Option SDD - Graph Analytics & Applications
École Centrale Casablanca, January 2026

Instructor: Fragkiskos Malliaros
Email: fragkiskos.malliaros@centralesupelec.fr **Date:** January 5, 2026

Project Description

One of the main goals of the course is to provide the fundamental tools and algorithms that one can use in order to analyze network data from various domains and applications. To this direction, the main component of the course is the final project, where you will have the opportunity to develop quantitative and qualitative skills on network analysis methods and algorithms, and to obtain practical experience working with software and tools for large-scale graph mining and network analysis. Furthermore, those of you that are interested in research, the project will give a taste of graph mining and network science research.

In particular, there are several types of projects that one case choose to work on (some of them are related to each other; also think of interesting combinations among them):

- Experimental evaluation of algorithms and models on an interesting graph dataset. For example, select a graph mining task that you are interested in, pick 2-3 different algorithms for this problem and do an empirical experimental evaluation (e.g., take three different community detection algorithms, find a good dataset or construct your own dataset and compare the performance of the algorithms).
- A new algorithm for a graph mining task and evaluation on real/artificial data (e.g., a new community detection algorithms).
- Formulation of a problem using graph-based modeling, propose an algorithm that solves the problem (or use/adapt an existing one) and experimental evaluation (e.g., how to deal with the keyword extraction problem in text mining using graphs).
- A theoretical project that considers an algorithm/model and derives rigorous results about it.
- Efficient implementation of a graph mining algorithm and experimental evaluation.
- A problem that is related to your own research. If you are already working on a research project that network analysis might be applicable to, then finding out how to apply network analysis techniques to it will often make a very good project topic.

Note that, the above list is not exhaustive and you can come up with other interesting types of projects. Ideally, projects will be combination of an interesting application, experimentation on real/artificial data and some theoretical analysis.

Moreover, our advice is to pick a project that you can get excited and passionate about. Do feel free to propose ambitious things that you are excited about. We will always be available to discuss with you on potential projects ideas.

The projects will be evaluated based on:

- ***Significance:*** Is the problem "real" and "interesting", or just a "toy" problem? How original, important and well defined are the questions posed? How novel is the approach? Is the work likely to be useful and/or have impact?
- ***Technical Quality:*** Is the approach and the methods appropriate and well described? Are sufficient details provided? Is the technical material correct? Are the proposed algorithms or applications creative and interesting? Are the methods and algorithms reproducible? Is the interpretation (discussion and conclusions) well balanced and supported by the data?
- ***Organization:*** Is the final project report well organized (e.g., following the structure of a paper published in a top data mining conference)? Is the write-up clear and easy to read? Are the results presented in the most appropriate manner? Are figures and tables used appropriately?

Students should work in teams of 2-3 people. Make sure to mention the names of all of your team members when submitting the project deliverables.

Deliverable: Project Report

The report should represent all the completed work, having the following structure:

- **Abstract:** Short (200-250 words) abstract of your project.
- **Introduction/Motivation:** What is the project about? What is the problem you are trying to solve? What are the questions you want to answer? Why the problem is important? What are a few potential applications?
- **Problem Definition:** Introduce notation, provide formal definitions as needed, define any constraints or restrictions, define what you try to optimize (e.g., maximize or minimize an optimization function, or an accuracy/error function). Describe the problem in a formal way. Describe the hardness of the problem in a formal way.
- **Related Work:** Position the problem among the body of existing research. How does your project relate to previous research? How is your project replicating/different/complimentary to previous research? References to papers you cite should be explicit followed by a comment that describes how it is relevant.
- **Methodology:** How did you address the problem? What are the steps you had to take? Describe the data collection process. Provide any mathematical background necessary for the methods. Describe any algorithms or variations of the methods. Describe limitations or difficulties with your approach. Formally describe any important algorithms used from the literature. Try to be as specific as possible.
- **Evaluation:** How did you evaluate your work? What experiments did you run? What datasets did you use? Describe clearly your findings.
- **Conclusions:** What are the conclusions of your work? Are there any highlights? What are some ideas for future work?
- **References:** In the final report, you should provide full list of references.

Note: Keep in mind however, that if there is a good reason why your project doesn't match the above description, we will take that into consideration when grading your report. For example, we recognize that purely theoretical or pure data analysis projects may not fit the rubric above perfectly, and that depending on your project you may want swap the ordering of certain sections.

The evaluation of the final report will be based on the following guidelines:

Outline	Weight
Introduction/Motivation/Problem Definition	15%
Related Work	10%
Model/Methodology/Algorithm	30%
Evaluation/Results	35%
Style and writing	10%

Formatting and Page Limits: The suggested length of the proposal is **8-9 pages** and should be in PDF format. **Please include all team member names (as authors) in both the report and gradescope platform**.

All reports should be formatted according to the ICLR template (available in Overleaf): https://www.overleaf.com/latex/templates/template-for-iclr-2021-conference-submission/mmpfhsxmqdkp

How to submit? Please, submit the PDF file in **gradescope** as a **group submission** (Entry Code: **YBR5XZ**): https://www.gradescope.com/courses/1209548

Also, send the code and data (in case you have not used publicly available datasets) by email to netsci.class.centralesupelec@gmail.com (or preferably, the link to a repository where you have uploaded the code/data).

Resources

Please, check the **Resources** section of the website for datasets, software and other material that may help you in the project. We are also very happy to discuss with you in any aspect of the project :-)

Acknowledgments: Ideas for the project reports were borrowed from J. Leskovec (Stanford University) and M. Papagelis (York University).