Selected Papers of the 6th Polish Symposium of Physics in Economy and Social Sciences FENS2012, Gdańsk, Poland

Network Structure of Phonographic Market with Characteristic Similarities between Artists

A. Buda[a] and A. Jarynowski[b,c,*]

[a]Wydawnictwo Niezależne, Oriona 15/8, 67-200 Głogów, Poland

[b]Smoluchowski Institute of Physics, Jagiellonian University, W.S. Reymonta 4, 30-059 Kraków, Poland

[c]Department of Sociology, Stockholm University, Universitetsvägen 10 B: plan 9, SE-106 91, Stockholm, Sweden

We investigate relations between bestselling artists in last decade on phonographic market and from perspective of listeners by using the Social Network Analyzes. Starting network is obtained from the matrix of correlations between the world's bestselling artists by considering the synchronous time evolution of weekly record sales. This method reveals the structure of phonographic market, but we claim that it has no impact on people who see relationship between artists and music genres. We compare "sale" (based on correlation of record sales) or "popularity" (based on data mining of the record charts) networks with "similarity" (obtained mainly from survey within music experts' opinion) and find no significant relations. We postulate that non-laminar phenomena on this specific market introduce turbulence to how people view relations of artists.

1. Motivation and data sources

1.1. History of phonography

People have played and listened to the music since ages but the origins of phonography came from Thomas Edison who developed many devices including the phonograph (1877). The gramophone record was one of the dominant audio recording formats throughout much of the 20th Century. The Italian tenor Enrico Caruso (1906) and The Beatles (1964) were the first who sold more than 1 000 000 copies of their records. The new formats (vinyl, cassette, compact disc, and mp3, etc.) were more and more common. Thus, Michael Jackson's "Thriller" has become the most popular record ever and sold over 110 000 000 copies [1].

1.2. Establishment of phonographic market

Phonographic markets are well defined complex systems almost as old as financial markets. The figures speak for themselves: while Americans spent 100 million dollars on records before the fatal 1929 stock market crash, this number had plunged to a merely 6 million by 1933 [2]. Nowadays, global music sales in 2009 fell by 7% to US $17 billion. This is disappointing, but there are some very positive points. Digital music goes global in 2011 while action on piracy gains momentum. With rapid expansion into new markets by services such as iTunes, Spotify and Deezer, the major international digital music services are now present in 58 countries, compared to only 23 at the start of 2011 [3].

Traditionally, the record charts are based on weekly record sales. According to the International Federation of the Phonographic Industry (IFPI), the world's largest
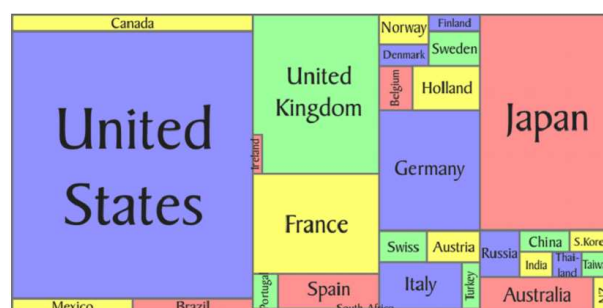

Fig. 1. According to the IFPI more than 95% of the total revenue of music in 2003 was derived from the 30 major countries in the proportion shown above, organized by geographic location.
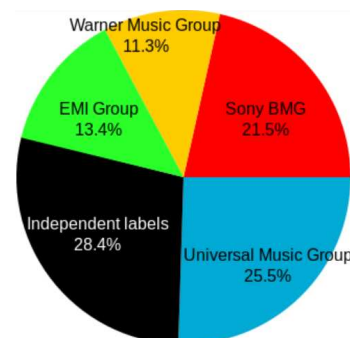

Fig. 2. World music market sales shares, according to IFPI (2005).

phonographic markets are [1, 4]: the USA, Japan, Great Britain, France and Germany (see Fig. 1). 80% of weekly record sales belong to the four biggest record companies (Universal, EMI, Sony BMG and Warner Bros (see Fig. 2)). All the world's most popular artists are signed to these companies [1]. Thus, since 2003 it is possible

———

*corresponding author; e-mail: andrzej.jarynowski@uj.edu.pl

to find their weekly record sales exactly [4]. Hence, the global phonographic market can be considered as a complex system and quantitative properties of phonographic markets have been already studied by economists and sociologists [5–8], and recently also by sociophysicists [9, 10].

