Control and Cybernetics

vol. **45** (2016) No. 2

Diffusion paths between product life-cycles in the European phonographic market*

by

Andrzej Jarynowski[1] **and Andrzej Buda**[2]

[1]Institute of Interdisciplinary Research, Wrocław, Poland
[2]Smoluchowski Institute of Physics, Jagiellonian University, Cracow, Poland
andrzej.jarynowski@uj.edu.pl
andrzejbudda@gmail.com

Abstract: We have investigated the product life-cycles of almost 17 000 hit singles having appeared on the 12 biggest national phonographic markets in Europe including: Austria, Belgium, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Switzerland, Sweden, and the United Kingdom. We have considered weekly singles charts from the last 50 years (1966-2015) in each country. We analyzed the spread of hit singles popularity (chart topping) as an epidemiological process, taking place in various European countries. Popularity of the hit singles is contagious in the sense of moving from one country to another. Thus, we consider time delays between the initial hit single release and reaching the highest position on consecutive national singles charts. We create a directed network of countries, this network representing the transmission of the hit singles popularity between countries. This is obtained by simulating the most likely paths and picking up the most frequent links. A country of initial hit single release is considered as a source of infection. Our algorithm builds up the spanning trees by attaching new nodes. The probability of attachment depends on infectivity of previous nodes from the tree corresponding to their:

1) market size;

2) distance in time between the new node and the potential spreaders.

Thus, we obtain a network of popularity spread with: a hub – the UK, a bridge – the Netherlands, and outliers – Italy and Spain. We have found a characteristic topology of hit singles popularity spread. On the top of this, the network of popularity spread has some typical properties of complex networks.

Keywords: epidemiology modeling, social network analysis, cultural diffusion, music charts

*Paper presented in its original version at the BOS2016 conference of the Polish Operational and Systems Research Society, Warsaw, October 2016. Accepted for publication in November 2016.

1. Introduction

Statistical properties of the global phonographic market have been already investigated well by the methods of econophysics and complex systems (Buda, 2012; Buda, Jarynowski, 2013). This theoretical approach allows as to study the system (phonographic market) composed of multiple interacting components (artists), where the value of an artist has been defined by the weekly albums sales. The most exciting property of such systems in general is the existence of emergent phenomena, which cannot be simply derived or predicted solely from the knowledge of the system' structure and the interactions between their individual elements. The methodology of complex system and econophysics proved helpful in the analysis of many issues of complex systems properties, including the collective effects and their coexistence with noise, long range interactions, the interplay between determinism and flexibility in evolution, scale invariance, criticality, multifractality, and hierarchical structure (Fronczak, Fronczak, 2008). The analyzed system is more predictable than financial markets, but is also more confusing, because of record labels, artists, mass media, and additional seasonal groups of collective customers that buy long-playing records. Thus, product life-cycles, are expressed by the trajectories of weekly albums sales that usually reach the maximum in the first week of album sales after the premiere and decrease affected by random processes (Jarynowski and Buda, 2014).

However, there is another part of phonographic market, represented by the songs released as hit singles – promoting the album and the artist. The spread of popularity of these hit singles across the countries reminds of the classical product life-cycles despite the fact that the official definition of popularity in case of songs is more complex. National charts have been usually based on weekly physical singles sales, but smaller markets have also compiled official national charts according to official airplays on the radio or TV. During the Internet era (actually, from 2003 until now), almost all national hit singles charts are mostly based on digital streaming of songs (mp3, ringtones, etc.) The popularity of songs, as a compilation of these factors, is a complex phenomenon, and so, in our research, we decided to measure this phenomenon from an economic point of view (the value of a song is measured weekly by its positions on the national charts). We have investigated trajectories of almost 17 000 hit singles, which have appeared on the 12 biggest national phonographic markets in Europe including: Austria (A), Belgium (B), France (F), Germany (D), Ireland (IRE), Italy (I), Netherlands (NL), Norway (N), Spain (E), Sweden (S), Switzerland (CH), and the United Kingdom (GB). We have considered weekly singles charts from the last 50 years (1966-2015) in each country. The process of information spread has been already investigated by us and described from the sociological and computational sciences' point of view, including the exploration of the European music charts network, indicating positive correlation with geographical and cultural distance network in various decades of music eras, referred, consecutively, to as: analog, digital and Internet (Buda and Jarynowski, 2015).

