Assignment 2: Complex Networks 2025

Antonio Lobo, Andras Kasa

March 2025

1 Introduction

Community detection is a central problem in network science, aiming to uncover groups of nodes that are more densely connected internally than with the rest of the network. These structures often correspond to functional units in social, biological, or technological systems. In this report, we evaluate the performance of several community detection algorithms on both synthetic and real-world networks.

The study is divided into two parts. First, we examine synthetic networks generated using a Stochastic Block Model (SBM), varying intra-community connection probabilities while keeping other parameters fixed. We assess the effectiveness of Infomap, Louvain, Greedy Modularity, and Girvan-Newman algorithms using structural metrics (e.g., modularity) and external validation indices (e.g., NMI, Jaccard Index). Second, we analyze a real-world face-to-face interaction network from a primary school, investigating the effect of edge weighting on detection quality.

This work aims to highlight how algorithmic assumptions, network structure, and the presence of edge weights influence the ability to recover meaningful community partitions. We also examine computational trade-offs and sensitivity to parameters, providing insights into the practical deployment of these methods in complex settings.

2 Methodology

2.1 Community Detection Algorithms

For the synthetic networks, we implemented four community detection algorithms:

- **Infomap:** This approach considers community detection as an issue of information theory. Information flow on a network can be explored through a random walker, and communities in the network are found by minimizing the description length of the walker's trajectory. In short, the communities in the network are found from the regions in the graph where the random walker is most likely to be "held" for a long time.
- **Louvain:** It is a fast hierarchical algorithm that maximizes modularity - the measure of the internal density of links within a community relative to the external links to other communities. It starts with each node as a separate community, and then iteratively combines communities in order to increase modularity. The algorithm is well known for its practical efficiency even in large networks.
- **Girvan-Newman:** Unlike the others, this is a divisive algorithm that works by progressively removing edges from the network rather than aggregating nodes. It targets edges with high "betweenness centrality": those that serve as bridges between different parts of the network. As these bridges are removed, the network naturally fragments into distinct communities. The process can be visualized as a dendrogram showing the hierarchical structure of communities. This algorithm is computationally intensive, with a time complexity of $O(m^2n)$ for a network with m edges and n nodes, making it less practical for large networks but valuable for its conceptual approach to community detection. For our implementation, we limited the algorithm to generate up to 10 communities and selected the partition with the highest modularity value. This approach significantly reduced computation time while still preserving the quality of the community detection. In our accompanying Jupyter notebook, we provide a parallelized version that runs all community detection algorithms concurrently, making the analysis more efficient. Despite these optimizations, the Girvan-Newman algorithm remains the most time-consuming component of our analysis.

For the real network, the weighted and unweighted graphs are analyzed using a modularity maximization approach (Louvain or Leiden) while keeping node positions fixed from the weighted network layout (using, for example, Kamada-Kawai with the inverse of the weight values).

2.2 Evaluation Metrics

To compare the community detection results with the known true partition, we utilized several established metrics that quantify the similarity between different community partitions:

- **Jaccard Index**: A measure of the similarity between two partitions based on the pairs of nodes. It computes the ratio of node pairs that are grouped together in both partitions to the total number of node pairs that are grouped together in at least one partition. The Jaccard Index ranges from 0 (no similarity) to 1 (identical partitions).
- **Normalized Mutual Information (NMI)**: An information-theoretic measure that quantifies how much information one partition provides about another. It captures the agreement between two partitions while accounting for random chance. NMI values range from 0 (no mutual information) to 1 (perfect agreement), with higher values indicating better alignment between the detected communities and the ground truth.
- **Normalized Variation of Information (NVI)**: A metric based on the concept of information variation that measures the distance between two partitions. It represents how much information is lost and gained when moving from one partition to another. Lower values indicate better matching between partitions, with 0 representing identical partitions.

Additionally, for each algorithm and for different values of prr (the intra-block connection probability), we tracked:

- **Number of Communities**: The total number of distinct communities identified by each algorithm. This allows us to compare how well the algorithms recover the true number of communities (5 in our stochastic block model).
- **Modularity (Q)**: A quality function that measures the strength of division into communities by comparing the density of connections within communities to connections between communities. Higher modularity values indicate better community structures, with values typically ranging from 0 to 1 for networks with strong community structure.

These metrics collectively provide a comprehensive assessment of how well different community detection algorithms perform across varying levels of community distinctiveness in the network.

3 Results

3.1 Synthetic Networks Analysis

3.1.1 Baseline Performance of Community Detection Algorithms

Our analysis of community detection algorithms reveals how their performance changes as the intra-block connection probability (prr) increases. When prr is extremely low (0.0), all methods struggle to identify the correct communities. Infomap fragments the network into 32 clusters, while the other methods deliver poor similarity scores compared to the ground truth. As prr increases to 0.1, both Louvain and Greedy make significant progress—achieving NMI scores around 0.5—whereas Infomap surprisingly declines by merging everything into a single community.

At $prr = 0.2$, we reach a critical threshold: Infomap, Louvain, and Greedy all nearly perfectly identify the true community structure, with NMI and Jaccard scores approaching 1.0 and accurately detecting all 5 communities. This high performance continues steadily up to $prr = 1.0$, with only a minor deviation in Infomap at $prr = 0.5$, where it identifies just 4 communities.

Table 1 shows the detailed performance of our constrained Girvan-Newman implementation.

Table 1: Performance of the constrained Girvan-Newman algorithm at different prr values

prr	**Metric**	**girvan_newmanc**
0.1	NMI	0.05
	VI	0.30
	Jaccard	0.19
	# Communities	10
0.2	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5
0.3	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5
0.4	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5
0.5	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5
0.75	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5
1	NMI	1.00
	VI	0.00
	Jaccard	1.00
	# Communities	5

Our constrained Girvan-Newman implementation, limited to exploring up to 10 community divisions and selecting the partition with highest modularity, shows interesting performance patterns. It struggles at $prr = 0.1$ with an NMI of only 0.05, but achieves perfect detection (NMI = 1.0) at $prr = 0.2$ and maintains this perfect performance through all higher prr values. This sharp transition aligns with what we observed in the visual analysis, where the constrained Girvan-Newman showed a more binary pattern—either failing to detect communities or perfectly identifying them—compared to the more gradual improvements seen in other algorithms.

The computational efficiency comparison reveals substantial differences among the algorithms. For networks at $prr = 1.0$, Infomap completed in just 0.04 seconds and Louvain in 0.09 seconds, while Greedy required 0.54 seconds. In stark contrast, our Girvan-Newman implementation required significantly more time—hundreds of seconds for the same task. These empirical results align with the theoretical complexity of the Girvan-Newman approach ($O(m^2n)$ for a network with m edges and n nodes) and illustrate why it becomes prohibitively expensive for analyzing larger networks without appropriate constraints.

These findings complement our earlier visual analysis, confirming the existence of a detectability threshold around $prr = 0.2$ and highlighting the varying strategies each algorithm employs when community structure is weak versus when it becomes well-defined.

3.1.2 Visual Analysis of Algorithm Performance

Let's look at how well each algorithm performs as we increase community strength (prr). Figures 1, 2, 3, 4, and 5 show us different ways to measure how accurately the algorithms detect communities.

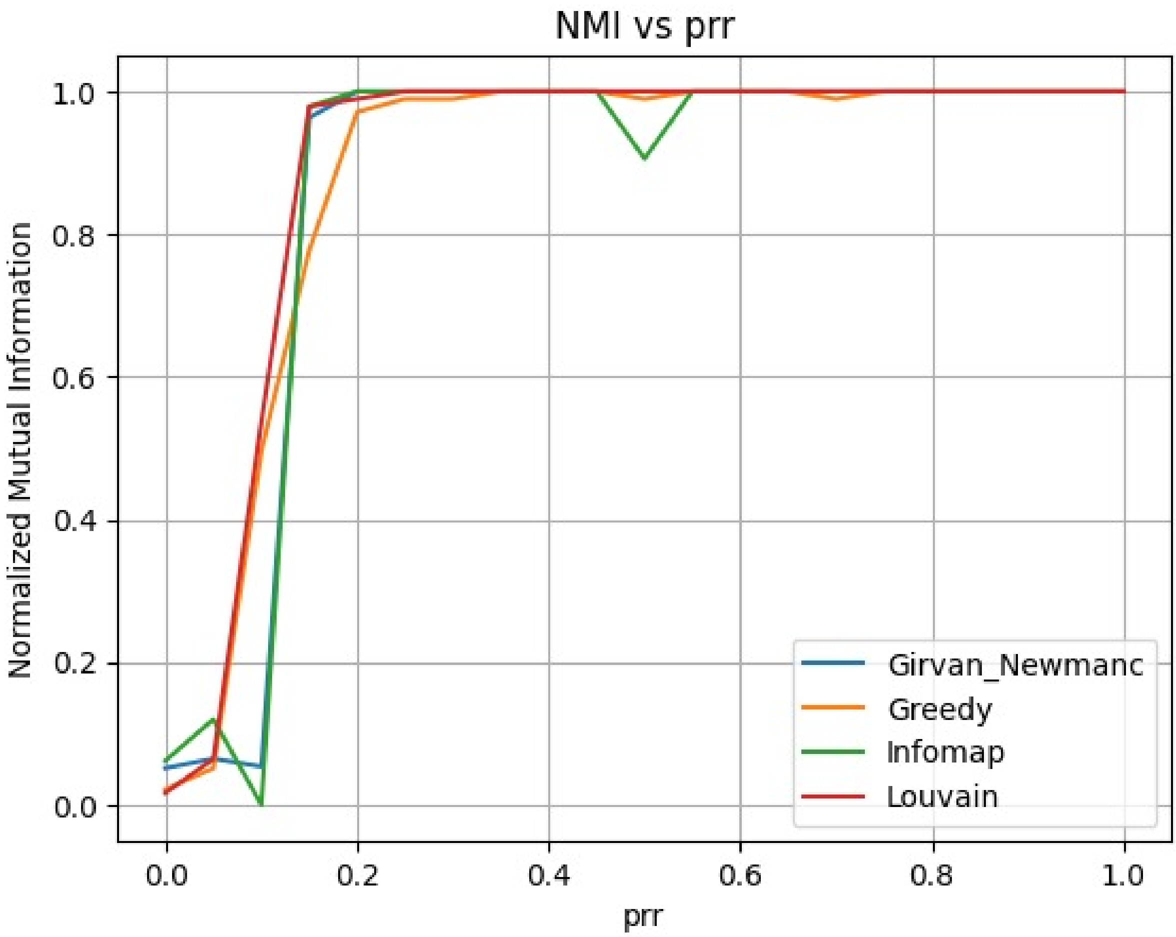


Figure 1: Normalized Mutual Information (NMI) for different community detection algorithms across varying intra-community connection probabilities (prr). Higher values indicate better alignment with the ground truth.

