Percolation: anthropological physics for social harnessing

Renato Fabbri *

*University of São Paulo

Draft

Percolation is a python package for harnessing the social networks of the user. It is based on the pilars of: social percolation, creation of audiovisual artifacts, resource recommendation, typologies and the generation of reports. The fields of complex networks and linked data give scientific support for the exploitation of the integrated (virtual) social space.

complex networks | software toolbox | anthropological physics

Abbreviations: RDF, resource description framework; BoW, bag of words; PyPI, python package index

The Percolation package consists of routines for social harnessing and can be directly hacked mainly through four packages: Gmane, Participation, Social, MASS. The text ahead consists of the core features of each of these five python packages.

Packages.

Gmane

The Gmane package is dedicated to exploring the Gmane database of email lists. Core features are:

- Download email messages from Gmane database.
- Load messages and make basic data structures.
- Make interaction networks.
- Take measures from interaction network.
- Make PCA from measures, with observance of each component formation.
- Observe Erdös sectors in the networks (see appendix).
- Histograms and circular statistics for time activity.
- Histograms for user activity.
- Facilities for network evolution of fixed window size, such as plotting timeline of measures and making video of the evolving network through Versinus [1].
- Facilities to output images and latex tables.

Participation

The participation package is dedicated to exploring social participation data. Core features are:

- Access to a starting set of participatory data (see appendix).
- Data integration through linked data principles (RDF data, OWL ontologies).
- Access to routines of participatory data translation from PostgreSQL, MySQL and MongoDB to RDF (triplification).
- Access to routines for delivering participatory OWL ontologies.
- Routines to raise ontologies from data, return OWL code and images. (This is a novel and pertinent method of making ontologies to fit data, a somewhat *data-driven ontology elaboration method*).
- Analysis of participatory data through complex networks (interaction and friendship networks) and text mining (BoW, Kolmorov-Smirnov histogram comparison and entropy).
- Resource recommendation, with explicit routines and potential uses.
- Recommendation of resources by both absolute criteria (ranking of users, words, etc.) and by criteria relative to a reference resource (e.g. recommend users given a reference post).
- Bootstrapping the basic structure of ontologies to HTML.
- Simplest web server to give HTTP access to data and methods.

Social

The social package delivers routines to access usual social network platforms data, such as Facebook, Twitter, LinkedIn and IRC. Core features are:

- Screen scrapping of Facebook data.
- Twitter search and streaming through multiple APP keys.
- Parsing IRC logs.
- Access data from LinkedIn (ToDo).

MASS

MASS is music and audio in sample sequences. Core features are:

- Synthesis routines for notes and noises.
- Calculations in 64 bit floating point.
- Parameters updated each PCM sample.
- Exact handle of duration and frequency parameters.
- ADSR envelopes.
- Table lookup.
- Four basic waveforms (sine, saw, square and triangle).
- Tremolo and vibrato implementations.
- Musical and DSP methods implemented according to [2].
- Predefined synthesis methods for other packages (Gmane, Social, Participation, Percolation).

Percolation

Percolation unites Gmane, Participation, Social and MASS packages to enable anthropological physics experiments and social harnessing. Core features are:

- Ease percolation in social systems by processes such as collection and diffusion of information.
- Enable knowledge about the networked self.
- Make abstract animations from social data.
- Make music from social data to fit animation as soundtrack.
- Verification of expected stability and differentiation on the social structures.
- Directions for agents and networks typologies, extending features from Gmane package.
- Integration of resources through RDF data and OWL ontologies.
- Routines for representing data from social networks (tweets, Facebook data, IRC logs, LinkedIn, Gmane lists) as RDF.
- Cross provenance resource recommendation, extending facilities from the Participation package.
- Generation of activity reports.

Real Data.

Current outcomes.

1. Fabbri, Renato. "Versinus: a visualization method for graphs in evolution." arXiv preprint arXiv:1412.7311 (2014).

Discussion

Materials and Methods

Appendix: Erdös sectors

Definition 1. *The Erdös Sectors S of the network N are defined as the three sectors provenient from the comperrisson of N to an Erdös-Renyi network with the same number of nodes and adges.*

$$equations \qquad [1]$$

Appendix: Data and ontologies on the participation package

Data: Participa.br, AA, Cidade Democrtica

Routines for data triplification: Participa.br, AA, Cidade Democrtica

Ontologies: OPa, OPS, OBS, VBS, OCD, Ontologiaa (old OPA?).

Routines for raising ontologies: OPa, OPS, OBS, VBS, OCD, Ontologiaa

ACKNOWLEDGMENTS. Author is grateful to CNPq (process 140860/2013-4, project 870336/1997-5), UNDP (contract 2013/00056, project BRA/12/018), SNAS/SGPR, and the Postgraduate Committee of the IFSC/USP.

2. Fabbri, Renato, et al. "Psychophysics of musical elements in the discrete-time representation of sound." arXiv preprint arXiv:1412.6853 (2014).