

Contents

Contents	vii
List of Figures	ix
List of Tables	x
1 Introduction	1
1.1 Problem formulation	1
1.2 Protein-Protein interaction network	1
1.3 Gene Ontology	3
1.4 Motivation	4
1.5 Objectives	5
1.6 Contribution	5
1.7 Organization	5
2 Literature review	6
2.1 Existing approaches	6
2.1.1 Graph based clustering approach	6
2.1.2 Gene ontology based clustering approach	8
2.2 Discussion	9
3 Proposed method	10
3.1 Symbol table	10
3.2 Definitions	10
3.3 Flowchart	11
3.4 Algorithm	12
3.5 Effect of input parameter p	13
4 Experimental results	14
4.1 GO semantic similarity score	14

4.2	Co-localization score	15
4.3	Precision, Recall and F-measure	16
5	Conclusion and Future work	19
5.1	Conclusion	19
5.2	Future work	19
	Bibliography	20

List of Figures

1.1	PPI network	2
1.2	Structure of Gene Ontology	3
3.1	Flowchart	11
3.2	F-measure comparison with respect to p using Gavin dataset	13
3.3	F-measure comparison with respect to p using Krogan_core dataset . . .	13
3.4	F-measure comparison with respect to p using Collins dataset	13
4.1	GO semantic similarity score comparison using Gavin dataset	15
4.2	Co-localization score comparison using Gavin dataset	16
4.3	F-measure comparison of various methods using Collins dataset	18
4.4	F-measure comparison of various methods using Krogan_core dataset . .	18

List of Tables

2.1 Comparison of the existing approaches	8
3.1 Symbol table	10
4.1 GO semantic similarity score comparison using Gavin dataset	14
4.2 Co-localization score comparison using Gavin dataset	15
4.3 Comparison using Gavin dataset	17
4.4 Comparison using krogan_core dataset	17
4.5 Comparison using Collins dataset	17
4.6 F-measure comparison using Collins dataset	17
4.7 F-measure comparison using Krogan_core dataset	17