TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

CENTRAL CAMPUS, PULCHOWK

THESIS NO.: 070MSCS659

Approximating k-center problem using Hierarchical Clustering

By

Narendra Maharjan

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN COMPUTER

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

LALITPUR, NEPAL

NOVEMBER 2015

Approximating k-center problem using Hierarchical Clustering

By

Narendra Maharjan

(070/MSCS/659)

Thesis Supervisor:

Dr. Subarna Shakya

Professor

A thesis submitted in partial fulfillment of the requirement for
the degree of Master of Science in Computer System and Knowledge Engineering.

Department of Electronics and Computer Engineering

Institute of Engineering, Central Campus, Pulchowk

Tribhuvan University

Lalitpur, Nepal

November 2015

COPYRIGHT

TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS
DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

Ananda Niketan, Pulchowk, Lalitpur, P.O. Box 1175, Kathmandu, Nepal.
Tel : 5534070, 5521260 extn. 315, Fax : 977-1-5553946, E-mail : doece@ioe.edu.np

Our Ref :

DEPARTMENTAL ACCEPTANCE

The thesis entitled “Approximating k-center problem using Hierarchical Clustering”, submitted by Narendra Maharjan in partial fulfillment of the requirement for the award of the degree of “Master of Science in Computer System and Knowledge Engineering” has been accepted as a bona fide record of work independently carried out by him in the department.

Dr. Dibakar Raj Pant
Head of the Department

TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS
DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

Ananda Niketan, Pulchowk, Lalitpur, P.O. Box 1175, Kathmandu, Nepal.
Tel : 5534070, 5521260 extn. 315, Fax : 977-1-5553946, E-mail : doece@ioe.edu.np

Our Ref :

CERTIFICATE OF APPROVAL

The undersigned certify that they have read, and recommended to the Institute of Engineering for acceptance, a thesis report entitled "Approximating k-center problem using Hierarchical Clustering" submitted by Mr. Narendra Maharjan in partial fulfillment of the requirement for the degree of Master of Science in Computer System and Knowledge Engineering.

____

Supervisor, Dr. Subarna Shakya
Professor
Department of Electronics and Computer Engineering

____

External Examiner, Mr. Om Bikram Thapa
Chief Technical Officer
Vianet Communications Pvt. Ltd.

____

Committee Chairperson, Dr. Shashidhar Ram Joshi
Professor
Department of Electronics and Computer Engineering

DATE OF APPROVAL: November 6, 2015

ACKNOWLEDGEMENT

I would like to express my sincere thanks to Department of Electronics and Computer Engineering (DOECE) for providing me the opportunity to explore my interest and ideas in the field of engineering through this Thesis.

I owe a debt of gratitude towards my Thesis supervisor, **Prof. Dr. Subarna Shakya**, for his proper guidance, encouragement and support.

Finally, I would like to thank all our teachers and friends who have helped me directly or indirectly with this thesis.

Narendra Maharjan

ABSTRACT

The K-center problem is a classical problem in facility location: given n cities and the distances between them, we wish to select k of these cities as centers so that the maximum distance of a city from its closest center is minimized. The problem is NP-hard problem. For this thesis the vertices or node must be in a metric space. In this thesis agglomerative hierarchical clustering with complete linkage is used which divides the graph of nodes into set of clusters. For each node within the cluster the maximum distance to its center will be minimum. Finding cluster center is another task named as 1-center problem which states as the smallest circle that contains all of a given set of points i.e. the maximum radius circle to include all nodes will be minimum. In this thesis k-center problem is solved in metric space graph first by clustering nodes of graph and then finding k center within each cluster. Finally the result is compared with existing methods for k center problem and 1-center problem.

