**Thesis Outline**

**Chapter 1: Introduction**

1.1 Introduction of proteins and domains

1.2 Need for domain identification

1.3 Literature survey

1.3.1 Sequence based methods

1.3.2 Structure based methods

1.3.3 Manual & Semi-automatic methods

1.4 Organization of thesis

**Chapter 2: Materials & Methods**

2.1 Introduction

2.2 Clustering

2.2.1 Hierarchical clustering

2.2.2 Agglomerative clustering

2.2.3 Divisive clustering

2.3 The k-means clustering algorithm

2.4 Domain identification by graph theoretic approach

2.4.1 Newman’s modularity method

2.4.2 Vishveshwara’s graph spectral analysis

2.4.3 Hari Krishna’s method

2.5 Identification of number of domains by using physical properties of a protein

2.6 K-means clustering algorithm for predicting domain boundaries

2.7 Implementation details

2.7.1 Construction of data set

2.7.2 Libraries & tools

**Chapter 3: Results & Discussion**

3.1 Comparative analysis of graph theoretic approach with clustering

**Chapter 4: Conclusion**