**Network-Based Separation and Mean Shortest Distance Calculation**

This repository provides Python scripts to calculate network-based distance and network separation between two sets of drug target genes on a given protein–protein interaction (PPI) network.

It includes automated scripts to run calculations across multiple pairs of gene sets.

## Repository Structure

* separation.py – Core script that calculates the mean shortest distances and network separation between two drug target sets.
* automated\_separation.py – Automates the analysis across all drug combinations, running separation.py for each pair automatically.
* interactome.tsv– Protein–protein interaction (PPI) network file in tab-separated format (gene1, gene2).
* \*.txt –
* output/ – Folder where results for all pairwise calculations are stored.

## Requirements

* Python 2.7 or 3.x
* Libraries:
  + networkx
  + numpy

Install required libraries via pip if needed:

pip install networkx numpy

Usage

1. Calculate Separation for a Single Drug Pair

**python separation.py --g1 DrugA.txt --g2 DrugB.txt -o output.txt**

**Parameters:**

--g1 : File containing the first drug’s target proteins (one per line).

--g2 : File containing the second drug’s target proteins.

-n : (Optional) Network file in tab-separated edge list format (default: interactome.tsv).

-o : Output file to save results (default: separation\_results.txt).

**Output includes:**

* Mean shortest distance within each drug target set (**d\_A**, **d\_B**).
* Mean shortest distance between the two sets (**d\_AB**).
* Network separation (**s\_AB**) between the two drugs.

### 2. Automate separation calculations for multiple pairs

Use the automated script to calculate separation for all possible **drug–drug combinations** automatically:

python automated\_separation.py

**Features:**

* Scans the current folder for all .txt drug target files.
* Calculates pairwise separation for all possible drug pairs.
* Saves the output files in the **output/** directory with names such as DrugA+DrugB.txt.

## Notes

* This approach quantifies how **close or distant** two drugs’ targets are within the PPI network.
* Negative separation values (**s\_AB < 0**) indicate that two drugs target **proximal regions** of the interactome, suggesting potential **mechanistic or synergistic overlap**.
* Positive separation values (**s\_AB > 0**) indicate **distinct or non-overlapping target regions**.
* The automated version is suitable for large-scale pairwise analysis (e.g., screening drug–drug combinations).