In the NMI plot (Figure 1), we see that all algorithms suddenly improve when prr reaches about 0.16-0.2. The Louvain algorithm (red line) improves first, followed by Girvan-Newman constrained (blue) and then Greedy (orange). Infomap (green) behaves a bit strangely, with a noticeable drop around $prr \approx 0.5$ where it temporarily gets worse.

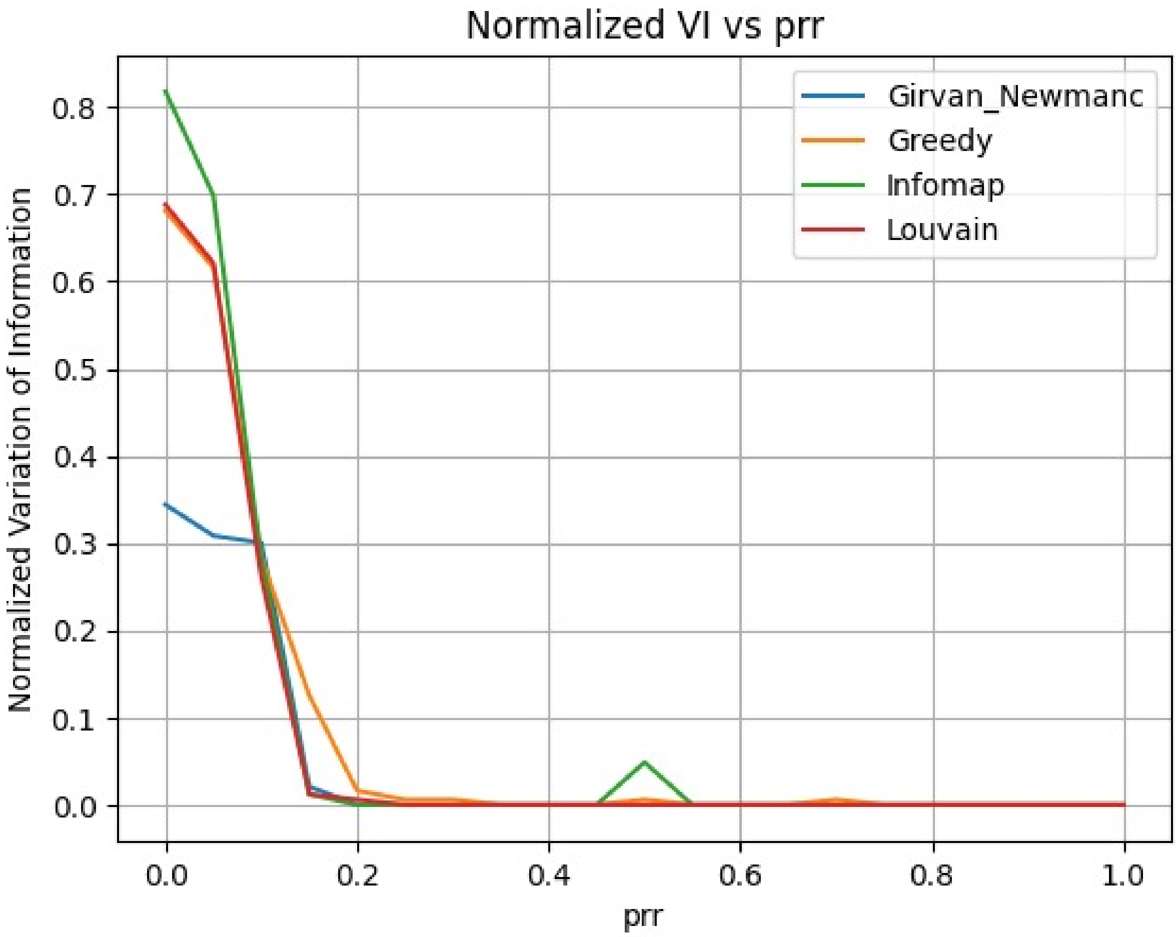


Figure 2: Normalized Variation of Information (VI) for different community detection algorithms across varying intra-community connection probabilities (prr). Lower values indicate better performance.

The Normalized VI plot (Figure 2) tells the same story but from another angle - here, lower values are better. We see a sharp drop around $prr \approx 0.2$, with most algorithms approaching zero (perfect detection) when $prr > 0.2$. Interestingly, Girvan-Newman constrained starts with the lowest VI value when prr is very low, suggesting that its simple approach of putting everything in one community actually works better than trying to find many communities when there isn't a clear structure.

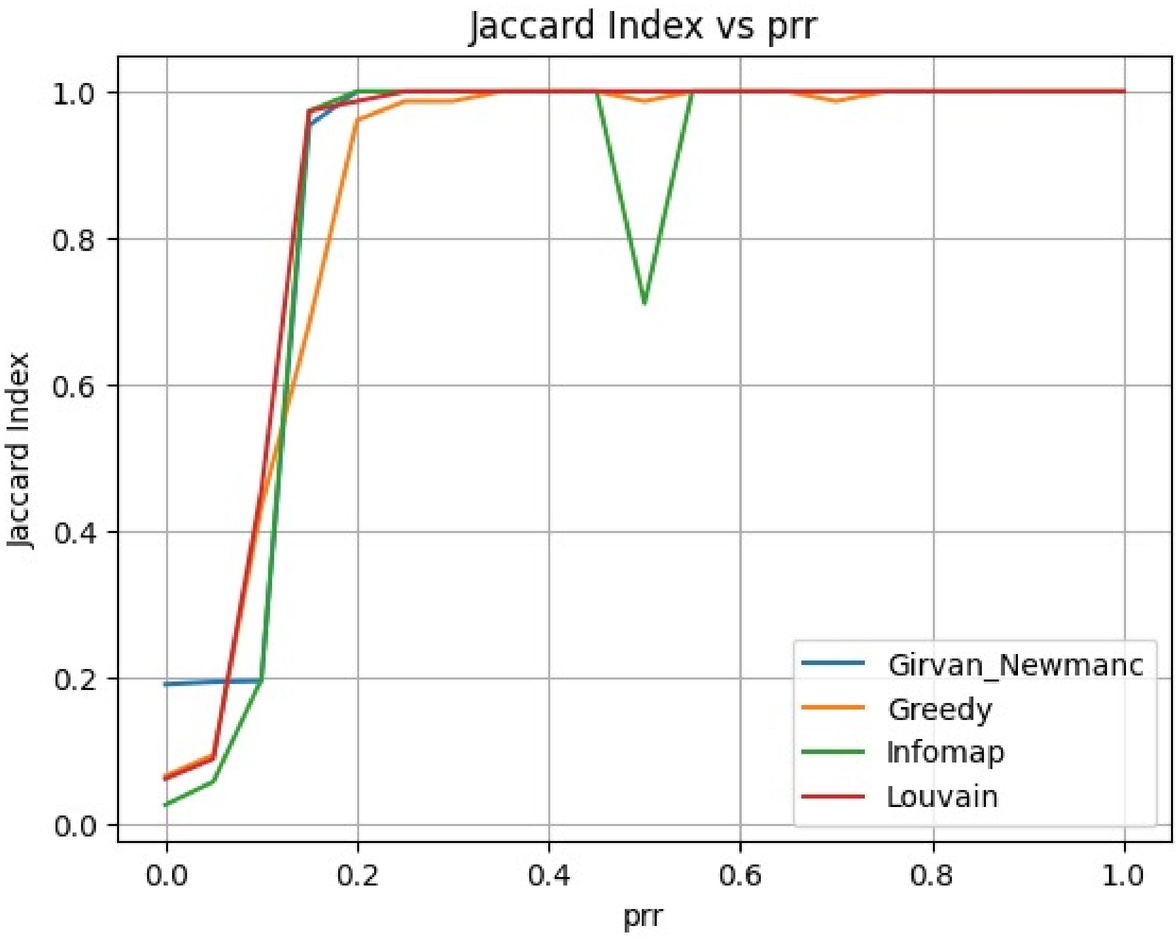


Figure 3: Jaccard Index for different community detection algorithms across varying intra-community connection probabilities (prr). Higher values indicate greater similarity to the ground truth communities.

