

Community Detection in Bacterial and Viral Networks

Jonas Ziegler

February 1, 2026

Abstract

This paper investigates community structure in bacterial and viral protein similarity networks using modularity-based community detection methods. We compare several algorithms and perform a sensitivity analysis on the resolution parameter using Adjusted Rand Index and Adjusted Mutual Information.

1 Introduction

- 1.1 Motivation**
- 1.2 Definition of a Network**
- 1.3 Mathematical Representation**
- 1.4 What Can We Do with Networks?**

2 Network Summary Statistics

- 2.1 Basic Metrics**
- 2.2 Degree Distribution**
- 2.3 Centrality Measures**

3 Similarity Measures for Community Comparison

- 3.1 Rand Index**
- 3.2 Adjusted Rand Index**
- 3.3 Adjusted Mutual Information**
- 3.4 Comparison of ARI and AMI**

4 Community Detection Methods

- 4.1 Modularity**
- 4.2 Edge Betweenness (Girvan–Newman)**
- 4.3 Louvain Method**
- 4.4 Leiden Method**

5 Data Description

- 5.1 Data Source**
- 5.2 Network Structure**
- 5.3 Preprocessing**

6 Experimental Setup

- 6.1 Implementation**
- 6.2 Hyperparameters**
- 6.3 Evaluation Criteria**

7 Results

- 7.1 Descriptive Comparison of Bac and Vir**
- 7.2 Method Comparison**
- 7.3 Hyperparameter Sensitivity Analysis**
- 7.4 Interpretation of Results**

8 Discussion

9 Conclusion