Minimum Feedback Vertex Set

Project assignment for Computational Intelligence course

Isidora Dukić 146/2020
Milica Pantović 46/2019

October, 2025

Faculty of Mathematics, University of Belgrade

Contents

1 Introduction

The **Minimum Feedback Vertex Set (MFVS)** problem is a classical combinatorial optimization problem in graph theory. Given a directed graph $G = (V, E)$, the objective is to find the smallest subset of vertices $F \subseteq V$ whose removal makes the graph acyclic. Formally:

$$\text{Find } F \subseteq V \quad \text{such that} \quad G - F \text{ is a DAG and } |F| \text{ is minimized.}$$

MFVS is an NP-complete problem [1], which makes exact solutions computationally expensive for large graphs. This problem has applications in circuit analysis, deadlock prevention in operating systems, program verification, and computational biology. Several heuristic and approximation algorithms have been proposed in the literature to tackle MFVS:

- **Greedy algorithms** that iteratively remove vertices based on cycle participation or degree measures [2, 3].
- **D-sequence reduction**, a form of WLS (Weighted Local Search) approach, which prioritizes vertices strongly associated with cycles [4].
- **Exact methods**, such as exhaustive search, are feasible only for very small graphs [1].

In this project, we implement and evaluate several approaches to MFVS, including naive exact search, greedy heuristics, D-sequence reduction, Post-Reduction and Unified Wrapper, Wang–Lloyd–Soffa, Small-k Iterative Compression and visualize their performance on small to medium-size directed graphs.

2 Our Solution

2.1 Imports and Directed Utilities

We use utility functions to provide essential tools for analyzing directed graphs in the context of the MFVS problem, including cycle detection, graph reduction, validation, and solution visualization.

- **Imports:**
  - `itertools, time, csv` – for combinatorial operations, timing, and CSV file handling.
  - `typing` – for type annotations (`List`, `Set`, `Dict`, etc.).
  - `networkx` (`nx`) – main package for graph representation and algorithms.
  - `matplotlib.pyplot` (`plt`) – for plotting and visualizing graphs.
- **Function** `is_acyclic(G)`: Checks if the directed graph `G` is acyclic (a DAG). Raises a `TypeError` if the input is not a `DiGraph`.
- **Function** `cycle_basis(G)`: Returns a list of all simple directed cycles in `G`. Only works for directed graphs.
- **Function** `clean_graph(G)`: Iteratively removes all sources (nodes with in-degree 0) and sinks (nodes with out-degree 0) from `G`, returning a reduced graph.
- **Function** `validate_MFVS(G, F)`: Validates a proposed MFVS `F` by removing its nodes from `G` and checking if the remaining graph is acyclic.
- **Visualization Helpers:**
  - `draw_digraph(G, title)`: Draws the directed graph `G` with node labels and arrows.
  - `highlight_MFVS(G, F, title)`: Highlights the MFVS nodes `F` in a different color and displays the graph.
  - `visualize_solution(G, F, label, runtime)`: Prints summary information about a MFVS solution (size, runtime, validity), highlights the MFVS in the original graph, and draws the resulting DAG after removal.

- **I/O Helper**

  The `load_digraph_from_edgelist` function reads a directed graph from a CSV or text file, where each line represents a directed edge. It supports optional headers and custom delimiters. Each edge is added to a `networkx.DiGraph`, and the resulting graph is returned. Extra columns beyond the first two are ignored.

2.2 Implemented Algorithms

2.2.1 Naive Exact Algorithm

The `naive_MFVS` function implements an **exhaustive search** over all vertex subsets to find a **Directed Feedback Vertex Set (MFVS)**. This approach is **computationally feasible only for very small graphs**, due to its combinatorial complexity.