The Jaccard Index (Figure 3) confirms what we've seen already. All algorithms improve dramatically around $prr \approx 0.2$. Louvain and Girvan-Newman constrained get it right slightly before Greedy does, while Infomap again shows that strange dip at $prr \approx 0.5$.

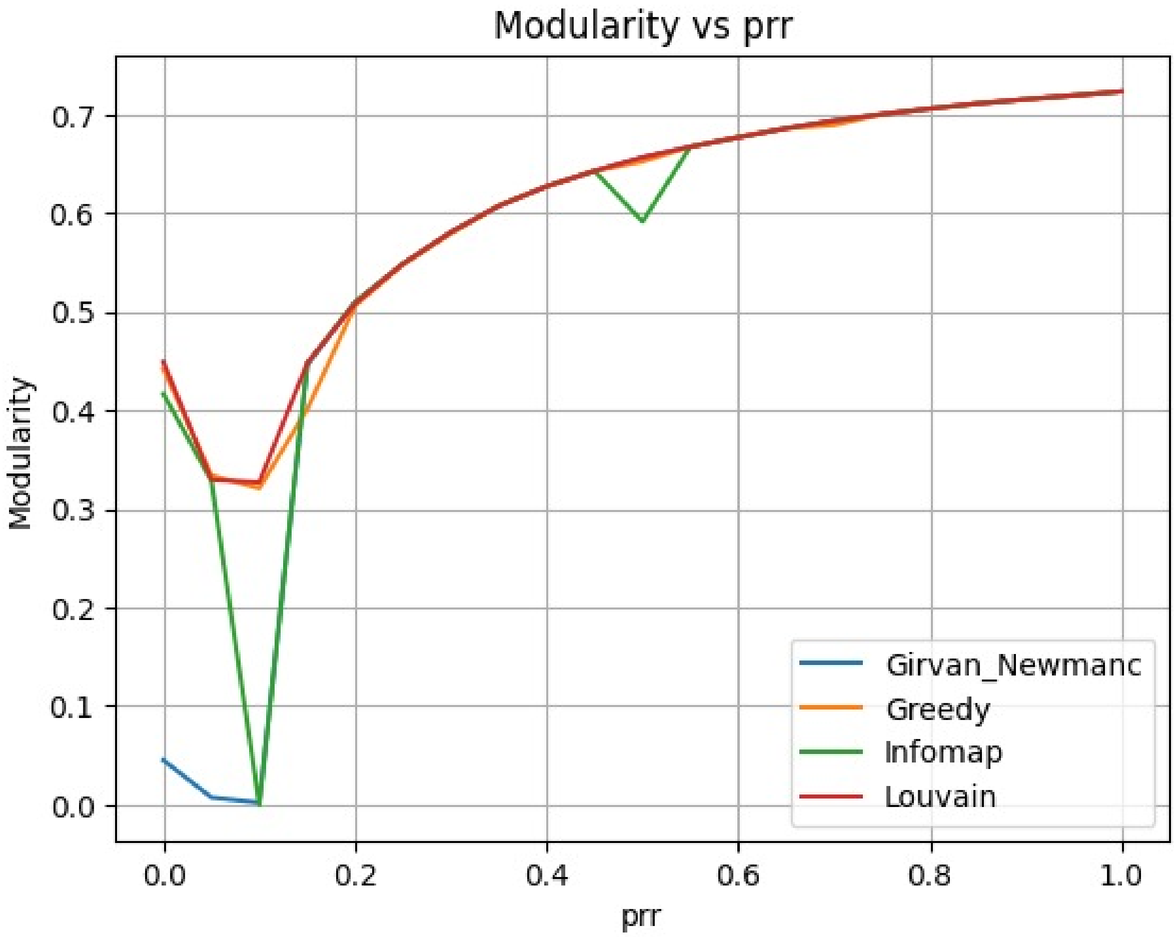


Figure 4: Modularity values for different community detection algorithms across varying intra-community connection probabilities (prr). Higher values indicate stronger community structure.

Figure 4 shows the modularity values - a measure of how well-defined the communities are. Surprisingly, even when prr is low, most algorithms (except Girvan-Newman) find divisions with fairly high modularity values, even though these divisions don't match the ground truth. This is an important warning: high modularity doesn't always mean correct communities! As prr increases, all algorithms eventually agree on similar modularity values.

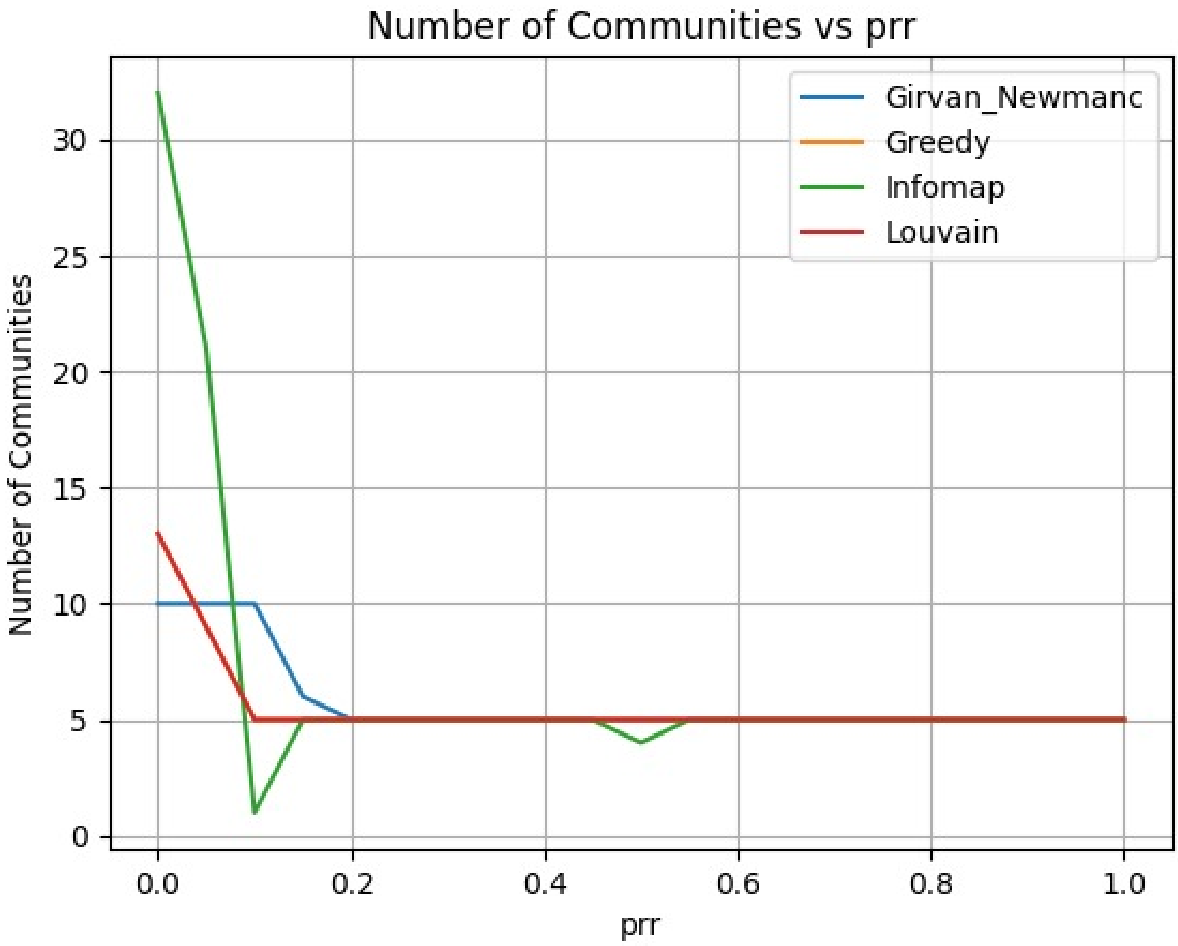
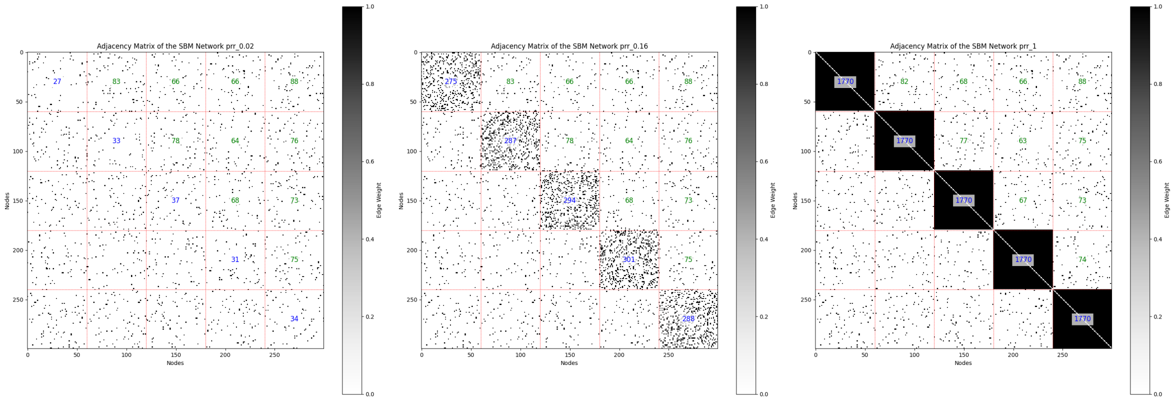


Figure 5: Number of communities detected by each algorithm across varying intra-community connection probabilities (prr). The ground truth contains 5 communities.

