Complex network approach to classifying classical piano compositions

Chen Xin, Huishu Zhang, and Jiping Huang
Department of Physics and State Key Laboratory of Surface Physics, Fudan University, Shanghai 200433, China

Abstract

Complex network has been regarded as a useful tool handling systems with vague interactions. In this paper we construct complex networks for 770 classical piano compositions of Mozart, Beethoven and Chopin based on musical note pitches and lengths. We find prominent distinctions among network edges of different composers. Some stylized facts can be explained by such parameters of network structures and topologies. Further, we propose two classification methods for music styles and genres according to the discovered distinctions. These methods are easy to implement and the results are sound.

Experiments & Simulations

- We know that each composition is composed of continuous musical notes with various length of each note. The fluctuations of notes constitute different melodies, and further, different compositions. We extract 770 pieces of piano music of Mozart, Beethoven, and Chopin.
- We choose to analyze piano music because the possible notes within any composition are fixed, since there are always 88 keys on a piano, which are labeled by their letter-names and octaves. So it is convenient to correspond note pitches to specific keys with their letter-names and octaves.
- Two main descriptions of a musical note are its pitch and duration. Here, we simply split all notes of a specific composer into long-duration notes and short-duration notes. Due to that numerous lengths of actual durations exist, this simplification is reasonable and helps to construct networks that are not too sparse.
- Following the aforementioned classification of notes, any one of the notes we analyze can be distributed into one of the 88*2=176 categories. Thus, for a melody, every time the jump from note i to note j occurs, the weight of the corresponding directed edge is increased by 1. By doing so, any composition can be chronologically expressed by end-to-end edges on unified network nodes. The superposition of all such networks of a composer's compositions constitutes the musical network for his/her own.

Result

- The node representing the note with the n th lowest pitch is labeled ``2n-1'' if its duration is short, or ``2n'' if its duration is long. Under such setup, we calculate the number of connected nodes, number of edges, clustering coefficient, and node degree and edge weight distributions of the networks of the three composers. When calculating the edge weights, we standardize the data by dividing the original edge weights by the total number of note fluctuations of the relevant composer. The outdegree distributions of nodes for the three networks are shown in fig. 3, and the intensities of edge weights for the three composers are shown as heat maps in fig. 4.

- Despite some universalities discovered by early researches, there are obvious distinctions among the three networks. This enlighten us to propose new classification methods for any pieces of music based on the connections between notes, since it is more precise to describe melodies by note jumps instead of mere notes.

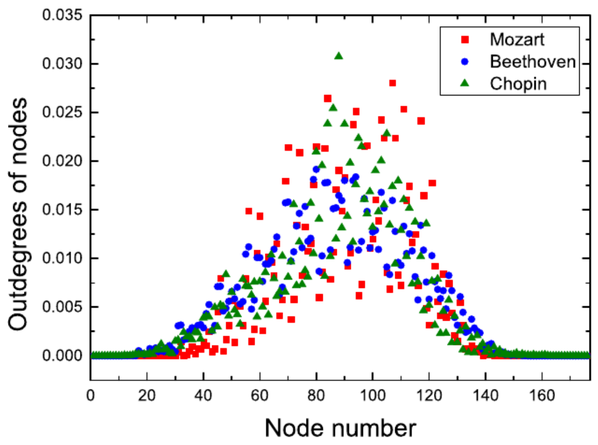

Fig. 3: Outdegrees of 176 network nodes for Mozart (red squares), Beethoven (blue dots), and Chopin (green triangles).

Classification

- The first classification method is called complete distance method, and is operated as follows: for any one of the standardized network, there are totally 176*176=30976 possible edges with arbitrary weights. Thus, for a particular composer, his/her style can be expressed by a point in a 30976-dimention space, with its coordinates being the weights of all the edges. We compare the three ``distances'' from the composition to the three composers, and the composition is classified as someone's style whose position is closest to this composition. The illustration and summary of the results of this method is shown in fig. 5.
- Another classification method is the presence-or-absent method. For any particular composer, there are specific and unique note jumps in his/her compositions acting as signatures, that do not appear in other's compositions. We are able to classify pieces of music by these specific characteristics. And if we combine the two classification methods, approximately at least 90% of the compositions can be classified ``correctly''.