In our research, we have chosen the portfolio of the world's most popular 30 artists (that sold more than 11 000 000 units) according to albums sales between September 2003 and September 2011. These artists were popular, but not always critically acclaimed. Those artists represent variety of styles, genders, and are signed to major record companies — see Fig. 2.

1.3. Role of experts in industry

Music writers also started treating pop and rock music seriously in 1964 after the breakthrough of the Beatles. Music press was more and more influential in discovering new artists. At the end of each year music magazines publish their critics' polls that include the best albums and singles of the year according to critics' opinions. These critics' polls were often subjective, but had impact on weekly record sales [12]. We decided to study critics' polls from 2003 to 2007. We chose the most influential music magazines (Mojo, NME, Pitchfork Media, Q, Rolling Stone, Spin, Uncut, Village Voice) and the Mercury Prize — the annual music prize awarded for the best album from the UK and Ireland. Nominations are chosen by a panel of musicians, music executives and journalists. Music recommendation authorities seem to have influence on people's vision of music [11], so we decided to add it to our studies.

1.4. Research questions

In this chosen cohort displayed in Fig. 3, artists often collaborate in the same projects, or made the cover versions of other artist's songs. A question appears: how do similarity in music and cooperation between artists

impact on connection in phonographic or popularity networks? What does influence on people's choices and decisions in buying records? Do people buy records like common good or like a luxury good? Do similar artist compete or are "surfing" together on trends waves? How critic's polls influence on phonographic market?

2. Introduction to analysis

The motivation of the present study is to find a kind of more general arrangement between artists and built networks and compare different approaches. The phonographic network is obtained starting from the matrix of correlations coefficient computed between the world's most popular 30 artists (this list corresponds to all analyzes in this paper). The value of an artist is defined by weekly record sales. The graph is obtained starting from the matrix of correlations coefficient computed between the world's most popular 30 artists by considering the synchronous time evolution of the difference of the logarithm of weekly record sales.
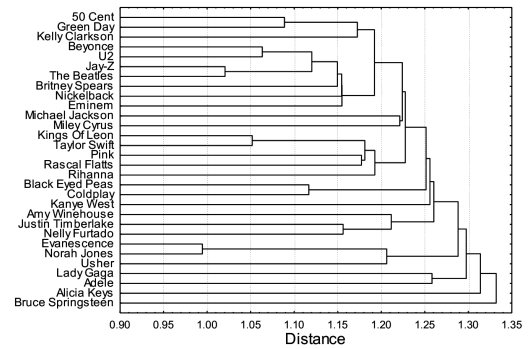
Fig. 4. Hierarchical structure of phonographic market (2003–2011). Distances given by formula (1).

Fig. 5. Significant ($\alpha = 0.05$) positive correlation between artist sales. The same network with colored nodes according to linked phonographic companies can be found in Fig. 6.

We show topological arrangement of market by hierarchical diagrams (Fig. 4) and networks (Figs. 5, 7). On the other hand, popularity of artist calculated from ranks
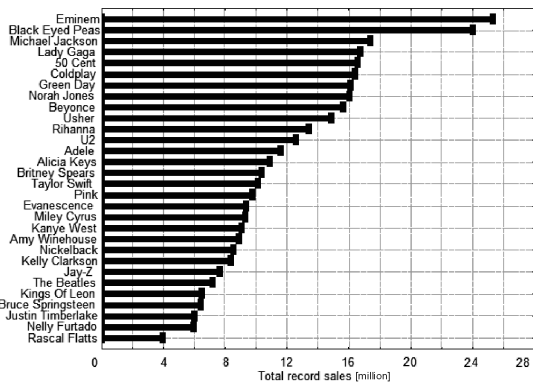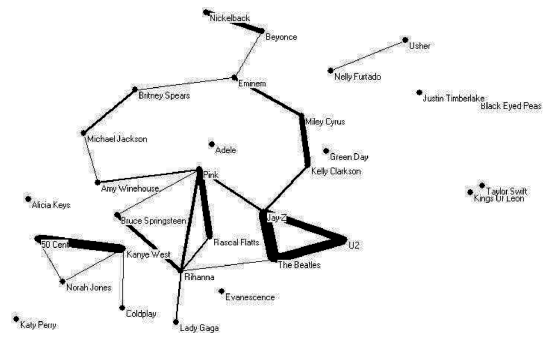

