

Community Structure in Jazz

Pablo M. Gleiser[†], Leon Danon^{*}

Departament de Física Fonamental, Universitat de Barcelona, Diagonal 647, Barcelona, Spain.

In the last years the physics community has devoted a strong effort to the study of social networks. The availability of large databases containing information on the collaborations between movie actors, scientists, etc. has allowed for many statistical properties of the networks to be characterized. These studies have revealed that some characteristics appear to be general for these kind of networks. In particular the average distance between any pair of vertices grows slowly, what is known as the small world property. It has also been observed that the degree distribution $P(k)$ follows a skewed distribution. In the particular case where $P(k)$ presents a power law tail the network is known as scale free.

An interesting point which has recently raised attention is the community structure of networks. Communities appear in networks when vertices join together in tight groups that have a few connections between them. By eliminating these connections it is possible to isolate the communities. In fact this is the main idea of the algorithm that Girvan and Newman have recently proposed¹.



Figure 1. Communities in the jazz bands network.

In this work we study and characterize the network of collaboration of early jazz musicians. Since musicians are already grouped in bands it is possible to study the collaboration network at different levels. First we construct the network from a 'coarse-grained' point of view. In this case each vertex corresponds to a band, and a link between two bands is established if they have at least one musician in common. Then we construct the network from a 'microscopic' point of view, the musicians network (MN). In this case each vertex corresponds

to a musician, and two musicians are connected if they have played in the same band. The community structure analysis reveals the presence of communities which have a strong correlation with the recording locations of the bands. It also shows the presence of a racial segregation between the musicians. In Fig 1 the different colors represent cities where a band has recorded. A clear correlation between communities and recording locations is observed.

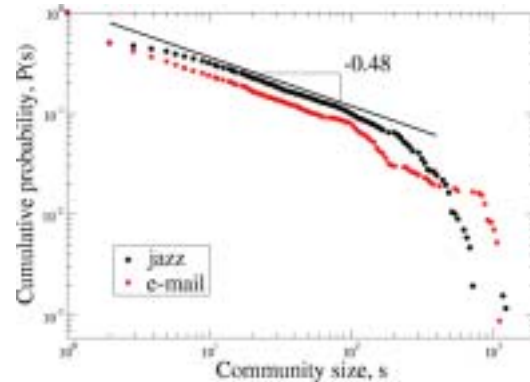


Figure 2. Community size distribution $P(s)$ as a function of community size s .

We characterize quantitatively the community structure by considering the cumulative community size distribution $P(s)$, that is, the probability of a community having a size larger or equal to s . We calculate this quantity as in², and the probability distribution is shown in figure 2. The distribution is heavily skewed, following a power law behavior $P(s) \sim s^{-\alpha}$ with $\alpha = 0.48$ between $s = 2$ and $s \simeq 100$ see Fig 2. The results for the jazz musicians network are plotted in full circles, while full triangles correspond to the e-mail network of University Rovira y Virgili². This suggests that a common principle is responsible for the organisation of both social networks.

[†] gleiser@ffn.ub.es

^{*} ldanon@ffn.ub.es

¹ M. Girvan and M. E. J. Newman, Proc. Nat. Ac. Sci. USA **99**, p. 7821 (2002).

² R. Guimerá, L. Danon, A. Díaz-Guilera, F. Giralt and A. Arenas, <http://arxiv.org/abs/cond-mat/0211498> (2002).