

# WMR 2013 Final Term

May 15, 2013

## Final Assignment

The Final assignment consists on the analysis of specific aspects of the interaction network collected by the WMR's students. Each group<sup>1</sup> has to solve exercises for *at least 9 points* (chosen from the proposed problems) and discuss results and methodologies in a written survey. Surveys have to be sent to pedre@di.unipi.it, giulio.rossetti@isti.cnr.it, lpappalardo@di.unipi.it in pdf format (using as subject [WMR2013 Final Term]).

Both Mid Term and Final Term projects will be discussed during an oral dissertation.

*Students who did not pass the Mid Term project must gain at least 12 points.*

**Oral Discussion:** Agree a date for the oral dissertation with the professor.

**Submission:** the project must be submitted within the exam session of July (at least 3 days before the date of the discussion).

**Page Limit:** 5 (images excluded)

---

<sup>1</sup>Preferably composed by the same students who have collaborated for the mid term.

## Exercises

### [Mandatory]

- **Network analysis [2 points]**: characterization of the network through the analysis of basilar measures. Given the multidimensional nature of the network the analysis must be made on the full graph as well as on the induced graph on each dimension (i.e. on the graph only based on the “real life” interaction).

### [To choose]

- **Tie strength [1+1 points]**
  - Analyze if (and how) the strengths associated to interactions, over the single dimension, evolve over time.
  - Exploiting the approach proposed in [1] study the multidimensional strength of the ties. Further readings [2]
- **Link Prediction [1+2 points]**
  - Partition the network in a training and a test set using the temporal annotations and apply some of the classical unsupervised link prediction approaches introduced in [3] (i.e. Common Neighbors, Adamic Adar, Jaccard, Preferential Attachment).
  - Given the multidimensionality of the network test some multidimensional variations of the former methods as proposed in [4].
- **Trust [1+1 points]**
  - Study the correlation among trust and strength values (for each dimension, both static and dynamic).
  - Study the correlation among trust and classical network measure (i.e. edge betweenness, Jaccard/Adamic Adar/Common Neighbors between the endpoint of the edge).
- **Community Discovery [3 points]** Apply on each single dimensions at least two community detection algorithms: K-Cliques, Girvan-Newman [5] and DEMON [6]. Discuss the obtained results (distribution of nodes in communities, number of communities etc. etc.).
- **Diffusion [4 points]** Implement and test the SIR, SIS or SIRS models explained in [7] (chapter 21: Epidemics). Alternatively you can apply a cascade model (as introduced in [7] chapter 19: Cascading behaviors in networks).
- **Multidimensional Network Analysis [2 points]** Present a multidimensional study of the network using measures introduced in [8].

## Code and libraries

In order to perform the analysis we suggest to use the networkx<sup>2</sup> Python library. A suggested Python oriented IDE (free for a trial version of 30 days) is Pycharm<sup>3</sup>.

Implementations of multidimensional graph measures, of the DEMON algorithm and of classic link prediction approaches (as well as all the anonymized network datasets<sup>4</sup>) can be downloaded from:

<http://www.giuliorossetti.net/about/ongoing-works/material/>.

## References

- [1] Luca Pappalardo, Giulio Rossetti, Dino Pedreschi: “How Well Do We Know Each Other?” Detecting Tie Strength in Multidimensional Social Networks. *ASONAM 2012*:1040-1045
- [2] Mark S. Granovetter: The Strength of Weak Ties. *American Journal of Sociology*, Volume 78, Issue 6 (May, 1973), 1360-1380
- [3] David Liben-Nowell, Jon M. Kleinberg: The link prediction problem for social networks. *CIKM 2003*: 556-559
- [4] Giulio Rossetti, Michele Berlingerio, Fosca Giannotti: Scalable Link Prediction on Multidimensional Networks. *ICDM Workshops 2011*:979-986
- [5] Girvan M. and Newman M. E. J.: Community structure in social and biological networks, *Proc. Natl. Acad. Sci. USA* 99, 78217826 (2002)
- [6] Michele Coscia, Giulio Rossetti, Fosca Giannotti, Dino Pedreschi: DEMON: a local-first discovery method for overlapping communities. *KDD 2012*:615-623
- [7] David A. Easley, Jon M. Kleinberg: *Networks, Crowds, and Markets - Reasoning About a Highly Connected World*. Cambridge University Press 2010
- [8] Michele Berlingerio, Michele Coscia, Fosca Giannotti, Anna Monreale, Dino Pedreschi: Foundations of Multidimensional Network Analysis. *ASONAM 2011*:485-489

---

<sup>2</sup>website: <http://networkx.github.io/>

<sup>3</sup>website: <http://www.jetbrains.com/pycharm/>. If you need an extension of the license send an email to [giulio.rossetti@isti.cnr.it](mailto:giulio.rossetti@isti.cnr.it), subject [WMR2013 Pycharm].

<sup>4</sup>Several network variants, with the same id mapping, are available. You can use them as they are or build your own derivative networks: in the latter case specify the process used to obtain them.