Fig. 3. Chosen cohort of the 30 bestselling artists and music bands. The bands total record sales also contain their members' contributions, e.g. Paul McCartney and John Lennon record sales are included in The Beatles category.
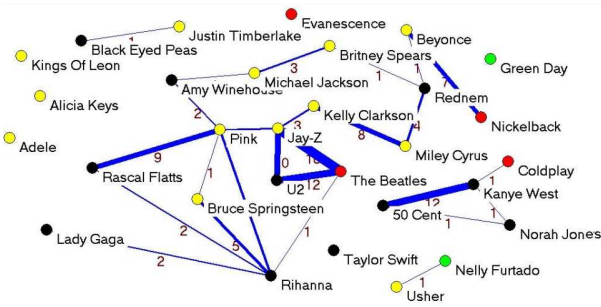
Fig. 6. Network of significant rescaled positive correlations with nodes colored by music companies (UNIVERSAL: black, SONY: yellow, EMI: red, WARNER: green). For the strongest correlations (rescaled value above 4) 7 links are mixed and only within the same company.
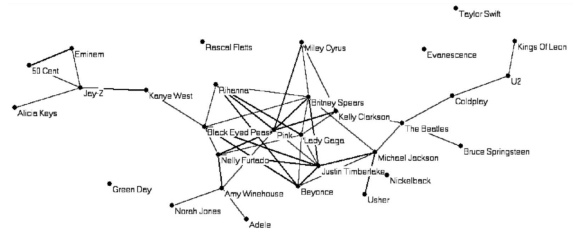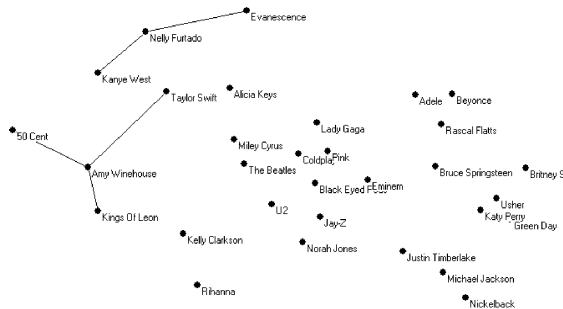

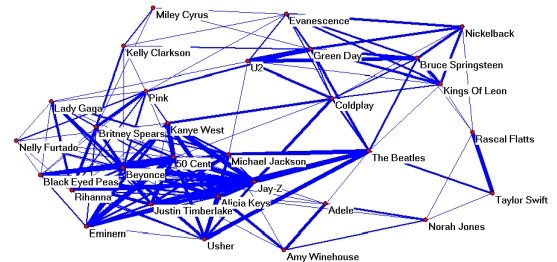Fig. 7. Negative correlations ($\alpha = 0.05$) between artists sales.

Fig. 8. Hierarchical structure of charts popularity (2003–2007). Distances given by formula (2).

Fig. 9. `allmusic.com` similarities.

Fig. 10. Expert's (Andrzej Buda) view.

in charts of their albums sales year by year was visualized in the hierarchical diagram in Fig. 8. Most influential music journal for different branches of music were investigated via data mining procedures to find names or aliases of chosen artists with their rank on charts. The link between artists appears when both artists exist on the same chart list in a chosen year. An indicated relation is stronger if their rank on such a list is similar. Weights of the link grow if those artists appear in more than one chart that year. On the other hand, links between artists can be defined as a similarity between them. This similarity can be understood in two ways: as an automatically annotated links found on `allmusic.com` website (Fig. 9) or as an expert (subjective people's view (Fig. 10)). On `allmusic.com` there are two categories to link artists: following or followed by. That gives opportunity to obtain directed network. Experts, however, could rank their subjective feeling of similarity between artists in survey. They were asked at the end, which were for them the most important factors to link artists. The last step was to compare such obtained networks (Tables I, II).

