

Comparison of Network Analytics and Significance Analysis on Spotify Artist Feature Collaboration Network

Learning From Networks - Mid-term report

Fabio Cociancich, Luca Fantin, Alessandro Lincetto

Master Degree in Computer Engineering - University of Padova

I. EXPERIMENTS

We have determined the feasibility of performing our experiments on the CAPRI cluster [1]. The project repository can be cloned and the necessary Python packages can be installed on our own profile, making it easy to efficiently translate the code development done on our own machines into this testing environment. For any computation, we will submit our jobs to the cluster via the SLURM work scheduler. This allows us to exploit the full computational power of CAPRI while also keeping track of execution time and resource usage.

II. STATISTICAL HYPOTHESIS TESTING

Our statistical hypothesis testing procedure will comprise of two steps. Because of the specificity of our definition of random graph, instead of analytically computing the distribution of the centrality metrics of our model, the first step will compute how similar this distribution is compared to a Gaussian distribution starting from the features computed on the actual generated graphs, through a *normality test*. We will use Shapiro-Wilk test [2], since it is considered the most powerful normality test available [3]. The second step will actually determine how likely it is for the features computed on the real graph to have been drawn from the same distributions as the random graph. The available tests will be determined by the output of the first step, since a lot of procedures are based on the assumption of a Gaussian population distribution.

REFERENCES

- [1] University of Padova Strategic Research Infrastructure Grant 2017. *CAPRI: Calcolo ad Alte Prestazioni per la Ricerca e l'Innovazione*. <https://capri.dei.unipd.it/>.
- [2] S. S. Shapiro and M. B. Wilk. "An Analysis of Variance Test for Normality (Complete Samples)". In: *Biometrika* 52.3/4 (1965), pp. 591–611. ISSN: 00063444, 14643510. URL: <http://www.jstor.org/stable/2333709> (visited on 12/17/2024).
- [3] Nornadiah Mohd Razali and Bee Yap. "Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests". In: *J. Stat. Model. Analytics* 2 (Jan. 2011).