

WMR 2015 Mid Term

April 2, 2015

Assignment

The Mid Term project consists in the collection and analysis of a sample of the Last.fm social networks. Each group¹ has to:

- i) crawl the data generated by at least 5000 users of the Last.fm service (social network and last 100 listenings per user);
- ii) Collect genre and hottness information for all the artist present at least once in the users listening using Echonest;
- iii) analyze the crawled Last.fm social network sample and compare the obtained statistics with the ones of a random graph having the same number of nodes and edges;
- iv) analyze the informations regarding users' listenings, artists' genre and hottness;
- v) discuss all the results in a short written report.

Reports, along with the code produced (if any) and the data collected, must be sent to pedre@di.unipi.it, giulio.rossetti@isti.cnr.it, lpappalardo@di.unipi.it (using as subject [WMR2015 Mid Term]).

Submission Deadline: 3 May 2015.

¹Composed **at most** of 3 students preferably having different backgrounds.

Code and libraries

The python scripts needed to perform data collection can be found on the WMA page of the didawiki². Follow the instructions provided in README.txt file to correctly crawl the required data.

To carry out the network analysis you can make use of Cytoscape³, Gephi⁴ as well as the networkx Python library⁵.

Accessory Informations

- The composition of the groups have to be communicated **as soon as possible** in the WMA15 Facebook group.
- Each group **must** register its own Last.fm⁶ and Echonest⁷ API key in order to crawl the required data⁸.
- The seed node for the data crawling **must** be different for each group.

²<http://goo.gl/dciezl>

³<http://www.cytoscape.org/>

⁴<https://gephi.org/>

⁵<http://networkx.github.io/>

⁶<http://www.lastfm.it/api>

⁷<https://developer.echonest.com/account/register>

⁸Detailed intruction can be found in the scripts README.txt file.