In this paper, we continue our research and focus on economic properties and interactions between national markets. The classical product life-cycle theory is defined by stages that depend on diffusion of innovations (including: innovators, early adopters, early majority, late majority and laggards) (Rogers, 1962). In case of hit-singles, the national charts are incomparable, because of such issues as different methods of compilations in each country, size of a market, etc. But if we consider the stages of popularity for a single, performed in various countries (that are supposed to represent innovators, early adopters and late adopters), the product life-cycle may be described epidemiologically by the SI models (Anderson and May, 1992). Suppose the simplest model, in which the population is divided into two classes: the susceptible (S) ones, who can catch the disease; and the infectives (I), who can transmit disease and have it. Hence, we use epidemiological notation and principles of modeling. In general, we consider only first stages of popularity for a hit single in a given country (until reaching the highest position on the chart) and then such a country becomes Infective so that it can infect other, Susceptible countries. This is similar to a single season SI model (there is no other infection circulating at the same time, and there is short enough time span of the spreading process) and the reconstruction paths of infections, which is a well known problem in such a system. Thus, we describe the stages of the product life-cycle in terms of epidemiology and consider time delays between the initial hit single release and reaching the highest position on consecutive national singles charts. Our aim is to show the paths of interactions between the 12 biggest national markets in Europe and to detect how one country affects another.

We provide the Social Network Analysis (SNA) technique to understand and model the spread of popularity and social communities creation (Chojnacki, Kłopotek, 2011). The SNA procedure enables us to identify the characteristic roles of various countries involved in the spreading process and to discuss what does it mean in a sociological perspective. Based on network properties many features of the system can be predicted (Morzy, Wierzbicki, and Papadopoulos, 2009), because most information concerning influence comes from delays in the times of adoption of a new single hit. Currently, SNA serves as a state-of-art methodology and set of tools enabling a multifaceted in-depth exploration of interacting systems.

Intra-European popularity spread in the domain of popular music has never been studied from the empirical perspective (up to our knowledge). Except for the Eurovision Song Contest (which, as this is commonly known, is extremely biased, due to many circumstances) there are no quantitative analyses of this issue. Moreover, we collect data from various sources and hence a lot of manual processing was needed. However, epidemiological processes, social diffusion, innovation spread and viral marketing were broadly investigated, and so we have placed our concept within the 'state-of-the-art' social network analysis methodology. In our research we have considered trajectories of the hit singles as they performed on the national charts in European countries as a product lifecycle. We treat the highest position on each chart as the key time point.

As a main data parameter t we consider the number of weeks (after the initial single release) when the single reaches the peak of the national record chart.

2. Data analysis

At the very beginning, it is necessary to define the level of popularity for a song, which will be understood here as the highest position on a national singles chart. For example, 'We Found Love', performed by Barbadian singer Rihanna in 2011, initially entered the singles charts on October the 8^{th}. Table 1 shows the climbing up the song on the national charts, and constitutes also a good illustration of the respective differences.

In this case, the source of infection was Belgium, but in the first week France and Norway had been already infected. Italy and Spain were the last countries that were affected by this spread of global popularity. If we consider the time delay between the first week of entering the market and the week of reaching the highest positions, we will obtain the map of Europe, as shown in Fig. 1, according to the investigated data set (1966-2015) of hit singles. Statistically, it is the UK, the Netherlands and Belgium that usually start the infection. We assume that the epidemic spread starts from a source country (single release and promotion of it), and further only directly between countries, which form different kind of social networks. Even if we know that the spread takes place by external source (non-European country), the selected European country is selected as a source.

