct communities (colored in red and respectively blue) of “similar” currencies from a topological point of view

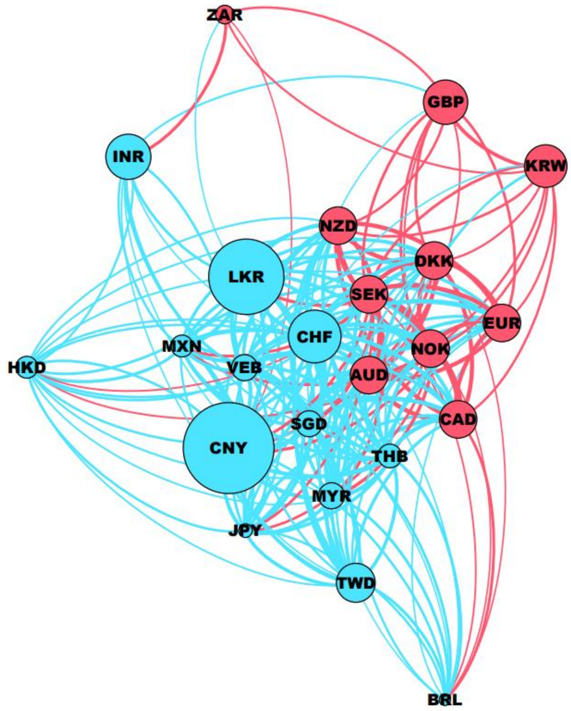


Figure 4. The correlation network with edges of corr < 0.3 filtered out. The two distinct communities can be seen on with different colors, while thickness of the edges is proportional with the strength of the correlation. The size of the nodes is proportional with the betweenness centrality of the node across the entire network. Of particular interest is the node BRL which has good correlation with nodes from both the communities

Running the community detection algorithm with the parameter *resolution=1.0* on the filtered graph we obtain the graph shown in figure 4. There can be seen two distinct communities consisting of red and blue nodes.

Next we oriented towards the spatial visualization of the network using specific tools and algorithms. Even if for the renderings discussed up to new we have used the Fruchterman-Reingold algorithm, one can observe that it is not very useful in analyzing the structural properties of the networks being “visually messy”.

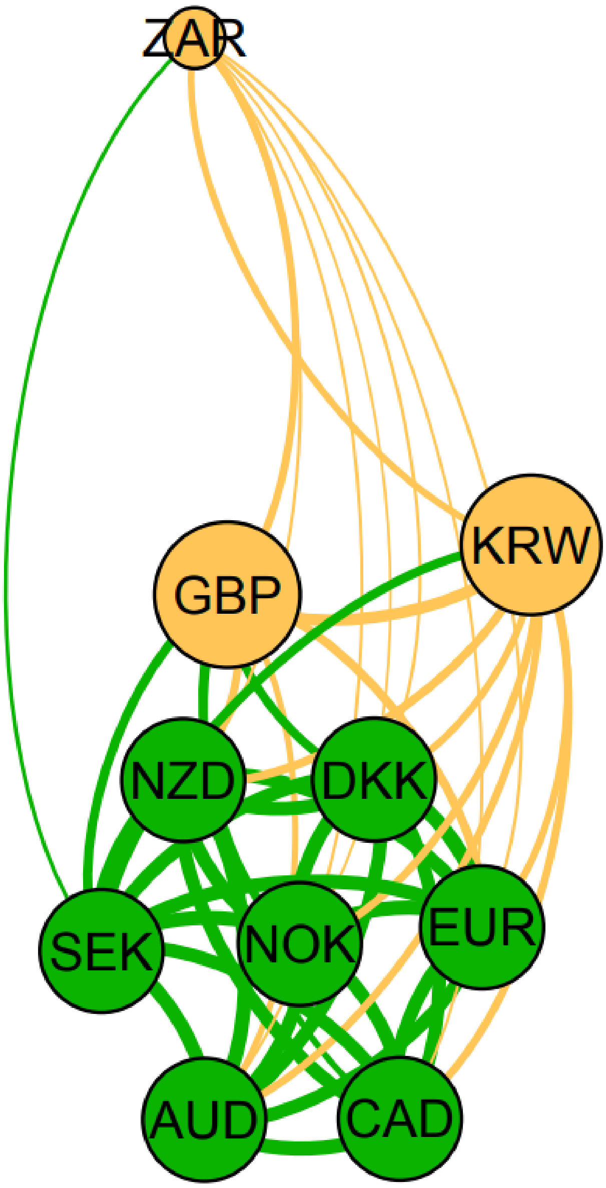
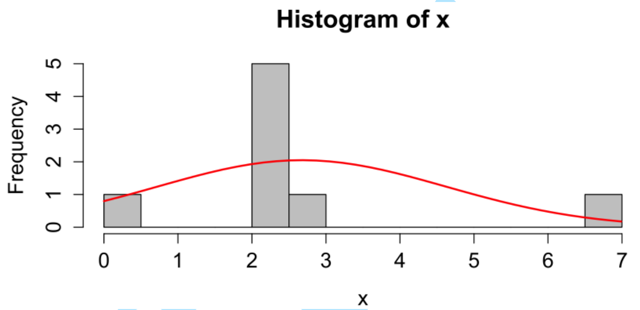


