Visualization of Lectures connected to Hydro Power

1. Introduction

Geography is a diverse field that spans both human and physical dimensions, each containing numerous sub-disciplines. At the University of Bern, these sub-disciplines are organized into research units to offer a comprehensive educational experience. Now in my master's program, I am focusing primarily on hydrology, with a specific interest in hydro power.

I have visited various courses available in geography, along with lectures in geology and business administration. This diversity can make it challenging to maintain a clear understanding of how each course aligns with my specific interest in hydro power. To address this, I have utilized network science to map out a visual representation of all the lectures I've attended that relate to hydro power. This approach has helped in creating a structured overview by categorizing these lectures into five core topics, thus facilitating a clearer path through my academic journey in geography.

On one hand, we have three core topics directly connected to hydropower: hydrology, natural hazards, and methods (involving lectures teaching skills in software such as GIS and Python). On the other hand, we have core topics like spatial planning and business administration, which, although not directly connected to hydropower, can support a career in the field and provide a broader understanding of related areas.

2. Methodology

This chapter outlines the methodology used to create network visualizations mapping the relationships between lectures and core topics related to hydropower within my academic curriculum. The process involved both manual data preparation and automated visualization techniques. Python scripts were employed to generate various visualizations, leveraging network science principles to provide clear, structured insights into the academic framework.

2.1 Data preparation

The data was manually collected and organized into a CSV file. This file categorizes lectures attended during my master's program into five core topics: Hydrology, Spatial Planning, Business Administration, Natural Hazards, and Methods.

The CSV file includes columns for each core topic and their corresponding lectures, along with additional columns for the European Credit Transfer and Accumulation System (ECTS) associated with each lecture. This structured format ensures comprehensive and accurate representation.

The allocation of lectures to core topics is based on the lectures content. While many lectures directly contribute to a core topic, some are only indirectly connected to their allocated core topic. New lectures must be manually added to this CSV file, ensuring the data remains current and accurate.

2.2 Network structure

Nodes in the network represent individual lectures or core topics, uniquely identified and labeled based on the lecture or core topic name. Edges represent the connections between core topics and lectures. An edge is drawn between a core topic and a lecture if the lecture is categorized under that core topic. Connections are established based on the inclusion of lectures under relevant core topics as defined in the CSV file. Additional connections are made between the main topic "Hydro Power" and all core topics to highlight its central importance.

2.3 Network visualization with Python

The automated part of the network visualization consists of Python scripts that read the CSV file and create a mind map-like network. Python libraries such as NetworkX, PyVis, and Matplotlib were used for creating static and interactive visualizations. The adjustText library was utilized to manage label overlaps. Nodes are color-coded based on core topics, with specific colors assigned to each core topic. Lecture nodes are colored light grey.

Node sizes are adjusted based on the ECTS credits associated with each lecture or core topic, and further weighted by the number of connections. Labels are applied to nodes, indicating the name of the lecture or core topic. Font sizes and weights vary to emphasize different elements.

A spring layout algorithm was used for positioning the nodes. Parameters were adjusted to ensure consistent and clear spacing between nodes, avoiding overlap and enhancing readability. For the interactive network visualization, PyVis was used to create a dynamic graph that allows users to hover over nodes to view detailed information such as the number of connections and ECTS credits.

2.4 Network types

Four different network visualizations have been created with the assistance of ChatGPT. The Python scripts include comments to explain the functionality of each code section. The goals and functionalities of each visualization are sketched within table 1.

Table 1: The four network visualization types.

