TASK 2:

1. Implement any two algorithms/ ML methods for community detection on the graph at any time T
    a. Louvain Method

    The Louvain method is a community detection algorithm designed to find modular structures in complex networks. It is based on the optimization of modularity, a metric that quantifies the quality of a partition of a graph into communities. The algorithm is iterative and consists of two phases: the "greedy" phase and the "aggregation" phase.

I.Greedy Phase:

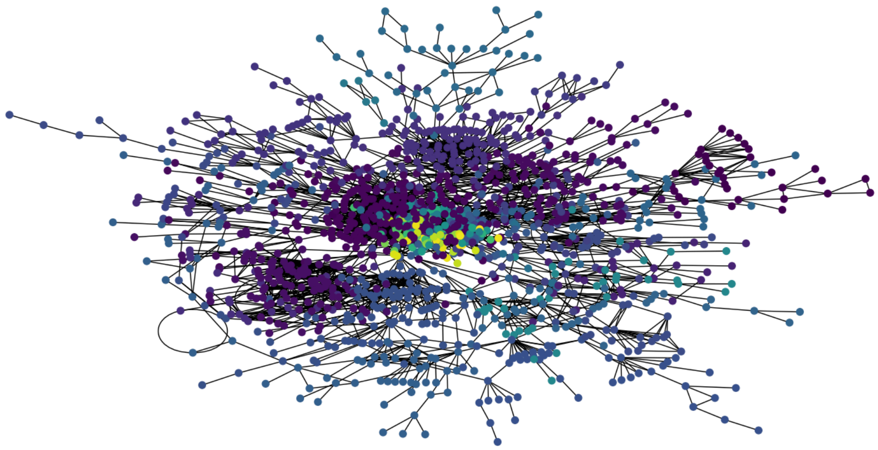
- Initialization: Each node starts as its own community.
- Iteration:
    - For each node, it evaluates the change in modularity if the node were to move to the community of one of its neighbors.
    - If the modularity increases, the node is moved to the community with the maximum modularity gain.
    - This process is repeated until no further improvement in modularity is possible.

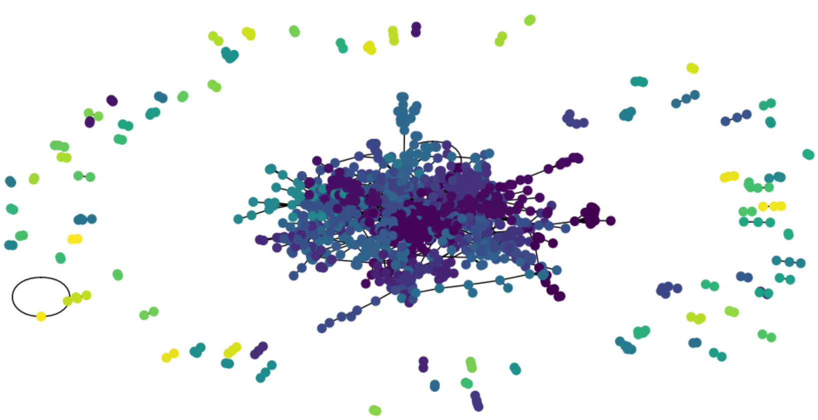
Ii. Aggregation Phase:

- The communities found in the first phase are considered as nodes in a new network, and edges between these nodes are weighted by the sum of the edge weights between nodes in the original network.
- The algorithm then goes through the greedy phase again on this new network, iteratively refining the partitioning.

This graph was made with the help of data at time 1994-01-01.