Table 1. Increasing positions on the national charts of Rihanna's 'We Found Love' during 8 weeks. After reaching the top position, cells are in black meaning a country is infected

Country	1^{st} week	2^{nd} week	3^{rd} week	4^{th} week	5^{th} week	6^{th} week	7^{th} week	8^{th} week
A	-	-	-	12	12	4		
B	3							
CH	3	3	2	2	2	1		
D	-	-	-	1				
E	15	14	14	9	9	9	5	3
F	1							
GB	-	1						
I	-	-	-	7	4	4		
IRL	-	3	1					
N	1							
NL	14	3						
S	2	2	2	2	1			

For each single hit we construct a time series, which is the sequence of countries taken along the successive points in time and for the here presented example this sequence looks like that of Table 1b.

Table 2. Time series of infection times (defined as peaking time on the national chart) of Rihanna's hit single 'We Found Love'

X - country	F	N	B	GB	NL	IRL	D	S	A	CH	I	E
t_x - delay	1	1	1	2	2	3	4	5	6	6	6	14*

* Spain's peaking time is not on the list in Table 1

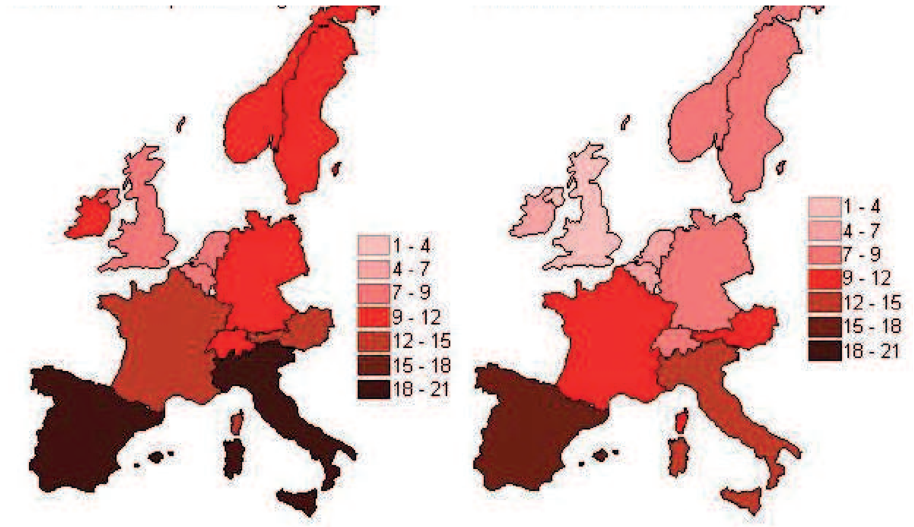

Figure 1. Mean (left) and median (right) time delay t_x (in weeks) of attaining the highest positions on the charts (Jarynowski and Buda, 2014)

As said, we assume that a country has been affected, when the highest position is reached. As the main data parameter of delay, t_x, we consider here the number of weeks (after the initial single release) when the single reaches the peak of the national record chart X.

For each song, it is possible to identify the Minimum Spanning Tree - MST (Fig. 2) according to the matrix of distances, which are defined using the individual time delays as follows:

$$d_{xy} = |t_x - t_y| \tag{1}$$

where t_x and t_y represent the numbers of weeks to reach the highest positions in countries X and Y (time to infection).

The deterministic construction of a chain of events in the form like that of MST does not lead, for the kind of data we consider here, to a unique realization.

This is, in particular, due to the fact that within the same time window more than one country can adopt – be infected by – a new single. In the ABBA's song case, illustrated in Fig. 2, there are at least two possible paths of transmission: either France (upper graph) or Switzerland (lower graph) continue the tree after Germany. Thus, the MST construction may lead to a Markovian process, where only a single country at a given time can be attached as a new node. The history of transmission between the previous nodes is going to be forgotten.

