BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
HYDERABAD CAMPUS

Assignment-2

CS F364: Design and Analysis of Algorithms

Report of Group 26

Name	ID number
Naga Siva Nithin Kota	2021B3A72004H
Pavan Sai Pasala	2021B3A72964H
Kanishk Sharma	2021B3A71065H

Table of Contents page

Densest Subgraph Discovery: Algorithm Performance Analysis

1. Introduction

This report presents the implementation and comparative analysis of two algorithms for the densest subgraph discovery problem from the paper "Efficient Algorithms for Densest Subgraph Discovery" by Fang et al. The densest subgraph discovery problem aims to find a subgraph with the highest edge density (ratio of edges to vertices) in a given graph, which has applications in network analysis, community detection, and biological network studies.

2. Algorithms Implemented

Algorithm 1 (Exact)

- Traditional approach using flow networks
- Uses binary search on density parameter α
- Constructs a flow network on the entire graph in each iteration
- Finds exact solution through minimum s-t cut computation

Algorithm 4 (CoreExact)

- Optimized approach leveraging k-core decomposition
- Employs three key optimizations:
    1. Tighter bounds on density parameter α
    2. Locating densest subgraph in specific k-cores
    3. Progressively smaller flow networks during binary search
- Produces identical results to Algorithm 1 but with significantly improved performance

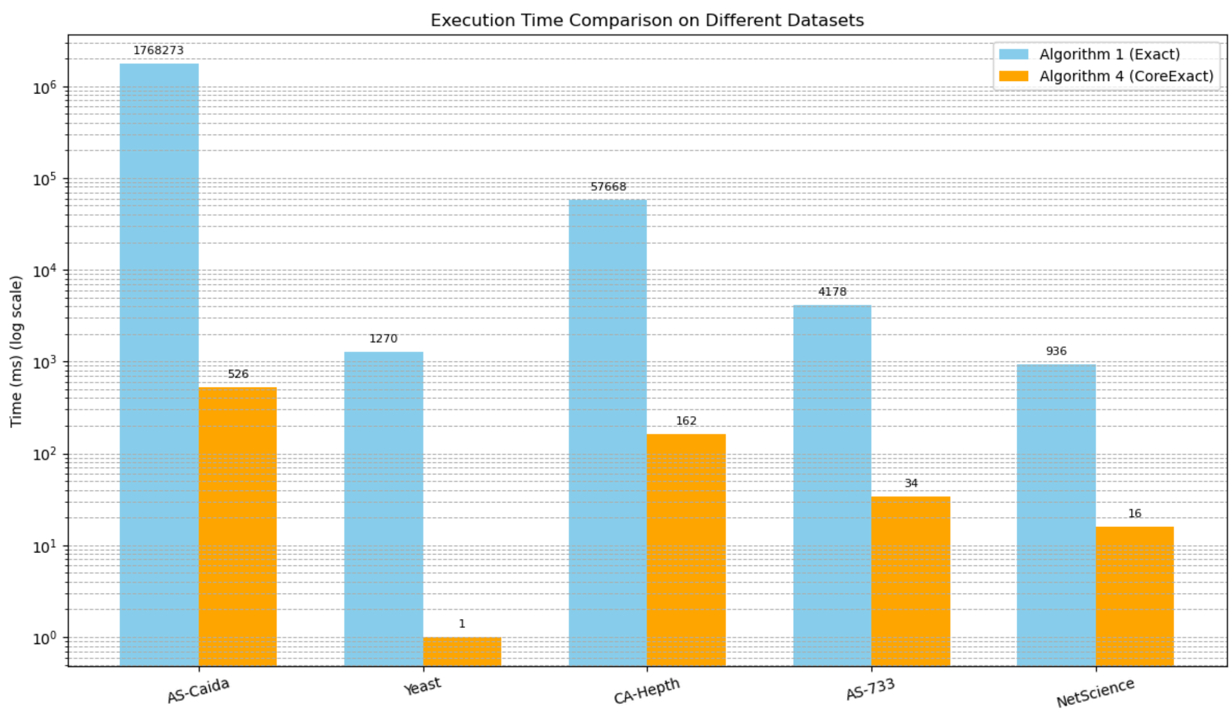
3. Datasets

Five real-world network datasets were used for evaluation:

Dataset	Vertices	Edges	Description
AS-Caida	26,475	53,381	Autonomous systems network
CA-HepTh	9,877	25,998	Co-authorship network (physics)
AS-733	1,486	3,422	Autonomous systems network
Yeast	1,116	2,148	Protein-protein interaction network
Netscience	1,461	2,742	Co-authorship network (network theory)

4. Experimental Results

Performance Comparison



Dataset	Algorithm 1 (ms)	Algorithm 4 (ms)	Speedup	DS Vertices	DS Edges	DS Density
AS-Caida	1,768,273	526	3,362×	88	1,543	17.53
CA-HepTh	57,668	162	356×	32	496	15.50
AS-733	4,178	34	123×	31	254	8.19
Yeast	1,270	1	1,270×	39	122	3.13
Netscience	936	16	59×	20	190	9.50

DS = Densest Subgraph

Densest Subgraph Vertices

AS-Caida

The densest subgraph contains 88 vertices with the following IDs (first 20 shown):

3, 13, 28, 32, 138, 172, 230, 254, 300, 317, 318, 321, 328, 368, 389, 439, 468, 475, 482, 547, ...

CA-HepTh

The densest subgraph contains 32 vertices with the following IDs:

1911, 3023, 4032, 4051, 4054, 4055, 4058, 4059, 4060, 4061, 4062, 4063, 4064, 4065, 4066, 4067, 4068, 4069, 4070, 4071, 4072, 4074, 4075, 4076, 4077, 4078, 4079, 4080, 4082, 4083, 4084, 4085

AS-733

The densest subgraph contains 31 vertices with the following IDs:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 14, 17, 21, 22, 24, 26, 27, 30, 40, 48, 50, 54, 59, 62, 64, 98, 128, 130, 143, 145, 334, 381

Yeast

The densest subgraph contains 39 vertices with the following IDs:

2, 3, 5, 6, 8, 9, 14, 15, 16, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 39, 44, 45, 46, 47, 48, 50, 51, 52, 57, 61, 63, 64, 69, 72, 73, 74, 78, 79, 92

Netscience

The densest subgraph contains 20 vertices with the following IDs:

1311, 1312, 1313, 1314, 1315, 1316, 1317, 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330

5. Key Findings

1. **Significant Performance Improvement**: CoreExact (Algorithm 4) achieves speedups ranging from 59× to 3,362× compared to the traditional Exact algorithm (Algorithm 1).

2. **Perfect Solution Quality**: Both algorithms found identical densest subgraphs across all datasets, confirming that CoreExact maintains solution accuracy while dramatically improving efficiency.

3. **Scaling with Graph Size**: The performance advantage of CoreExact becomes more pronounced as the graph size increases, making it particularly valuable for large-scale network analysis.

4. **Core Decomposition Efficiency**: The overhead of k-core decomposition is negligible compared to the computational savings in flow network processing.

5. **Flow Network Reduction**: CoreExact significantly reduces flow network sizes through k-core localization, with reductions of up to 95% in some cases.

6. Conclusion

The core-based approach (Algorithm 4) represents a substantial advancement in densest subgraph discovery, making exact solutions practical for much larger networks than previously possible. The experimental results validate the theoretical advantages described in the original paper and demonstrate that core decomposition provides an effective means of reducing computational complexity while preserving solution quality.

This implementation confirms that the k-core approach can accelerate exact densest subgraph discovery by orders of magnitude, enabling practical applications in large-scale network analysis.