This graph contains communities that were cut using Louvain Method





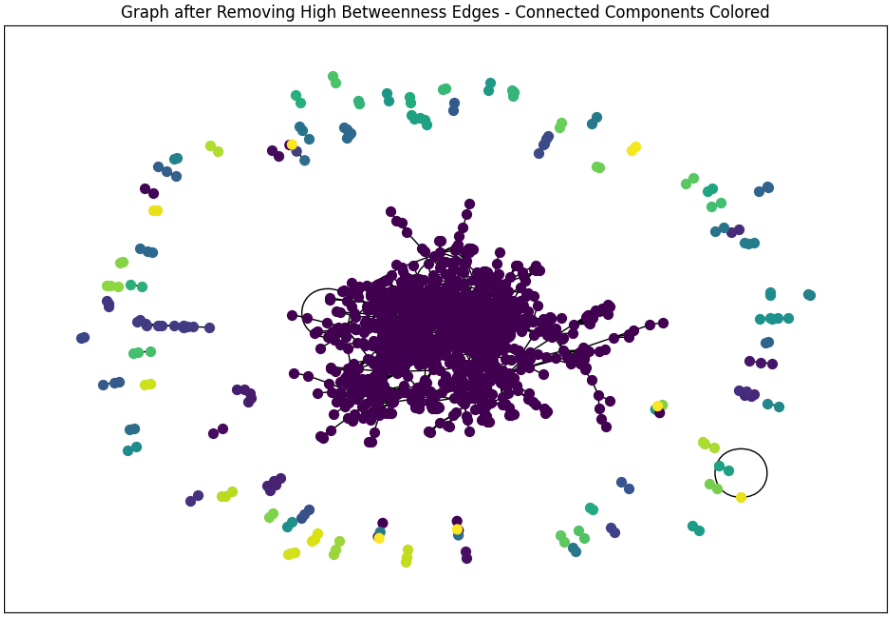
b. Using Betweenness Centrality:

Steps include:

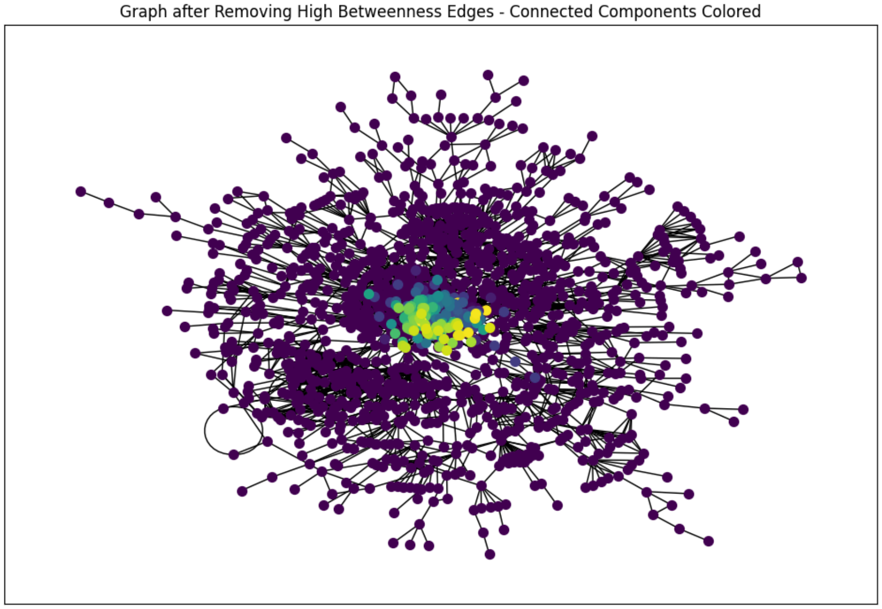
- **Setting Up the Network**:
  - Begin with a graph that illustrates relationships or citations.
- **Node Selection Based on Date**:
  - Create a subgraph by filtering nodes according to a specific date.
- **Computing Betweenness Centrality**:
  - Evaluate the betweenness centrality for each node in the subgraph.
- **Identifying Nodes with High Betweenness**:
  - Nodes surpassing a set threshold in betweenness centrality are deemed influential.
- **Eliminating Edges Connected to High-Betweenness Nodes**:
  - Disconnect nodes associated with high-betweenness by removing their edges.
- **Visualizing the Altered Network**:
  - Observe the network's modified form post-edge removal to uncover potential clusters.
- **Improving Cluster Separation**:
  - Enhance the distinction between clusters for better interpretability.
- **Simplifying the Network and Reducing Noise**:
  - Streamline the network by minimizing noise introduced by influential nodes.
- **Exploring Community Detection Possibilities**:
  - Disconnected subgraphs may indicate the presence of distinct communities.
- **Gaining Insights into Network Dynamics**:
  - Understand alterations in connectivity and potential vulnerabilities within the network.
- **Fine-Tuning and Iterating**:
  - Optionally, adjust parameters and repeat the process for further refinement.