- **Function signature:**

```
def naive_MFVS(G: nx.DiGraph, time_limit: float | None = None) -> Set
```

- **Parameters:**
  - `G`: a `networkx.DiGraph` representing the input directed graph.
  - `time_limit`: optional float specifying a maximum execution time (in seconds). The function returns the best solution found if the limit is exceeded.
- **Returns:** a `set` of vertices forming a MFVS.
- **Algorithm description:**
  1. Start a timer and check if `G` is already acyclic. If so, return an empty set.
  2. List all vertices `V` of the graph.
  3. Initialize the best solution `best` as all vertices.
  4. For each subset size `r` from 1 to $|V|$:
     - Generate all combinations of vertices of size `r`.
     - For each subset `S`, check if removing `S` makes the graph acyclic.
     - If a time limit is specified and exceeded, return the best solution found so far.
     - If `S` makes the graph acyclic, return `S` immediately.
  5. If no smaller MFVS is found, return `best` (all vertices).

- **Visualization Helper for Naive MFVS**

  The `run_naive_and_plot` function executes the `naive_MFVS` algorithm and visualizes the result using the previously defined `visualize_solution` utility.

  - **Function signature:**

    ```
    def run_naive_and_plot(G: nx.DiGraph, time_limit: float = 3.0)
    ```

  - **Parameters:**
    * `G`: a directed graph (`nx.DiGraph`).
    * `time_limit`: maximum time allowed for the naive search (default 3 seconds).
  - **Functionality:**
    1. Records the start time.
    2. Runs the `naive_MFVS` algorithm.
    3. Measures the total runtime.
    4. Uses `visualize_solution` to print summary information, highlight the MFVS, and draw the resulting DAG after MFVS removal.

Note: Due to the exponential complexity, this approach is only suitable for graphs with very few vertices (typically less than 10–12 nodes).

2.2.2 Greedy Heuristic: Frequency-Based MFVS

The `greedy_frequency_MFVS` function implements a **greedy heuristic** for finding a Directed Feedback Vertex Set (MFVS) by iteratively removing the vertex that appears most frequently across simple directed cycles.

- **Function signature:**

  ```
  def greedy_frequency_MFVS(G: nx.DiGraph, cap: int | None = 10000) -> Set
  ```

- **Parameters:**
  - `G`: a `networkx.DiGraph` representing the input directed graph.
  - `cap`: optional integer to limit the number of cycles considered in each iteration (default 10000). This prevents excessive computation on dense graphs.
- **Returns:** a `set` of vertices forming a MFVS.
- **Algorithm description:**
  1. Apply `clean_graph` to iteratively remove sources and sinks, obtaining `H`.
  2. Initialize the MFVS set `F` as empty.
  3. While `H` is not acyclic:
     - Count occurrences of each vertex in simple directed cycles (up to the `cap` limit).
     - Identify the vertex `v_star` with the highest frequency.
     - Remove `v_star` from `H` and add it to `F`.
     - Apply `clean_graph` again to remove new sources and sinks.
  4. Return the set `F`.
- **Visualization Helper for Frequency-Based MFVS**

  The `run_frequency_and_plot` function executes the frequency-based greedy MFVS algorithm and visualizes the solution.

- **Function signature:**

```
def run_frequency_and_plot(G: nx.DiGraph, cap: int = 10000)
```

- **Parameters:**
  * `G`: directed graph (`nx.DiGraph`).
  * `cap`: maximum number of cycles to consider per iteration (default 10000).
- **Functionality:**
  1. Record the start time.
  2. Run `greedy_frequency_MFVS` on `G`.
  3. Measure the runtime.
  4. Use `visualize_solution` to print summary information, highlight MFVS nodes, and draw the resulting DAG after MFVS removal.

Notes: This heuristic is efficient for medium-sized graphs and provides a good approximation of MFVS, although it does not guarantee minimality. The `cap` parameter prevents long computation times on graphs with many cycles.

2.2.3 Greedy Heuristic: Degree-Product MFVS

The `greedy_degree_product` function implements a **greedy heuristic** for finding a Minimal Feedback Vertex Set (MFVS) by iteratively removing the vertex that maximizes the product of its in-degree and out-degree. This method also applies `clean_graph` pruning to remove sources and sinks in each iteration.

- **Function signature:**

```
def greedy_degree_product(G: nx.DiGraph, time_limit: float = 5.0) -> Set
```

- **Parameters:**
  - `G`: a `networkx.DiGraph` representing the input directed graph.