TABLE I

Correlations between networks 2003–2011. Those significant ($\alpha = 0.05$) are in bold.

Networks	2003–2011				
	Expert (Buda)	Experts	All_music	Market_corr	Market_sq
Expert (Buda)	1.00	**0.79**	**0.34**	−0.01	0.02
Experts	**0.79**	1.00	**0.41**	−0.01	−0.01
All_music	**0.34**	**0.41**	1.00	0.02	0.02
Market_corr	−0.01	−0.01	0.01	1.00	**0.93**
Market_sq	0.02	−0.01	0.02	**0.93**	1.00

Correlations between networks 2003–2007. Those significant ($\alpha = 0.05$) are in bold.

Networks	2003–2007			
	Experts	All_music	Critics	Market_corr
Experts	1.00	**0.41**	−0.00	−0.02
All_music	**0.41**	1.00	−0.05	0.00
Critics	−0.00	−0.05	1.00	0.01
Market_corr	−0.02	0.00	0.01	1.00

3. Hierarchical analysise of phonographic market and charts

Initially, we analyzed correlations network of artist's record sales. The correlation coefficient defines a degree of similarity between the synchronous time evolutions of a pair of assets, where we took of underlying[†] sales value. By definition, the matrix of the Pearson correlation coefficients is a symmetric matrix with ρ_{ij} and the $n(n-1)/2$ correlation coefficients characterize the matrix completely. Every correlation from that matrix based on two vectors containing P_i and P_j: the weekly record sales of the artist i and j for every week for given time interval. The correlation coefficient reflects similarity between assets. It can be used in building the hierarchical structure in financial markets and finding the taxonomy that allows isolating groups of assets. The distance between artists is defined by:

$$d(i, j) = \sqrt{2 - \rho_{ij}}. \tag{1}$$

Such a distance (1) is a base to obtain Fig. 4. In Fig. 8 the similar formula (2) applies with link parameter l_{ij} described manually via data mining procedures

$$d(i, j) = \sqrt{2 - l_{ij}}. \tag{2}$$

We looked at correlation coefficient in whole period as well as in one year's times windows, because of seasonality of artist fame and chance to compare it with others. Year window will also correspond to popularity measure, which is based on yearly charts.

A typical division of correlation, useful in finding the life-time of correlations between stocks in financial markets [13] where almost all correlation coefficients are positive, is not valid anymore because of different properties of the system.

Most of phonographic correlations are insignificant (85% of cases with $\alpha = 0.05$) but positive. However, some of the significant correlation coefficients in the phonographic markets are negative when artists essentially compete over the same group of customers.

Coincidence in presence on the same chart list in specific year was introduced as a measure of popularity. Accumulated coincidences matrix for 2003–2006 can be also presented as a hierarchical tree (Fig. 8) which seems to be different than sale tree (Fig. 4), but still reveals "stars" cluster described more in chapter about social network analyses (SNA).

Phonographic diagram can be also represented as a network. Correlation coefficient has then meaning of a weight of a link. Due to that definition, we would obey always full connected network. To avoid that, we decided to state some threshold. The best, natural one, in our opinion, for correlations is P-value above certain $\alpha = 0.05$ and we are using the same rule everywhere in this paper when we are projecting a correlation matrix into a network.

Correlations between binary networks obtained from record sales in next years from 2003 to 2011 with full period. Those significant ($\alpha = 0.05$) are in bold.

Correlations	full	2004	2005	2006	2007	2008	2009	2010	2011
Full	1.000	−0.013	**0.163**	**0.120**	**0.151**	−0.069	0.064	−0.013	0.073
2004	−0.013	1.000	0.029	−0.035	−0.039	−0.024	−0.056	0.010	−0.032
2005	**0.163**	0.029	1.000	0.037	−0.013	−0.045	−0.031	0.035	−0.016
2006	**0.120**	−0.035	0.037	1.000	−0.019	−0.003	**0.110**	−0.048	−0.007
2007	**0.151**	−0.039	−0.013	−0.019	1.000	0.017	−0.018	0.082	−0.015
2008	−0.069	−0.024	−0.045	−0.003	0.017	1.000	−0.034	−0.004	0.035
2009	0.064	−0.056	−0.031	**0.110**	−0.018	−0.034	1.000	**0.103**	**0.115**
2010	−0.013	0.010	0.035	−0.048	0.082	−0.004	**0.103**	1.000	**0.101**
2011	0.073	0.032	0.016	0.007	0.015	0.035	**0.115**	**0.101**	1.000