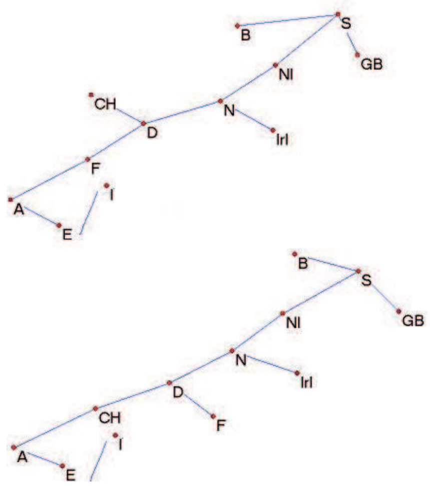

Figure 2. The Minimum Spanning Tree based on the matrix of time-delay-based distances between countries for ABBA's single 'Dancing Queen' (1976). The source of infection was in Sweden

Instead of directly constructing the MST from the data, we provide an algorithm to retrieve the set of the feasible trees, where each connection (edge in the tree) depends on the whole history.

The periodicity and the evolution of musical trends (Grabowski & Kosinski, 2010) in various eras of recorded and broadcasted music (analog, digital and Internet) makes observations more complicated. To cope with that, let us investigate the process at two levels:

- local – to get a deeper insight into the evolution of chosen songs; the time delays (t_x) extraction procedure was shown for the case of Rihanna's advance on the charts (Table 1); time delays transformed into proper distance measures allow us to form the spanning trees representing possible paths of infection and some of Minimum Spanning Trees for ABBA's hit single were also presented (Fig. 2); this approach provides an opportunity for a reader to understand in a more qualitative way what is happening in the system;
- global – to get a general view and to find the general properties, we observe the whole history of charts in Europe (1966-2015) and it is only this

level of comparisons between countries that gives us the opportunity to draw significant conclusions, concerning the differences, in the quantitative manner, see Figs. 1, 3, and 4.

3. Results

In our research, we have manually detected the sources of infections (initial countries) for all the 50 most popular hit singles that finally affected the whole of Europe in the period 1966-2015, and 50 the best singles of each year according to given meta-analysis (Buda, 2006). Then, we have identified the most likely paths of infection for each single hit. Our algorithm builds up the spanning trees for each hit-single 1000 times by attaching new nodes. It follows the algorithm described below, and must be repeated many times, due to the dimensions of the space of possible path's realizations, so extensive statistics is needed. For each single hit, the time series of ordered time delays of 12 national charts is constructed, $T=[t_1,.., t_x,..,t_{12}]$, where $t_1 = 1$ (the week of release) and $t_i \geqslant t_{i-1}$. Further, for every realization, we follow:

1) We choose the source of infection – the country of initial release (e.g. ABBA's song 'Dancing Queen' comes from Sweden), see Fig. 2;

2) We attach next in the row node(s) (with shortest distance) to the source (e.g. Belgium and the UK are simultaneously attached to Sweden, see Fig. 2);

3) The next ones (X) are added to one of the previous nodes, but not any more in the deterministic manner in formation of the MST, but in a stochastic way, according to a probability of attachment. This step is repeated until all countries are attached to the tree. The probability of attachment is inversely proportional to the difference between time delays and proportional to the size of a market in a previous node, according to the formula:

$$P_{x-y} \sim \frac{1}{t_x - t_y} * \frac{N_Y}{N_{AVG}} \tag{2}$$

where:

P_{x-y}- probability that new node X will be attached to node Y;

N_Y - market size of the node Y;

N_{AVR} - average market size in Europe, for the operational definitions of N, see below.

The sum of P_{x-y} over all pervious nodes Y is normalized to 1 for each X.

We sneak in a preview of epidemiologic notation of popularity spread. We could define viruses as the hit singles that affect one country after another. Node's infectivity (the force of transmission of infection) is proportional to the size of a local phonographic market.