Simple Static Network
• **Goal:** Provide a foundational overview of lecture connections to core topics. • **Functionality:** Reads the CSV file, constructs a graph with nodes for lectures and core topics, and draws edges between them. Nodes are color-coded and sized uniformly for simplicity.
Static Network with ECTS Weighting
• **Goal:** Incorporate ECTS values into the visualization. • **Functionality:** Adjusts node sizes based on ECTS credits, providing a visual representation of the academic weight of each lecture and core topic.
Static Network with ECTS and Connections Weighting
• **Goal:** Include both ECTS and the number of connections for node size adjustment. • **Functionality:** Node sizes are adjusted based on ECTS credits and the number of connections, offering a comprehensive view of the lecture's importance within the network.
Interactive Network with ECTS and Connections Weighting
• **Goal:** Create an interactive and user-friendly network visualization. • **Functionality:** Uses PyVis to generate an interactive graph where users can hover over nodes to see detailed information. Nodes are sized based on ECTS credits and connections, and core topic nodes are weighted similarly.

3. Results and Discussion

In this chapter, the four networks are presented and analyzed based on their visual appearance and effectiveness in offering a quick overview of the lectures and core topics related to hydropower. Additionally, the networks are examined to assess the value each lecture and core topic contributes to understanding hydropower and their potential impact on a career in the field.

3.1 Simple Static Network

The Simple Static Network visualization provides a foundational overview of how various lectures are connected to core topics. This visualization helps in identifying the structure and immediate connections between lectures and core topics, providing a clear but simple visual representation of the academic framework related to hydropower. One significant advantage of this network is its clarity and simplicity, making it easy to understand the direct relationships between lectures and core topics. With minimal text clustering, the network maintains good readability. However, the lack of detail is a notable disadvantage as it does not show the relative importance of each lecture, which could be misleading. Additionally, this network does not incorporate any weighting, which limits its depth in understanding the significance of each lecture. The clustering of the three core topics (Hydrology, Natural Hazards, and Methods) demonstrates a high amount of shared lectures. Spatial Planning only shares three lectures with the other core topics, while Business Administration is isolated, sharing no lectures.

