Graphs represent relationships. Some relationships can be represented as a deterministic graph while others can only be represented by using probabilities. Mining structures from graphs help us to find useful patterns in these relationships, which can be then applied towards solving important real world problems. Dense substructures are an important pattern to mine in graphs. It has applications in wide areas like social network analysis, bioinformatics etc. With the advent of "big data", real world graphs have become massive. However, finding dense substructures in massive graphs is an open question. In this thesis, we attempt to address some of the problems in the area of dense substructure enumeration in large deterministic or uncertain graphs.

In particular we look at the problems of Maximal Clique Enumeration (MCE) and Maximal Biclique Enumeration (MBE) from a large graph. Both these problems are a central task to many data mining problems arising in social network analysis and bioinformatics. We present novel parallel algorithms for MBE on top of the MapReduce framework, and an experimental evaluation using Hadoop MapReduce. Our algorithm is based on clustering the input graph into smaller subgraphs, followed by processing different subgraphs in parallel. Our algorithm reduces redundant work and increases load balance thus enabling it to scale to large graphs. We show theoretically that our algorithm is work optimal i.e. it performs the same total work as its sequential counterpart. We present a detailed evaluation which shows that the algorithm scales to large graphs with millions of edges and tens of millions of maximal bicliques. To our knowledge, this is the first work on maximal biclique enumeration for graphs of this scale. We also provide an experimental evaluation of algorithms for the MCE problem.

Finally we consider the problem is MCE on an Uncertain Graph, which is a probability distribution on a set of deterministic graphs. For parameter $0 < \alpha < 1$, we consider the notion of an α-maximal clique in an uncertain graph. We present matching upper and lower bounds on the number of such α- maximal cliques possible within a (uncertain) graph. We present an algorithm to enumerate α-maximal cliques whose worst-case runtime is near-optimal, and an experimental evaluation showing the practical utility of the algorithm.