We checked how such a network evolves in time (Table III). To do so, we divided time series into years intervals (there are some economic reasons why one year is the best interval and explanations comes from seasonal waves of selling, e.g. before Christmas or Valentine's Day

average sales grow up rapidly). This network is not stable. If we look at correlations between binary networks (link exists if correlation between these two artists is significant in given time interval), there is only small memory, because only 29% of correlations between following

years are significant (10% for all cases). Those binary networks do not consider sign of correlation at all. That brings curiosity what should be taken as a weight of the link. In standard analysis [12] correlation coefficient is transformed to Euclidean distance, and this transformation does not care if coefficient is positive or negative. To have in mind that sign of correlation does matter, let us introduce two measurements: pure coefficient and its square. A square has another good property — it is underestimating small correlations, which plays a role of noise reductor.

4. Product lifecycle in phonographic market

The Black Eyed Piece record sales (Fig. 11) represents typical behavior of music label policy. Firstly, we should realize what is happening in the date of a premiere of a new album of a famous artist. The rapid increase of record sales in such a week is obvious, but there is also a small increase just before the premiere. Fans are waiting for a new album and they are buying older ones. This is the stage of low growth rate of sales as the product is newly launched in the market. With a premiere of an album the first single of that album comes to promote it. After some time, next singles are coming and if it has got popularity the record sale increases. A growth comes with the acceptance of the innovation (new album) in the market and profit starts to flow. The last stage of product lifecycle is maturity. In this end stage of the growth rate, sales slowdown as the product has already achieved its acceptance in the market. Therefore, new artists start experimenting in order to compete by innovating new models of the product [15, 16]. Thus, The Black Eyed Peas should release another album to be still on the top of the record charts. It is clearly shown in their record sales history.
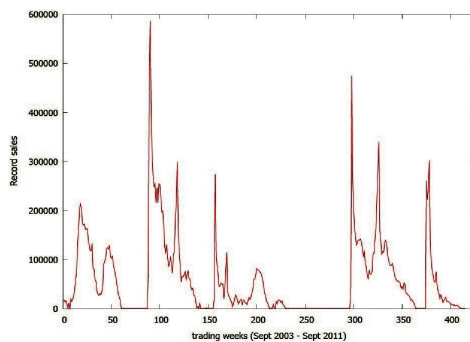

Fig. 11.   The Black Eyed Peas record sales history.

5. Network analysis

For that, we have built the network for portfolio of artists. It provides an arrangement of assets, which selects the most relevant connections of each point of the

set. Although the network structure in financial markets reflects the classifications of stocks in the industry sectors and sub-sectors reported in the Forbes annual reports on American Industry [17], SNA of correlations between artists does not always fit to music genres classified by the Billboard and other music magazines devoted to music industry.

The SNA reveals (Fig. 5) sectors that belong to rap (Kanye West, 50 Cent), but does not show the sector for pure pop music. Instead of pop, the main community, we have a celebrity sector that contain Lady Gaga, Rihanna, Bruce Springsteen, Pink and superstars Jay-Z, The Beatles and U2. What do these artists have in common? Although they represent various styles and genres, the only common thing they have is fame, high record sales, popularity and a place in music history.

Negative correlations network is more difficult to analyze and reflects random events. There is Amy Winehouse, who had mainly negative correlations (Fig. 7) because there was no new album releases when she died in July 2011. Thus, the posthumous peak of her record sales happened when other artists record sales decreased. This phenomenon kicked off system from a normal state.

6. Methods of comparison

We compared sales networks with such networks that explain similarity. allmusic.com portal allows us to build a directed graph, because it provides two kinds of relations: following and followed by. All other networks are undirected so in comparison with them, we took accumulation of arrows as a weight of the links. This is standard action to project directed network into undirected one [18].

Additional to that, we ask some music experts to annotate their ideas of links between artists. We used surveys to do so. To obtain expert view, we took an average network from every entry to survey[‡] we were running. Experts were also asked how they rank relations. For example, Buda based on this network on his book about music history (Fig. 10).

