ICTP SAIFR
International Centre for Theoretical Physics
South American Institute for Fundamental Research

SCHOOL ON COMPLEX NETWORKS AND APPLICATIONS TO NEUROSCIENCES

MUSIC NETWORKS

JAVIER M. BULDÚ

UNIVERSIDAD REY JUAN CARLOS (MADRID, SPAIN)
CENTER FOR BIOMEDICAL TECHNOLOGY (MADRID, SPAIN)

WELCOME TO THE MOST

SUBJECTIVE,

INCOMPLETE,

AND **FOLKLORIC**

LECTURE YOU WILL EVER FIND ABOUT

MUSIC NETWORKS!

(GOOD NEWS: NO EQUATIONS…
OK, MAY BE ONE OR TWO…)

OVERVIEW

I.- Creating Music Networks

II.- Note Networks

- Context
- Music vs Language

III.- Song Networks

- Affinity

IV.- Artist Networks

- Similarity vs Collaboration

V.- User Networks

- Recommendation

Creating Music Networks

(I-V)

MUSIC & NETWORKS: DO THEY HAVE ANYTHING IN COMMON?

Music is nice. Network theory is nice. Let's join them!

ONE POSSIBLE CLASSIFICATION OF MUSIC NETWORKS

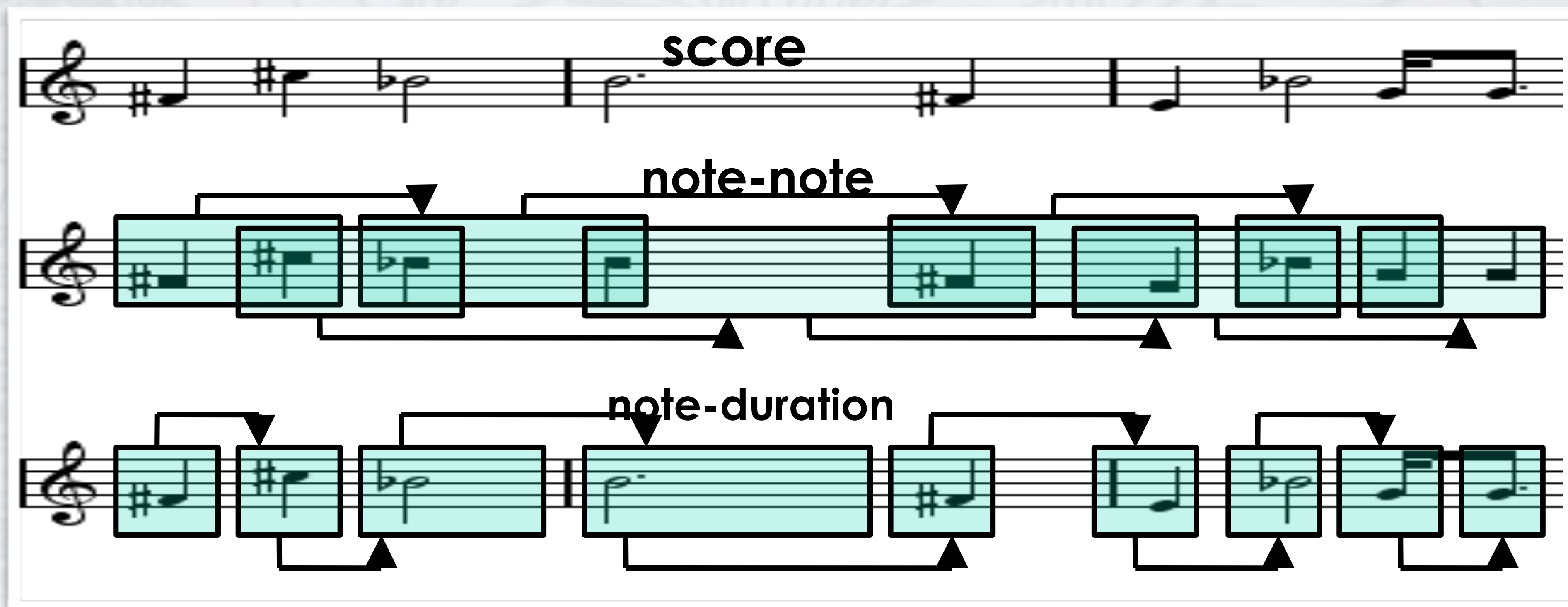
There is a diversity of networks related to music. According to the nature of the nodes, one possible classification is:

- Note Networks
- Song Networks
- Artist Networks
- User Networks

Other kind of classifications are possible, e.g. based on the nature of the links

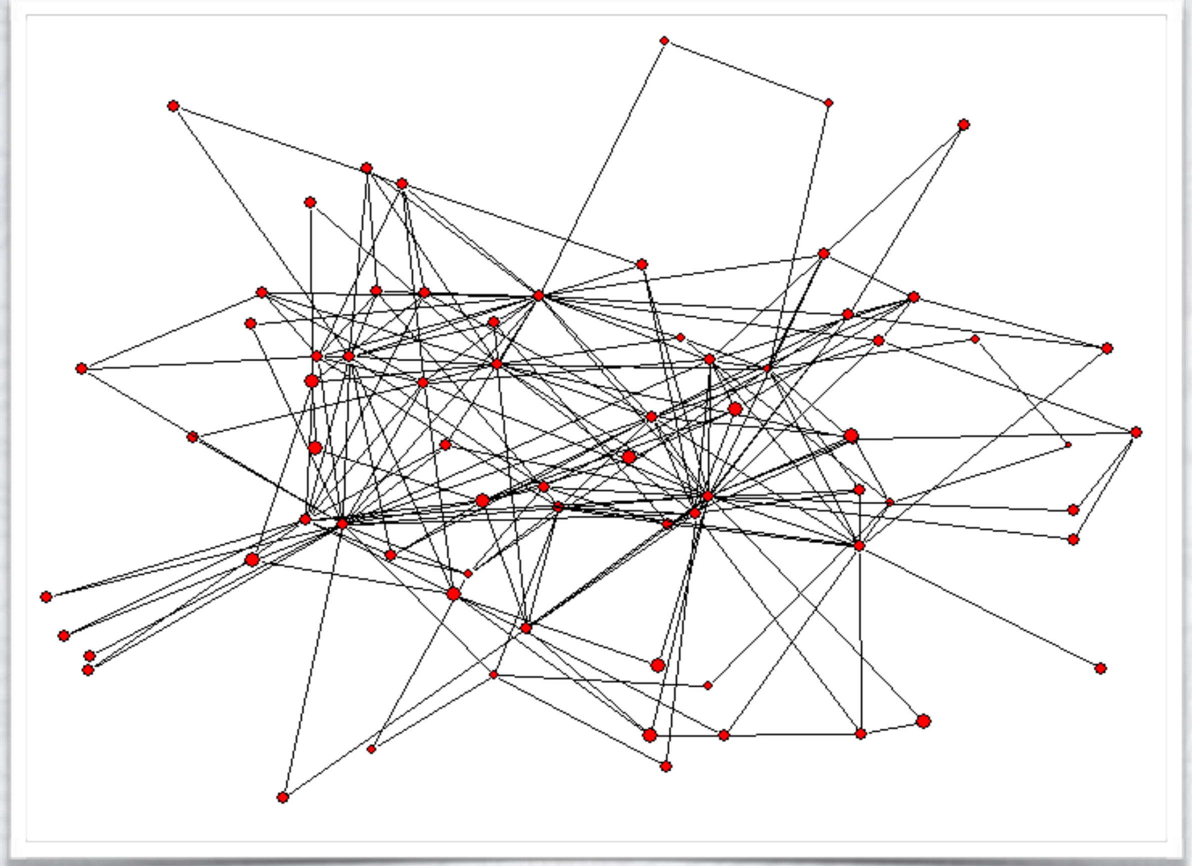
NOTE NETWORKS

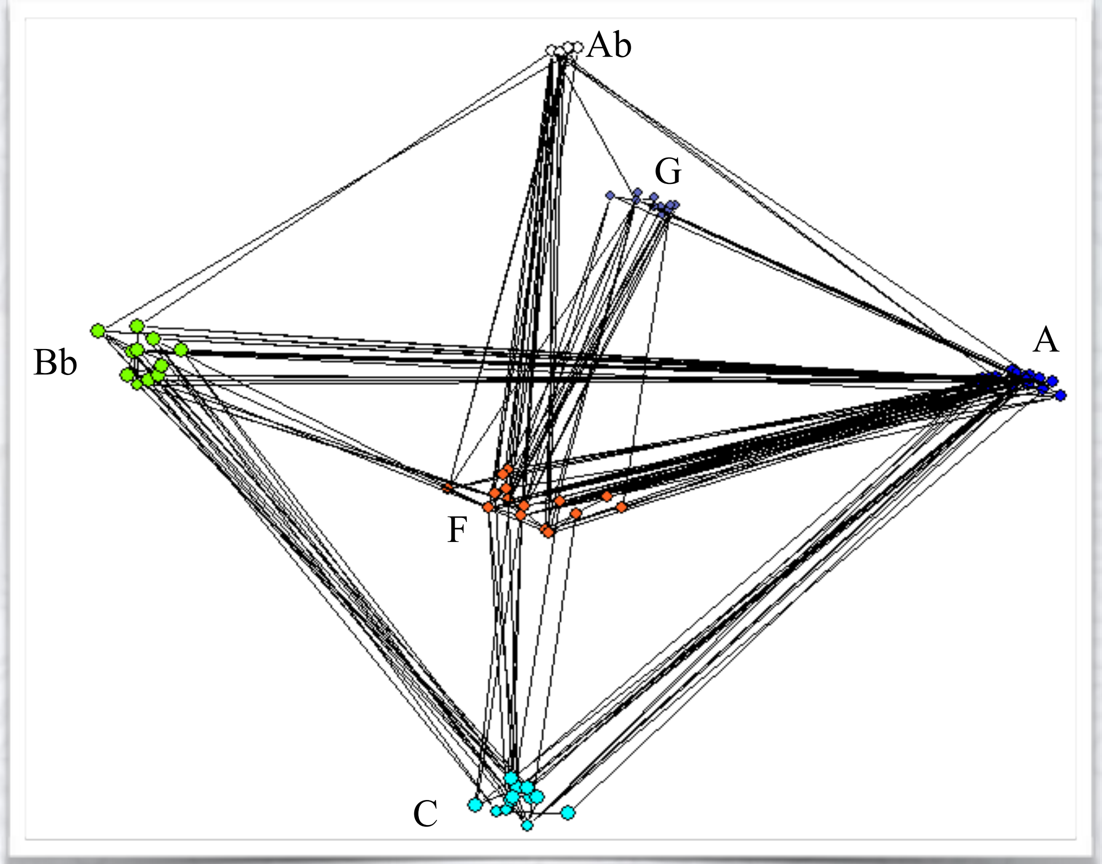
In this kind of music networks, the notes are the nodes, which are linked by proximity:



NOTE NETWORKS

Let's see an example; a note-duration network:
(guess artist and song!)

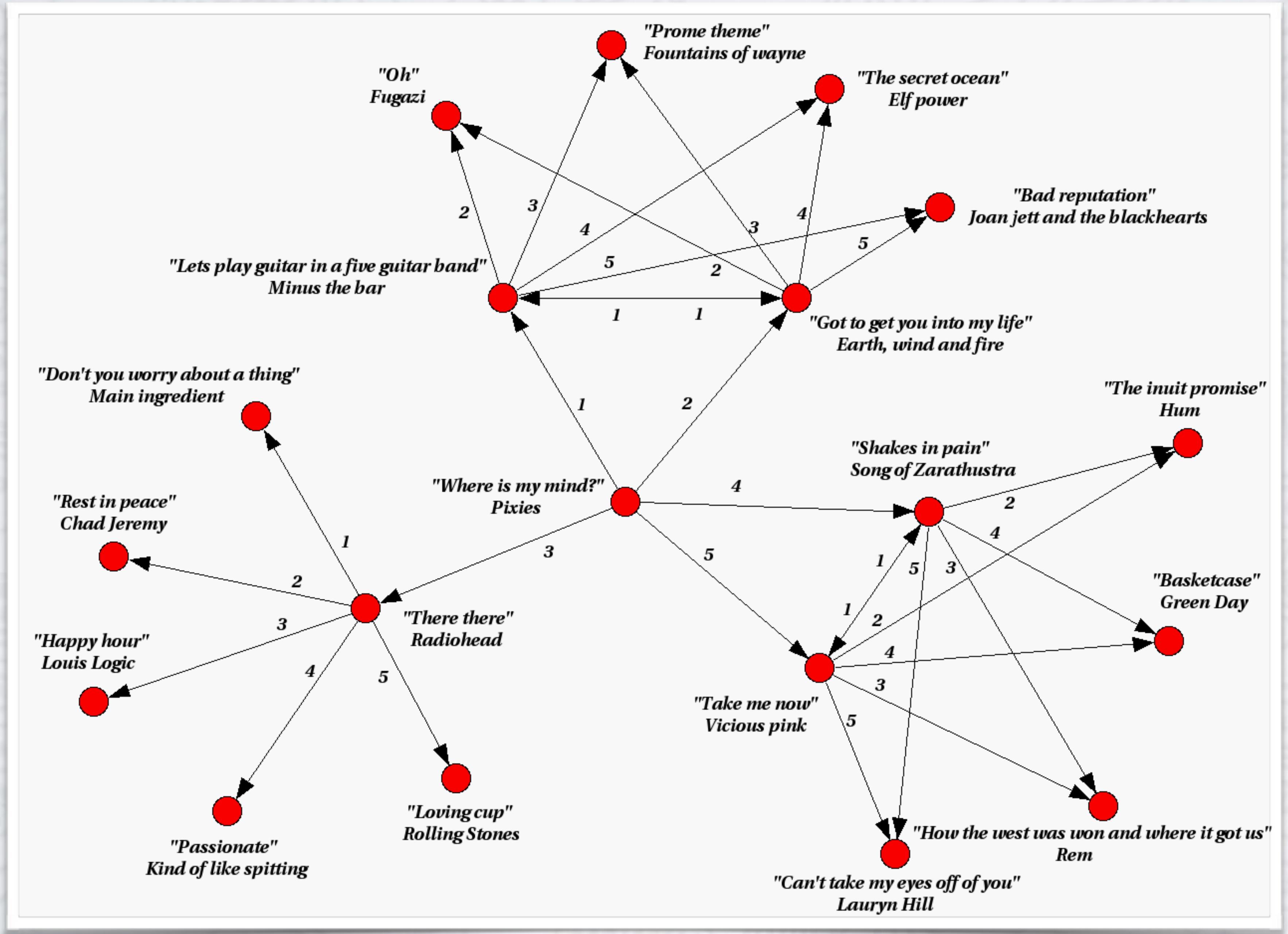




On the left: a note-duration network. On the right, the same network organized by notes.

SONG NETWORKS

In this kind of music networks, songs are the nodes, which may be linked according to different relations:

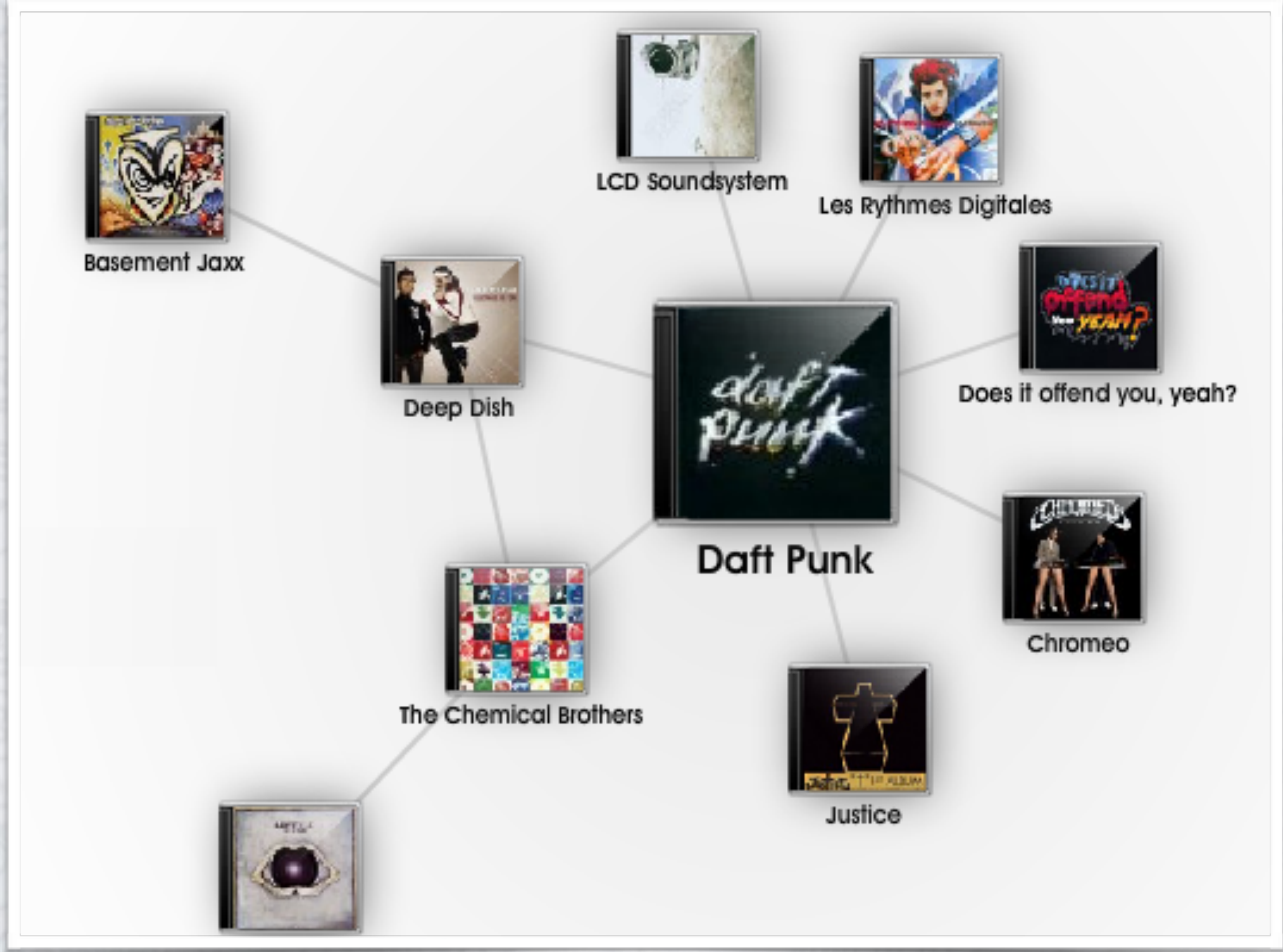


In this example, songs are extracted from a playlist Dataset. Only the top M=5 co-ocurrences are displayed, leading to a weighted and directed network.