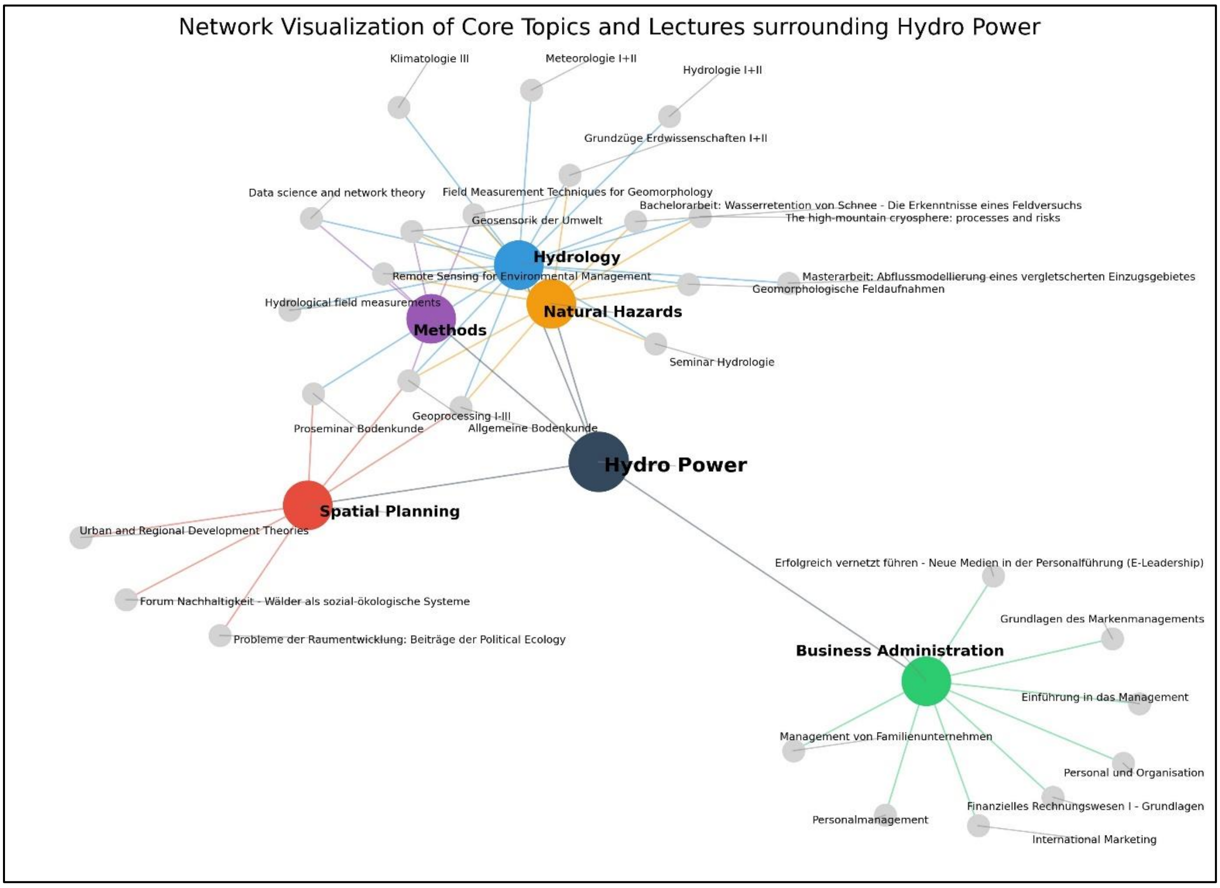

Figure 1: Simple Static Network

3.2 Static Network including the Weighting of ECTS

Building upon the Simple Static Network, the Static Network including the Weighting of ECTS incorporates the European Credit Transfer and Accumulation System (ECTS) values. The Python script adjusts node sizes based on the ECTS credits associated with each lecture and core topic. This enhanced visualization helps to understand the academic weight of each lecture and core topic in the context of hydropower studies, providing a clearer picture of their contributions to the curriculum. This network visualization allows one to instantly grasp which lecture contributes more based on ECTS values. Important modules like the “Masterarbeit”, “Bachelorarbeit”, or “Geoprocessing I-III” become more visible. However, the readability has decreased slightly as the node distribution was automatically adjusted due to the change in node size. This automatic adjustment can affect the distribution of nodes, making the network less readable. Additionally, some lectures may appear more important purely due to higher ECTS credits, which might not always align with their actual relevance to hydropower.