If we compare all of naturally weighted networks (Table I), we could see that there is no relation between similarity networks and phonographic ones (does not matter whether networks are based on a pure correlation coefficient or its square).

A lack of correlation between sales and similarity can be understood easily from the point of view of economics. On the other hand, a lack of a relation between sales and popularity (Table II) was a surprise (both networks are establishing music market). Unfortunately, sale and critic (chart) networks do not cover the same time window and precision is different (sales record — every week,

[‡]The survey (gulakov.cba.pl/artysci.php) is still open for anybody who wishes to participate in future projects.

charts — every year). To imitate a year precision in the case of sales, we calculated correlations between artist's sales for every year and treat networks as binary ones (if correlation is significant, then the link exists) and accumulate them over years (2003–2007), for which we were able to check ranks in charts.

7. Conclusions

We conclude that from point of view of SNA, there are no significant correlations between relations of artists' record sales on phonographic market with the view how people see music genres that artists belong to (Tables I, II). Different groups of consumers, who built the market, do not buy in the same basket of records of similar artists, its extension to music genre is not allowed. The analysis of phonographic market based on the SNA revealed heterogeneity in population of customers. We postulate that the first one buys music because of genres, and the second one buys record because of an artist.

There is a different mechanism of consumer's decision making. There could be at least two possible explanations. The first come from classical theory of economy — marginal utility and product lifecycle management. Buying a new similar product has lower utility than previous one. It is well known that the main goal of record companies is to increase record sales. Thus, they often insist on releasing album on more convenient time, which maximizes income. Such a strategy could explain turbulence meaningful to the market (see Fig. 4). Unfortunately, the provided method to compare networks has some limitations. The biggest one is a lack of comparable time interval between quantitative algorithmic finding relations in "popularity" and "sales" network based on time interval (even those intervals differ), with "similarities" based on human memory.

On the other hand, the peaks of record sales caused by new releases display analogy to turbulence in the fluid mechanics [10] rather than financial markets [19, 20]. This issue is worth investigating in future papers.

References

[1] *The IFPI Report Recording Industry in Numbers*, IFPI 2010.

[2] *The Perfect Blues Collection* [booklet], Sony 2011.

[3] *The IFPI Record Industry in Numbers 2012 Edition*, IFPI 2012.

[4] www.musicharts.net/ .

[5] P. Audley, M. Boyer, *The Value of Music to Commercial Radio Stations*, Society For Economic Research On Copyright Issues, Berlin 2007.

[6] *United Nations Developed Programme*, Human Development Report, 2007/2008.

[7] A. North, A. Shilcock, D. Hargreaves, *Environment Behaviour* **35**, No. 5 (2003).

[8] *Valuing the Use of Recorded Music*, PriceWaterhouse Cooopers Report, Australia 2008.

[9] J. Park, O. Celma, M. Koppenberger, P. Cano, M.J. Buldú, *Int. J. Bifurcat. Chaos* **17**, 2281 (2007).

[10] A. Buda, *Physica A* **391**, 5153 (2012).

[11] A. Buda, *History Of Rock, Pop And Hip-hop According To Critics Polls*, Wydawnictwo Niezależne, Wrocław 2006 (in Polish).

[12] P. Cano, O. Celma, M. Koppenberger, *Chaos* **16**, 013107 (2006).

[13] J. Kwapień, S. Drożdż, *Phys. Rep.* **515**, 115 (2012).

[14] A. Buda, A. Jarynowski, *Life-Time of Correlations and Its Applications*, Wydawnictwo Niezależne, Wrocław 2010.

[15] W.E. Cox, *J. Business* **40**, 375 (1967).

[16] B.C. Twiss, *J. Prod. Innovat. Manag.* **1**, 19 (1984).

[17] R. Mantegna, *Eur. Phys. J. B* **11**, 193 (1999).

[18] M. Tsvetovat, A. Kouznetsov, *Social Network Analysis for Startups. Finding Connections on the Social Web*, O'Reilly Media, Sebastopol 2011.

[19] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, P. Scherrer, *Nature* **381**, 767 (1996).

[20] R. Mantegna, H.E. Stanley, *Nature* **383**, 587 (1996).