From "The complex network of musical tastes", J.M. Buldú, P. Cano, M. Koppenberger, J.A. Almendral and S. Boccaletti, New J. Phys. 9, 172 (2007).

ARTIST NETWORKS

In this kind of music networks, musical artists are the nodes, who may be linked according to different relations:

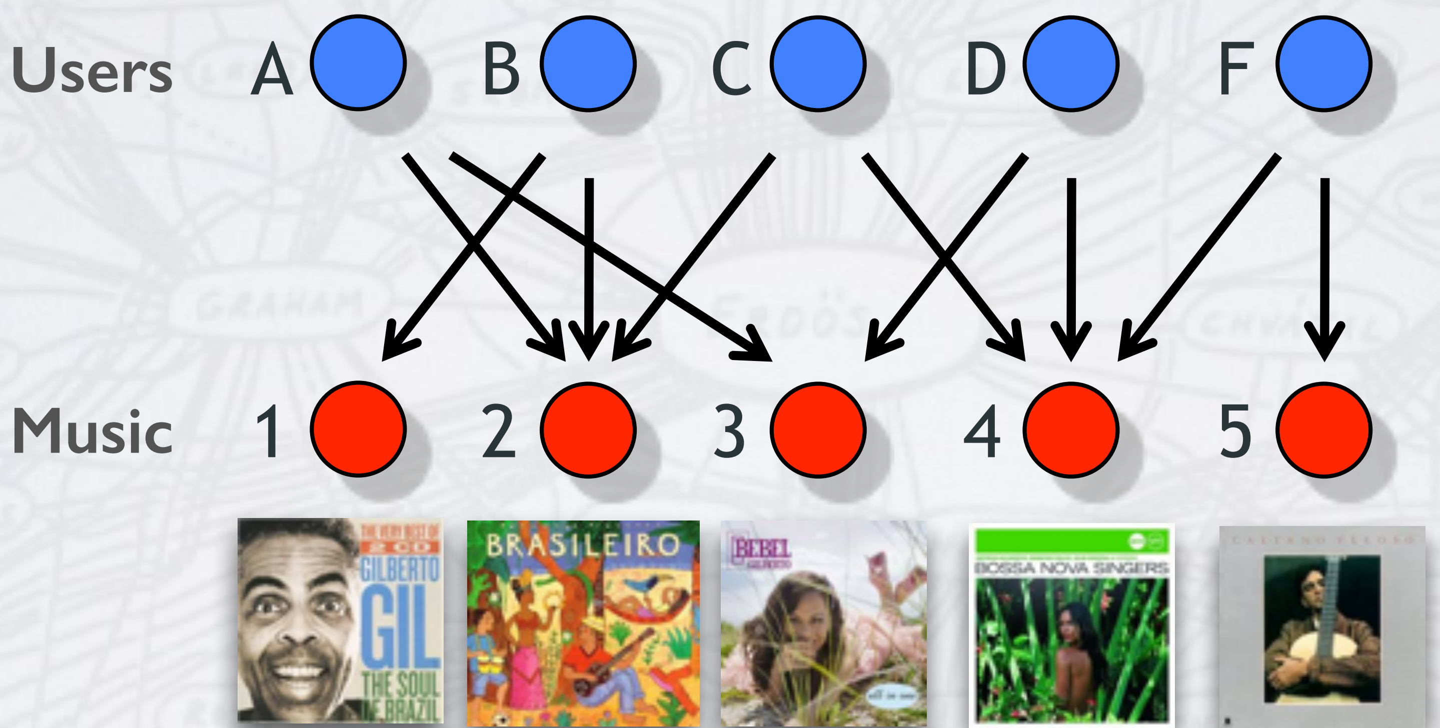


Links?

- Similarity
- Collaboration
- Affinity
- ….

USER NETWORKS

In this kind of music networks, users that have consumed a musical product are the nodes:



This kind of networks are extremely useful for designing **recommendation** systems.

Note Networks

(II-V)

ZIP'S LAW IN MUSICAL SEQUENCES

First, let's have a look at language:

Zip's law: the ranking of words follows a power law

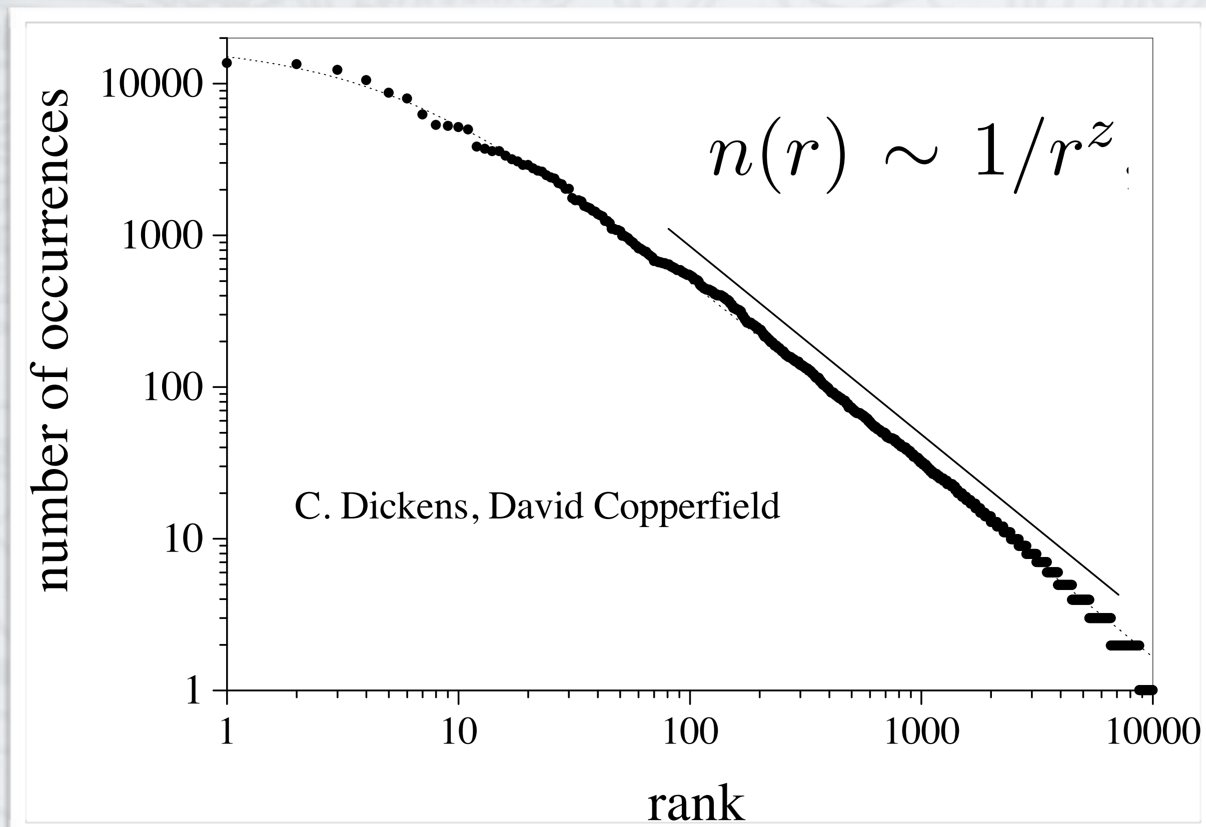


Figure 1: Zipf's plot (number of occurrences n versus rank r) for Dickens's David Copperfield. The number of different words is $V = 13,884$, and the total number of words is $T = 362,892$. In this double-logarithmic plot the straight line manifests the power-law dependence of $n(r)$ for large r. The dotted curve is a least-square fitting with the prediction of Simon's model, equation (1).

Simon model: "as words are successively added to the text, a context is created. As the context emerges, it favors the later appearance of certain words –in particular, those that have already appeared– and inhibits the use of others."

$$n(r) = \frac{1}{(a + br)^z}$$

Depend on the context

Depends on semantics

Zanette, Musicae Scientiae **10**, 3-18 (2006)

ZIP'S LAW IN MUSICAL SEQUENCES

In music, context is determined by a hierarchy of intermingled patterns occurring at different time scales (harmonic progressions, melody, tone, rhythm…):

Zipf's law and the creation of musical context

Damián H. Zanette
Consejo Nacional de Investigaciones Científicas y Técnicas
Instituto Balseiro, 8400 Bariloche, Río Negro, Argentina

Musicae Scientiae **10**, 3-18 (2006)

Does the Simon model apply to the **musical context**?

midi file
↓
note+duration
↓
network?

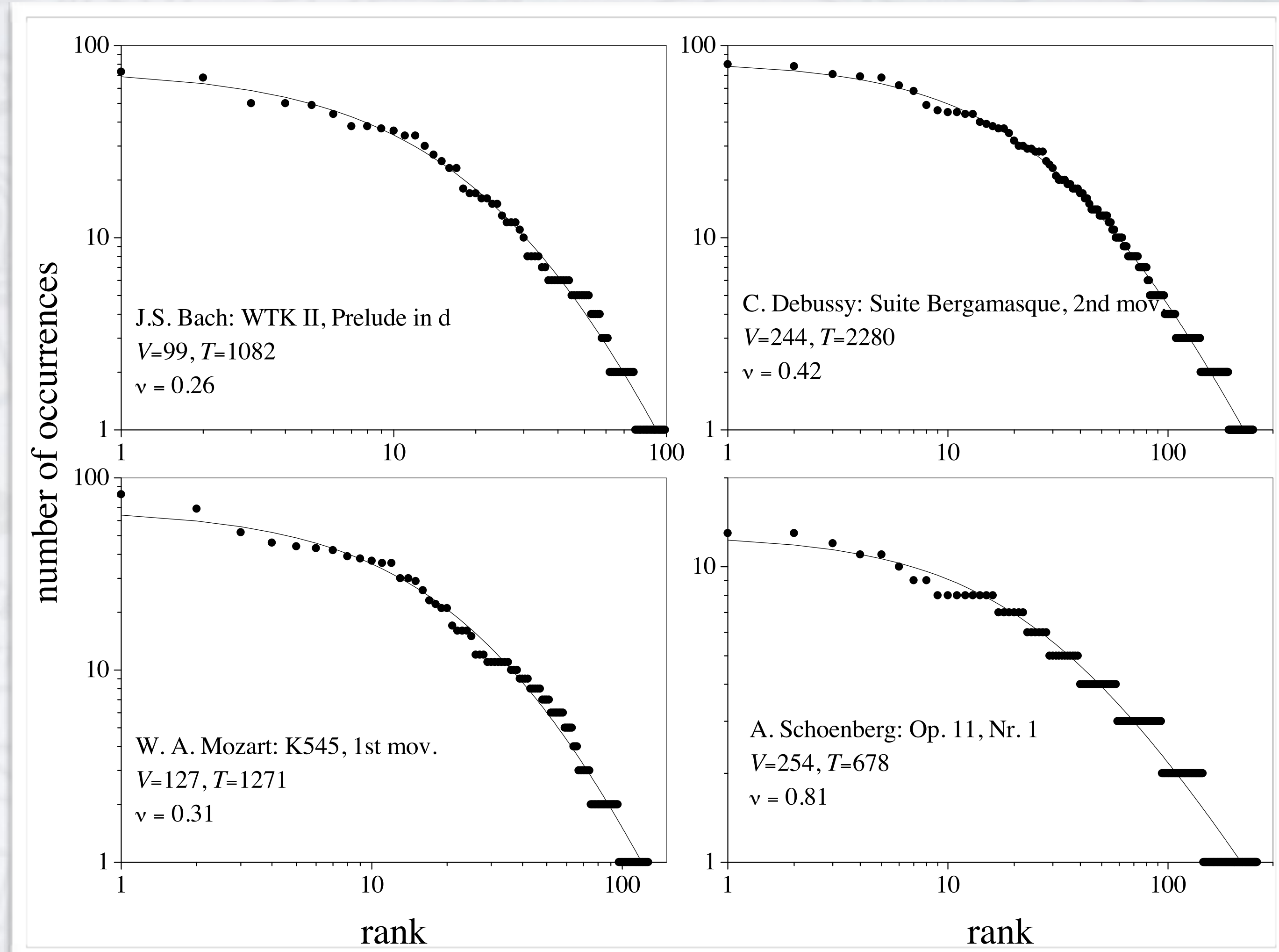
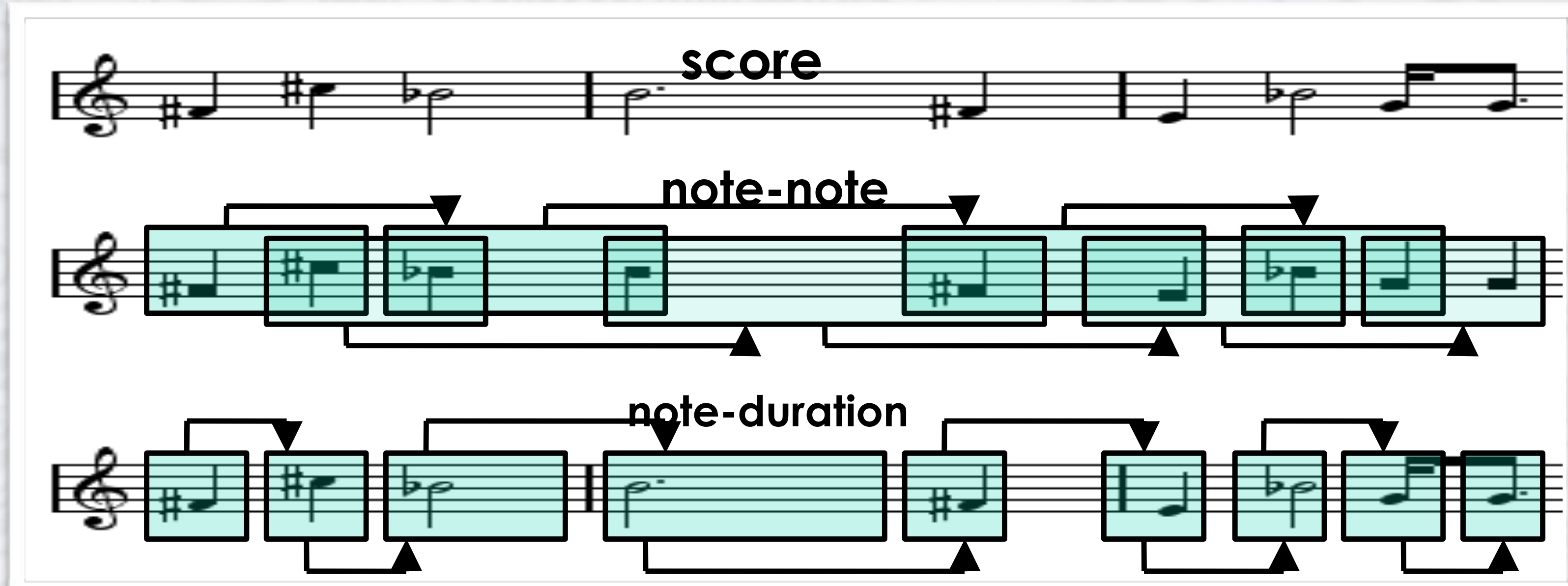


Figure 2: Zipf's plots for single notes in four musical compositions for keyboard. Their titles, as well as the corresponding value of V and T, are indicated in each panel. Curves stand for least-square fittings with the prediction of Simon's model, equation (1). The resulting exponent ν, which provides a quantitative measure of context definiteness, is given with each plot.