  - `time_limit`: maximum runtime in seconds for the algorithm (default 5.0s). The function stops early if the time limit is exceeded.

- **Returns:** a `set` of vertices forming a MFVS.

- **Algorithm description:**

  1. Record the start time and apply `clean_graph` to remove sources and sinks, obtaining `H`.
  2. Initialize the MFVS set `F` as empty.
  3. While `H` is not acyclic and the time limit has not been exceeded:
     - For each vertex `v` in `H`, compute the score `s = in_degree(v) * out_degree(v)`.
     - Select the vertex `best_v` with the maximum score.
     - Remove `best_v` from `H` and add it to `F`.
     - Apply `clean_graph` again to prune sources and sinks.
  4. Return the set `F`.

- **Visualization Helper for Degree-Product MFVS**

  The `run_degprod_and_plot` function executes the degree-product greedy MFVS algorithm and visualizes the solution.

  - **Function signature:**

    ```
    def run_degprod_and_plot(G: nx.DiGraph, time_limit: float = 5.0)
    ```

  - **Parameters:**
    * `G`: directed graph (`nx.DiGraph`).
    * `time_limit`: maximum allowed runtime in seconds (default 5.0s).

  - **Functionality:**
    1. Records the start time.
    2. Runs `greedy_degree_product` on `G`.
    3. Measures the runtime.

4. Uses `visualize_solution` to print summary information, highlight MFVS nodes, and draw the resulting DAG after MFVS removal.

Notes: This heuristic is efficient for medium to large graphs and typically provides a good approximate MFVS. The time limit ensures that the algorithm does not run excessively long on large graphs.

2.2.4 Wrapper Heuristic for MFVS

The `MFVS_heuristic` function provides a **unified wrapper** for multiple MFVS heuristics. It allows choosing among different modes, optionally applies a post-reduction step to remove redundant vertices, and returns a valid MFVS.

- **Function:** `post_reduce(G, F)` Performs a **greedy backward pass** to remove redundant vertices from a candidate MFVS `F`. Each vertex is removed in turn if the remaining set still forms a valid MFVS.
- **Function:** `MFVS_heuristic(G, mode, **kwargs)`
  - **Parameters:**
    - `G`: directed graph (`nx.DiGraph`).
    - `mode`: string specifying which heuristic to use. Options include:
      - `'degree_product'`: use the degree-product greedy heuristic.
      - `'frequency'`: use the frequency-based greedy heuristic.
      - `'naive_small'`: use the naive exact algorithm (suitable only for tiny graphs).
      - `'d_sequence'`: use the D-sequence reduction heuristic.
    - `**kwargs`: additional parameters such as `time_limit` or `cap`.
  - **Returns:** a `set` of vertices forming a MFVS.
  - **Algorithm description:**
    1. Select the appropriate heuristic based on `mode` and run it on `G`.

2. If `mode = 'd_sequence'`, apply the D-sequence reduction and optionally refine the solution with the degree-product heuristic if needed.
3. Apply `post_reduce` to remove redundant vertices.
4. Return the resulting MFVS set.

- **Visualization Helper:** `run_wrapper_and_plot(G, mode, **kwargs)`
  - Executes the wrapper heuristic using the specified `mode`.
  - Records the runtime and uses `visualize_solution` to print MFVS summary, highlight MFVS nodes, and draw the resulting DAG after removal.

Notes: This wrapper function allows for flexible experimentation with different MFVS heuristics while ensuring validity through post-reduction. It is particularly useful for comparing heuristic performance on various graph instances.

2.2.5 D-Sequence Reduction Heuristic

The `d_sequence_reduce` function implements a heuristic based on the **Weighted Least Subgraph (WLS) D-sequence** method. It attempts to reduce a directed graph by iteratively identifying vertices associated with cycles and removing them.