Figure 5 shows how many communities each algorithm finds. When prr is very low, the algorithms behave very differently: Infomap creates many small communities (over 30), Louvain finds about 13, and Girvan-Newman constrained stays around 10 before finding the right number. Once $prr > 0.2$, all algorithms correctly identify 5 communities, except for Infomap which briefly gets confused at $prr \approx 0.5$.

3.1.3 Visual Representation of Community Structure

Figure 6 shows the adjacency matrices for $prr = 0.02$, $prr = 0.16$, and $prr = 1.00$, illustrating how the block structure becomes more defined as prr increases.



(a) $prr = 0.02$ (b) $prr = 0.16$ (c) $prr = 1.00$

Figure 6: Adjacency Matrix obtained for different prr values.

To further investigate how each algorithm performs at different community strengths, we visualize the communities detected by each method. Figures 7 through 10 show the community assignments for each algorithm across three representative prr values. In all visualizations, node positions are fixed using the Kamada-Kawai algorithm applied to the $prr = 1.00$ network, and colors indicate community assignments.

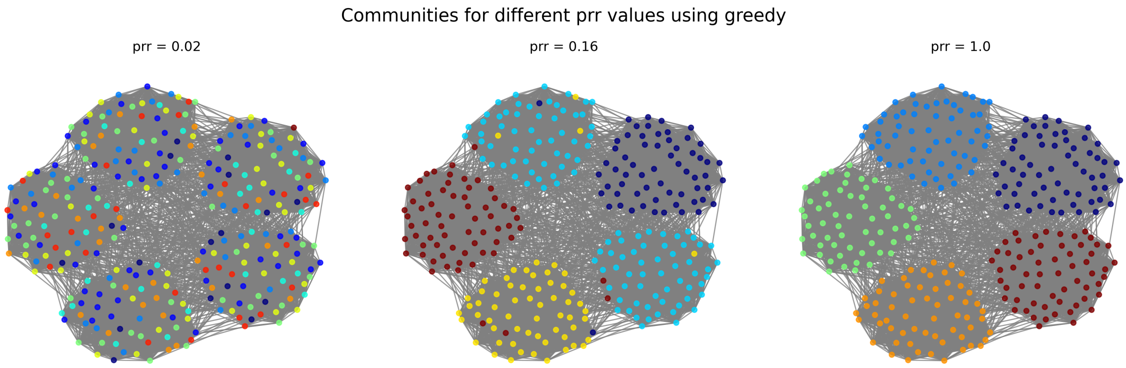


Figure 7: Communities detected by the Greedy Modularity algorithm at different prr values.

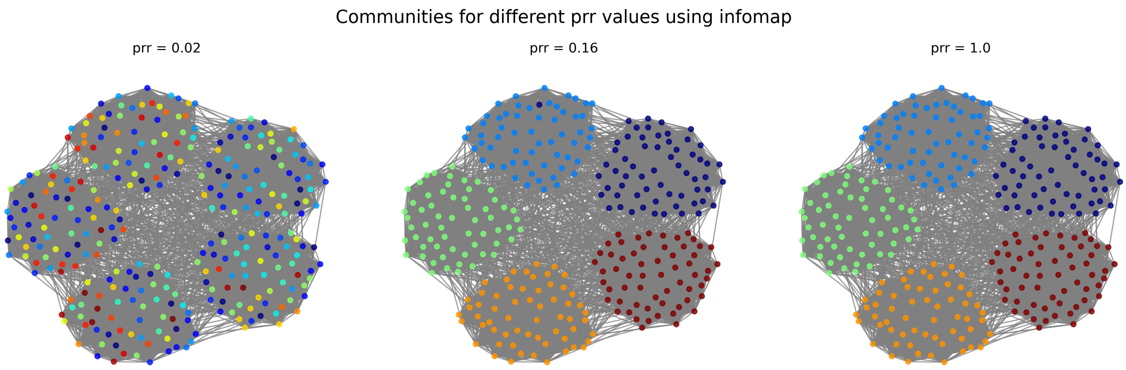


Figure 8: Communities detected by the Infomap algorithm at different prr values.

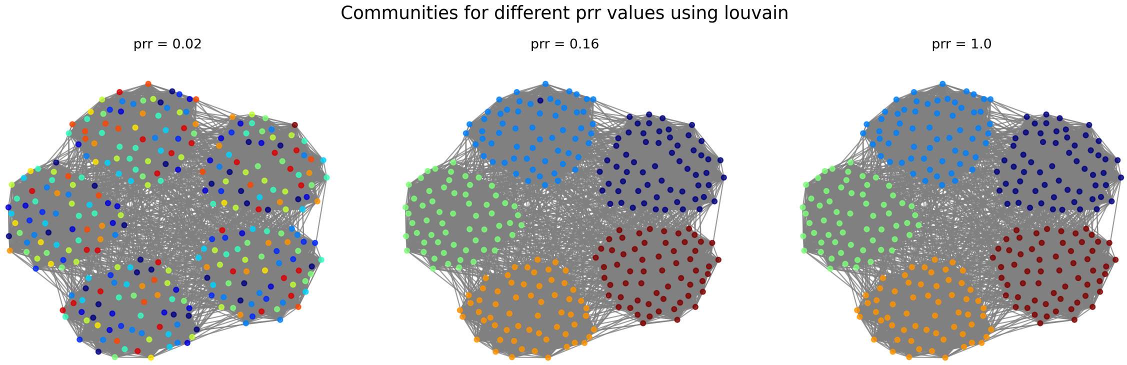


Figure 9: Communities detected by the Louvain algorithm at different prr values.

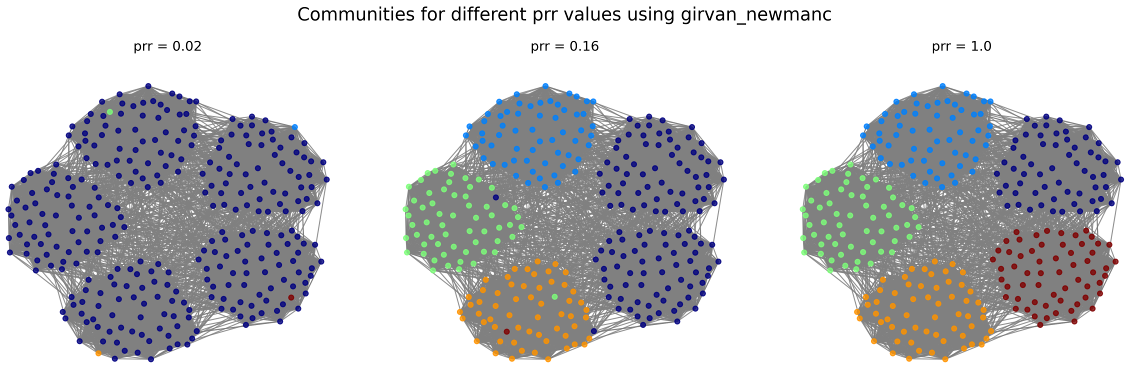


Figure 10: Communities detected by the constrained Girvan-Newman algorithm at different prr values.

These visualizations reveal distinct behaviors among the algorithms. At $prr = 0.02$, Greedy, Infomap, and Louvain algorithms produce fragmented communities, while the constrained Girvan-Newman algorithm identifies just one community (all blue nodes). At $prr = 0.16$, near the detectability threshold, most algorithms begin to identify meaningful communities, with Infomap, Louvain, and constrained Girvan-Newman showing similar five-community structures. Notably, the Greedy algorithm identifies only four communities at this threshold, merging two of the true communities. At $prr = 1.0$, all algorithms perfectly recover the true five-community structure.

The constrained Girvan-Newman algorithm (limited to exploring up to 10 community divisions and selecting the highest modularity partition) shows a more binary detection pattern compared to the other methods, which display more gradual transitions as community strength increases.

Figure 11 provides a complementary view showing the ground truth community structure for comparison.

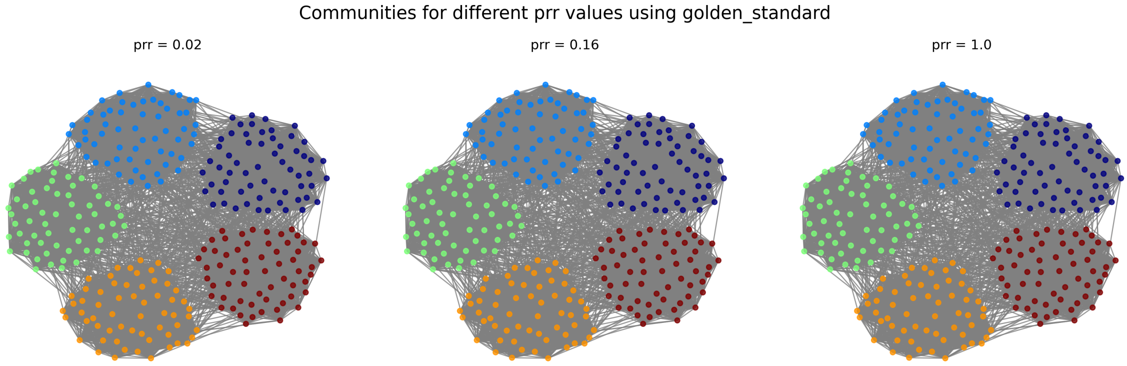


Figure 11: Ground truth community structure of the synthetic networks at different prr values.

3.1.4 Discussion on Modularity and Algorithm Differences

Limitations of Modularity: Although modularity is a widely used metric for community detection, it suffers from the resolution limit. For example, a network with $Q = 0.4$ may or may not have a meaningful community structure depending on network size and density. Our results indicate that even with $Q \approx 0.4$, the underlying block structure might not be well-captured if the intra-block links are sparse or if the algorithm merges small communities.