ORGANIZATION OF (NOTE) MUSIC NETWORKS

Beyond the existence of a musical context, how are (note based) music networks?:

Dataset: Music database consists of over 13000 western contemporary music pieces (in MIDI format) covering a broad spectra of music styles. Two different lexicons are defined by means of note duplets and note-duration pairs.

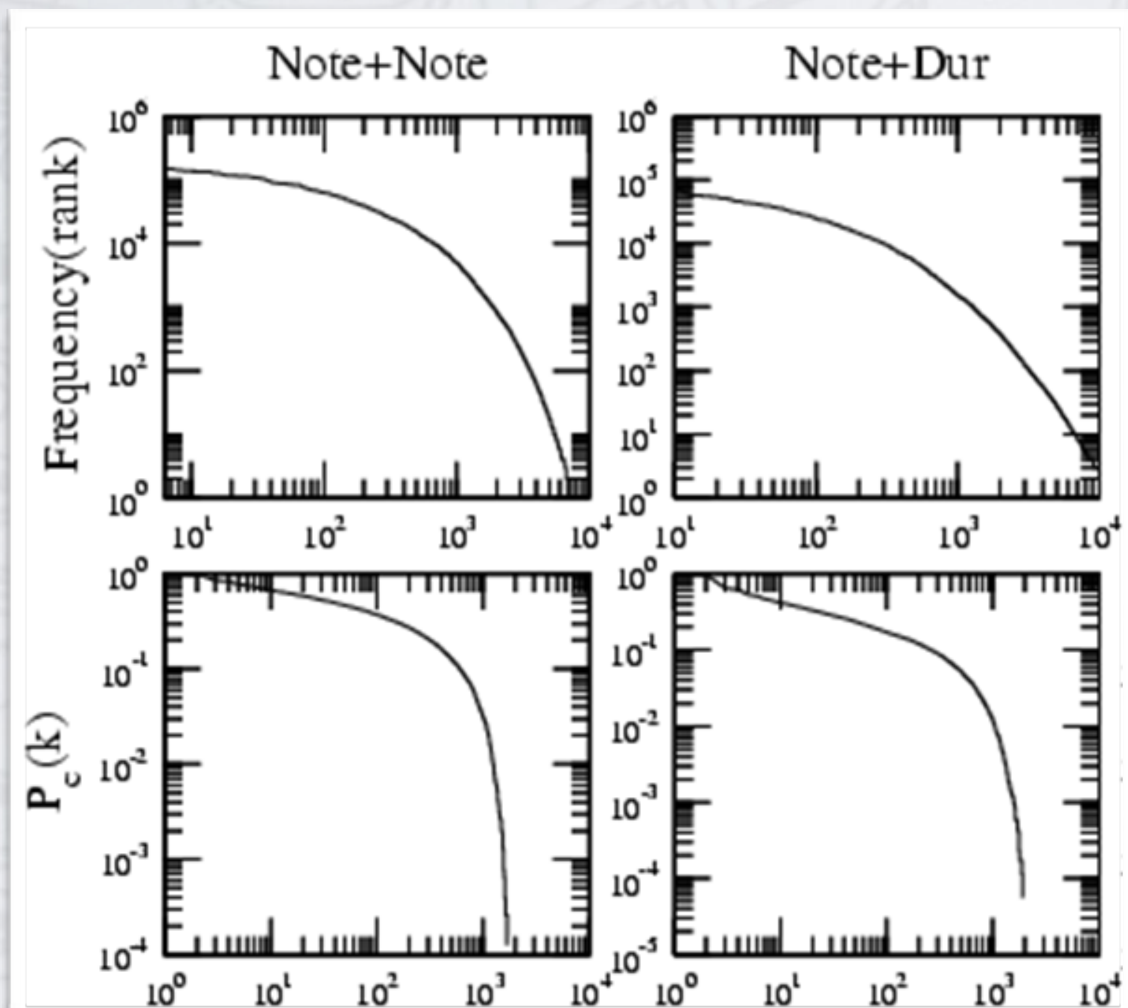


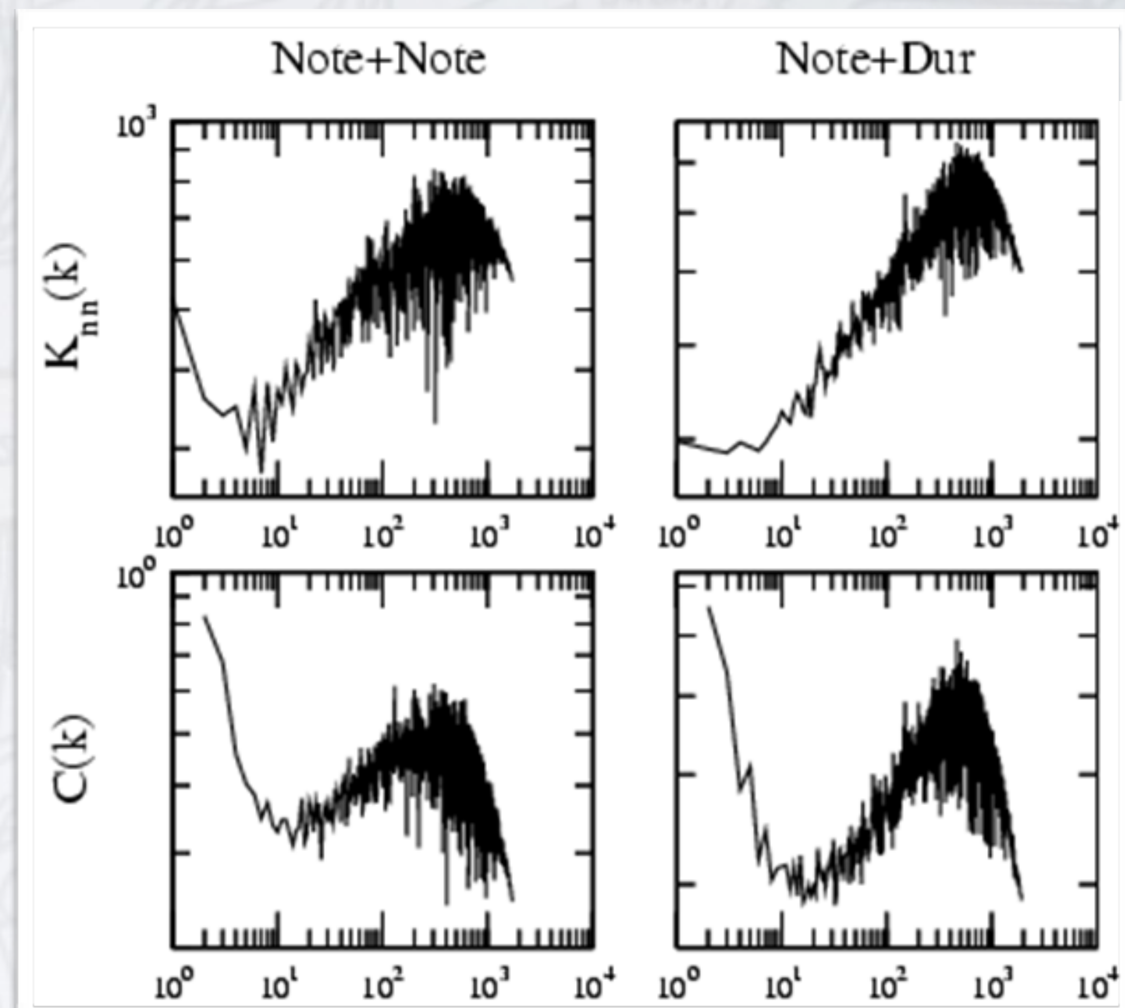
	n	m	$<k>$	C
Note + Dur	37198	2453937	107.3	0.29
Note + Note	7712	940449	180.8	0.47

P. Cano, M. Kopenberger, unpublished

ORGANIZATION OF (NOTE) MUSIC NETWORKS

Let's have a look of their basic topological properties:

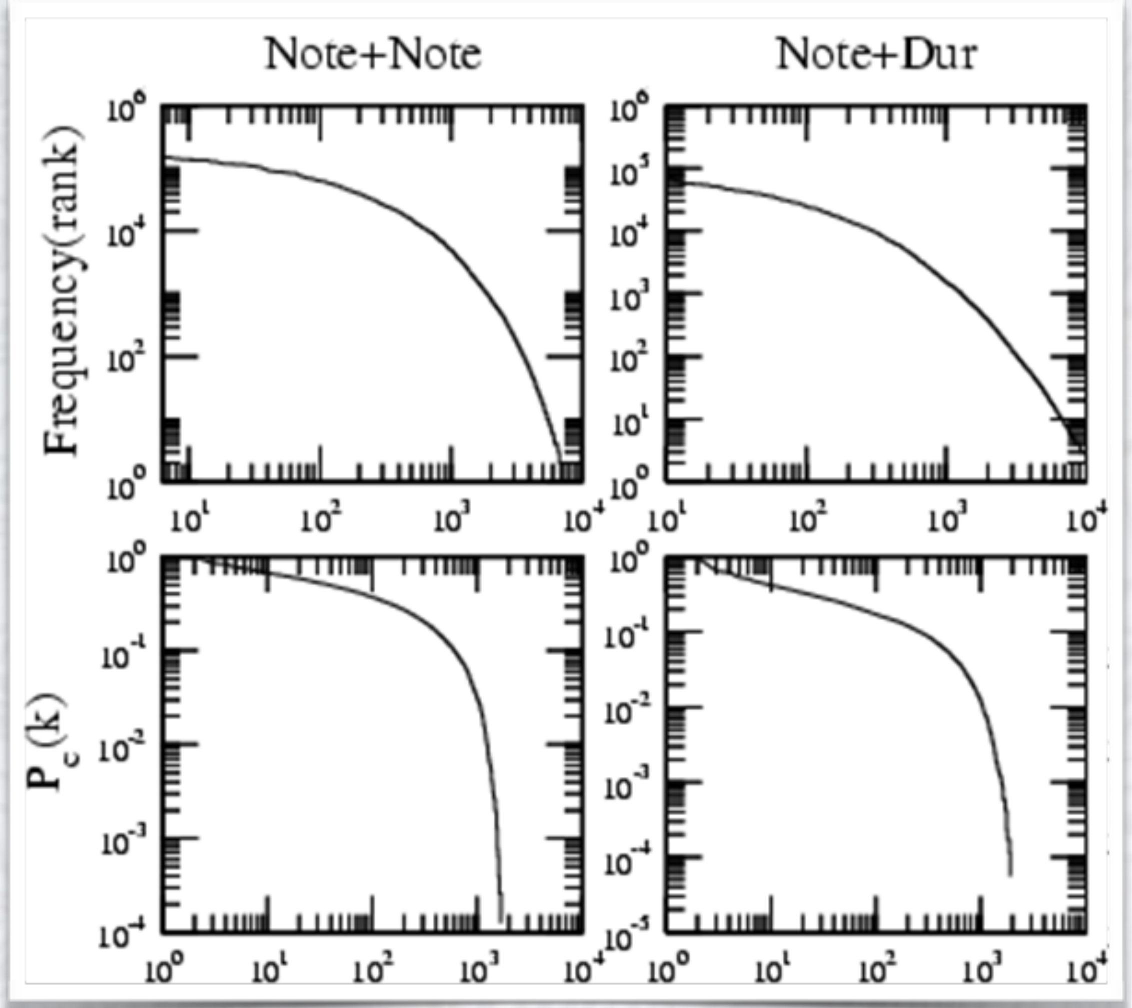




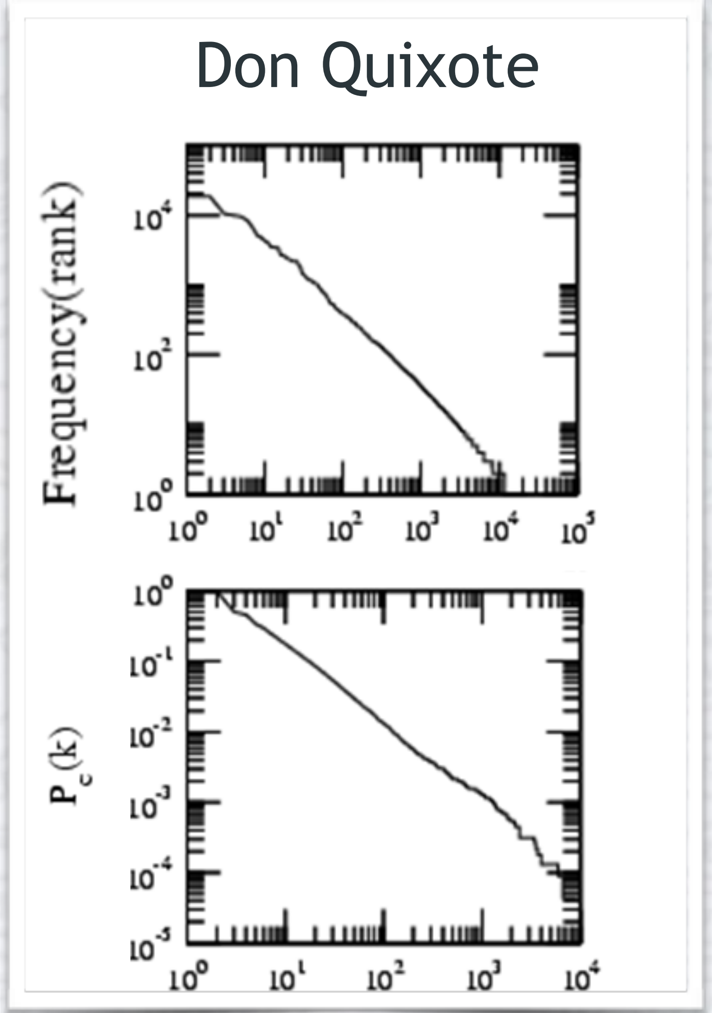
ORGANIZATION OF (NOTE) MUSIC NETWORKS

What about comparing music (note) networks with language networks?

music



text



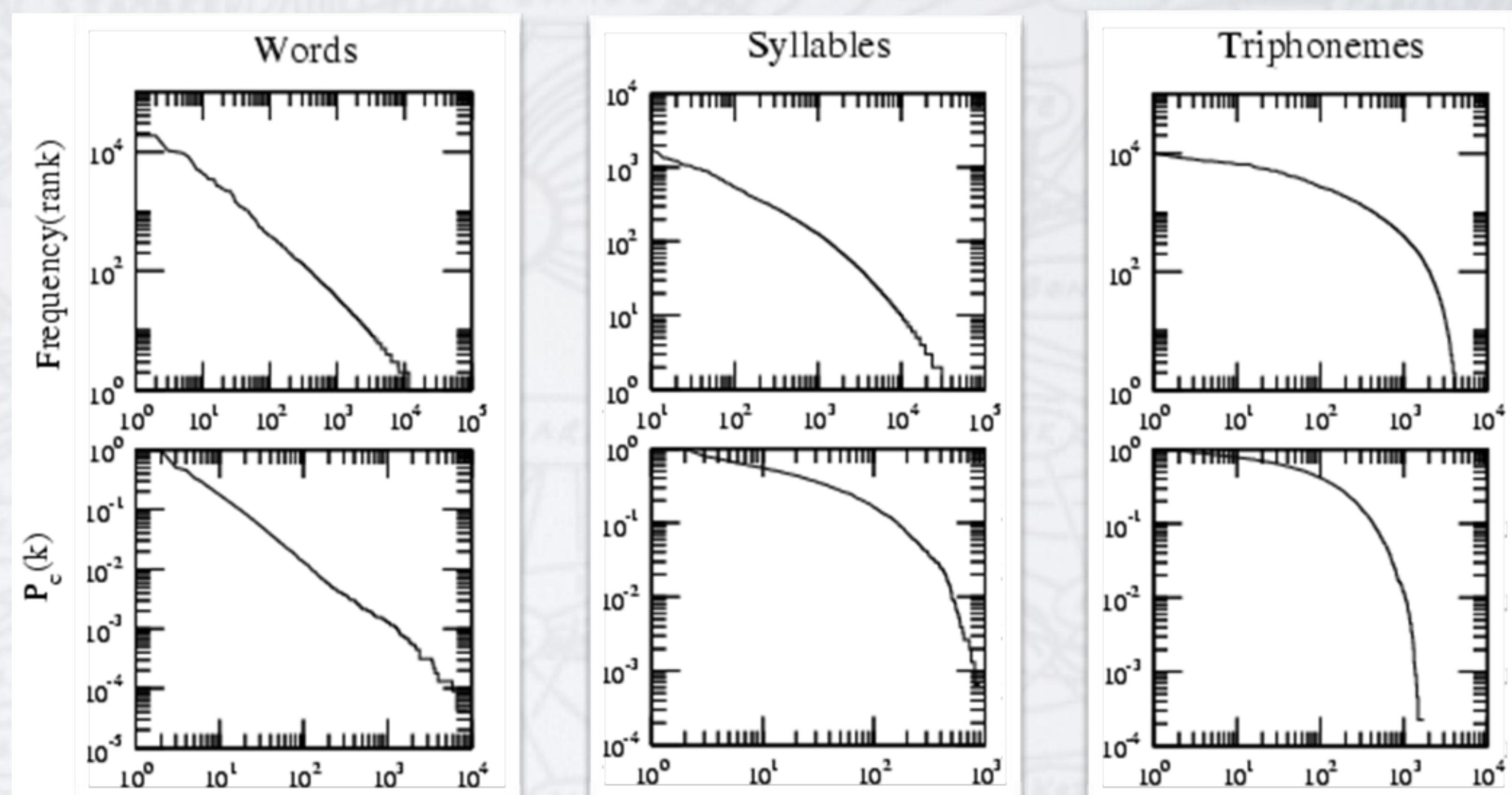
We obtain different frequency and probability distributions.