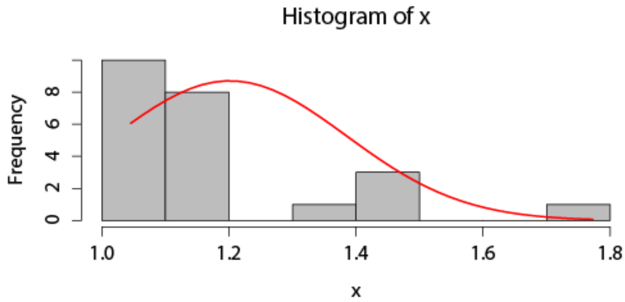
Figure 5. Applying the methodology in a recursive fashion for one of the communities we can observe the hierarchical behaviour of our network. Here we can observe the second level of analysis on the "red-community" from Figure 4. The former cluster of "European currencies" is now subdivided into two more distinct communities with a much better definition of the borders. The "green-community" is consisting of pure anglo-saxon currencies while the orange one is consisting of British Empire currencies

Of greater interest for our investigations is the Force Atlas 2 algorithm which is implemented as a standard plug-in in Gephi. Force Atlas 2's layout is based on the analogy of gravitational forces which attract the nodes and is solving the problem as a many body system in mechanics. There are various parameters to fine tune the behavior of the algorithm, but in our case the most important one is Edge Weight Influence and based on the semantics of our data we swept this parameter into the range from 1 to 10, setting it for the discussion at the value of 4. The obtained rendering can be seen in figure 5↑ and one can perceive the spatial clustering of the two communities with only few "rogue" nodes. The economical implications of these facts are going to be discussed in the next section.

The clustering coefficient is also a specific metric introduced in complex network analysis. Watts et al. defines it in [WS98] as "*a measure of the degree [...] nodes in a graph tend to cluster together*". For the filtered currency rate network the clustering coefficient is *0.865* with is an almost scale free network.



(a) Distribution of the node betweenness centrality with the normal curve overlapped



(b) Distribution of the node closeness centrality with the normal curve overlapped

Figure 6. The distribution of the node betweenness and closeness centralities are one of the key metrics in characterising from a qualitative point of view complex networks. In our case booth of the centralities follow a pretty normal distribution

The betweenness is one of the centrality metrics defined for complex networks, being applied to booth nodes and edges. When applied to nodes, the betweenness if computed as the number of shortest paths that cross that specific nodes. So, the shortest path between any two nodes is computed and for each node is counted how many of the paths pass through those specific nodes. Also, the distribution of betweenness is a key indicator of the type of complex network [New03]. Power-law distributions are considered the signature of social networks and uniform distribution as the one of random network [WS98]. In or case as it can be seen from figure 6b there is an almost normal distribution, consistent with the type of networks we are working with. In the same time, the distribution of the closeness centrality which is another metric of centrality based on the topological distance between each node and all the other nodes of the network exhibits a similar normal distribution.

4. DISCUSSIONS AND CONCLUSIONS

We carried the investigations presented in Section 3 in order to provide a new insight into the world financial markets, seen as a complex networks. The input data consisted of *23* time series covering the currency exchange rate in report to the US Dollar for a time span of 14 years, from January 2003 and up to August 2017. Computing the correlation index of the data we first done a statistical analysis of the data which can be observed in figure 2 and after analyzing the distribution of the data we built a correlogram as a specific visualization technique for plotting the correlation of various indices.

The core of our research was geared towards using complex network analysis and this represents the content of Section 3.2. After we defined the mechanism of building the correlation network we built the associated complete graph and the corresponding subgraphs (by imposing various thresholds for the correlation indices).

After we apply the 0.3 threshold on the data we observe that still *79.84%* of the data series are still correlated which mean that there is a statistical correlation between most of the major world currencies independent of their country or continent of origin. Computing complex network specific metrics is another step we carried out in analyzing our data. By using common metrics we are able to compare various geographical clusters in terms of their currency exchange rate which in term means the structure of the economical markets of the specific countries.