Differences Across Algorithms: The results obtained by Infomap, Louvain, and Girvan-Newman vary slightly in terms of the number of communities detected and the modularity values. These differences arise because each algorithm optimizes a different quality function or uses different heuristics. For instance, modularity maximization methods can suffer from the resolution limit, while Infomap is more sensitive to information flows. This explains why, for some prr values, the number of communities is not identical, and why the similarity measures (NMI, Jaccard Index) vary across methods.

3.1.5 Discussion on Modularity and Algorithm Differences

Limitations of Modularity: Modularity remains one of the most widely used objective functions for community detection due to its intuitive formulation: communities are identified as subgraphs with higher-than-expected internal edge density. However, its usefulness is hindered by the well-known resolution limit, which causes modularity-based methods to overlook small yet well-separated communities in large networks. In our experiments, even when $Q \approx 0.4$, the actual partitioning quality varied significantly. This is particularly evident at $prr < 0.1$, where Louvain and Greedy produced moderately high modularity values but failed to accurately recover the ground truth, as reflected in low NMI and Jaccard scores. Thus, a high modularity score alone does not guarantee faithful recovery of planted partitions—especially when intra-community densities are low or when community sizes are heterogeneous. This highlights the need to complement modularity with external validation metrics when ground truth is available.

Differences Across Algorithms: The differences in performance across Infomap, Louvain, Greedy, and Girvan-Newman stem from the distinct quality functions and heuristics each algorithm employs. Louvain and Greedy both maximize modularity, but their greedy strategies can converge to different local optima, especially in ambiguous regions of the parameter space. Infomap, in contrast, relies on compressing the description length of a random walk and is sensitive to information flow dynamics rather than edge density. This explains its tendency to fragment the network at low prr and its occasional instability, such as the merging of blocks 2 and 5 at $prr = 0.5$. Girvan-Newman, based on edge betweenness, does not optimize modularity directly and lacks an inherent stopping criterion, which leads to systematic underpartitioning unless externally constrained. The observed variation in NMI, Jaccard Index, and the number of detected communities across algorithms reflects these methodological differences, reinforcing the idea that algorithm choice should be informed by the specific structure and objectives of the analysis.

On the Use of Multiple Algorithms: Given the inherent biases and optimization goals of each algorithm, relying on a single method can offer only a partial view of the community structure. For a robust and comprehensive analysis, it is advisable to apply multiple algorithms and compare their outputs using external validation metrics such as NMI and Jaccard Index, as well as intrinsic indicators like modularity. This multi-algorithmic approach enables the identification of consistent patterns across methods and highlights cases where algorithmic disagreement may point to ambiguous or overlapping community boundaries—features that are often informative in their own right. Especially in exploratory studies or when ground truth is unavailable, triangulating results from different community detection paradigms can lead to more reliable and nuanced insights into network organization.

3.2 Analysis on Real-World Networks

To address the second task, we analyze a real-world dataset: the Primary School Network (PSN)[1]. This analysis serves two main objectives: (1) To evaluate how community detection algorithms perform on real-world data, which may deviate significantly from the structured patterns found in synthetic graphs (e.g., those generated by the Stochastic Block Model). (2) To investigate the effect of edge weights on community detection, where weights represent interaction strength and provide additional semantic value.

We begin by examining the structure of the PSN. Figure 12 shows the unweighted adjacency matrix and corresponding block-summed heatmap, with nodes sorted by school group. This visualization helps reveal the general block structure of the network.

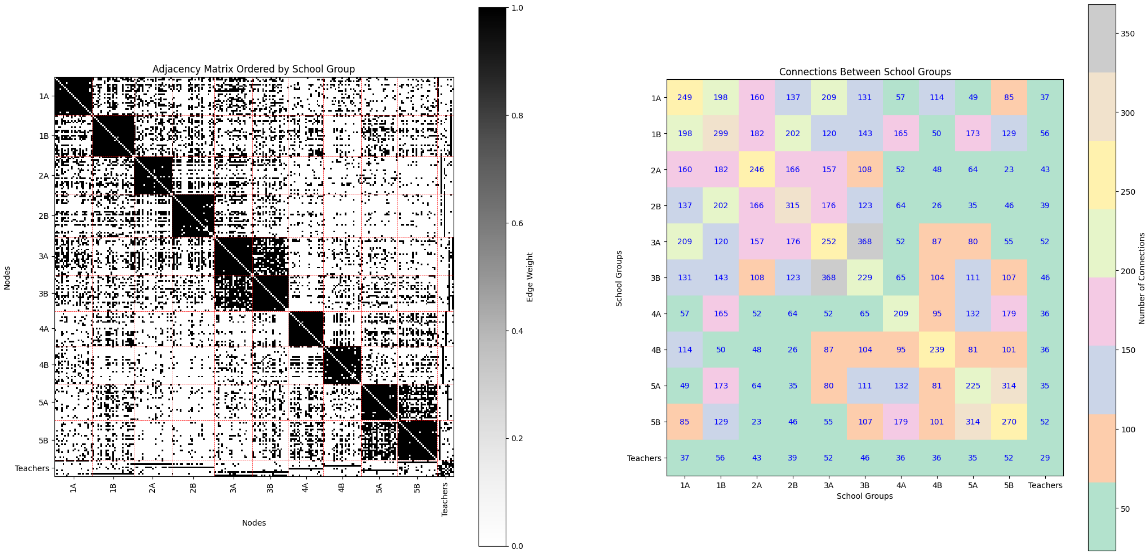


Figure 12: Adjacency matrix and heatmap of the unweighted PSN, with nodes sorted by school group.

Next, we incorporate interaction weights. Figure 13 shows the weighted adjacency matrix and heatmap. The contrast in edge intensities becomes clear: a small number of interactions dominate, while many are relatively weak—an effect visible in the grayscale colormap.

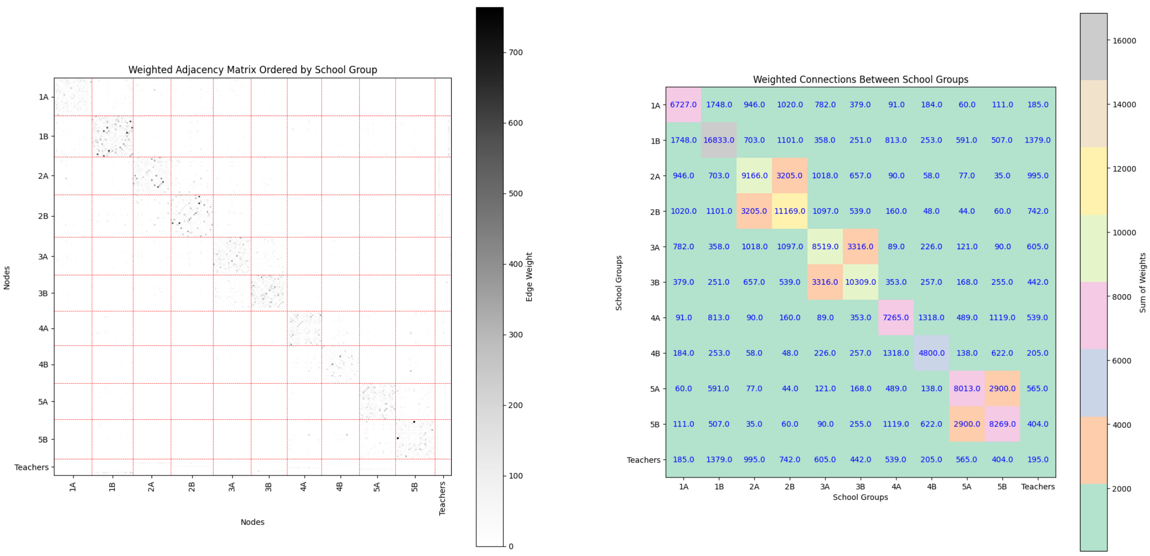


Figure 13: Adjacency matrix and heatmap of the weighted PSN, with nodes sorted by school group.

[1]Primary school contact network

3.2.1 Weighted vs. Unweighted Community Structure

For the PSN, community detection was performed on both the unweighted and weighted versions. Table 2 summarizes the results:

Table 2: Community Detection Metrics for the Primary School Network

Method	Network Version	# Communities	Modularity (Q)	NMI	nVI	Jaccard Index
Infomap	Weighted	8	0.66	0.87	0.10	0.63
Infomap	Unweighted	1	0.00	0.00	0.43	0.09
Louvain	Weighted	6	0.67	0.80	0.15	0.49
Louvain	Unweighted	6	0.28	0.80	0.15	0.46
Greedy Modularity	Weighted	6	0.67	0.81	0.14	0.50
Greedy Modularity	Unweighted	3	0.23	0.43	0.33	0.17

Although not initially prescribed for this task, Infomap was included in the comparison due to its distinct flow-based approach to community detection, which offers an alternative perspective on how information propagates through the network. Its inclusion proves insightful: in the weighted setting, Infomap performs competitively with modularity-based methods, detecting 8 communities with a modularity of 0.66 and the highest NMI (0.87), indicating strong alignment with the reference partition. However, its sensitivity to edge weights is also evident—when weights are removed, it collapses to a trivial single-community solution with zero modularity and negligible overlap with any meaningful structure. This behavior contrasts with Louvain and Greedy Modularity, which, while affected by the loss of weights, still manage to extract some level of community organization.