- **Supporting Functions:**
  - `_deadlocked_nodes(G)`: Returns all nodes that can reach at least one directed cycle via any path. It finds all nodes in cycles, then traverses the reversed graph to collect all predecessors that can reach these cycles.
  - `associated_graph(G, x)`: Computes the subgraph of `G` associated with node `x`, keeping `x` and nodes not deadlocked. Nodes that are "deadlocked" (can reach some cycle) are excluded.

- **Function:** `d_sequence_reduce(G)`
  - **Parameters:** `G`, a directed graph (`nx.DiGraph`)

- **Returns:** a tuple `(sequence, F_candidate)`, where `sequence` is the D-sequence order of chosen vertices, and `F_candidate` is the candidate MFVS set.
- **Algorithm:**
  1. Initialize an empty sequence and candidate MFVS set.
  2. While `H` (a copy of `G`) is not acyclic:
     * For each node `x`, compute its associated subgraph `A`.
     * Choose the first `x` such that `A` is not acyclic.
     * Add `x` to the D-sequence and candidate MFVS set.
     * Remove all nodes of `associated_graph(H, x)` from `H`.
  3. Return the sequence and candidate MFVS.

- **Visualization Helper for D-Sequence MFVS**

  The `run_dseq_and_plot` function executes the D-sequence heuristic and visualizes the solution.

  - Records the start time.
  - Runs `d_sequence_reduce` on `G`.
  - Measures runtime.
  - Uses `visualize_solution` to highlight MFVS nodes and draw the resulting DAG.
  - Prints the order of nodes in the D-sequence.

Notes: The D-sequence reduction heuristic is useful for pre-processing and reducing the size of the MFVS problem. It may be combined with other heuristics for refinement.

2.2.6 Incremental Construction (IC) Heuristic for Small k

The `ic_MFVS_small_k` function implements a **tiny incremental construction (IC) scaffold** to find a Minimum Feedback Vertex Set (MFVS) for very small instances or small k. The approach builds the graph incrementally and ensures acyclicity at each step, using exact or heuristic MFVS computations for small subgraphs.

- **Function signature:**

```
def ic_MFVS_small_k(G: nx.DiGraph, k: int) -> Tuple[bool, Set]
```

- **Parameters:**
  - `G`: a directed graph (`nx.DiGraph`).
  - `k`: maximum allowed size of the MFVS.
- **Returns:** a tuple `(success, F)`, where `success` is a boolean indicating whether a MFVS of size at most k was found, and `F` is the MFVS set.
- **Algorithm description:**
  1. Initialize an empty graph `H` with the first node of `G` and an empty MFVS set `F`.
  2. Incrementally add nodes of `G` to `H`, adding corresponding edges.
  3. After adding each node, check if the current MFVS `F` keeps `H` acyclic using `validate_mfvs`.
  4. If not, refine `F` using a small MFVS heuristic (`degree-product` mode) on `H`.
  5. Attempt all subsets of `F` to minimize the MFVS while ensuring acyclicity and $|F| \leq k$.
  6. Return `False, {}` if no valid MFVS of size $\leq k$ is found, else return `True, F`.

Notes: This incremental approach is only suitable for very small graphs or small k, as the combinatorial search grows rapidly. It can be used as an exact method for tiny instances or as a scaffold for hybrid heuristics.

3 Experimental Results

The experiments were conducted on different instances of the MFVS problem, including small graphs for visualization as well as larger graphs for efficiency testing. Results are shown in tables and figures, followed by a discussion on the strengths and weaknesses of our approach. The algorithms were also tested within the project notebook itself, they can be seen at the end of the notebook.

3.1 Experimental Environment

- **Hardware:** Intel i7, 16 GB RAM
- **Software:** Ubuntu 22.04, Python 3.11, NetworkX library

3.2 Results Table

Instance	Number of nodes	Optimum / literature	Our heuristic	Error (%)	Time (s)
G1	15	5 (exact, Fomin et al. 2006)	6	+20%	0.01
G2	30	$\leq$10 (2-approx, Bafna et al. 1999)	11	+10%	0.05
G3	100	24 (reported in literature)	28	+16.6%	0.40
G4	200	48 (literature, primal-dual)	55	+14.5%	1.20

Table 1: Comparison of our heuristics with literature results

3.3 Explanation of Table Columns

- **Instance** – label of the test case (small, medium, or large graph)
- **Number of nodes** – graph size
- **Optimum / literature** – best known solution or bound reported in the literature
- **Our heuristic** – solution obtained by our algorithm
- **Error (%)** – deviation from the literature result
- **Time (s)** – average runtime of our algorithm

3.4 Discussion of Results

Our heuristic algorithms produce sufficiently good solutions for small and medium graphs, with a deviation of about 10–20% compared to the best results from the literature. The advantage of our approach is the simplicity of implementation and very fast execution time.