The clustering was done on the same data as above ie 2 years

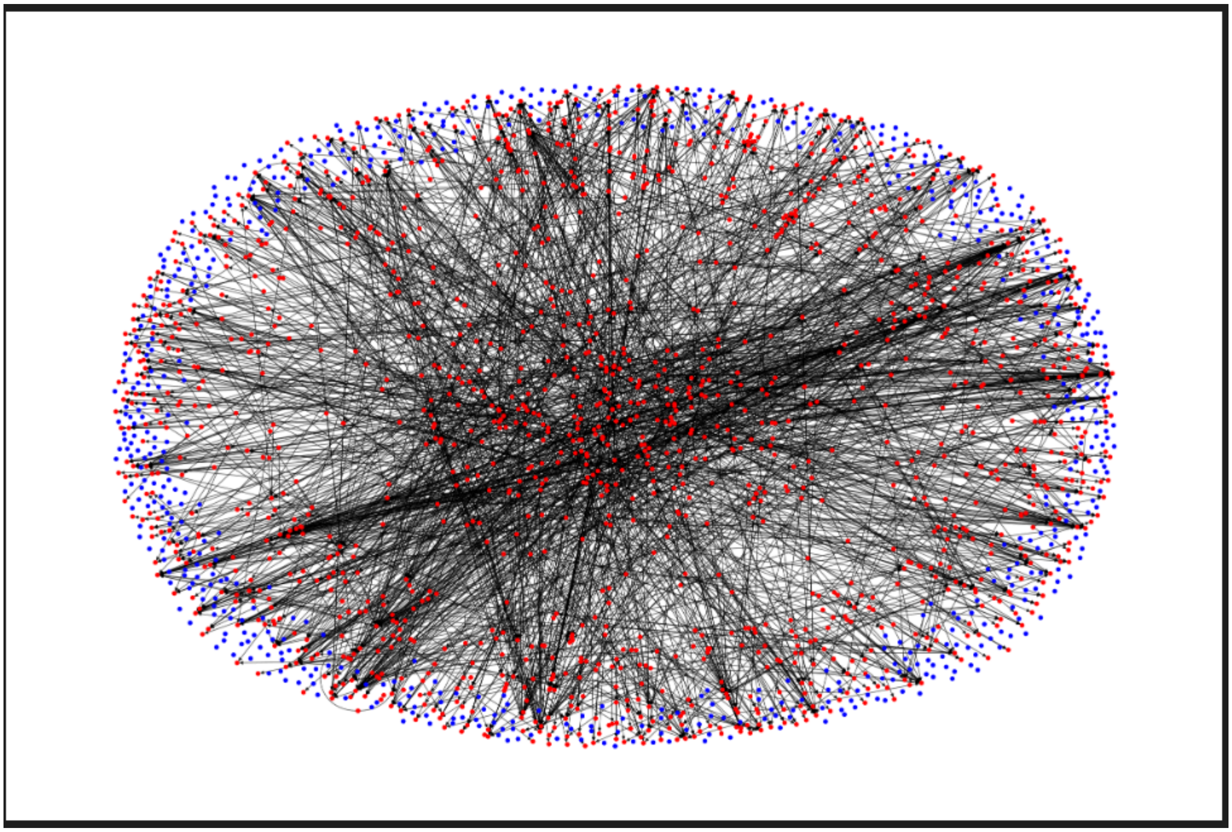
Graph after Removing High Betweenness Edges - Connected Components Colored



Graph after Removing High Betweenness Edges - Connected Components Colored



2.Analyzing the communities



This is the same graph without community detection,

Community detection algorithms, such as the Louvain Method, aim to identify groups or communities of nodes within a graph based on patterns of connectivity. The results of community detection may provide insights into the underlying structure or organization of the network. Analyzing why a community formed as it did can involve considering several factors:

