ECS637U/ECS757P - Digital Media and Social Networks (DMSN) Group Project: Assignments 4 and 7

This DMSN group-based coursework consists in conducting network analyses applied to a real-world topic using graphs generated from a dataset. You will present and discuss the project orally and in a written paper report. The coursework is comprised of the two following assignments:

- **Assignment 4: Project presentation [20% of the grade]**
- **Assignment 7: Project paper (report) [50% of the grade]**

The **deadlines** can be found in the table below:

DMSN GROUP PROJECT	Assignment 4	Assignment 7
Deliverable	Project presentation	Project paper (report)
Grade weight	20%	50%
Submission format/delivery	**PDF slides** (QMplus) and **oral presentation** during online lecture/lab session	**PDF paper** (QMplus)
Deadlines	**Slides**: Fri. 5 March (2pm) [end of Week 6] **Oral presentations**: Tue. 9 March (4-6pm) or Weds 10 March (9-11am) [Week 7]	**Paper**: Fri. 16 April (2pm) [Week 12]

Group project

This coursework should be done in **groups of 4 students in the same academic level**. This means that groups can either be composed of **UG students only**, or **PG students only**.

The group set up process has been introduced during the lab session in Week 2. You can start a group, join an existing group, or wait to be allocated to a group by the teaching staff if you prefer.

If you would like to be in a specific group, you have until Wednesday 17 February at 2pm to register your group on QMplus; this must be done using the Project Group Registration tool that can be found in the Group Project section on QMplus. **After this date, students without groups will be assigned to one.** Note that there are two separate Project Group Registration tools, one for UGs and one for PGs, however you should only be able to access the one which is appropriate for your level.

Preliminary work: Forging your skill set for the project

Ensure that you stay up-to-date with the course topics and that you read regularly the required chapters from the Easley & Kleinberg book listed at the end of each week's lectures. Also read the recommended papers discussed during the lectures. By combining theoretical knowledge and understanding of real-world applications, you will be able to propose and assess relevant network analyses for your project.

Once you have a group, you can start to meet with your team members to discuss the theoretical concepts and applications reviewed during the course. When you read academic papers, it is useful to take a critical analysis approach. Questions you may want to consider are:

- What is the technical content of the paper? How are the network analyses conducted?
- What are the strengths and weaknesses?
- What other kinds of datasets could be analysed using these analyses?

Project support

Tools

A [Discord](#) server has been set up for collaborative work and support for your DMSN projects. The Discord server was introduced during the lab session in Week 2. If you have not joined the DMSN Discord server yet, you can use the following invitation link: https://discord.gg/efcjN6mn

If the link does not work, please send an email to Teo Danneman (t.dannemann@qmul.ac.uk) and the module organiser to get a new invitation link, and add your details to the following spreadsheet:

- [List of students not yet registered on the DMSN Discord server](#)

Once registered, you can use the DMSN Discord server to find other students to form or join a group. The following spreadsheets can be used to find other students to form groups and then set up meetings based on their availability:

- [DMSN students in GMT timezone and their availability](#)
- [DMSN students in non GMT timezone and their availability](#)

Each group will get access to a Discord channel to be able to work collaboratively. The Discord server will also be used for project support with the demonstrators and lecturers. We strongly encourage you to use Discord for all collaborative work with your team members (this can also be useful if your group encounters difficulties, as the teaching staff will be able to see the exchanges on the shared channel).

Mentors

Each group will have a **mentor** who will provide guidance through online meetings on Discord. Mentors will be assigned to groups in Week 4.

You can either choose **one of the project topics listed on QMPlus** (see the **Project Topic list in the QMPlus Group Project section**) **or propose your own topic**. If you choose to propose a topic, please check with your mentor that the topic is suitable.