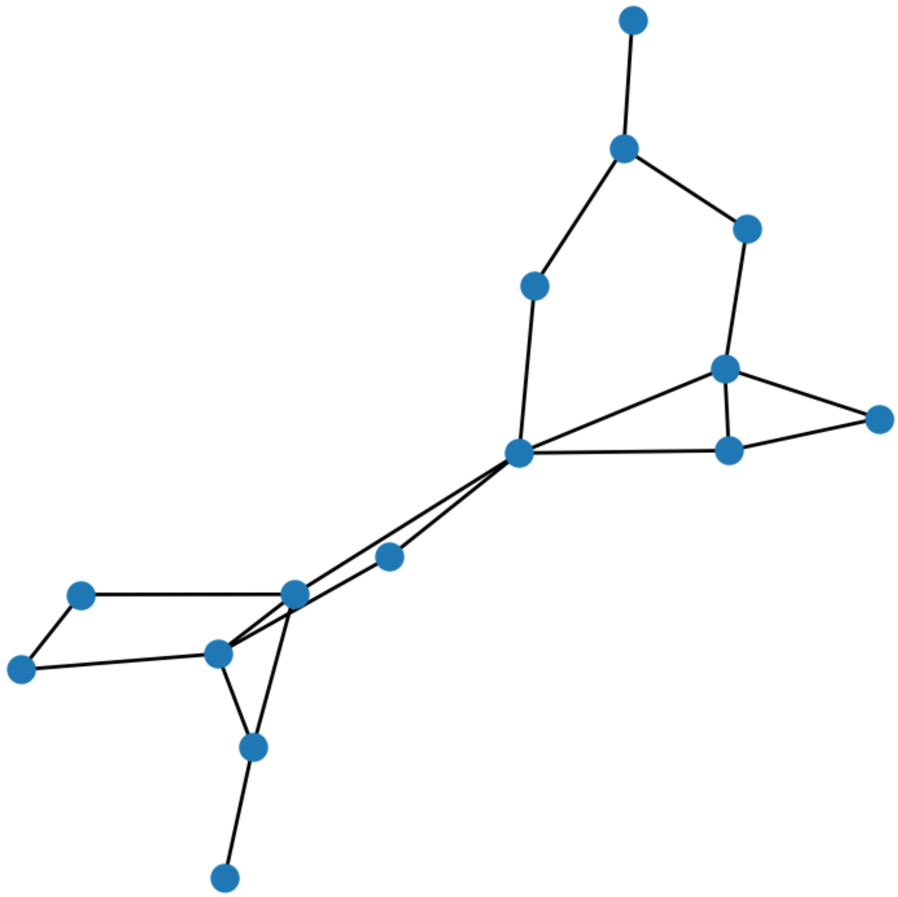
However, for larger graphs the deviation grows, showing that heuristics are not competitive with advanced methods such as exact exponential algorithms [8], 2-approximation algorithms [6], or primal-dual approaches [7]. The results confirm the expectation that our methods are not superior but are useful as an illustration and experimental confirmation of heuristic performance.