In Western European countries, the phonographic market size N_Y may be taken approximately as equivalent to the population, Y, according to regions (see Table 2). The probability of attachment increases in inverse proportion to time delays (memory effect). We weigh each link in inverse proportion with

Table 3. Estimates of node's infectivity – market size (approximate population numbers)

Size estimation of the market in M	Euro Region	Countries
30	"Benelux"	NL, B
90	"Extended German"	D, A, CH
60	"British Empire"	GB, IRL
50	"Iberia"	E
80	France	F
60	Italy	I
30	"Scandinavia"	S, N
57	Europe average	All

regard to the consecutive steps of simulation (new node's attachment from 1 to 12), in order to appropriately value the musical fashion adoption in the early steps of the process considered (early adopters).

Table 4. Node degrees calculated via simulation of popularity spread

degree	out	in
A	23	36
B	60	65
CH	45	63
D	75	53
E	16	27
F	43	50
GB	192	36
I	17	36
IRL	43	89
N	49	65
NL	74	94
S	55	78

We obtain the directed cumulative network from 100 000 trees, representing each singles hits (100) and numerous realizations of the graphs (1000). This network of popularity spread (Zbieg et al., 2012) consists of the following special elements: a hub – the UK, a bridge – The Netherlands, and the outliers – Italy and Spain (see Fig. 3). In this manner, we have identified a characteristic topology of hit singles popularity spread. It is clearly visible on the basis of this topology from which country viruses can infect another country. There is no surprise that the UK has the strongest ability to create (but not receive) viruses (due to its highest out degree – Table 3) among all of the European countries. The Netherlands is most likely to adopt (be infected) by the new fashion (due to its highest in degree – see also Table 3).

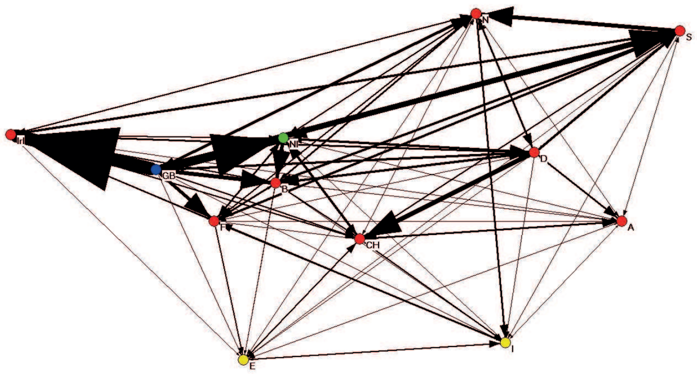

Figure 3. The directed European network of popularity spread based on matrix of time-delays between countries. Network of popularity spread with: a hub – the UK, a bridge – the Netherlands and the outliers – Italy and Spain

Clustering and community detection (modularity), which can be performed following various available methods (see further on for some comments on respective algorithms) altogether bring forth two subnetworks (Fig. 4), which are subjectively classified in brackets:

– Scandinavia, Benelux, UK, Ireland, and Germany (Susceptible – early adopters)

– Spain, Italy, Switzerland, and Austria (Resistant – late adopters).

The first group contains the countries that easily catch all infections, while the second one is composed of the more conservative ones. However, France is classified in some calculations in the Susceptible community, and in some other ones – in the Resistant one. The position of France in the directed networks, see Fig. 4, and so in the groups distinguished here, depends on the algorithm applied.

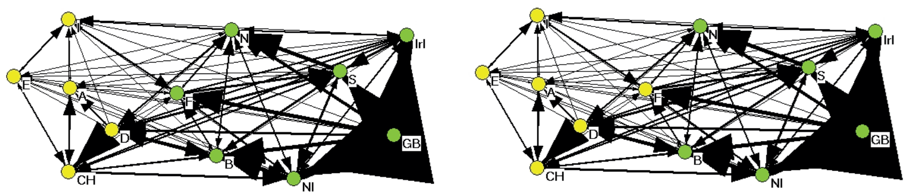

Figure 4. France classified either in the Resistant community, or in the Susceptible community. Left: Louvain algorithm. Right: VOS Clustering