Assignment 4: Project presentation [20%]

The **project presentation** will consist in a short oral pitch during which you should articulate your topic in a quick, insightful, and clear manner (lightning talk format). Your presentation should be supported by a **maximum of 5 or 6 slides** including the following content:

- Title side (study title, names, module and coursework information, year)
- Introduction of the topic (describe the context, motivations, and challenges)
- Presentation of the dataset (how it was collected, what it represents, how networks are formed, their size, etc.)
- Basic network statistics (present basic network statistics relating to your network(s))
- Proposed network analyses for the Assignment 7 Project paper (describe which analyses you have planned, why and how you will conduct them, in light of academic works); please note that you do not need to have conducted the analyses for this deliverable.

The slides should be submitted in the PDF format on QMplus by Fri. 5 March (2pm). Each team member must submit the slides on QMplus.

The talk should last 4 minutes. You should rehearse your presentation before your oral session to ensure that you keep time. All members of the group should speak. **Oral presentations will take place during the online lecture and lab sessions in Week 7 using Zoom**. The slot of your oral session in Week 7 will be communicated in due course. Presentations will be shared with the class so that you can practice presenting in front of an audience and learn from each other.

Examples of basic network statistics for Assignment 4 include degree distribution, network clustering coefficient, modularity, centrality measures, but other metrics relevant to your topic are also possible. Illustrate your analyses with figures (e.g. Gephi, graphs, code snippets, etc.).

You can use tools such as Gephi which doesn't require prior coding skills, or the Python packages for network analysis such as NetworkX for those of you who have Python skills. These tools will be introduced during lab sessions.

All *attending* members of the project team will receive the same grade. Members not present for the oral presentation without a valid EC will receive 0.

The marking criteria and estimated number of study hours for this assignment can be found on the following page.

Marking criteria for the Assignment 4 Project presentation (slides and oral presentation)

Introduction	10%
Dataset presentation	20%
Basic network statistics	30%
Proposed network analyses	30%
Clarity	10%

Estimated number of study hours for this coursework deliverable including lectures and tutorials: **30 hours**

Assignment 7: Project paper [50%]

This assignment consists in conducting several network analyses on graph(s) generated from the dataset(s) for the topic you introduced in Assignment 4. You can choose which network analyses to conduct, however these should be relevant for the topic and the research problem to be addressed. You can follow some of the proposed tasks listed in the topic description on QMplus and/or propose other approaches based on the theory or the literature (e.g. academic papers discussed during the lectures). Please discuss the choice of tasks with your mentor. In the topic descriptions on QMplus, a task's difficulty is reported (Easy/Medium/Hard) along with suggestions of suitable academic levels for the task (UG/PG).

Examples of network analysis tasks include real-world and random network comparison, robustness analysis, PageRank analysis, and cascade analysis (see the lectures/labs and QMplus Project resource page for information on these).

You should write as a group a **short paper of 6 pages excluding references** using the IEEE conference format (https://www.ieee.org/conferences/publishing/templates.html). The paper should follow the **following structure**:

Title (*an original title specific to your work*)

Authors' names

Abstract (*a short summary of the paper*)

1. **Introduction**

   Describe the context of the problem, define the research problem, discuss the motivations and challenges. Also briefly introduce the structure of the remainder of the paper.

2. **Related work**

   A literature review related to your topic and the analyses that you have conducted

3. **Dataset and network presentation**

   Describe the data and network(s) that you have analysed and how they are structured

4. **Network analysis methodology**

   Describe your network analysis methodology based on academic literature and justify why the approaches you proposed are appropriate to address the research problem described in the introduction

5. **Results and discussion**

   Present your results and discuss them in a critical way. Provide reflections on the implications of the results for the problem addressed. You should use figures to report your results where appropriate (graphs, tables, screenshots, etc.).

6. **Conclusions and perspectives**

   Summarise the work and provide perspectives on how it could be extended

7. **References**

   Include a list of references following the IEEE referencing guidelines

Writing style

The paper should be written following an [academic writing](#) style (formal prose articulating a rigorous reasoning using evidence and references to other academic works, and prioritising such reasoning over emotional feelings). The paper should demonstrate [critical thinking](#). The paper should **not** read like a work diary or a newspaper article. You should stay objective and factual, or otherwise clearly specify when you are being speculative or making assumptions.