Figure 14 shows the color-coded layout of the network (with positions fixed from the weighted network using the Kamada-Kawai algorithm with inverse weights). The visual differences in community assignment between the weighted and unweighted versions are apparent, and the metrics in Table 2 support these observations.

When edge weights are included, all three algorithms—Infomap, Louvain, and Greedy Modularity—detect community structures with high modularity ($Q \approx 0.66$–0.67), indicating a clear separation between dense subgraphs. Additionally, high values of NMI and Jaccard Index in the weighted cases show strong agreement with the ground-truth partition, derived from metadata.

In contrast, the unweighted versions yield significantly worse results. Infomap, in particular, fails entirely in the unweighted setting, identifying only a single community (Q = 0.00), with NMI and Jaccard scores near zero. This demonstrates Infomap's sensitivity to the presence of weight information, which it uses to model information flow. Louvain and Greedy Modularity are more robust to the removal of weights, but their performance also drops considerably—modularity and agreement metrics decline, and in the case of Greedy Modularity, the number of communities is halved.

Overall, these findings highlight the critical importance of considering edge weights when analyzing real-world networks. Weighted edges capture interaction intensity—vital for realistic community detection in social systems. Omitting this information leads to oversimplified and less meaningful partitions. Other approaches, such as keeping edges over a threshold probably lead to better results.

3.2.2 Community Composition by School Groups

In this subsection we will review and comment the composition of the communities found by each algorithm. With this purpose we have plotted a heatmap and pie charts that represent the relationship between communities and real school groups. We will go through the compositions we have obtained for each of the algorithms one by one.

Infomap: As observed previously, Infomap performs well only on the weighted graph. In contrast, its performance on the unweighted graph is notably poor: the pie charts (Figure 15) show a single, undivided distribution that mirrors the original class proportions, indicating no meaningful community separation.

In the case of the weighted graph, however, Infomap successfully identifies communities that closely correspond to the original groups—except for a few deviations. Specifically, community 2 merges groups "3A" and "3B", while community 1 combines "4A" and "5B" (Figure 16). The corresponding pie charts (Figure 17) reveal that Teachers are almost uniformly distributed across the eight detected communities, suggesting a limited correlation between community structure and teacher assignment.

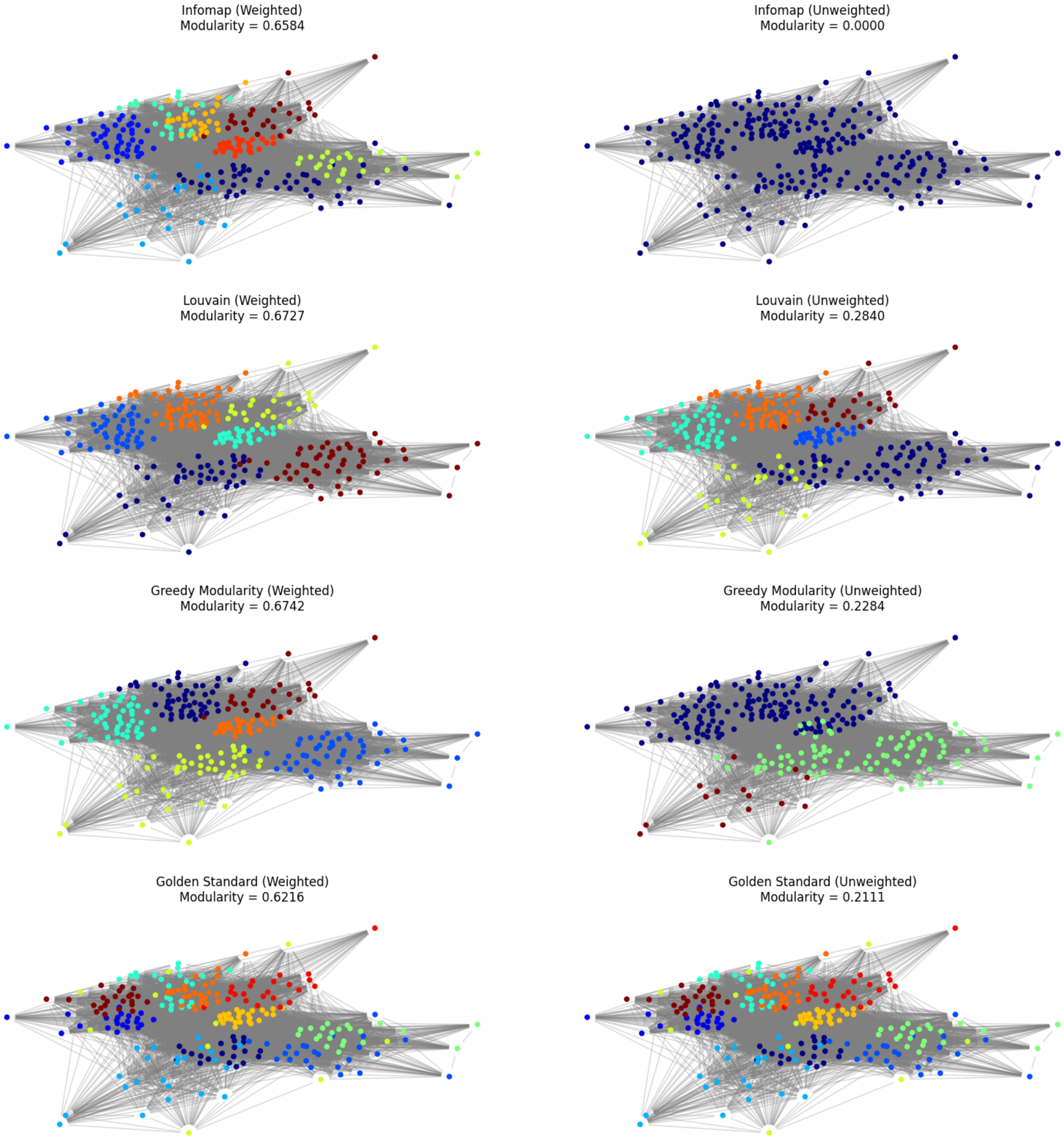


Figure 14: Color-coded layout for the primary school network. Node colors indicate community membership; the layout is fixed from the weighted network. In the ground-truth metadata, yellow nodes represent teachers.

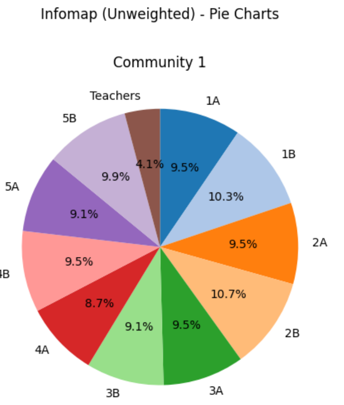


Figure 15: **Infomap (Unweighted)** – Pie charts illustrating the distribution of detected communities by school groups.

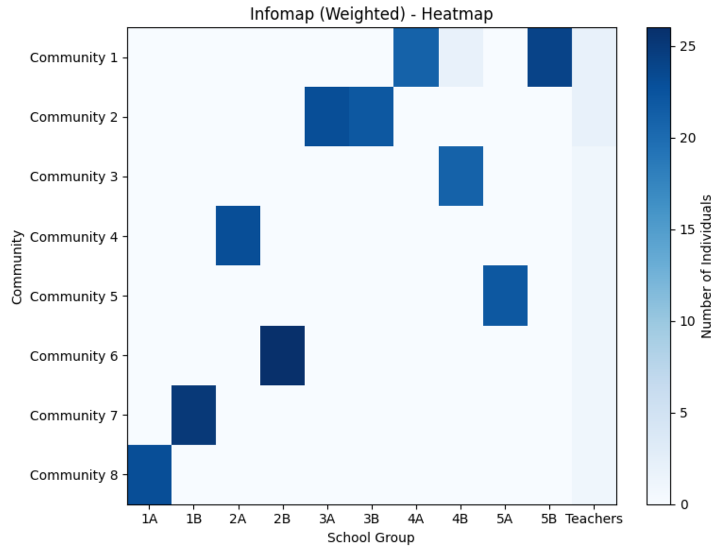


Figure 16: **Infomap (Weighted)** – Heatmap illustrating the distribution of detected communities by school groups.

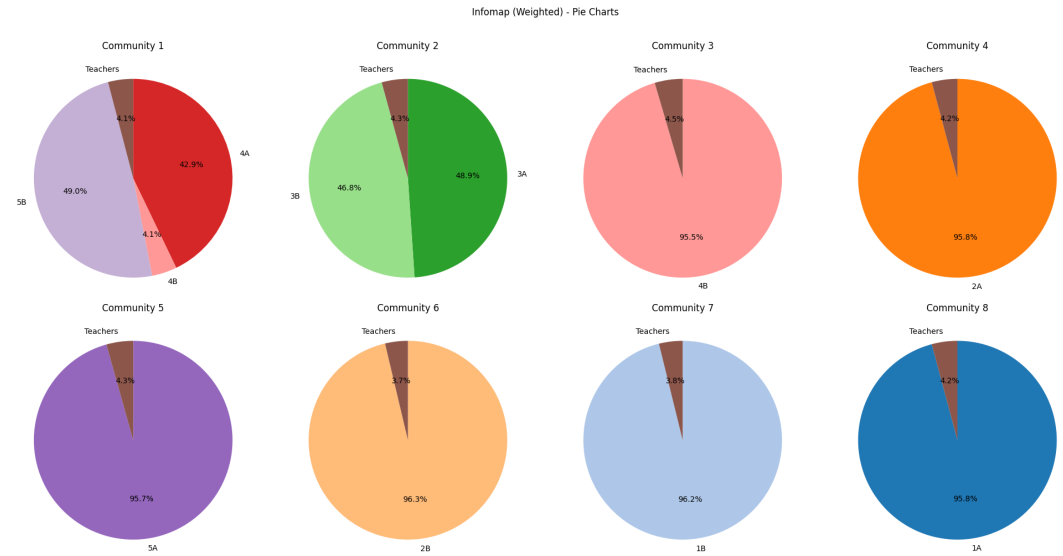


Figure 17: **Infomap (Weighted)** – Pie charts illustrating the distribution of detected communities by school groups.