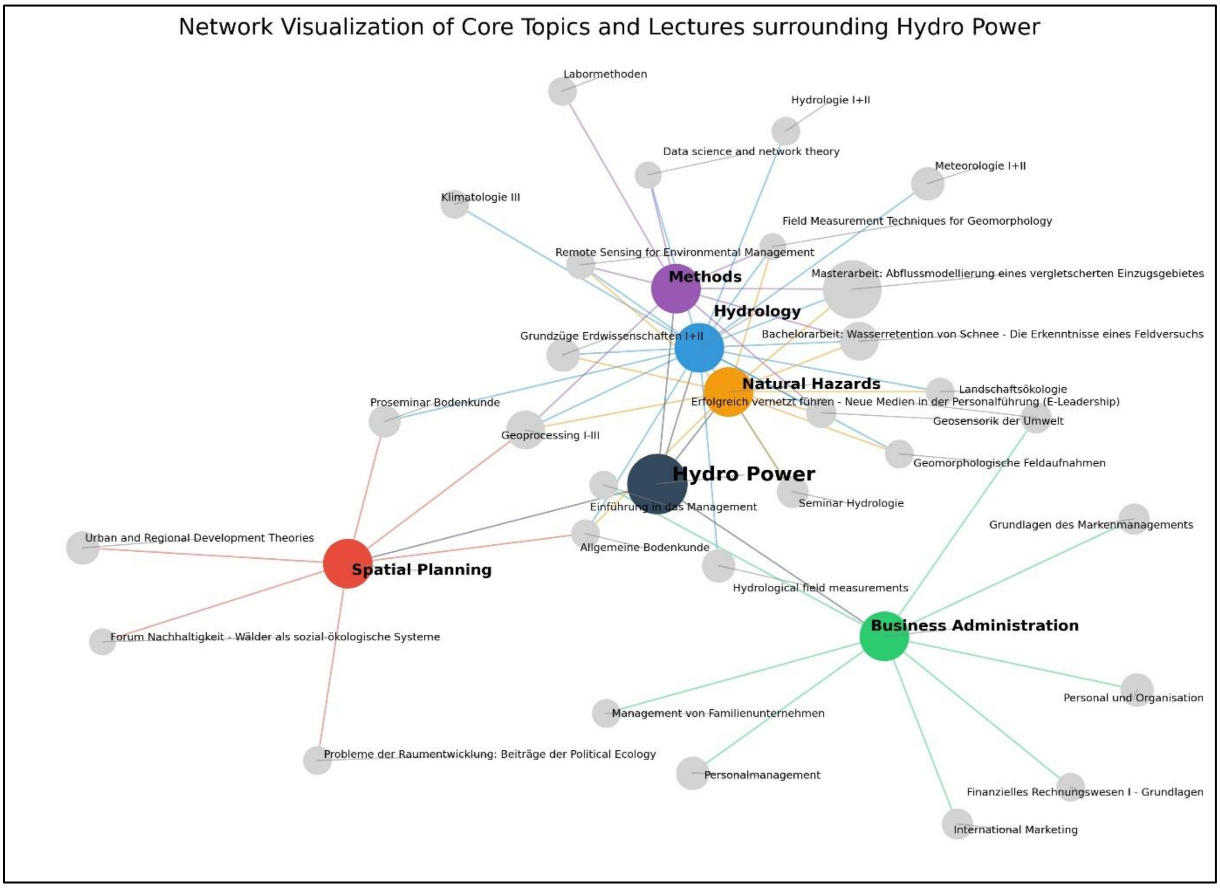

Figure 2: Static Network including the Weighting of ECTS

3.3 Static Network including the Weighting of ECTS and Connections

In addition to the previous network, the Static Network including the Weighting of ECTS and Connections incorporates the number of connections to adjust the weighting of the lectures. This provides a more detailed view of how interconnected each lecture is, alongside its academic weight, offering deeper insights into the significance of each lecture within the network. Lectures like “Allgemeine Bodenkunde” or “Geosensorik der Umwelt”, which are connected to three core topics and therefore arguably more relevant than a lecture which only brings value in a specific core topic, become more visible. The automated distribution of nodes in this plot improves readability. However, the added complexity might make the network harder to interpret at a glance. Finding a balance between ECTS weighting and connections can also be challenging, potentially leading to misinterpretation.