The most fascinating result arise from the community detection algorithm we ran on the graph. We hypothesized that only by using topological characteristics of the network we could be able to derive the characteristics of the financial markets and the communities identified in Figure 3↑ and Figure 4↑ come co confirm our assumptions. The first step with good enough results can be observed in Figure 5↑. There we have the “red-community” consisting of *10* nodes and the “blue-community” having the rest of *13* nodes. The Blue-community is consisting of Asian currencies while the Red-community is a bit more mixed having currencies preponderantly from the Eurozone, but also a few intruders such as *{AUD, CAD, ZAR, KRW}*. We consider this a being the expression of the economical trade between the countries. For example KRW is the South Korean Won, and in the same time South Korea is one of the major players in the electronic and industrial markets. The South African Rand (ZAR) is exhibiting a similar behavior explainable through the fact South Africa being former British Colony (at least in part) and in the same time the South Korean diaspora in South Africa being the largest one in Africa which explains the higher economic exchange between the two countries [BG13].

On the second level of investigation we took into consideration the possible hierarchical structure of the data and continued by applying in a recursive fashion the algorithm for the first-level communities identified in Figure 2↑. For example the Red-community is subsequently divided into two more granular communities with a better definition of the boundaries. The Orange-community from now is the one discussed above consisting of British, South Korean and South African currencies while the Green-community of composed of traditional European exchange currencies.

We consider our research as a novel one in analyzing and visualizing the spatial relationships between world economies using the new field of complex network analysis. We took into consideration the currency exchange rate as a global metric of the country’s economy but a similar approach can be used for various other investigations. Our data cover a time span of 14 years with take into consideration also some more troubled time periods and as a direction of further investigation could be considered taking some “slices” in time and comparing the evolution of the graph in various time periods, putting emphasis on the transition between “steady-state” economy and the periods of financial crisis. Also we want to quantize the “quality” of the economic exchange by using a superposition of network metrics.

REFERENCES

[AB02] **Reka Albert, Albert-Laszlo Barabasi -** *Statistical mechanics of complex networks*, Reviews of modern physics, pp. 47, 2002.

[AB08] **Franklin Allen, Ana Babus -** *Networks in finance*, Wharton Financial Institutions Center Working Paper No. 08-07, 2008.

[BHJ09] **Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy -** *Gephi: an open source software for exploring and manipulating networks*”, ICWSM, 2009.

[BG13] **Brittany A. Baumann, Kevin P. Gallagher -** *Post-Crisis Capital Account Regulation in South Korea and South Africa*, 2013.

[Car13] **Petre Caraiani** - *Using Complex Networks to Characterize International Business Cycles*, PloS one, pp. e58109, 2013.

[Fri02] **Michael Friendly** - *Corrgrams: Exploratory displays for correlation matrices*, The American Statistician, pp. 316—324, 2002.

[H+07] **César A Hidalgo, Bailey Klinger, A.-L. Barabási, Ricardo Hausmann** - *The product space conditions the development of nations*, Science, pp. 482—487, 2007.

[KR07] **Raja Kali, Javier Reyes** - *The architecture of globalization: a network approach to international economic integration*, Journal of International Business Studies, pp. 595—620, 2007.

[K+10] **Dror Y. Kenett, Michele Tumminello, Asaf Madi, Gitit Gur-Gershgoren, Rosario N. Mantegna, Eshel Ben-Jacob** - *Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market*, PLoS one, pp. e15032, 2010.

[K+11] **Dror Y. Kenett, Yoash Shapira, Asaf Madi, Sharron Bransburg-Zabary, Gitit Gur-Gershgoren, Eshel Ben-Jacob** - *Index cohesive force analysis reveals that the US market became prone to systemic collapses since 2002*, PLoS one, pp. e19378, 2011.

[L+11] **Kyu-Min Lee, Jae-Suk Yang, Gunn Kim, Jaesung Lee, Kwang-Il Goh, In-mook Kim** - *Impact of the topology of global macroeconomic network on the spreading of economic crises*, PloS one, pp. e18443, 2011.

[Man99] **Rosario N. Mantegna** - *Hierarchical structure in financial markets*, The European Physical Journal B-Condensed Matter and Complex Systems, pp. 193—197, 1999.

[New03] **Mark E. J. Newman** - *The structure and function of complex networks*, SIAM review, pp. 167—256, 2003.

[T+05] **Michel Tumminello, Tomaso Aste, T. Di Matteo, R. N. Mantegna** - *A tool for filtering information in complex systems*, Proceedings of the National Academy of Sciences of the United States of America, pp. 10421—10426, 2005.

[WS98] **Duncan J. Watts, Steven H. Strogatz** - *Collective dynamics of 'small-world networks'*, Nature, pp. 440—442, 1998.

ACKNOWLEDGEMENT

This paper is supported by the Sectoral Operational Programme Human Resources Development POSDRU/159/1.5/S/137516 financed from the European Social Fund and by the Romanian Government