Agglomerative Greedy Algorithm: The performance of the Greedy algorithm exhibits a clear distinction between the unweighted and weighted versions of the graph. On the unweighted graph, the algorithm yields three broad communities (Figure 18), which partially reflect the original groupings. Community 1 aggregates most lower-numbered school groups (1A–3B), while Community 2 comprises primarily upper groups (4A–5B), and Community 3 isolates group 4B entirely. This is further illustrated in the pie charts (Figure 20), where each community contains a mixture of school groups, although with some internal consistency. Notably, Community 3 is entirely composed of group 4B, suggesting a highly cohesive subgroup.

On the weighted graph, the Greedy algorithm demonstrates markedly improved resolution, identifying six well-defined communities (Figure 19). Each community maps closely to one or two school groups, indicating strong alignment with the underlying structure. For instance, Community 1 groups "2A" and “2B”, Community 3 matches “3A” and “3B”, and Community 4 includes “4A” and “4B”. The corresponding pie charts (Figure 21) confirm this pattern, showing minimal overlap among groups. Teachers are present in all communities but form only a small proportion, reinforcing the idea that the detected communities are primarily driven by student group interactions. The use of weights enables the network to capture more nuanced interactions, resulting in a distinct community for each school year—except for the first year, likely due to the random composition of groups and the persistence of connections among students who were previously classmates.

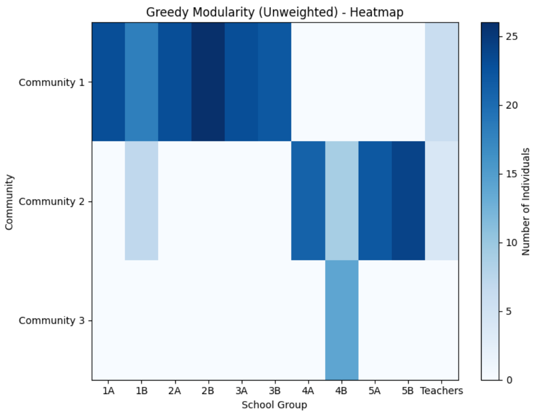


Figure 18: **Greedy (Unweighted)** – Heatmap showing the distribution of detected communities across school groups.

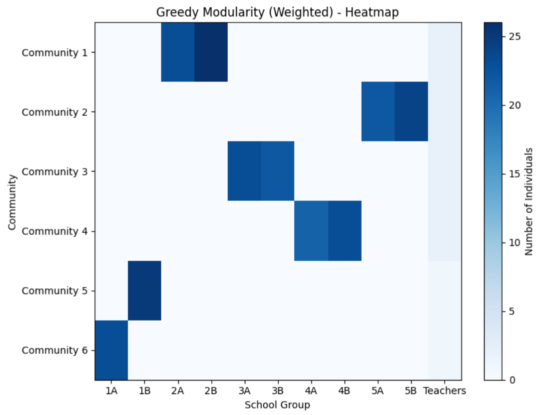


Figure 19: **Greedy (Weighted)** – Heatmap showing the distribution of detected communities across school groups.

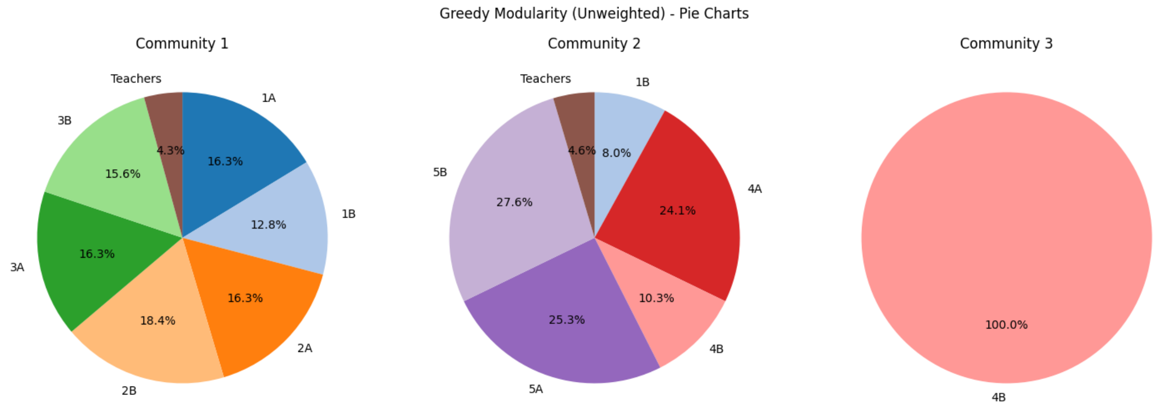


Figure 20: **Greedy (Unweighted)** – Pie charts illustrating the distribution of school groups within each detected community.

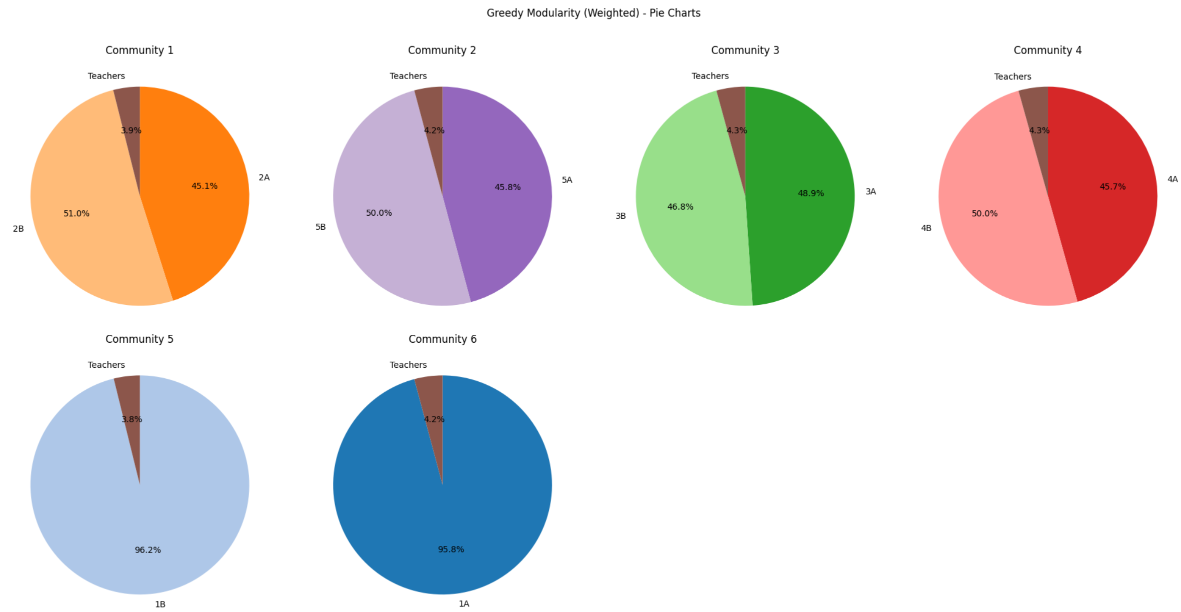


Figure 21: **Greedy (Weighted)** – Pie charts illustrating the distribution of school groups within each detected community.

Louvain: The Louvain algorithm yields consistently high-resolution community structures on both the unweighted and weighted versions of the graph.

On the unweighted graph, Louvain identifies six communities (Figure 22) that largely reflect the original school groupings. Each community appears fairly well-separated, and the corresponding pie charts (Figure 24) confirm that communities are composed of closely related groups. For example, Community 2 combines groups "2A" and "2B," and Community 3 includes "3A" and "3B." Notably, Community 1 and Community 5 consist almost entirely of single groups (1B and 1A, respectively), indicating the algorithm's ability to isolate tightly knit subgroups.

On the weighted graph, Louvain again detects six distinct communities (Figure 23), each strongly aligned with one or two original school groups. As seen in the pie charts (Figure 25), the detected communities display minimal overlap, and the majority of communities are nearly pure in their group composition. For example, Community 1 consists almost entirely of group "1B," and Community 3 is predominantly group "1A." As with previous methods, Teachers are distributed sparsely and evenly, further suggesting that community boundaries are not significantly influenced by teaching assignments. Similar to the Greedy algorithm, the first-year groups are split into distinct communities, likely due to the random formation of groups and the continuation of friendships from earlier years.

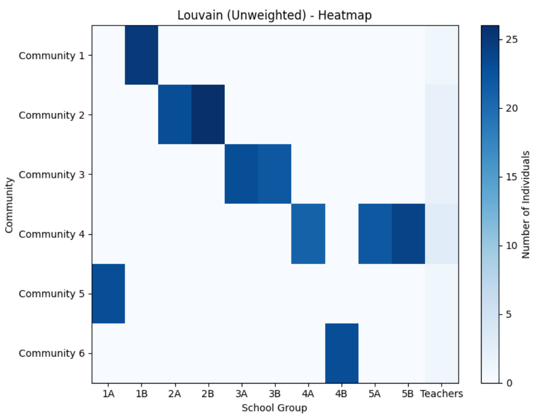


Figure 22: **Louvain (Unweighted)** – Heatmap showing the distribution of detected communities across school groups.