ORGANIZATION OF (NOTE) MUSIC NETWORKS

But, what happens if we have a look at **“other scales”**?

Redes-Complejas ⟶ Re-des-com-ple-jas

	n	m	$<k>$	C
Quijote (triphonemes)	4461	368882	153.7	0.33
Quijote (syllables)	1531	51124	55.7	.93
Quijote (words)	22518	152581	12.6	0.74

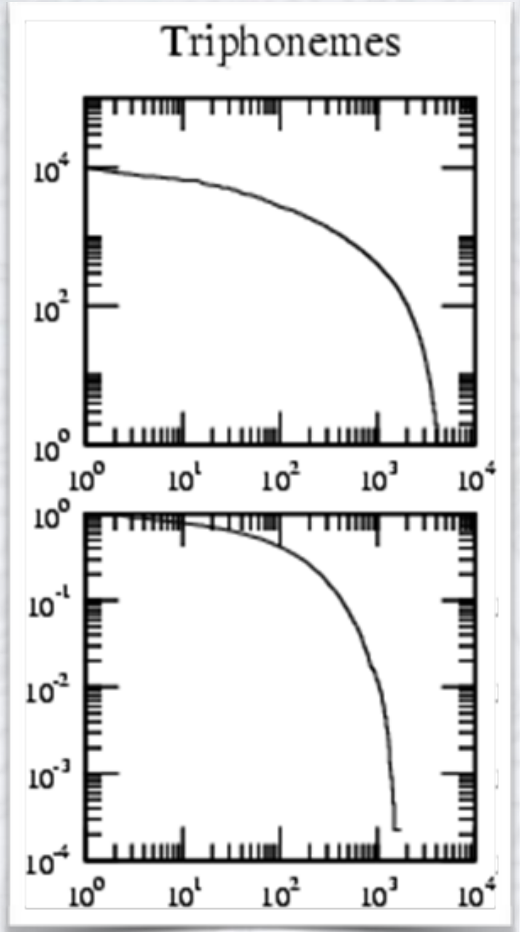


The distribution changes when the semantics is lost.

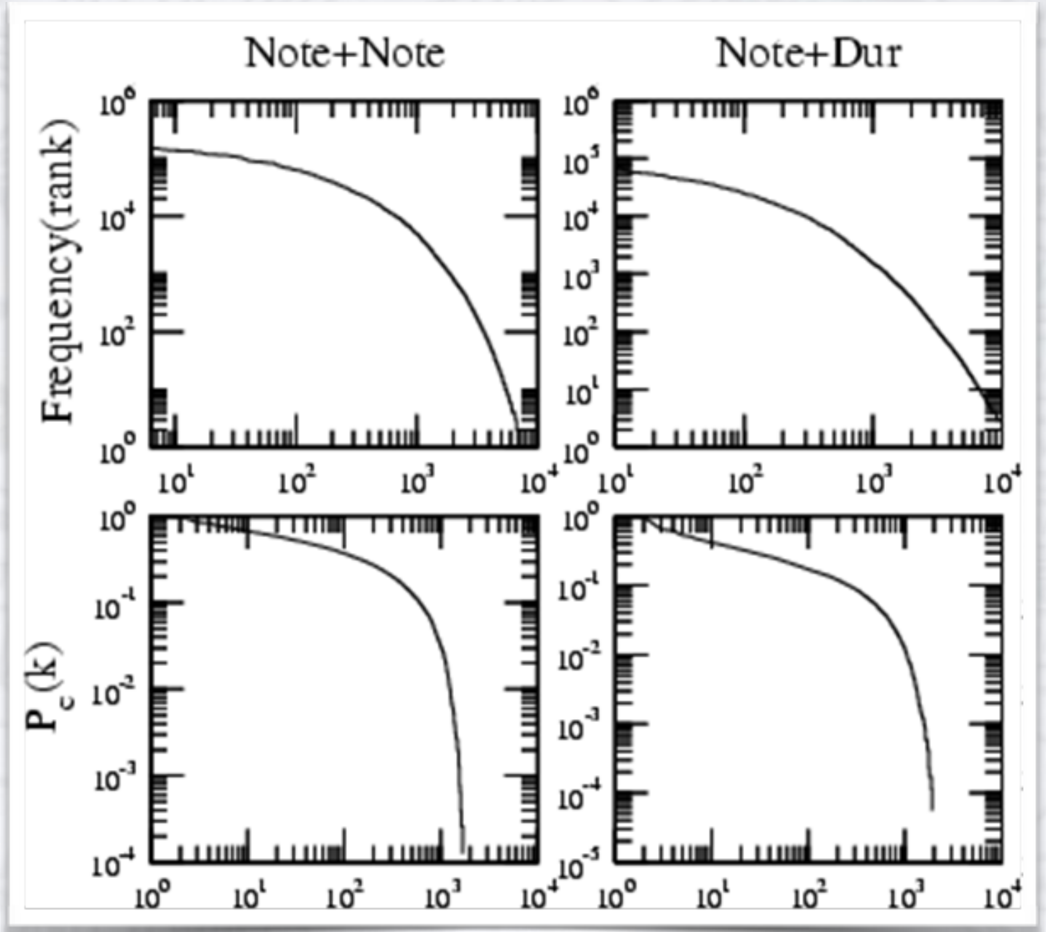
ORGANIZATION OF (NOTE) MUSIC NETWORKS

When semantics is lost, frequency and degree distributions become similar:

text



music

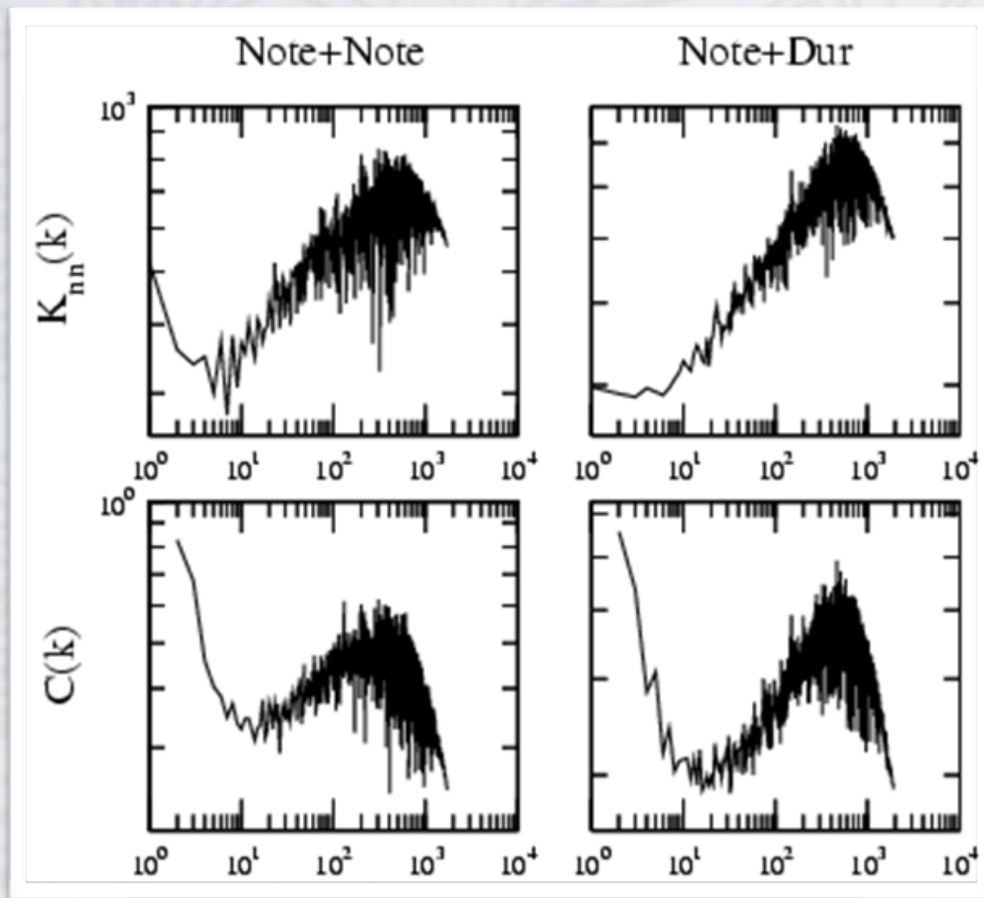


What about the other topological properties?

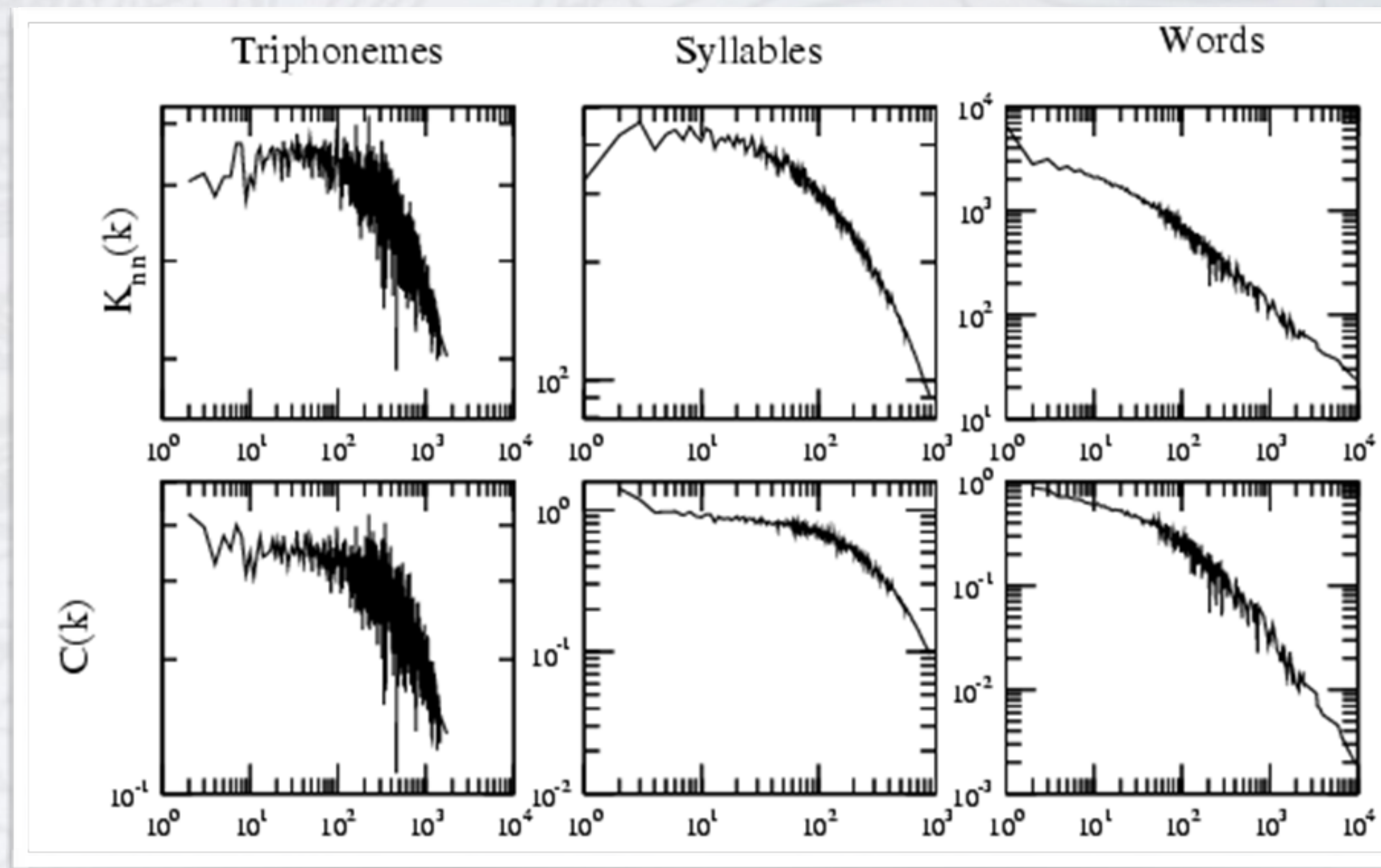
ORGANIZATION OF (NOTE) MUSIC NETWORKS

What happens with the degree-degree correlations and clustering?:

music



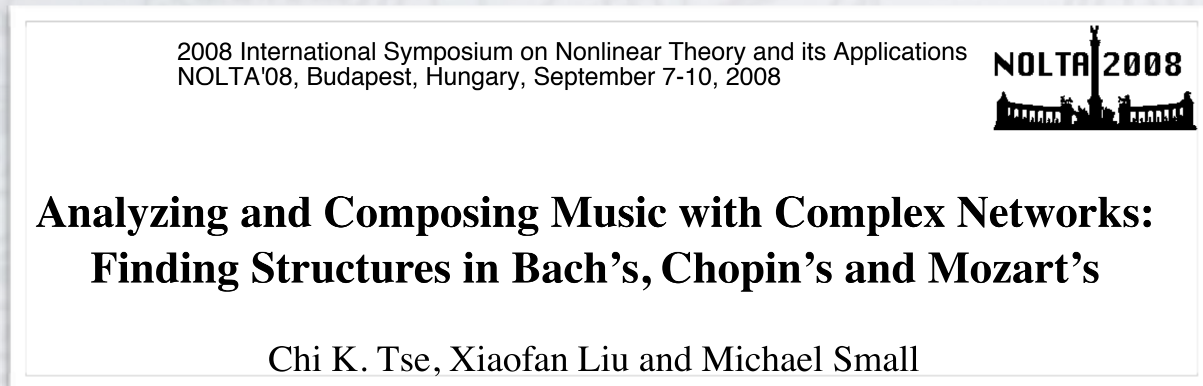
text



Music (note) networks are **assortative** with a particular clustering distribution

ORGANIZATION OF (NOTE) MUSIC NETWORKS

A part of investigating their properties, we can use music networks for other purposes:

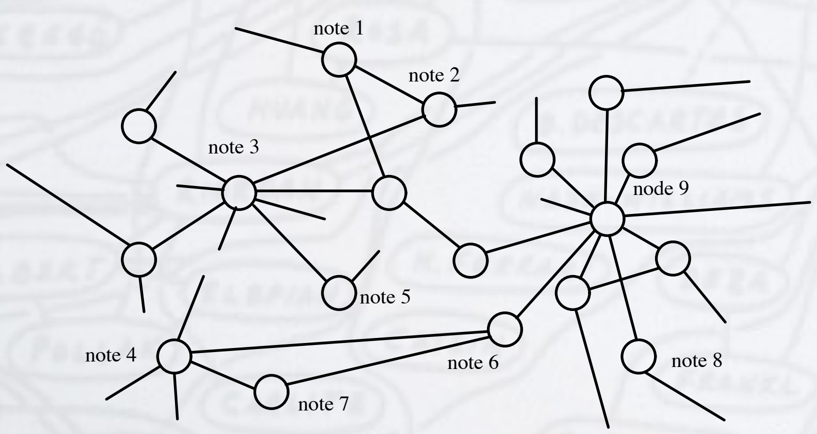


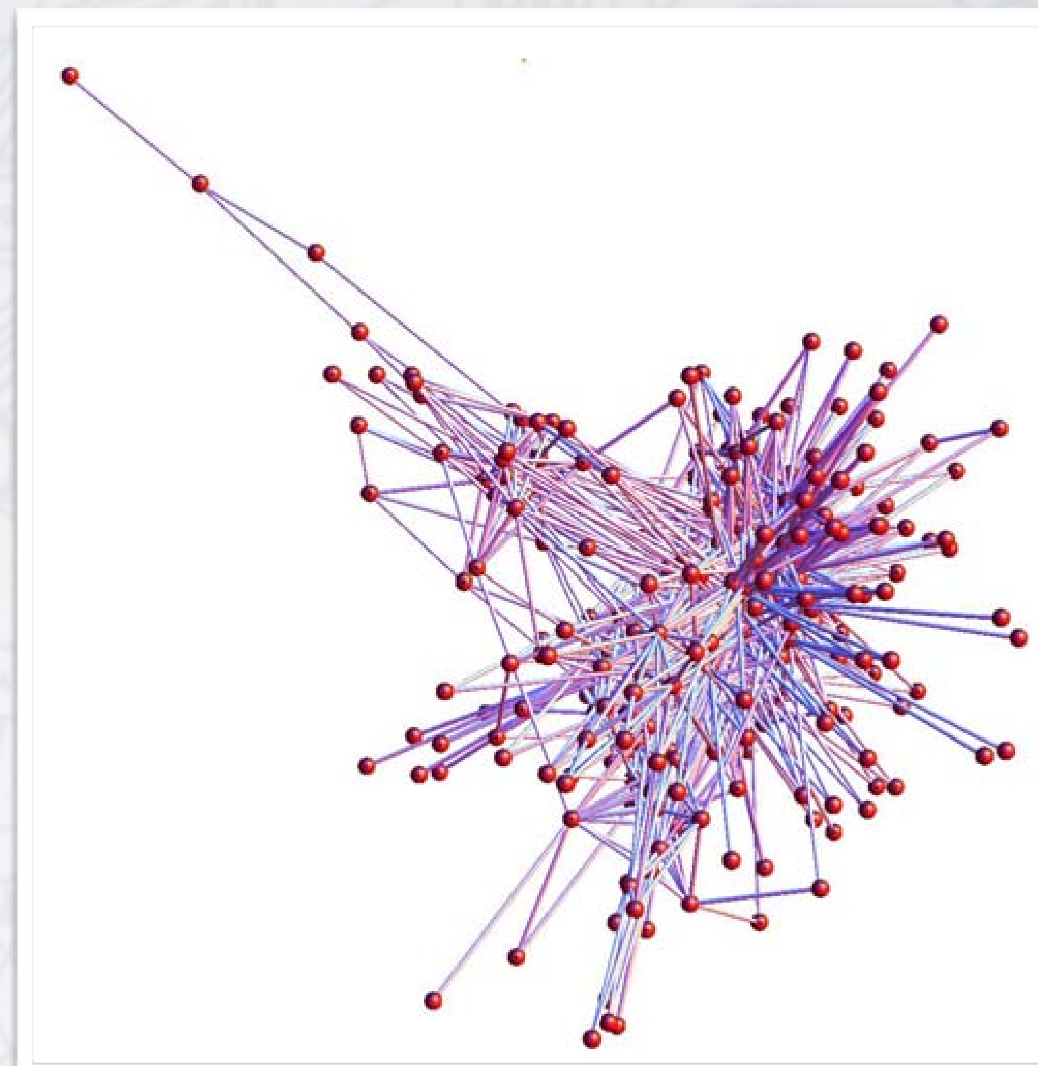
2008 International Symposium on Nonlinear Theory and its Applications
NOLTA'08, Budapest, Hungary, September 7-10, 2008

Analyzing and Composing Music with Complex Networks: Finding Structures in Bach's, Chopin's and Mozart's

Chi K. Tse, Xiaofan Liu and Michael Small

It is possible to use the network properties to create music with the use of guided random walks





This is a Bach's Sonata. Not joking

Song Networks

(III-V)

SONG NETWORKS

In this kind of music networks, the songs are the nodes. They are interesting for different reasons:

- For analyzing their **structure** (paths, modules,…)
- For the detection of the most **influential songs**
- For **classification** purposes (labeling)
- For designing efficient (automatic) **recommendation** systems

SONG NETWORKS

We are going to overview **an example** about how songs networks are related to musical tastes:

The complex network of musical tastes

Javier M Buldú[1], P Cano[2], M Koppenberger[2], Juan A Almendral[1] and S Boccaletti[3]

[1] Departamento de Física, Universidad Rey Juan Carlos, Tulipán s/n, 28933 Móstoles, Madrid, Spain
[2] Music Technology Group, Universitat Pompeu Fabra, 08003, Barcelona, Spain
[3] CNR-Istituto dei Sistemi Complessi, Via Madonna del Piano, 10, 50019 Sesto Fiorentino (Florence), Italy
E-mail: javier.buldu@urjc.es

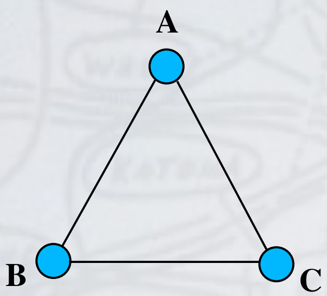
New Journal of Physics **9** (2007) 172
Received 13 October 2006
Published 28 June 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/6/172

www.theartofthemix.com

- AoM consists of a website where users upload and exchange **playlists** of their favorite music.
- The **songs,** somehow, **fit in those lists**, even though they do not need to belong to the same country, decade or musical genre.
- In this way, a certain connection results between songs of the list, whose origin is based on the **musical taste** of the playlist author.
- We create **networks** where songs are the nodes and **co-occurrence** in a playlist gives rise to links between them.

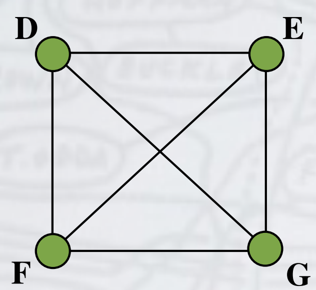
SONG NETWORKS

We create networks that, interestingly, evolve in time:



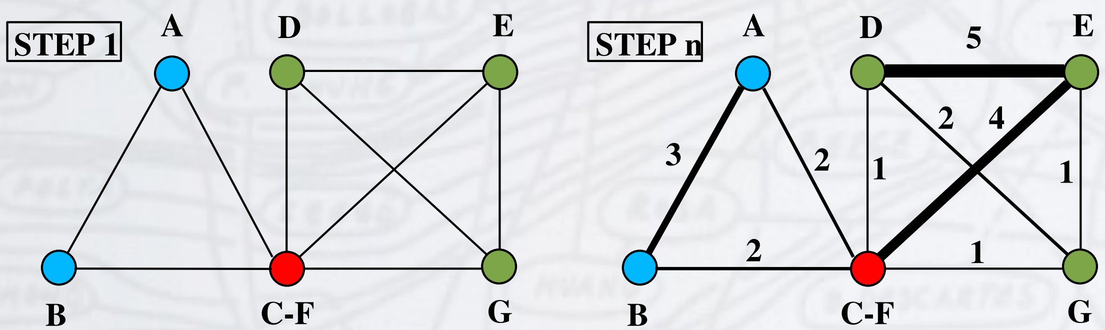
PLAYLIST I

TRACK	ARTIST	SONG
A	Jevetta Steele	Somewhere Over the Rainbow
B	Desmond Dekker	Israelites
C	**Johnny Cash**	**Ring of Fire**



PLAYLIST II

TRACK	ARTIST	SONG
D	Adam Ant	A Place in the Country
E	Hal Sirowitz	Chopped Off Arm
F	**Johnny Cash**	**Ring of Fire**
G	Concrete Blonde	Close to Home



	The Art of the Mix							
Year	1998	1999	2000	2001	2002	2003	2004	2005
n (nodes)	9450	26 223	60 673	127 519	240 157	360 034	457 660	482 856
m (links)	54 789	204 277	614 644	1 711 053	4 115 893	7 278 256	9 946 715	10 602 036
S of GCC	58.6%	84%	90%	92.9%	93.9%	93.7%	92.8%	92.4%
$\bar{d}$ (d_{max})	6.65 (15)	5.24 (13)	4.70 (11)	4.37 (12)	4.22 (12)	4.13 (13)	4.12 (15)	4.12 (15)
C	0.958	0.906	0.870	0.843	0.828	0.820	0.819	0.819

Table 1. Summary of several network parameters as a function of year: number of nodes n, number of edges m, relative size S of the GCC, precisely, its percentage among all nodes, mean geodesic path $\bar{d}$ inside the GCC, diameter d_{max} of the GCC and the average clustering coefficient C of the network.

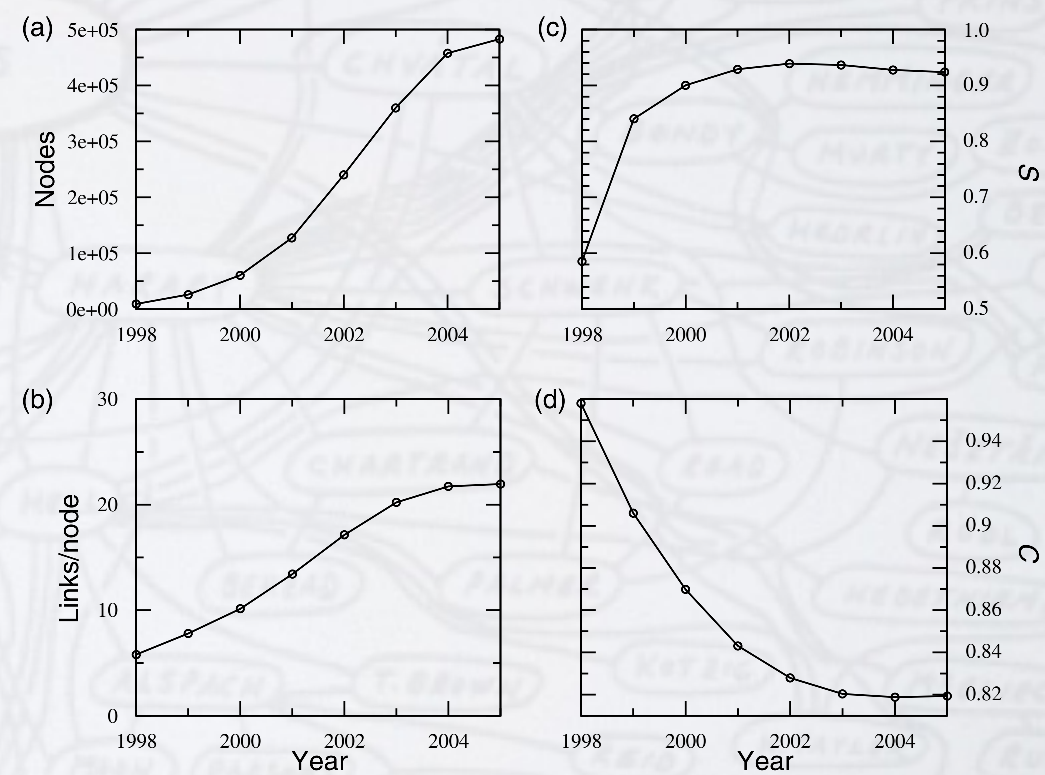


Figure 2. Evolution of the network: (a) number of nodes and (b) number of weighted links per node from 1998 to 2005. In (c) and (d) we compute, respectively, the relative size of the giant component S and the evolution of the mean clustering coefficient C.

SONG NETWORKS

We can analyze the interplay between songs, quantify co-occurrences and detect regions of influence:

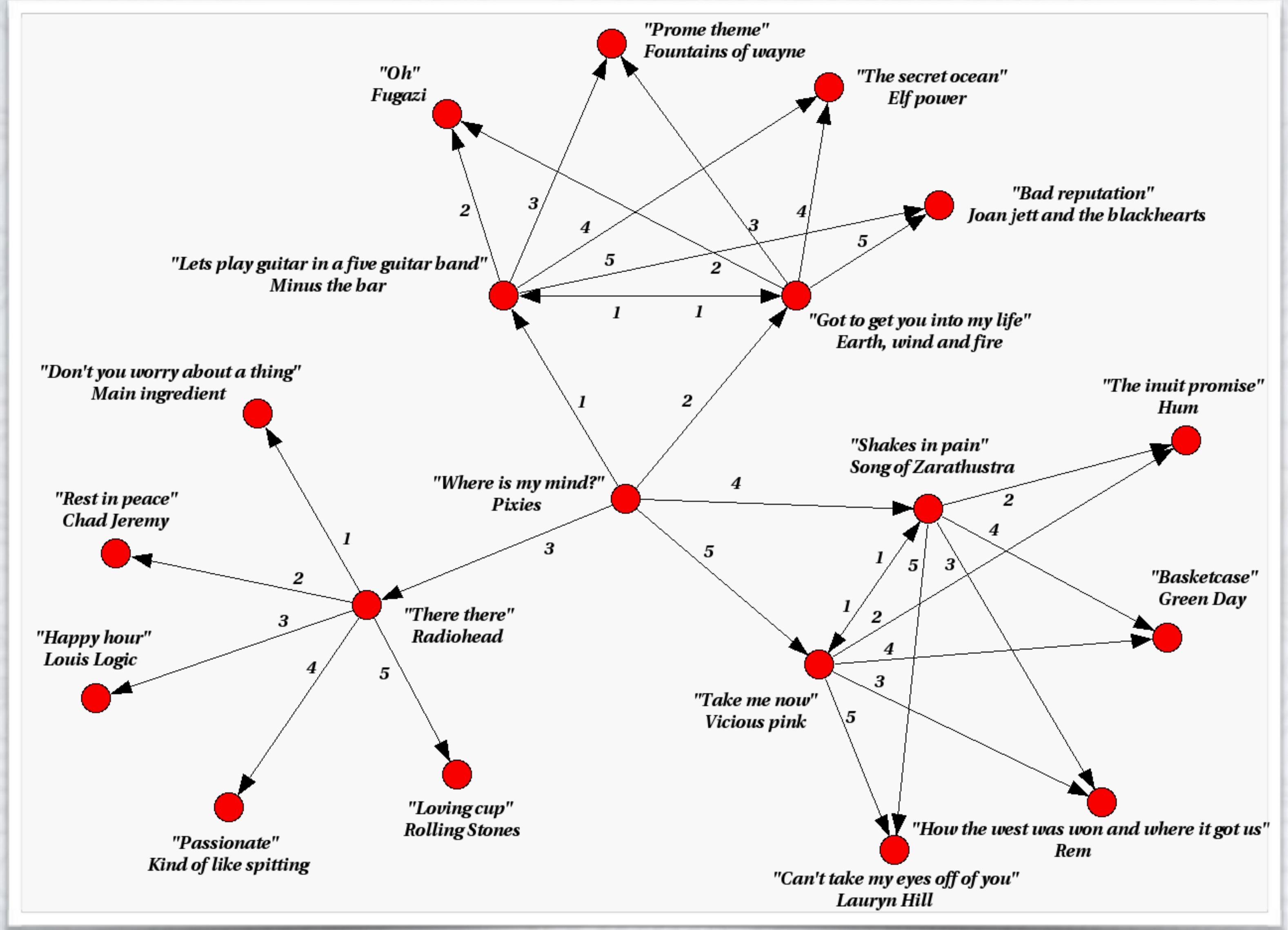


Figure 8. Example of a recommendation network obtained from the affinity matrix. Only the part of the network surrounding the song 'Where is my mind?' is shown. Numbers correspond to the ranking of affinity of the outgoing links.