All project submissions will be checked by plagiarism detection software (Turnitin), so please remember to upload a preliminary version of your paper well before the deadline if you want to check the Turnitin plagiarism score assigned to your paper. This will give your group adequate time to address any issues before submission.

Team members' contributions

The **contributions of each author** to the project as a whole must be specified in the author listing at the top of the paper. You should indicate a brief summary of the tasks conducted by each author and a contribution percentage representative of the work that has been conducted.

Individual grades will be adjusted according to the authors' contributions.

We expect that in most cases the project work will be divided equally amongst team members. Equal contributions for each author should then be used (25% for each author in groups of 4). If the contributions of team members were different, please agree on different contribution percentages for each author within your team well before the submission; for example:

Jane Doe	Bill Smith	Leanne Kains
Jane.doe@qmul.ac.uk	bill.smith@qmul.ac.uk	leanne.kains@qmul.ac.uk
Problem formulation, algorithm implementation, report writing.	Data acquisition and pre-processing, graph drawing, report writing.	Running tests, comparing algorithms, tabulating results, report writing.
35%	25%	40%

Discuss any issues with your mentor and/or the module organiser/lecturers long before the submission date so that we can help you resolve them.

PG students will be expected to provide more in-depth network analyses than UG students. More in-depth analyses can be conducted for example by (i) addressing a more challenging problem, (ii) providing a more comprehensive literature review, (iii) completing network analyses flagged as "Hard" in the topic description on QMplus (or similar ones), and (iv) providing a more detailed discussion on the implications of the results.

Network analysis tools

You are free to use any network data analysis tools, e.g. Gephi which doesn't require prior coding skills, or Python packages for network analysis such as NetworkX for those who have Python skills. These tools will be introduced during lab sessions. Other network analysis tools can be found in R and Matlab. You can calculate certain metrics by hand, however this approach will likely only be suitable for very small networks.

Submission

Each team member must submit the following on QMplus:

- Your **6-page long paper (excluding references)** in the PDF format. The filename of your PDF paper should follow this convention:

<Level>_Group_<Group number>_<Short title>.pdf

For example: `UG_Group_1_An Analysis of ties between UK and EU.pdf`

- Include in your paper an Appendix section ***only if*** you implemented your own dataset collection and/or network analysis algorithm (please include the code and descriptions). If you need an Appendix, the paper can go beyond 6 pages. QMPlus will only accept **one** PDF document, so you need to include the Appendix in the paper.

Project support resources

In addition to the information provided in the lectures and labs, **project support resources** can be found in the Group Project section on QMplus (within specific project topic pages and on the project resources page). These include references to network analysis software, tutorials, public network datasets, and bibliography.

Examples of projects from last year will also be added to the Group Project section on QMplus, however please note that the project guidelines were different last year (e.g. different topics, different types and number of tasks, the paper length is now shorter).

If you propose your own and topic and dataset, the selected or collected dataset must include over 200 nodes. It can either be collected by your group or be chosen from a public dataset such as those provided in the Stanford Large Network Dataset Collection (http://snap.stanford.edu/data/), the Index of Complex Networks (https://icon.colorado.edu/#!/networks?platform=hootsuite), and the Network Repository (http://networkrepository.com/). Please discuss your ideas, datasets (etc.) with your project mentor.

The marking criteria and estimated number of study hours for this assignment can be found on the following page.

Marking criteria for the Assignment 7 Project paper

Introduction and problem definition	10%
Related work	20%
Dataset and network presentation	10%
Network analysis methodology	10%
Results and discussion	40%
Conclusions and perspectives	10%

Estimated number of study hours for this coursework including lectures and tutorials: **75 hours**