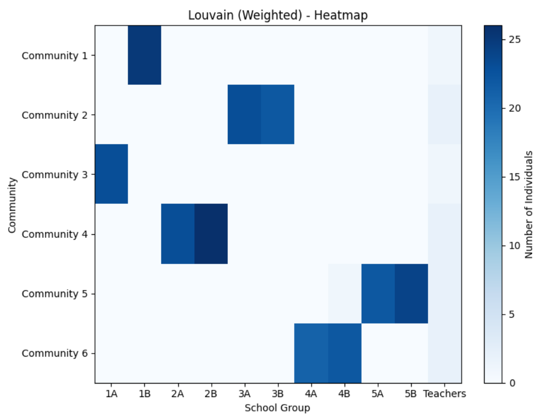


Figure 23: **Louvain (Weighted)** – Heatmap showing the distribution of detected communities across school groups.

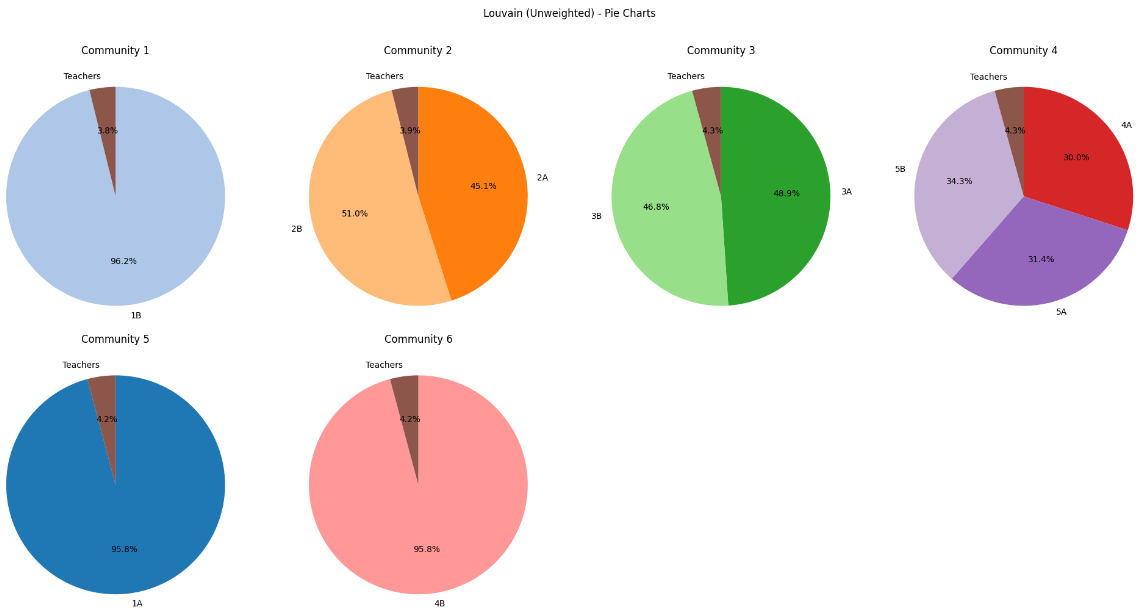


Figure 24: **Louvain (Unweighted)** – Pie charts illustrating the composition of each detected community by school group.

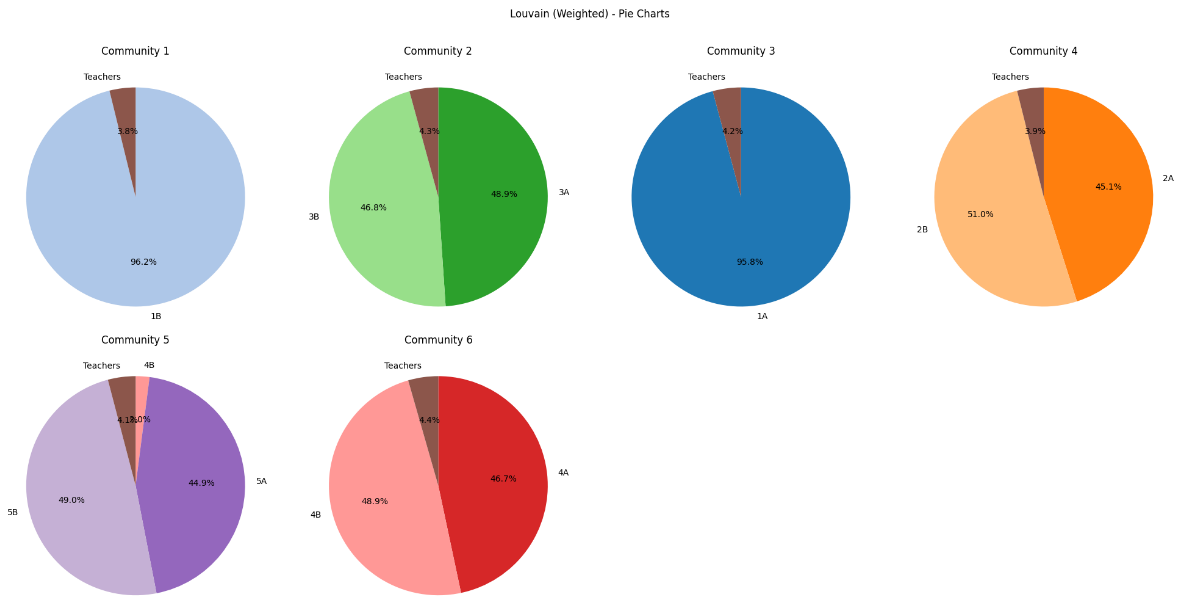


Figure 25: **Louvain (Weighted)** – Pie charts illustrating the composition of each detected community by school group.

3.2.3 Discussion on the Relevance of Weights

The comparative analysis across **Infomap**, **Aglomerative Greedy Algorithm** and **Louvain** consistently demonstrates the critical role of edge weights in revealing meaningful community structure in real-world networks. Weights, in this context, capture the intensity or frequency of interactions—information that is essential for accurately identifying tightly knit groups.

Infomap, in particular, exemplifies this sensitivity. In the weighted setting, it performs exceptionally well, detecting 8 coherent communities and achieving the highest NMI score. However, when weights are removed, it entirely collapses into a single community, reflecting its inherent reliance on weighted flow dynamics for partitioning. This collapse underscores the risk of oversimplification when weights are ignored.

Louvain and Greedy Modularity also exhibit substantial differences between their weighted and unweighted results. In the weighted graph, both algorithms consistently find communities that align closely with the known school groups, with modularity values around 0.67 and high external validation scores. However, in the unweighted graph, their performance deteriorates—modularity drops, community resolution decreases, and the alignment with metadata weakens.

Importantly, while Louvain and Greedy Modularity are more robust than Infomap in the absence of weights, their capacity to capture fine-grained structures still significantly benefits from weighted information. In particular, groups with lower intra-group interaction (e.g., "4B") are more accurately identified only when edge strength is considered in Greedy.

These findings strongly suggest that, for real-world networks such as the primary school contact network, edge weights should not be discarded. They carry essential information that enhances community detection algorithms, yielding more interpretable and structurally consistent partitions. Simplifying a network to its binary form by removing edge weights may lead to the loss of critical insights—especially in settings where tie strength represents underlying relational dynamics.

Overall, incorporating edge weights leads to more granular, accurate, and meaningful community detection across all algorithms analyzed. This highlights the importance of designing or selecting algorithms that leverage weighted information effectively when analyzing complex, real-world interaction networks.

Finally, if an algorithm for unweighted graphs want to be employed keeping only edges above a threshold seems a more sensible answer than keeping all of them and discarding the weights.

4 Conclusions

This study investigated the performance of several community detection algorithms on both synthetic and real-world networks, with a focus on evaluating accuracy, structural insights, and computational efficiency.

In the synthetic setting, results demonstrate that the detectability of communities improves markedly as the intra-block connection probability (prr) increases. A critical threshold around $prr = 0.2$ emerges, beyond which most algorithms, especially Louvain and Infomap, achieve near-perfect recovery of ground-truth communities. At very low connection strengths ($prr = 0.02$), most algorithms try to find communities but end up creating random groupings that don't match the true structure. The constrained Girvan-Newman algorithm takes a different approach, putting all nodes in a single community when connections are this weak. The Girvan-Newman algorithm, though conceptually illustrative, lags in both accuracy and efficiency. Our optimization approach of limiting it to 10 communities and selecting the partition with highest modularity value offered a good compromise between computational tractability and detection quality, though it still remained the most computationally intensive algorithm in our analysis.

In the real-world case study of a primary school contact network, the inclusion of edge weights significantly enhanced the quality of detected communities. Algorithms performed more accurately and produced partitions more consistent with known metadata when weight information was retained. Infomap in particular was highly sensitive to weight presence, collapsing to a trivial solution without them. Louvain and Greedy Modularity showed more robustness, though still benefited substantially from weighted input.

Our results show that using multiple community detection algorithms provides a more complete picture of network structure, as each algorithm has different strengths and weaknesses, especially when community boundaries are not clearly defined.

Overall, our findings highlight the critical role of both algorithm choice and network representation in community detection. When available, incorporating edge weights is essential for uncovering meaningful structure in real-world networks. Moreover, algorithmic assumptions—such as resolution limits or the need for predefined community counts—must be carefully considered in interpreting results.