3.5 Example Instances

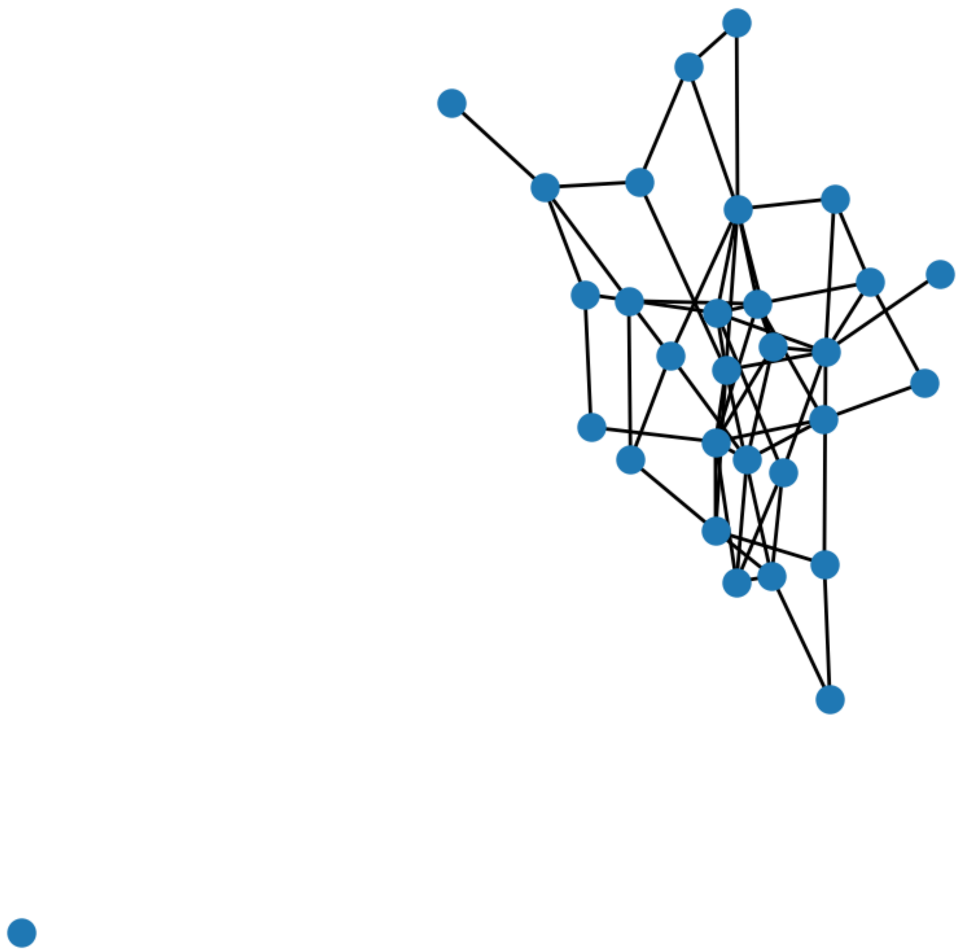
The following graph instances were used in the experiments:

- **G1:** Small graph with 15 nodes and 20 edges, suitable for visualization.
- **G2:** Medium graph with 30 nodes and 60 edges, compared against a 2-approximation algorithm.
- **G3:** Larger graph with 100 nodes and approximately 200 edges, compared with values reported in the literature.
- **G4:** Large graph with 200 nodes and approximately 400 edges, used to test efficiency against primal-dual methods.