A. Density of Connections: Communities often form because nodes within a community have a higher density of connections among themselves compared to connections outside the community. This suggests a more cohesive and interconnected subgroup.
B. Modularity Score: The Louvain Method optimizes modularity, a measure of the quality of community structure within a network. Modularity is calculated based on the difference between the observed and expected connections

within communities. A higher modularity score indicates a better-defined community structure.
C. Node Similarity: Nodes within a community may share similar attributes, such as node degree, centrality, or other topological features. Analyzing the characteristics of nodes within a community can help understand why they are grouped together.
4. Edge Betweenness: The Louvain Method often considers the edge betweenness, which measures the number of shortest paths passing through an edge. Communities may form around edges with lower betweenness, indicating that these edges act as "bridges" between communities.
5. Graph Topology: The overall structure of the graph, including its size, connectivity, and the distribution of nodes, can influence community formation. For example, densely connected subgraphs may naturally form distinct communities.
6. Dynamic Processes: Consider any dynamic processes that might have influenced community formation, such as the evolution of the network over time or the influence of external factors.
7. Domain-Specific Knowledge: Depending on the nature of the graph, domain-specific knowledge can provide additional insights into the reasons behind community formation.

In Summary,

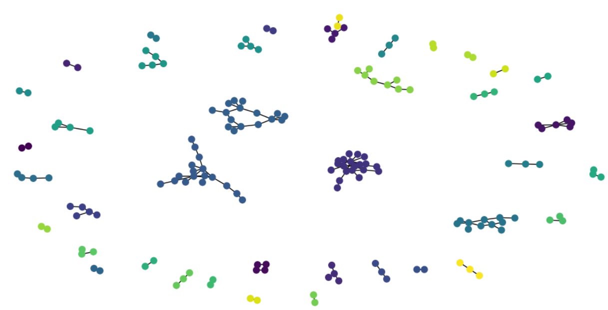
Using community detection in this graph can help us separate domains of research, As nodes which are similar or more tend to be closely connected are separated out in a community. The Louvain Method could not cut a very discrete line between nodes , but help nodes in the same cluster separate out. The Edge-Centrality Method on the other hand could not separate out nodes that are somewhat similar but can help create absolute lines between domains of the research.

Another approach could be applying the centrality algorithm and then applying the Louvain Method of Clustering. This could help in separating out absolute domains and also help us see the minor separations between subtopics of the same domain.

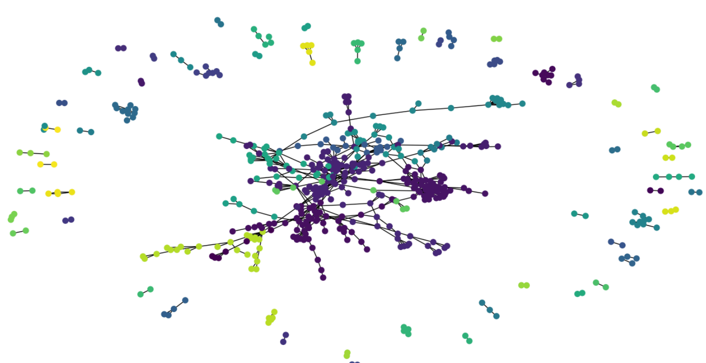
In conclusion, the combination of community detection and centrality algorithms offers a nuanced exploration of the graph, providing insights into both overarching domains and finer distinctions within interconnected research topics.

3. Temporal Community Detection

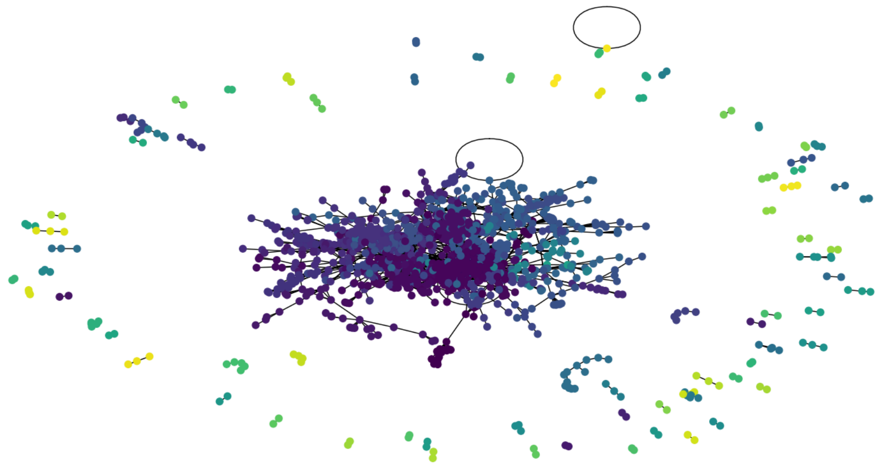
Below are the phases the graph evolved through and how communities evolved with time.