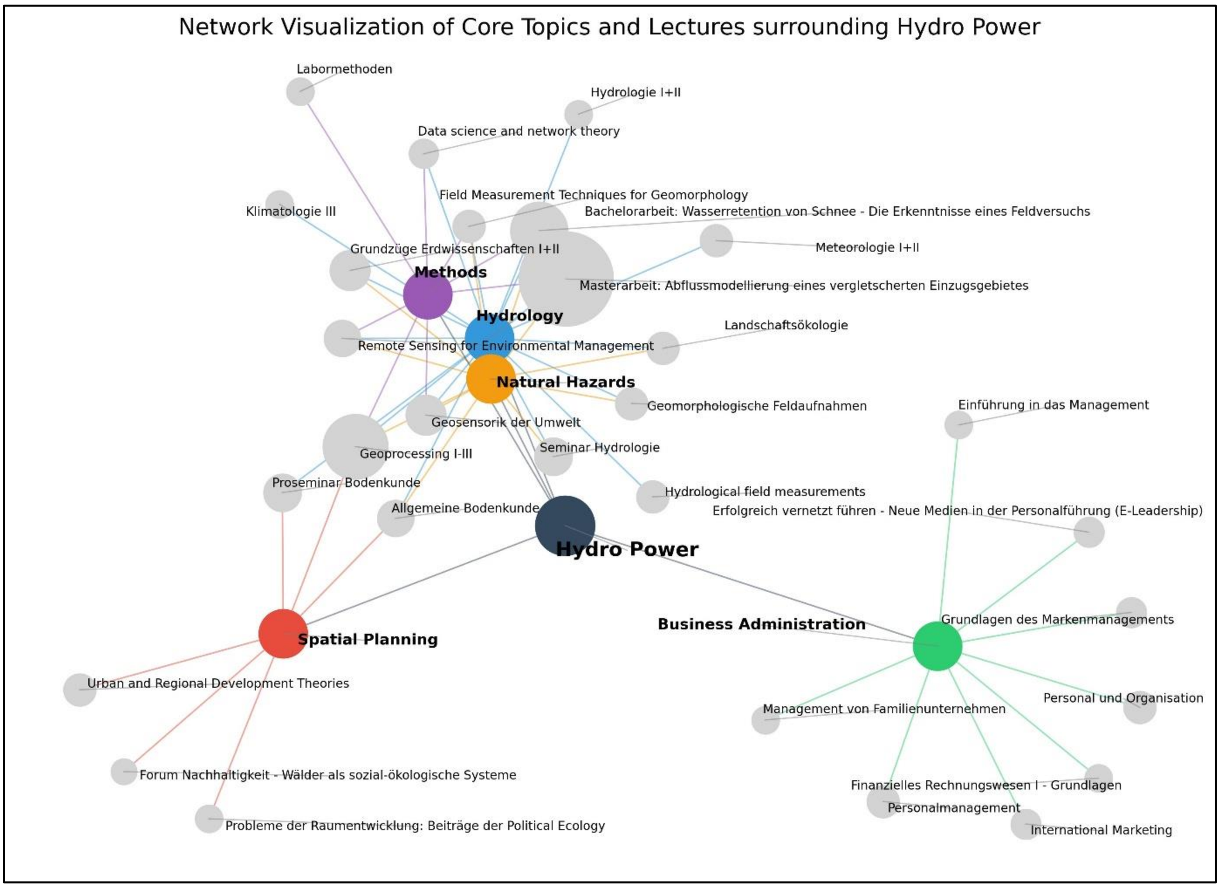

Figure 3: Static Network including the Weighting of ECTS and Connections

3.4 Interactive Network including the Weighting of ECTS and Connections

The Interactive Network visualization offers a dynamic and user-friendly way to explore the connections between lectures and core topics. This visualization allows users to interact with the network by hovering over nodes to see specific information about connections and ECTS credits. It is the most visually appealing and informative, providing detailed insights and a better distribution of nodes than static visualizations. In addition to the previous networks, this visualization also adapts the node size of the core topics based on the number of connected lectures and how much ECTS they provide. An interesting observation from this compared to the other visualizations is that Business Administration, based on credits and connections, adds more value to a future career in hydropower than Spatial Planning, although Business Administration has no shared lectures. The interactivity allows users to explore the network in greater detail, but it requires a higher level of engagement and might be less accessible for quick overviews. The downside of the interactive plot is that it cannot be properly used as a graphic in a document since it is a HTML file.