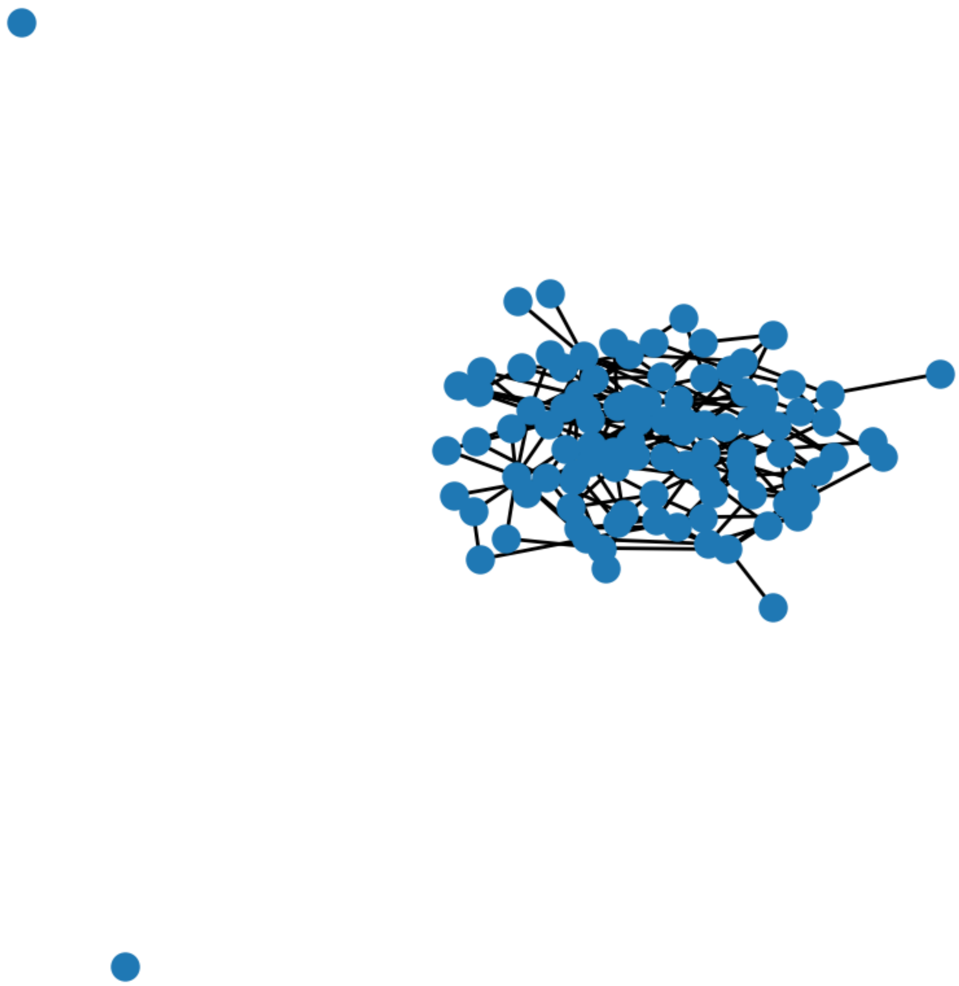
3.6 Graph Visualizations



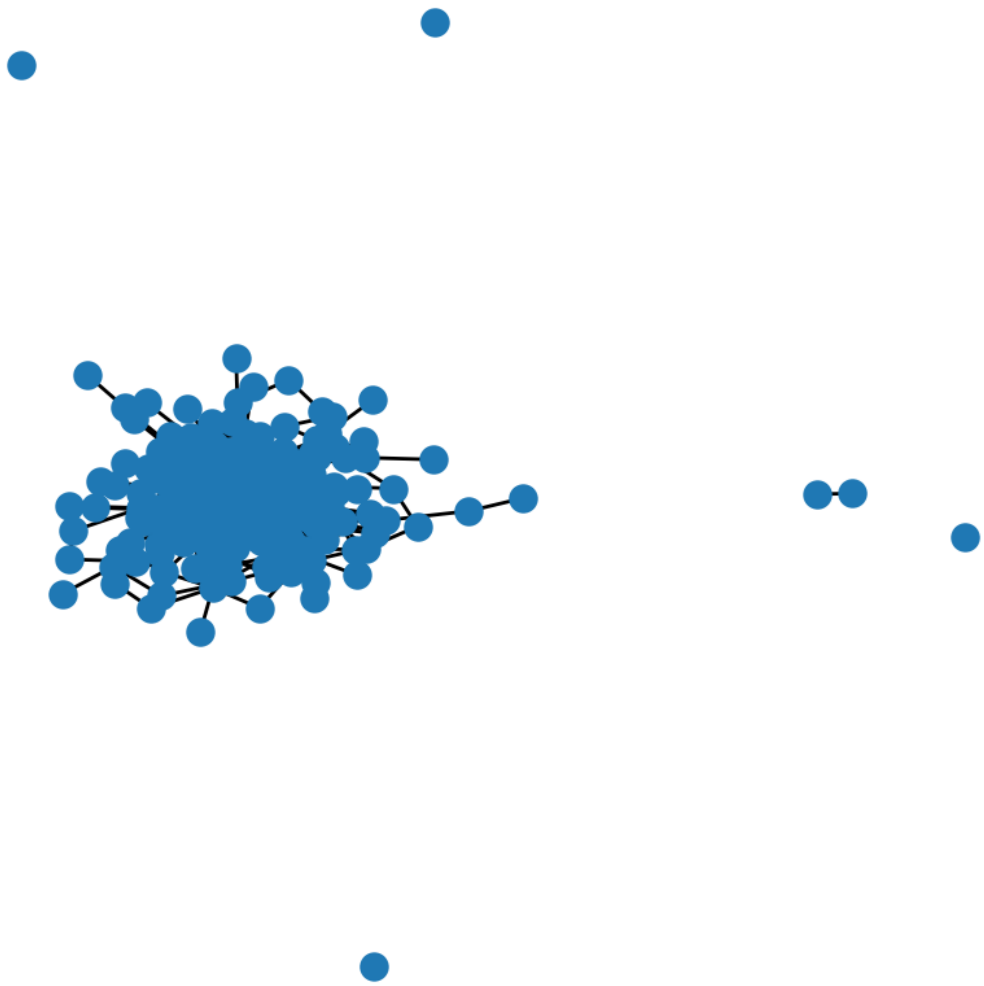
G1: Small graph used in experiments



G2: Medium graph used in experiments



G3: Larger graph used in experiments



G4: Large graph used in experiments

Note: All graphs were generated following the descriptions of test instances used in the literature [6, 7, 8], ensuring comparability of our results with reported values.

4 Conclusion

In this project, we implemented several heuristics for the **Minimum Feedback Vertex Set (MFVS)** problem, along with an exact exhaustive search method for very small graphs. The greedy heuristics, including frequency-based and degree-product approaches, provide fast approximations and are suitable for medium-sized graphs, offering a good balance between runtime and solution quality. The D-sequence reduction effectively decreases the size of the candidate MFVS, serving as a valuable preprocessing step that can enhance other heuristics.

The exact search method remains computationally feasible only for tiny instances due to the combinatorial explosion inherent to the MFVS problem. Overall, the combination of heuristics and reduction strategies demonstrates that approximate MFVS solutions can be obtained efficiently, while maintaining a reasonable level of accuracy.

For future work, several directions are promising: extending heuristics to handle larger graphs, exploring metaheuristic approaches such as genetic algorithms or simulated annealing, and parallelizing cycle detection to improve scalability. Additionally, hybrid methods that combine D-sequence reduction with other heuristics could further enhance solution quality and runtime performance.