A (crisp, exclusive) partition of the popularity network engenders a classification of the countries such that each country is assigned to exactly one of two communities. Community is a social structure with connectivity within the subgraph, corresponding to it, that is higher than connectivity with the rest of the graph*. There are several ways to perform partitioning of a set and it is a computationally, as well as intellectually demanding task. Algorithms for finding communities can be of different type, namely, for instance, such as: the Minimum-cut method, Hierarchical clustering, Girvan–Newman algorithm, Modularity maximization, Statistical inference, or Clique-based methods. Here we choose, just for visualization purposes, two very common community detection algorithms. Both the Louvain Method and VOS Clustering, when used for community detection are methods to extract communities based on function optimization (De Nooy, Mrvar, Batagelj, 2011). In the first case the modularity indicator is maximized (that measures the density of edges inside the communities relative to the edges outside the communities). In the second case, the specific VOS quality score, when high enough, allows to distinguish communities, and it is a well-known technique of multidimensional scaling. In the analyzed network, there is a clear division into the Resistant and Susceptible sub-networks, which, apparently can be deduced from modularity. However, the dividing line in not obvious and one node - France - can be assigned either to the Resistant or to the Susceptible community.

4. Conclusions

The study reported focuses on the influence of geographical and cultural proximity on several aspects of hit singles, such as the time it takes to appear on a certain chart. In this paper we have shown the structural dependencies between various local European phonographic markets, regarding this particular aspect. We have built an epidemiological simulation setup to extract the most important viral components of the system. We obtained a network of popularity spread with a hub – the UK, a bridge – The Netherlands, and the outliers – Italy and Spain. The most influential country is the United Kingdom, infecting the others in the early steps of propagation, as witnessed by the highest out-degree (Table 3). However, UK is immune to foreign hit-singles, especially from non-English speaking countries. This node has one of the lowest in-degrees in the whole of the here analyzed Europe (see also Table 3). On the other hand, The Netherlands, while not being the initial source of infection, can, apparently, adopt foreign hit-singles quickly, no matter where they come from. The Dutch phonographic market is similar to a European bridge (broker) (Przybyła and Weron, 2014) because of its high in-degree. These results reveal interactions between local phonographic markets in Europe, and show that the process of popularity spread has some typical properties of complex networks (e.g. the evidence of a hub). Concerning such processes of diffusion and spread in the networks, they have been investigated for the empirical and random networks (Leskovec et al., 2007; Boss et al., 2004).

*It should be noted that this, apparently simple and intuitive definition, does not lead to unambiguous determination of the respective structures, see Owsiński (1981), ed.

The current issue of UK exiting European Union in 2016 could be analyzed from this perspective for fundamental understanding all of the possible permutations and challenges if phonographic connectivity between UK and Continental Europe would be reduced. The UK is very resistant to foreign culture, but Continental Europe is susceptible to British hits. Moreover, continental Europe relies very much on the direct links with the UK (high out-coming degree of the UK), which allows for a stronger influx of English influence. Pop music typically uses the British or US, English language, and, in addition, the peoples of Europe have established the habitus of musical preferences and are somehow ‘addicted’ to music distributed via the UK. On the other hand, Great Britain has very low in-degree, so it does not depend on Continental Europe at all in the domain considered. It seems that, at least form the cultural perspective, Continental Europe can lose more on cutting the links with the UK, than the other way around.

We observe that the network of countries breaks down into two well-defined clusters (Fig. 4) of early and late adopters, like other systems, which feature the associated meaningful taxonomies (Paulus and Kristoufek, 2015). The positive correlation between our network and geographic or cultural grid-map of Europe is still observed, but some additional features are much more characteristic. In our case, the clustering procedure also ends up with meaningful division, but here it follows more marketing than economic product life-cycle classification. However, the role of France in the European market is still worth examining and explaining. According to our results, France belongs either to the ‘infectious’ community, or to the ‘resistant’ community. This behaviour may be a result of broadcasting regulations in France, where the spread of popularity through radio or TV for the non-French songs is strictly limited by the barrier of 50% in broadcasting. This limit has an influence on the hit-singles sales, as well. Thus, the French phonographic market is resistant, but it is also able to adopt relatively quickly the most popular foreign songs.