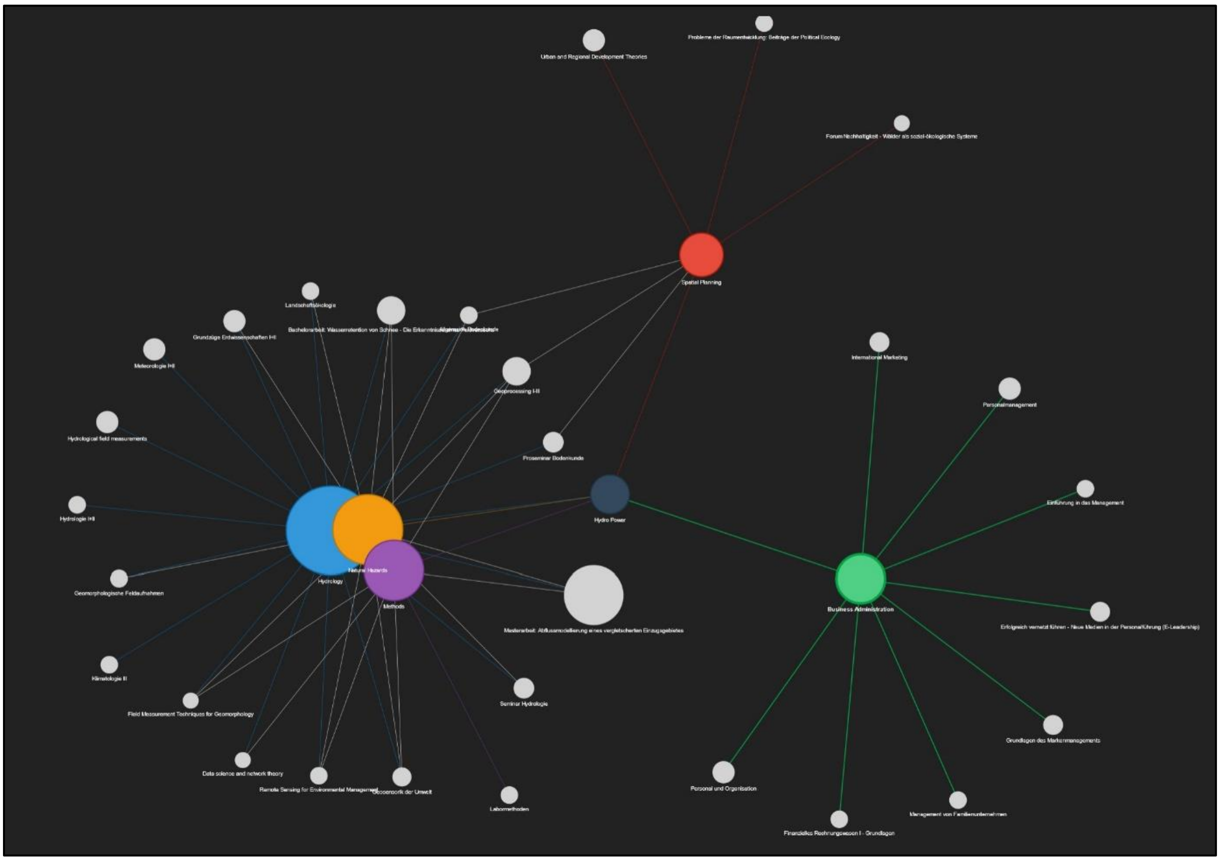

Figure 4:Picture of Interactive Network including the Weighting of ECTS and Connections.

4. Conclusion

These network visualizations provide a structured overview of the lectures and core topics related to hydropower, facilitating a clearer path through the academic journey. The visualizations help in identifying key lectures and their contributions to the overall curriculum, making it easier to plan and understand the academic focus areas. While the node sizes don't necessarily reflect the actual value of lectures or core topics regarding hydropower, the analysis broadly agrees with personal insights about the importance and helpfulness of the individual lectures and modules.

The networks are easy to update and can be shared with future employers, providing them with a clear understanding of my studies. Given the broad scope of geography, this helps potential employers better appreciate my specific skills and knowledge areas, ensuring they know exactly who they are hiring. Therefore, the goal of creating a visual overview of my studies in relation to a possible career in hydropower has been successfully achieved. This approach ensures that the visualizations are accurate, informative, and valuable tools for academic planning and analysis, offering multiple perspectives on the curriculum and its relevance to hydropower.

References

OpenAI. (2024). ChatGPT. Available at: https://www.openai.com/chatgpt has been used to assist the writing of the python scripts.