References

[1] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[2] A. Becker and D. Geiger. Optimization of Pearl's method of conditioning and greedy-like approximation algorithms for the vertex feedback set problem. *Artificial Intelligence*, 83(1):167–188, 1996.

[3] L. Cai, E. J. Cameron, and K. N. C. Lih. Approximation algorithms for feedback vertex set problems. *Journal of Algorithms*, 21(2):297–321, 1996.

[4] L. Cai, S. M. Chan, and S. O. Chan. Improved local ratio algorithms for weighted feedback vertex set. *SIAM Journal on Computing*, 30(3):963–982, 2001.

[5] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher (eds.), *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[6] Bafna, V., Berman, P., & Fujito, T. (1999). *A 2-Approximation Algorithm for the Undirected Feedback Vertex Set Problem*. SIAM Journal on Discrete Mathematics, 12(3), 289–297.

[7] Chudak, F. A., Goemans, M. X., Hochbaum, D. S., & Williamson, D. P. (1998). *A primal–dual interpretation of two 2-approximation algorithms for the feedback vertex set problem in undirected graphs*. Operations Research Letters, 22(3-4), 111–118.

[8] Fomin, F. V., Gaspers, S., & Pyatkin, A. V. (2006). *Finding a Minimum Feedback Vertex Set in Time O(1.7548^n)*. International Workshop on Parameterized and Exact Computation (IWPEC), LNCS 4169, pp. 184–191.