The paper concerns the problem of recognizing the patterns of popularity of hit singles appearing on weekly charts in 12 European countries. We are aware, however, that the time delays in national musical popularity charts reflect a definite ambivalence and sensitivity to external factors (each single has its own history, also in a particular country). No universal method is available for modelling and reconstructing the ‘standard’ or ‘average’ path of popularity spread. Our research, therefore, not only explored the paths, along which a hit single might be able to ‘infect’ consecutive countries, but also raised the question of how Europe is connected within the perspective of ‘pop-culture’. We assumed that the ‘viral infection’ can be passed from country to country. Therefore, we excluded any ‘field like’ effects and limited ourselves only to ‘direct interaction’. In our research, we are limited to direct paths of interactions without considering any external fields (like the influence from the US).

Acknowledgements

The authors would like to thank Jarosław Kwapień (Institute of Nuclear Physics, Polish Academy of Sciences, Cracow), Andrzej Grabowski (Central Institute for Labour Protection – National Research Institute in Warsaw) and Luba Jakubowska (The Witelon State University of Applied Sciences in Legnica) for comments, as well as the reviewer for remarks.

References

Anderson, R.M., May, R. (1992) *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford.

Boss, M., Elsinger, H., Summer, M., Thurner, S. (2004) Network topology of the interbank market. *Quantitative Finance*, **4**(6), 677-684.

Buda, A. (2006) *Historia rocka, popu i hip-hopu: według krytyków 1974-2006* (*The history of rock, pop and hip-hop: according to the critics*; in Polish). A. Wydawnictwo Niezależne.

Buda, A. (2012) Does pop music exist? Hierarchical structure in phonographic markets. *Physica A* **391** (21).

Buda, A., Jarynowski, A. (2013) Network structure of phonographic market with characteristic similarities between artists. *Acta Physica Polonica A*, **123** (3).

Buda, A. Jarynowski, A. (2015) Exploring patterns in European singles charts. *Network Intelligence Conference (ENIC), 2015 Second European*, 135-139.

Chojnacki, S., Kłopotek, M. A. (2011) Random graph generator for bipartite networks modeling. *Control & Cybernetics*, **40**(3).

De Nooy, W., Mrvar, A., & Batagelj, V. (2011) *Exploratory Social Network Analysis with Pajek. Structural Analysis in the Social Sciences*, **27**. Cambridge University Press.

Fronczak, A., Fronczak, P. (2009) *Świat sieci złożonych: od fizyki do Internetu* (*The world of complex networks: from physics to Internet*; in Polish). Wydawnictwo Naukowe PWN.

Jarynowski, A., Buda, A. (2014) Dynamics of popstar record sales on phonographic market – stochastic model. *Acta Physica Polonica* B (PS) **2** (7).

Leskovec, J., Adamic, L. and Huberman, B. (2007) The dynamics of viral marketing. *ACM Trans. Web*, **1**, 1.

Morzy, M., Wierzbicki, A., and Papadopoulos, A. N. (2009) Mining online auction social networks for reputation and recommendation. *Control and Cybernetics*, **38**(1), 87-106.

Owsiński J.W. (1981) Intuition vs. formalization: local and global criteria of grouping. *Control and Cybernetics*, **10** (1-2), 73-88.

Paulus, M. and Kristoufek, L. (2015) Worldwide clustering of the corruption perception. *Physica A: Statistical Mechanics and its Applications*

428, 351-358.

Przybyła, P., Sznajd-Weron, K. and Weron, R. (2014) Diffusion of innovation within an agent-based model: Spinsons, independence and advertising. *Advances in Complex Systems*, **17**(01), 1450004.

Rogers, E.M. (1962) *Diffusion Of Innovations* (1st ed.). Free Press of Glencoe, New York.

Zbieg, A., Żak, B., Jankowski, J., Michalski, R. and Ciuberek, S. (2012) Studying Diffusion of Viral Content at Dyadic Level. *ASONAM 2012.* Istanbul.