- The **affinity network** projects co-occurrences into a network.
- Each song is linked to the other *M* songs that have **co-occurred** the most with it.
- We obtain a **weighted directed network.**
- The analysis of this network, at **different scales**, gives information about the interplay between songs

SONG NETWORKS

Interestingly, it is also possible to track the evolution of the "greatest hits":

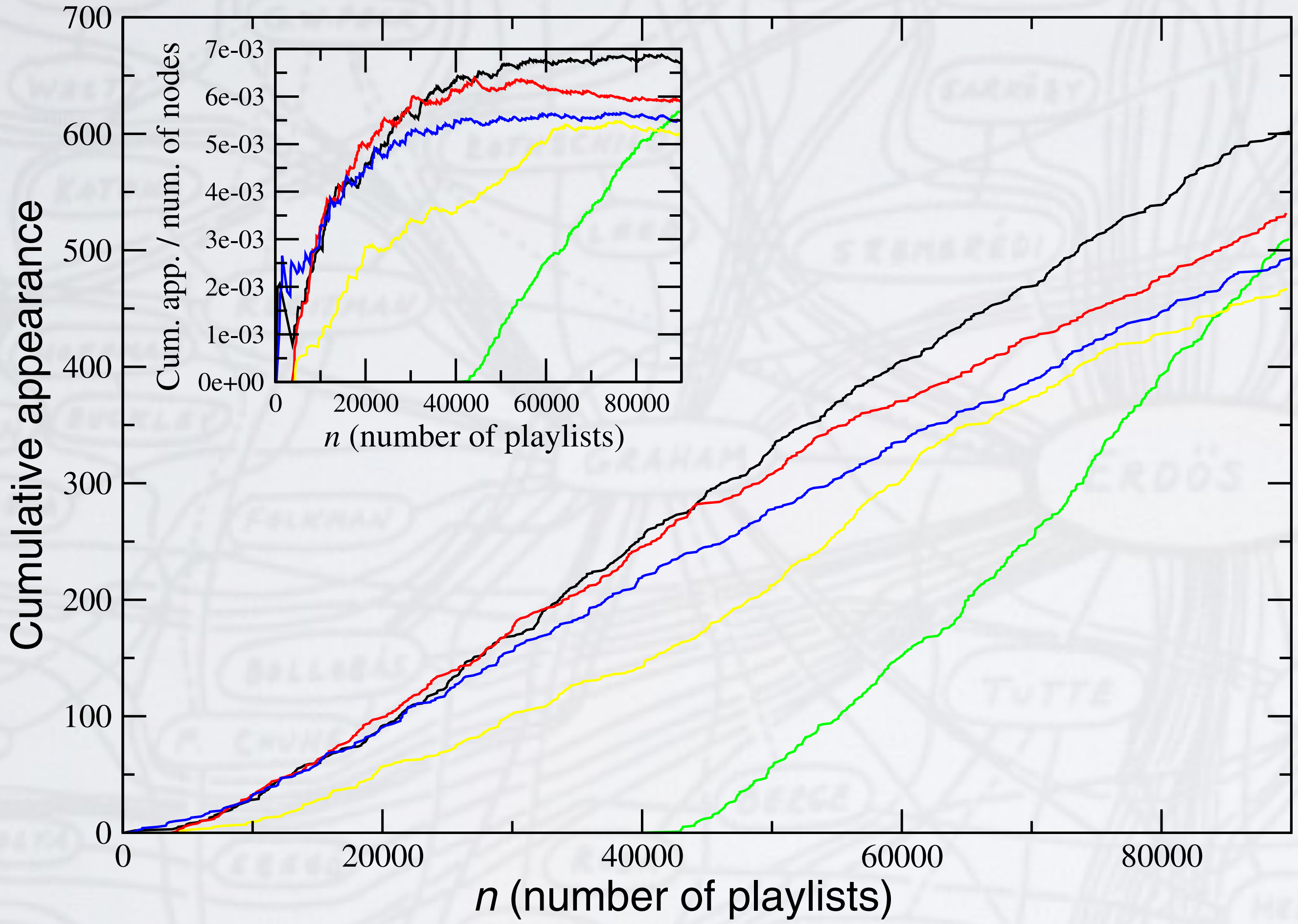


Figure 9. Temporal evolution of the three highest connected songs in the whole network (i.e., that obtained at year 2005), where n is the total number of playlists at a certain date. In the ordinate axis, we measure the accumulated appearances of a certain song among all playlists. In the inset, we evaluate the rate of appearance by measuring the number of appearances per playlist.

- It is possible to **track the rate** of appearance of the songs.
- It is possible to generate **models** describing the evolution of these songs.
- We can **classify** songs according to their rate of appearance.
- We can **predict** the behavior of a song and advance its decay.

Artist Networks

(IV-V)

ARTIST NETWORKS

In this case, **musical artists are the nodes**, which are linked according to a certain interplay between them:

- **Similarity:** Two artists are linked if the play similar music (links are normally created by musical editors).
- **Collaboration:** Two artists get connected if they have ever played together.
- **Affinity:** Two artist are linked due to a certain affinity, such as appearing in the same playlist or having a disc bought by the same person.
- Any other you may think about…

ARTIST NETWORKS

Let's see an example of how to use community detection to extract information about (artist) music networks:

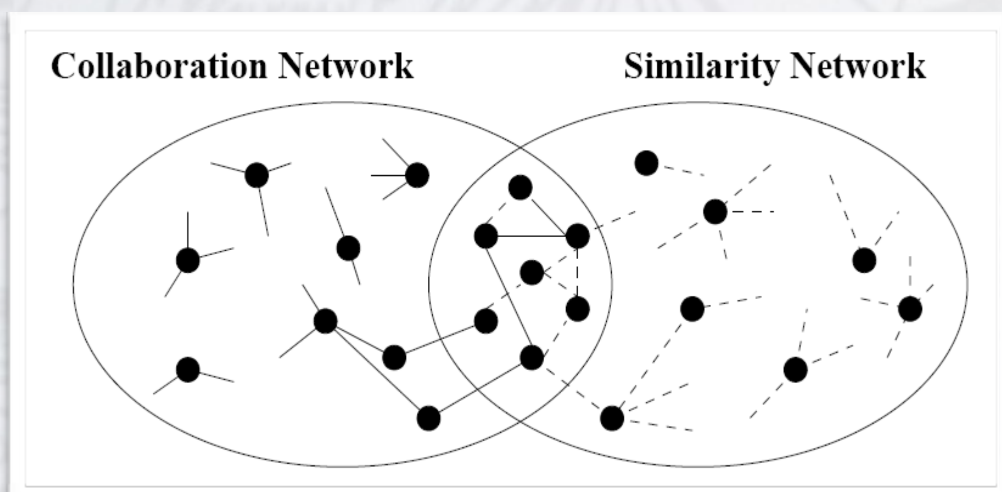
CHAOS **18**, 043105 (2008)

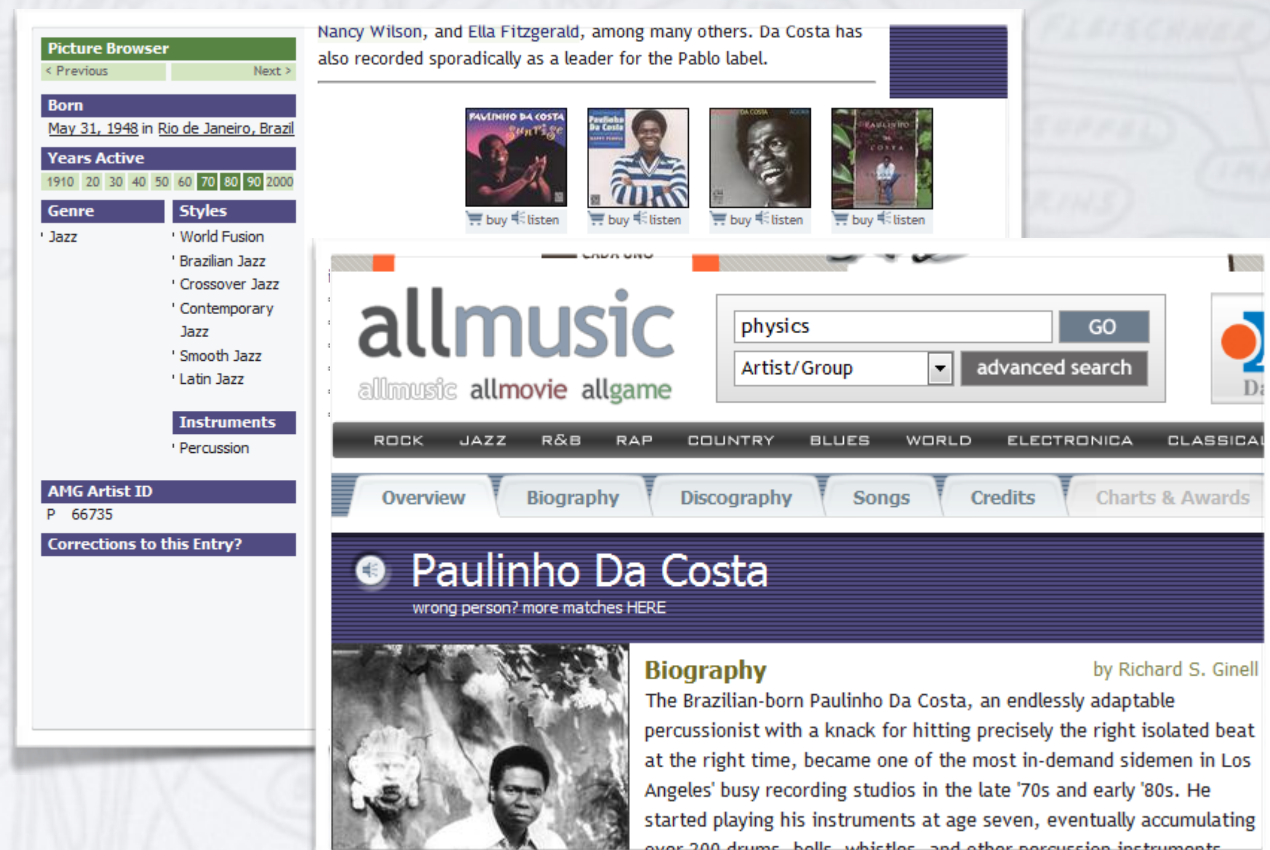
Community structures and role detection in music networks

T. Teitelbaum,[1] P. Balenzuela,[1] P. Cano,[2,3] and Javier M. Buldú[4]

[1]*Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina*
[2]*Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain*
[3]*BMAT, Barcelona Music and Audio Technologies, 08018 Llacuna 162, Barcelona, Spain*
[4]*Complex Systems Group, Universidad Rey Juan Carlos, Tulipán s/n, 28933 Móstoles, Madrid, Spain*

(Received 3 July 2008; accepted 3 September 2008; published online 14 October 2008)



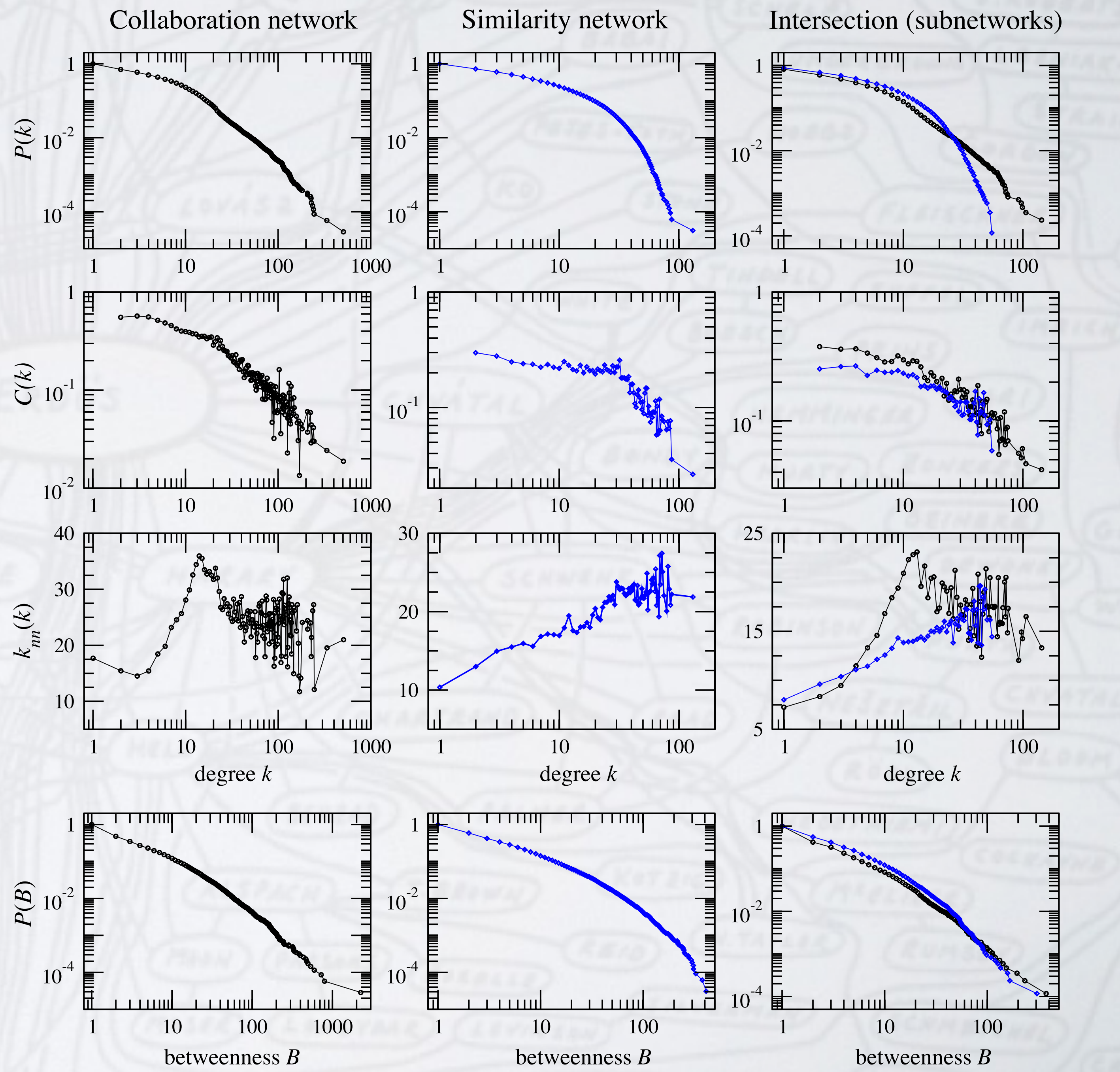


Both datasets are obtained from AllMusicGuide database (http://www.allmusicguide.com)

ARTIST NETWORKS

Despite having the same number of nodes, networks do not necessary have the same structure:

	Similarity network		Collaboration network	
	entire	intersection	entire	intersection
n	32 377	8 509	34 724	8 509
m	117 621	24 950	123 122	20 232
size of S_0	30 384 (94%)	7 219 (85%)	30 945 (89%)	6 054 (71%)
$\bar{d}$ ($d_{\max}$)	6.5 (22)	6.0 (20)	6.4 (23)	6.3 (19)
C	0.185 (18.5%)	0.178 (17.8%)	0.182 (18.2%)	0.171 (17.1%)
$k_{\max}$	131 R.E.M.	55 Eric Clapton	508 P. Da Costa	143 P. Da Costa R. Van Gelder
highest-betweenness artist	Sting	Sting	P. Da Costa	P. Da Costa

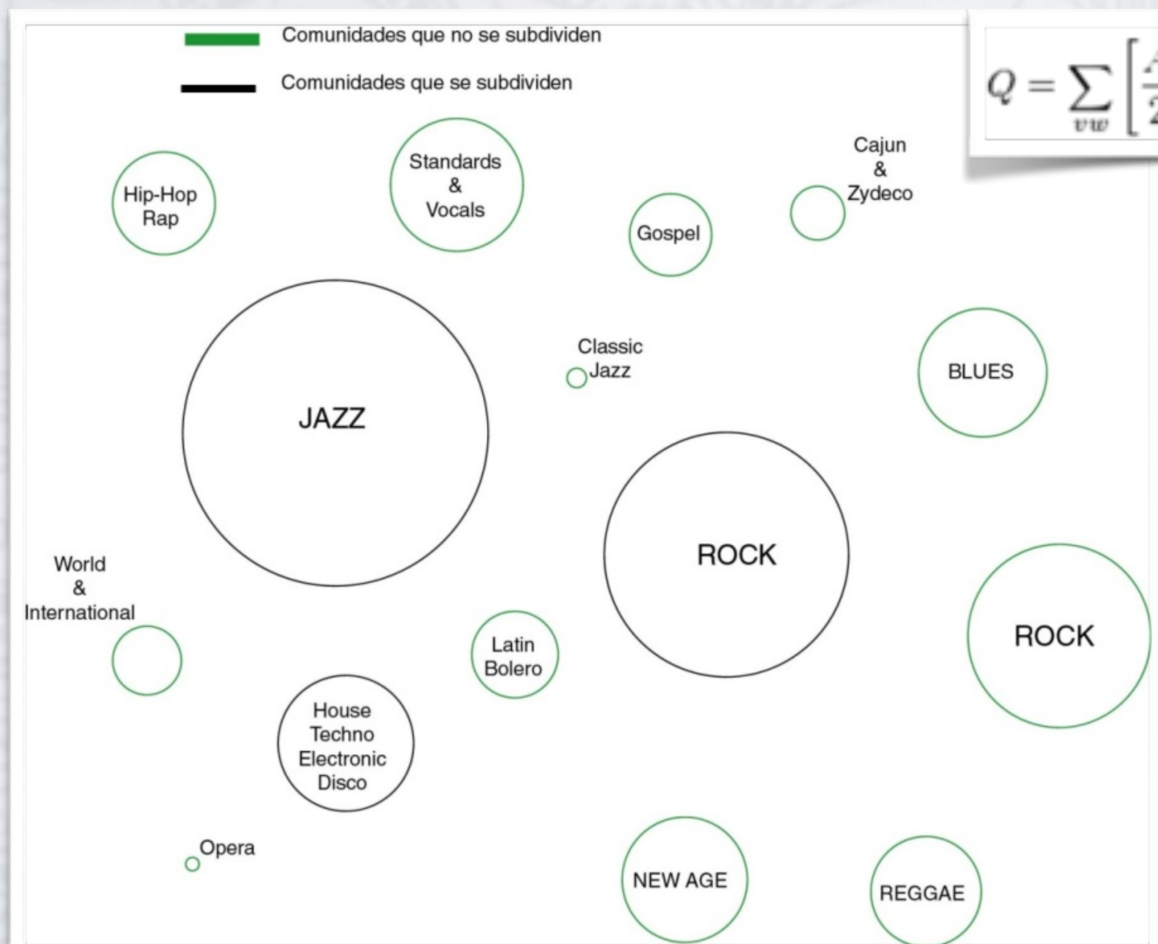


ARTIST NETWORKS

Let's see an example of how to use community detection to extract information about (artist) music networks:

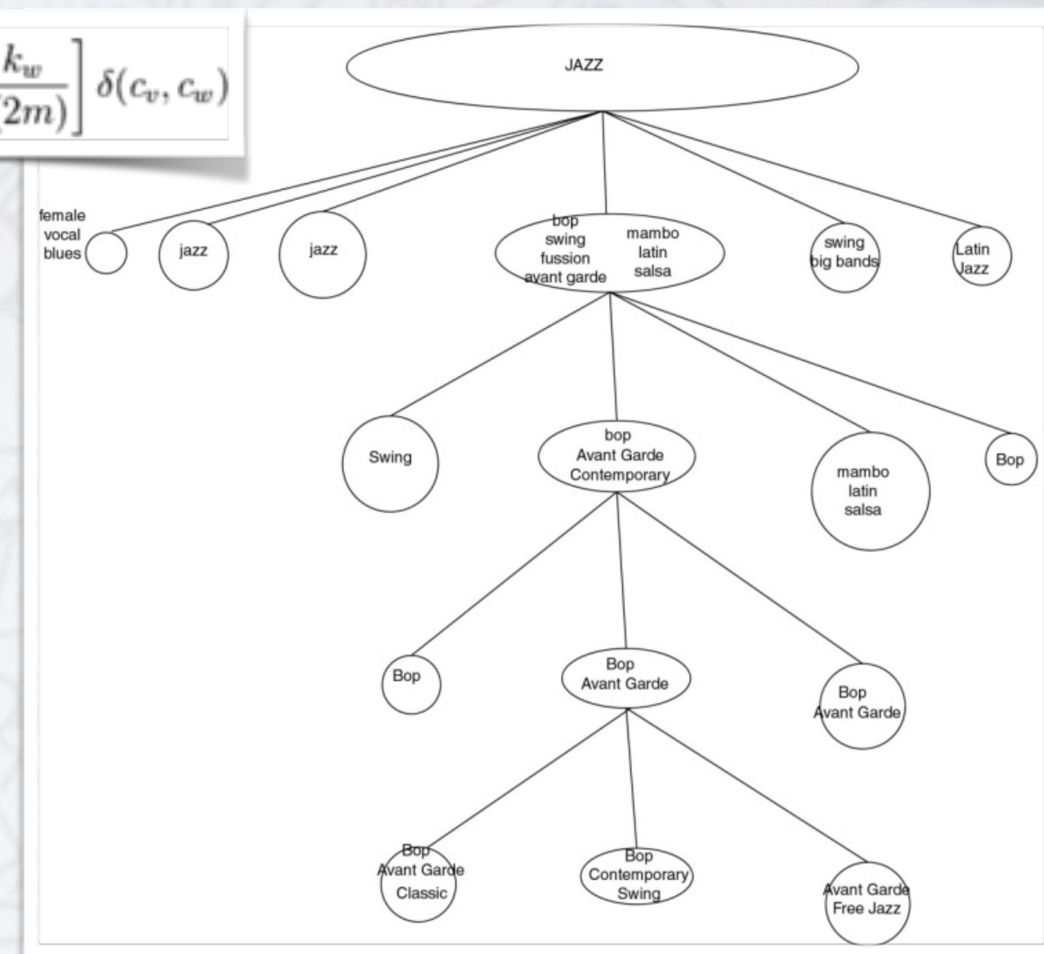
We can split networks into communities using the modularity as the reference

$$Q = \sum_{vw} \left[\frac{A_{vw}}{2m} - \frac{k_v * k_w}{(2m)(2m)} \right] \delta(c_v, c_w)$$



Division of the similarity network

Each community can, in turn, be divided into more sub-groups



Divisions are hierarchical.

ARTIST NETWORKS

It is possible to detect the hubs of each network:

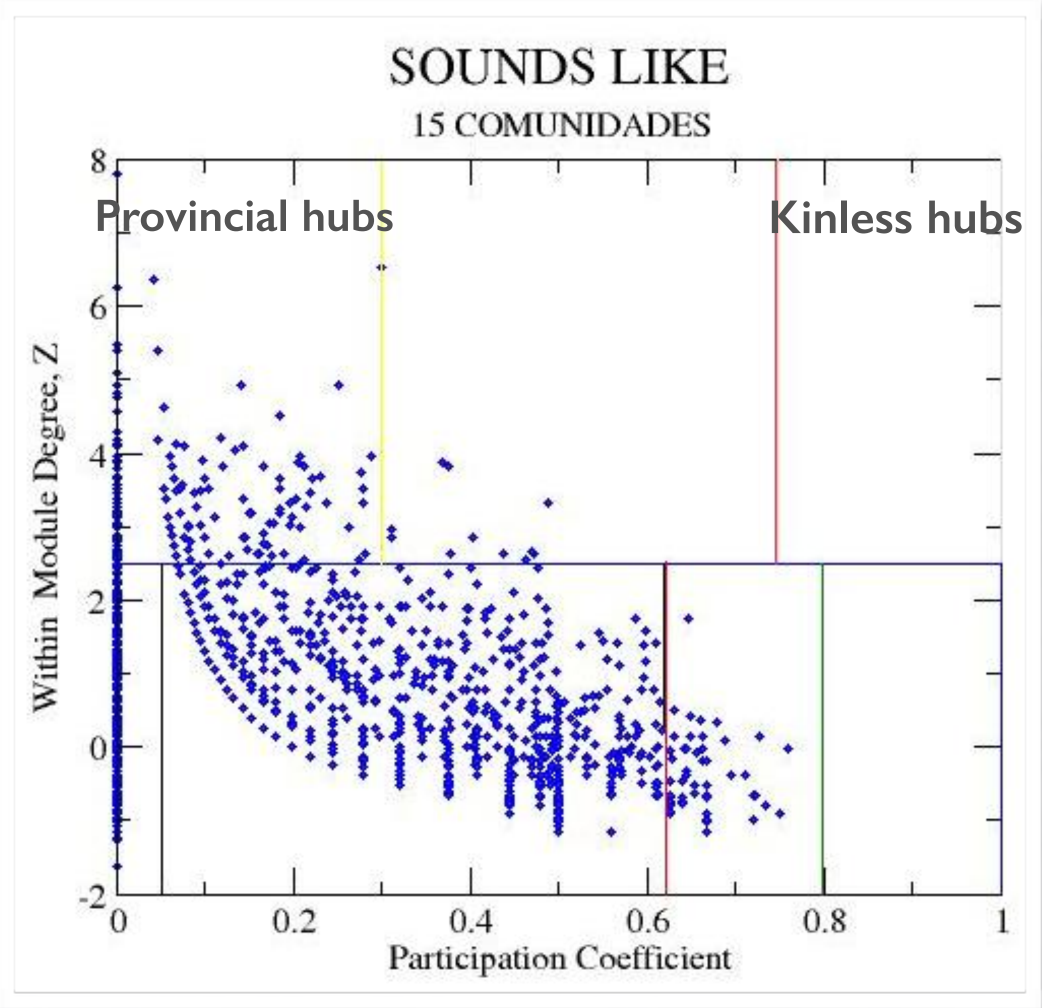
Within-module degree*:

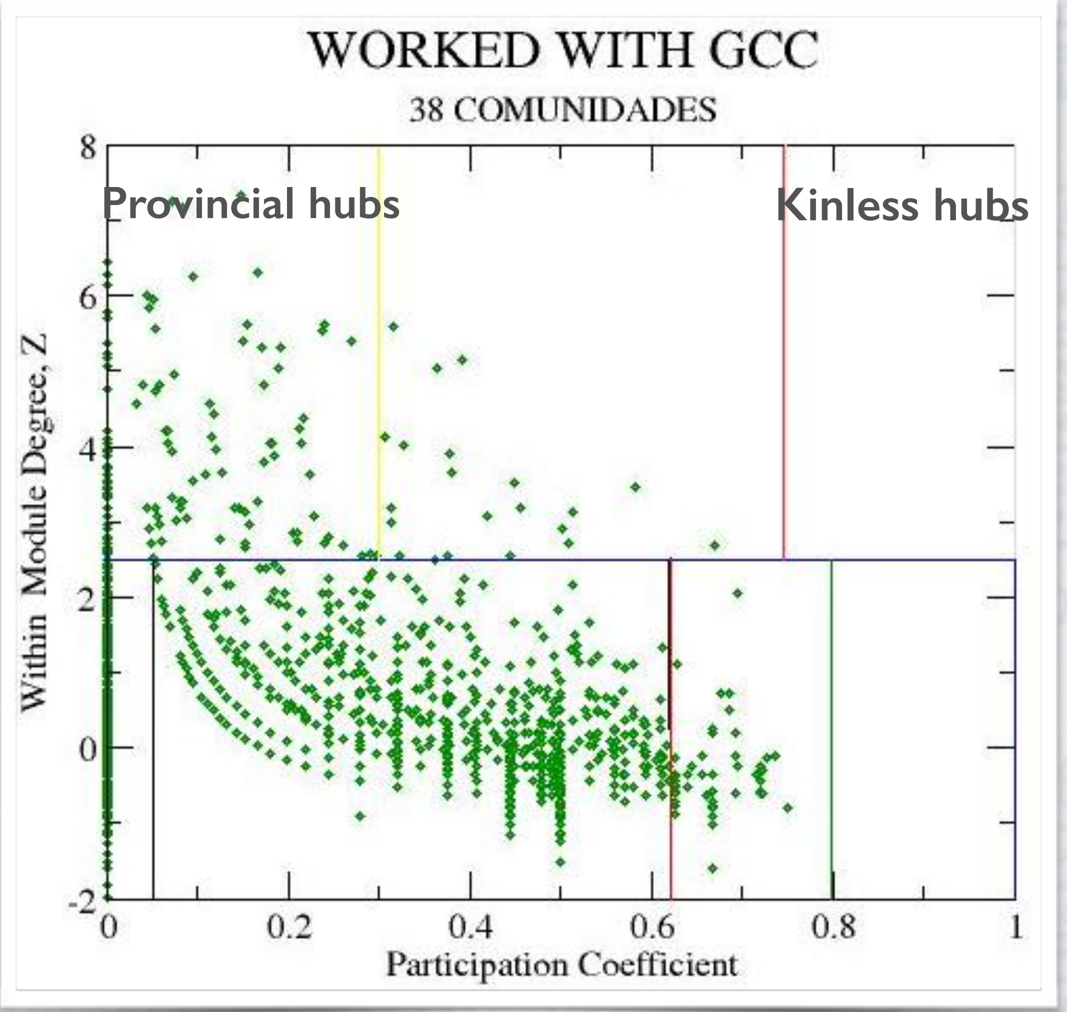
$$z_i = \frac{\kappa_i - \bar{\kappa}_{s_i}}{\sigma_{\kappa_{s_i}}}$$

Participation coefficient*:

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{\kappa_{is}}{k_i} \right)^2$$

* (Read more at R. Guimerà et al., Nature 433, 895 2005)

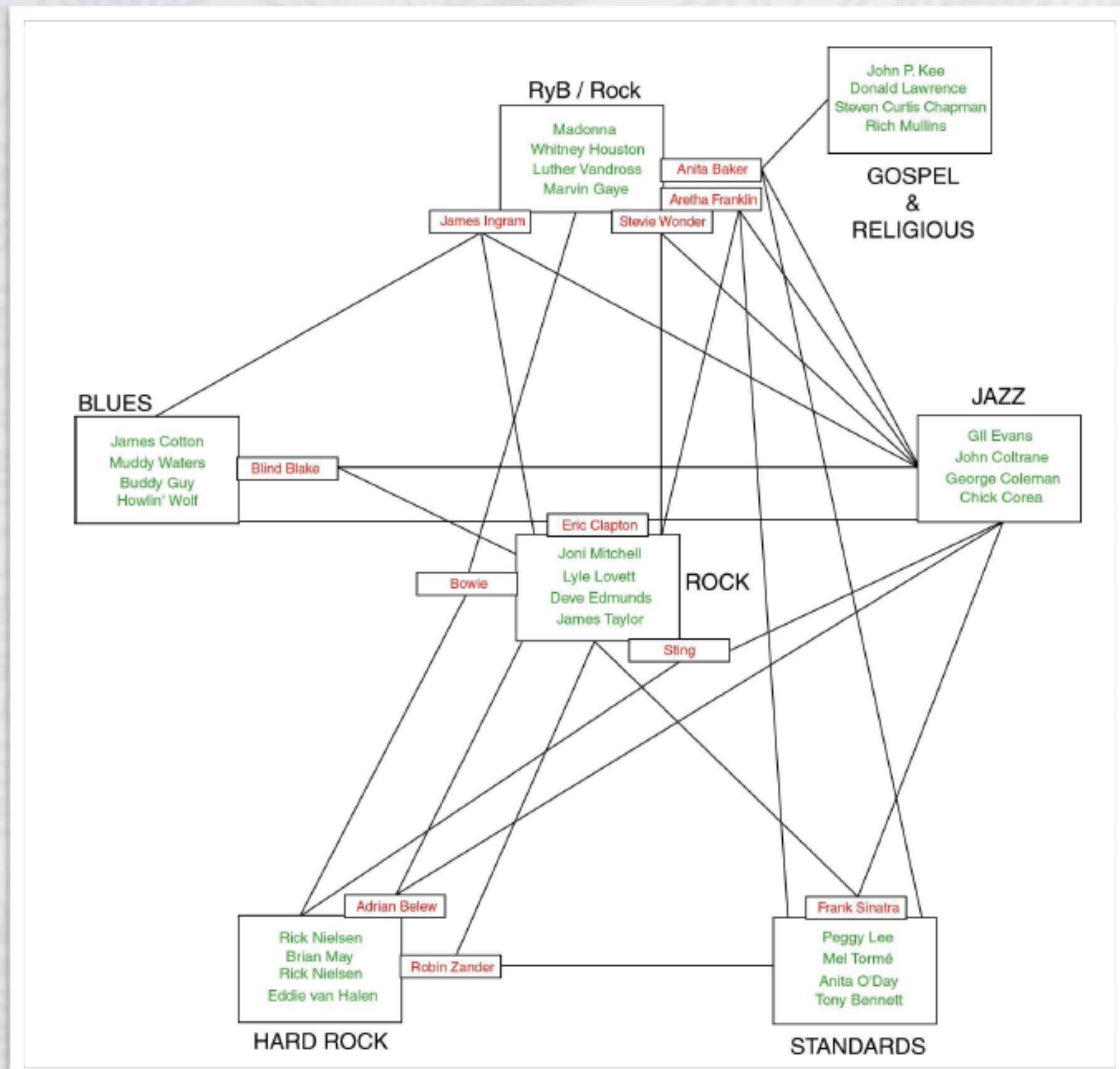




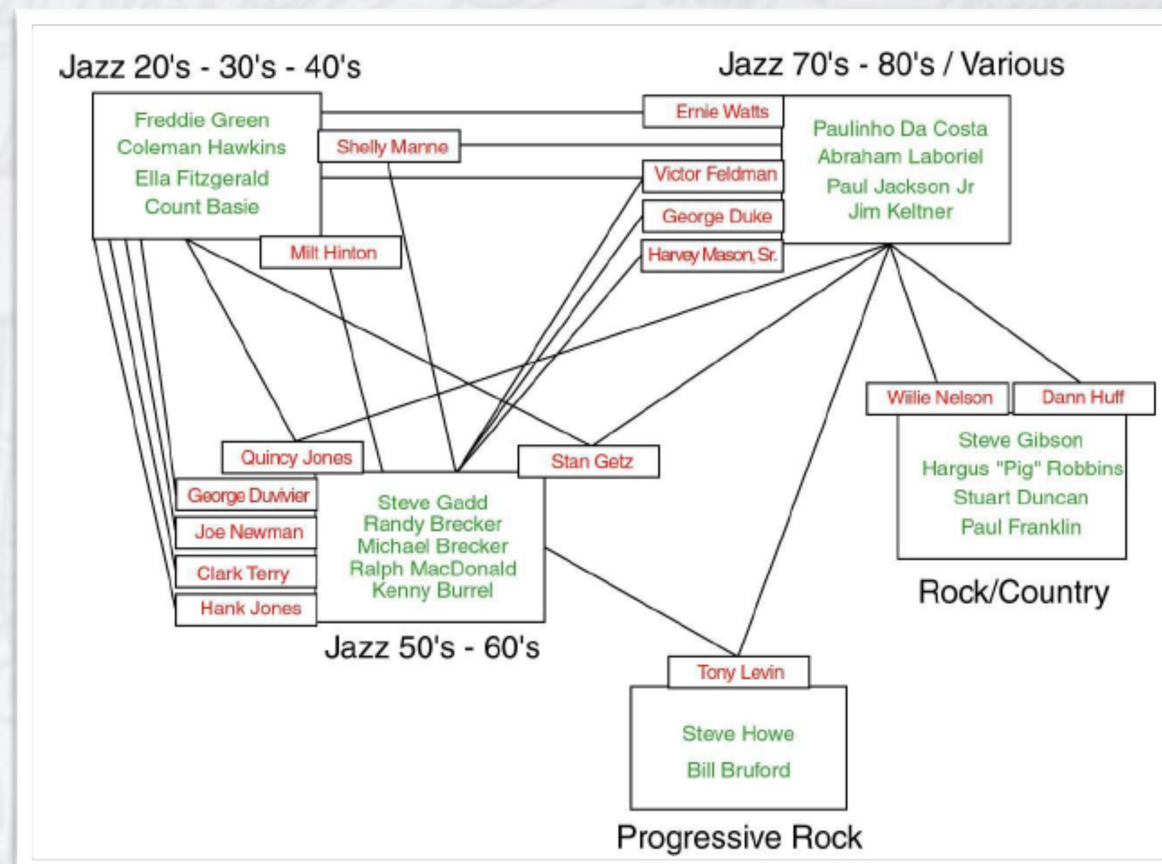
ARTIST NETWORKS

In this way, it is possible to create a cartography of the hubs existing in both networks:

