## Introduction

- This repository contains code related to the paper “Multi-objective community detection for bipartite graphs” (see untouched/BipartiteCommunityPremium(2).pdf)

- A Research IT work has been conducted to test the code, clean and rewrite it in a way to be shared with the community.

## File/directory description and usage

- **‘untouched’ directory** contains the conference paper versions, other literature on the topic, and all original (unaltered) notebooks which were used to generate the paper experiments (non-shuffled versions, see below). Please note that they are not working as they use previous versions of pymoo, condor, and igraph packages **(igraph 0.9.1, pymoo 0.4.2.2, condor 1.1)**. They mainly serve documentation and archival purposes. For running versions of the 3 notebooks (data generation, contestants, and multicriteria approach), please see below.

- Notebooks 1. **Data Generation.ipynb** and **Data Generation (shuffled).ipynb**, 2. **Contestant.ipynb** and **Contestant (shuffled).ipynb**, 3. **Multicriterion approach.ipynb** are the working notebooks adapted from the aforementioned ones (the shuffled tag denotes notebooks which shuffle graph vertex indices during data generation, they are not used in the paper results). The code remains the same, the modifications include updating them to work with the current package versions **(igraph 0.9.9, pymoo 0.5.0, condor 2.0)**. Please note that condor package has been modified to be able to run BRIM algorithm properly, the updated condor code is include within this repository (see below).

- Files **multicriterion\_3d.py** and **multicriterion\_2d.py** are similar to **Multicriterion approach.ipynb** but they allow running the legacy code in 3d or 2d modes.

- **‘condor’ directory** contains the updated condor package to make the code run. If one is interested in using the old version fo condor, one needs to update the code for the older condor interface and import/use condor\_1.1.py file (included) instead.

- **‘moo’ directory is the actual code package that replaces the legacy code**. It contains the updated code for data generation (data\_generation.py), contestant algorithms (contestant.py), multicriteria approach (multicriteria.py) and a utility module (utils.py) which provides functionality for writing/reading graphs into various file formats, writing graphs and reading graphs from the format used in the legacy code, etc. The usage of the new code (package moo) is explained by example in the notebooks (see below)

- Notebooks **01\_Data Generation.ipynb 02\_Contestants.ipynb 03\_Multicriteria Approach.ipynb paper\_figures.ipynb** show many examples of how to use the package. Their usage is recommended.

## Notes

- The code can be used as it is to reproduce the paper results (see notebooks 02\_Contestant.ipynb, 03\_Multicritera Approach.ipynb, and paper\_figures.ipynb). It can also be used with 3rd party data (graphs and groud truth data separated) by following the example in notebook 01\_Data Generation.ipynb.

- Legacy code does not guarantee reproducibility. Please use the new code instead (allows passing a random number generator seed).

- Pending work consists in adapting the code to return communities as results, in the case of missing ground truth.