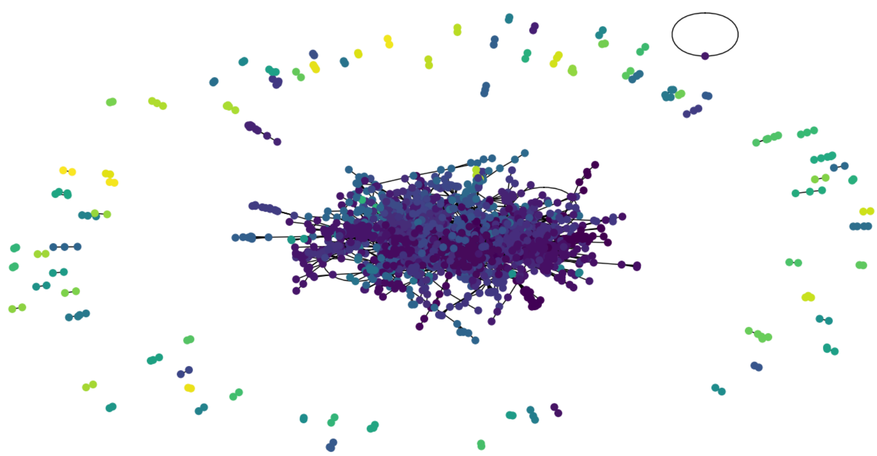
Phase 1



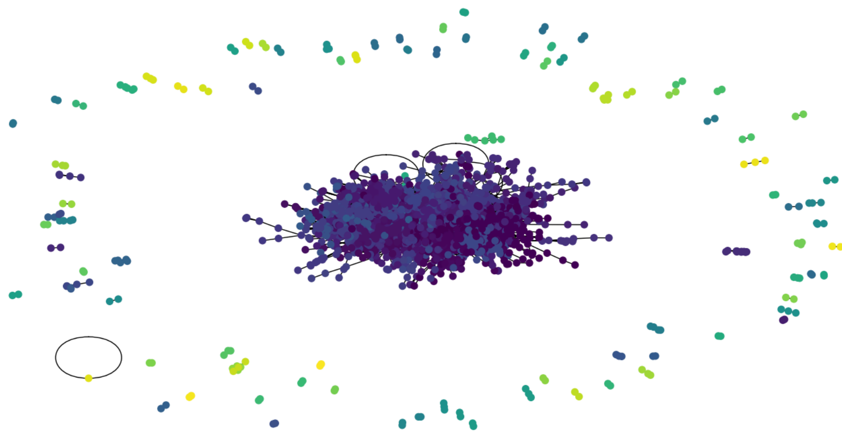
Phase 2



Phase 3



Phase 4



Phase 5

These are the evolving Phases of graph 2.5 yrs, each phase took 6 months to develop.

Interesting Insights

1. Phase 1 shows how initially different domains boomed in research independently.
2. Phase 2 shows how these different domains started linking through research that was a part of both, Though the topology of the graph was sufficient enough to differentiate these domains.
3. Phase 3 shows how these links between some domains increased so much in number that they were soon considered the parts of the same community.
4. While some nodes that were originally clustered into the same community grew branches that they got separated into different communities.
5. This community analysis would be beneficial enough to understand that things like “Wave Nature” and “Particle Nature” that were

initially studied as different topics are now considered under the same domain of “Quantum Physics”
The graph temporal Community Detection Algorithms will help us to find such more topics and study how physics evolved through time.

6. Another such analysis that can be drawn from such algorithms is how topics that were originally researched under the same domain separated out with the course of time.
Such examples include, that the “study of particles at atomic level” was considered a domain that people researched in. Later this thing got separated out and it broke into two communities “Nuclear Studies” and “Nano Technology” that have nothing in common now other than the original roots.
This is an example of how a community broke into 2 different communities with time and new and new research progressed on the particular sides of the node.so vast in number