Similarity network



Collaboration network

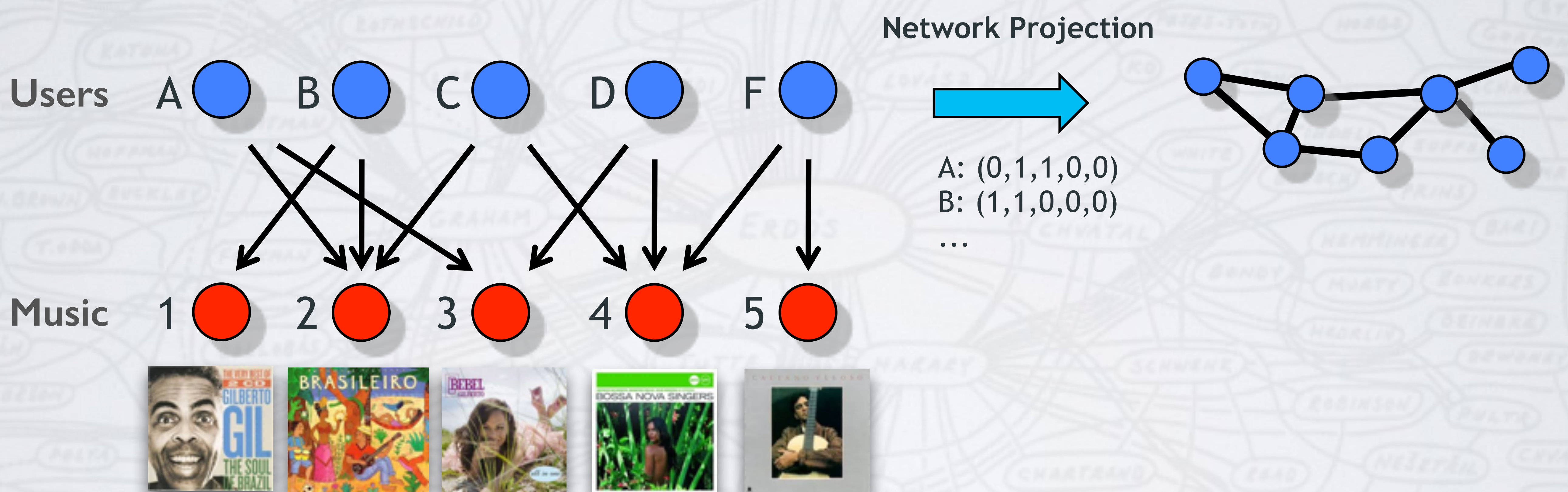


Users Networks

(V-V)

USER NETWORKS

In this kind of music networks, users that have consumed a musical product are the nodes:



This kind of networks are extremely useful for designing **recommendation** systems.

USER NETWORKS

Why should we care about user networks?
Recommendation algorithms

Internet has changed the way music is sold.

USER NETWORKS

Can we trust recommender systems?

CHAOS **16**, 013107 (2006)

Topology of music recommendation networks

Pedro Cano,[a] Oscar Celma, and Markus Koppenberger
Music Technology Group, Universitat Pompeu Fabra, Ocata 1, 08003 Barcelona, Spain
Javier M. Buldú[b]
Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Colom 11, E-08222 Terrassa, Spain

Four online music recommendation pages were analyzed:

Launch Yahoo!
Amazon
MSN Entertainment
AllMusicGuide

USER NETWORKS

How is the topology of the networks?

	n	<k>	C	d	d_r	r	γ_{in}	γ_{out}
MSN	51,616	5.5	0.54	7.7	6.4	-0.07	2.4±.01	
Amazon	23,566	13.4	0.14	4.2	3.9	-0.06	2.3±.02	2.4±.04
Yahoo!	16,302	62.8	0.38	2.7	2.3	-0.21		
AMG	29,206	8.15	0.20	6.2	4.9	0.18		

Small World

Good navigation properties

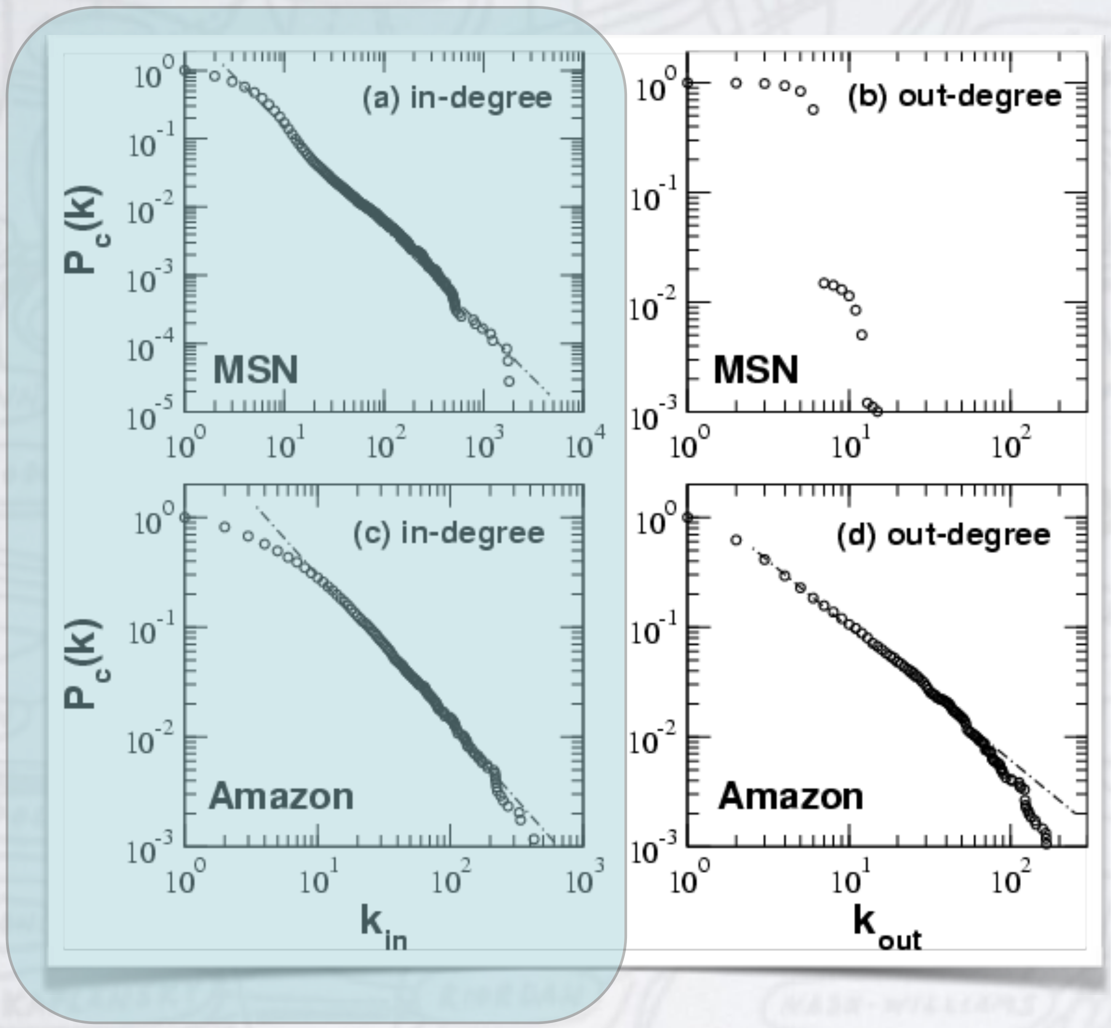
Kleinberg, Nature 406:845 (2000)

de Moura..., PRE 68, 036106 (2003)

USER NETWORKS

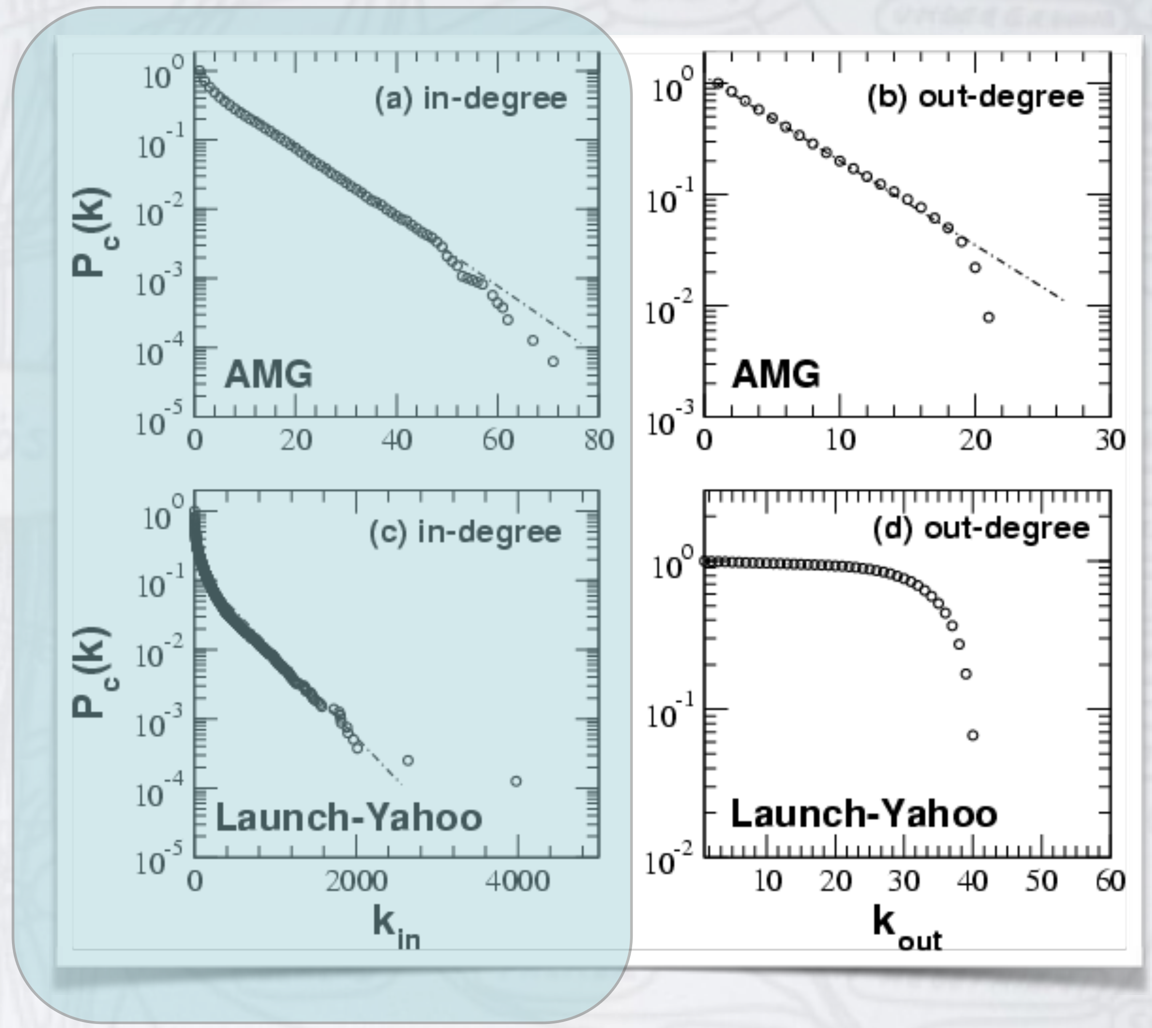
Interestingly, we find two different kind of networks:

in-degree



scale-free

in-degree



exponential

USER NETWORKS

Why are distributions so different? We asked how they were designed…

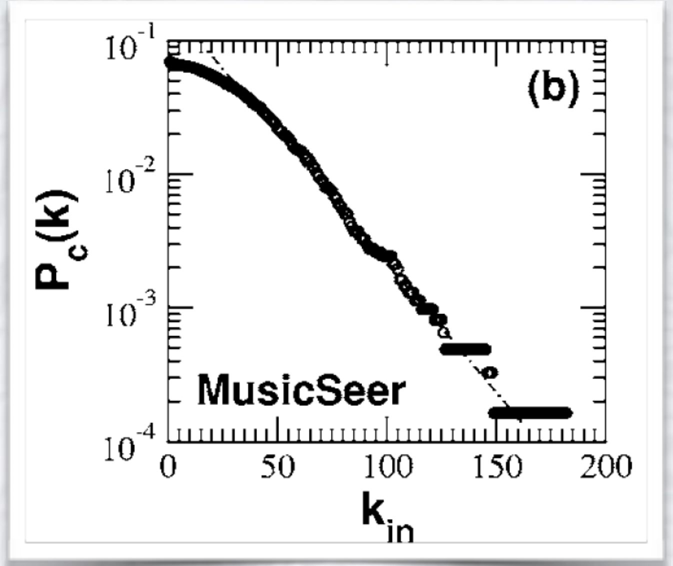
Power law { MSN, Amazon } ?? → Both MSN and Amazon networks were constructed thanks to collaborative filtering.

Exponential { All Music Guide, Launch Yahoo! } ?? → All Music Guide was created under the supervision of expert musical editors, who introduced links between artists. Launch Yahoo! Didn't give us any information about how the network was created.

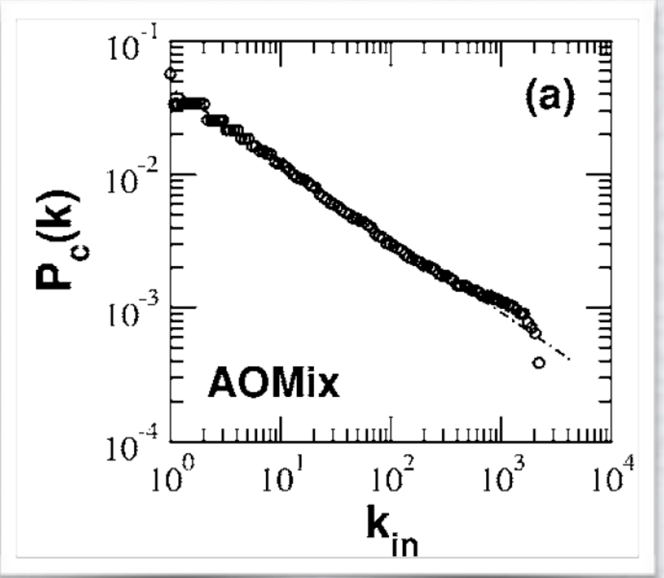
USER NETWORKS

We compared our results with another studies:

Music Seer:
- Experiment on the web. Similarity between musical artists is evaluated by means of directed surveys.
- People had to choose the similarity of an artist with a list of 10 editor selected artists.
- Only musical similarity is evaluated: We obtain an exponential decay…



Art of the Mix:
- Web page where users upload playlists of their favourite songs.
- Other factors enter the game: Trendiness, fame, musical tastes. We obtain a power law decay…



USER NETWORKS

What did we learn from this kind of study?

- In all cases, music recommendation networks are **small-world networks**, which is good news for navigation through them.
- Networks obtained by means of **collaborative filtering**, which are probably influenced by popularity or commercial trends, show **scale-free** structure.
- Networks obtained by means of musical editor supervision (or guided by), which guarantees that the **similarity criterion** is fulfilled, show **exponential** decay at their probability distribution.

Conclusions

TAKE HOME MESSAGE

- There is a **diversity** of ways for projecting (music) data into a (music) network
- Think about the **question** and then try to obtain the most adequate network
- … and, more importantly, these conclusions go **beyond music networks!**