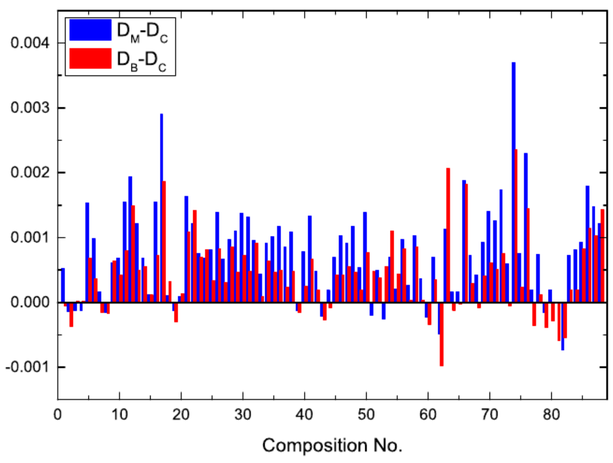

Fig. 5: The differences between the distances from Mozart and Chopin (blue columns) and from Beethoven and Chopin (red columns) for the 88 Chopin's compositions.

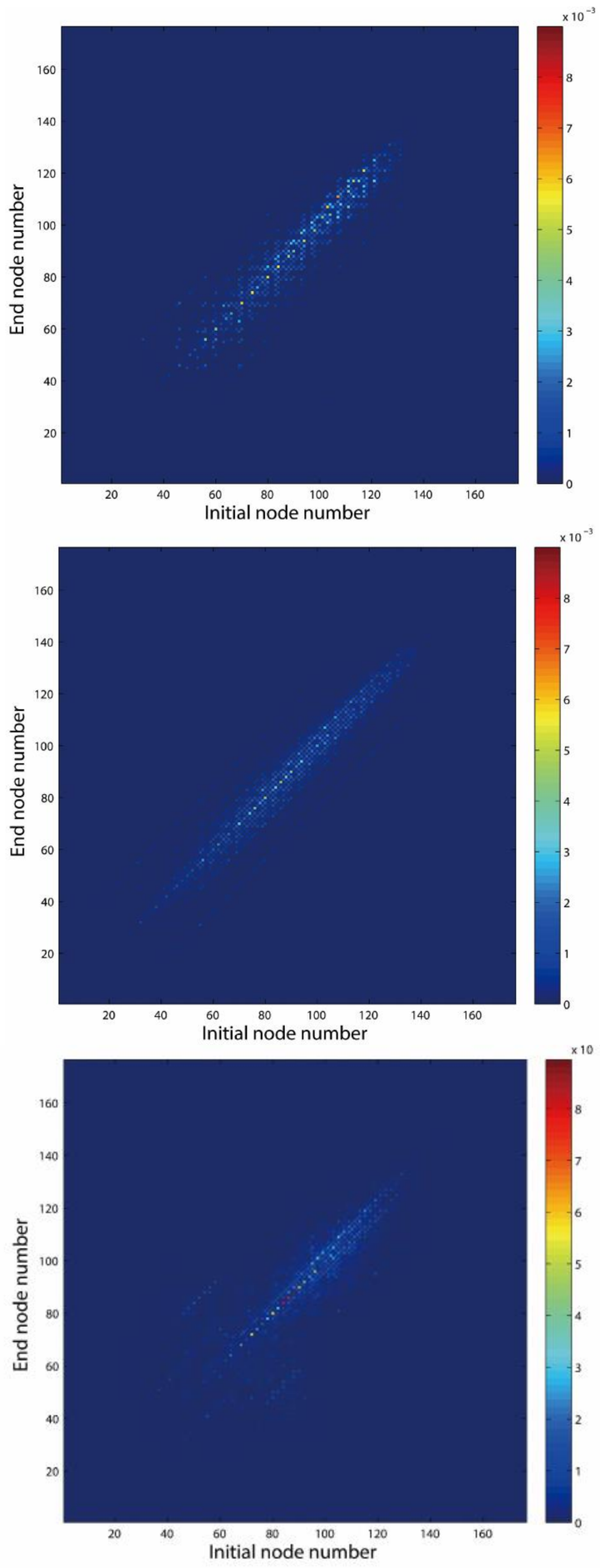

Fig. 4: Edge weights of the network constructed for Mozart, Beethoven, and Chopin as heat maps. The horizontal axis represents the initial node of the edge, and the vertical axis represents the end node of the edge. The colors in the maps indicate the occuring probability of the corresponding edges (*i.e.*, adjacent note jumps).

This work is published as **EPL**, volume 116, 18008 (2016) .

References:
[1] Liu X. F., Chi K. T. and Small M., Physica A, 389 (2010) 126.
[2] Barabasi A. L., Albert R. and Wood A., Science, 286 (1999) 509.
[3] Liu L., Wei J., Zhang H., Xin J. and Huang J., PloS ONE, 8 (2013) e58710.
[4] Kern website. Available: http:// kern.humdrum.net
[5] Opsahl T. and Panzarasa P., Soc. Networks, 31 (2009) 155.