Paper choose:

G. Karypis, E.-H. Han, and V. Kumar. CHAMELEON: A Hierarchical Clustering Algorithm Using Dynamic Modeling. COMPUTER, 32(8): 68-75, 1999.

Problem 1:

Problems:

1. Existing hierarchical schemes do not make use of information about the nature of individual clusters being merged.

2. Schemes ignore the important information such as the aggregate interconnectivity of items in two clusters, or whereas the closeness of two clusters.

Solutions:

1. Model the data items as a graph, represent data using a k-nearest neighbor graph Gk.

2. Dynamically model the similarity between the clusters by computing their relative inter-connectivity and relative closeness.

3. Use graph partitioning to obtain the initial fine-grain clustering solution.

4. Use the relative inter-connectivity and relative closeness to repeatedly combine together the sub-clusters in a hierarchical fashion.

Problem2:

Pros and Cons:

1. Handle any type of data since it is not specific to a particular domain

2. The complexity of the algorithm O(nm + n log n + m^2log m)

3. Different domains require different measures for connectivity and closeness, etc.