Keywords

K-center, 1-center, Agglomerative Hierarchical Clustering

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
INTRODUCTION

1. Introduction

1.1. Background

The K-center problem is a classical problem in facility location: given n cities and the distances between them, we wish to select k of these cities as centers so that the maximum distance of a city from its closest center is minimized. The problem is NP-hard and Hochbaum and Shmoys present a 2-approximation algorithm for graphs with edge weights obeying triangle inequality [1]. Suppose that we wish to locate facilities at k out of n cities such that the maximum distance of a city to its closest facility is minimized. That closest facility is considers as center for k out of n cities. The facility location problem is a well-studied problem in the operation research literature and recently has received a lot of attention in the computer science community. For a company, the facility location problem provides more strategic decisions than just giving importance to locate the lowest cost space for storing it's products. While identifying the best location for the company's distribution center, it must consider several things like the cost of new structure, new freight, inherent risk and the viability involved in choice of these facilities. Several variations of facility location problem can be formulated depending on the constraints on the nature of the facilities, and objective functions, e.g. the cost reduction, demand capture, fast response time etc. For locating the emergency facilities, such as hospitals, fire-fighting stations etc. covering a region using minimum radius circles is a natural mapping of the corresponding facility location problem where the objective is to minimize the worst-case response time. The placement of multiple facilities (or centers) in such a way that the each vertex can cover maximum number of demand nodes and the distance of all centers should be as minimum as possible, it is one of the biggest issues for researchers. This problem is known as k-center problem [2].

In graph theory this means finding a set of k-vertices for which the largest distance of any point to its closest vertex in the k-set is minimum. The vertices must be in a metric space, or in other words a complete graph that satisfies the triangle inequality. An instance of the k-center problem consists of a complete graph G = (V,E) with edge weights $w_e \geq 0$, $e \in E$ and w(u,v) = 0, $v \in V$. The problem is to find a subset $S \subseteq V$ of size at most k such that $w(S) = \max_{i \in v} \min S_{j \in s} W_{(i,j)}$ is minimized.

Hierarchical Clustering of n data points is a recursive partitioning of the data into 2,3,4... and finally k, clusters. Each intermediate clustering is made more fine-grained by dividing one of its clusters [3]. Rather than partitioning the data from single initial cluster, here agglomerative Methods is implemented where each objects is taken as cluster at starting and merging of similar cluster to one is carried out in each step finally forming one single cluster of desired number of k clusters. It is the bottom up approach starting with a separate cluster for each point and then progressively merging the two 'closest' clusters until only a single cluster remains. The different schemes are distinguished by their notion of closeness single-linkage clustering, complete-linkage clustering, Average-linkage Clustering among them Complete-linkage clustering will be implemented in this thesis, it is the distance between their farthest pair of points (thus complete linkage is explicitly trying to minimize the diameter) [4]. So complete linkage helps to group or cluster nodes together minimizing the bounding radius of that cluster, which can be further illustrated by comparing these three cost factors in this thesis.

Clustering gives clusters, but center has to be calculated explicitly. Finding center in each cluster is called 1-center problem which is classical combinatorial optimization problem in operational research: given a set of n demand points, a space of feasible locations of a facility and a function to calculate the transportation cost between a facility and any demand point, find a location of the facility which minimizes the maximum facility-demand point transportation cost [5].

1.2. Problem Statement

The K-center problem is about finding k centers in metric space so that the maximum distance for each objects to its nearest center is minimized. It is NP hard problem. So Approximation algorithm is implemented. In this thesis, using hierarchical clustering and solving 1-center problem, k centers are approximated.

1.3. Objectives

i. To implement Hierarchical clustering with complete linkage
ii. To approximate K centers in metric space graph.
iii. To compare with existing method for k-center problem and 1-center problem.

1.4. Scope

In this thesis k centers are approximated in metric space minimizing the maximum distance from n object to its nearest k center. There are various applications of this thesis mainly in facility location problem. Placement of Wi-Fi hotspot within geographical region to optimize coverage area for Wi-Fi or wireless zones.

Chapter 2
LITERATURE REVIEW

2. Literature Review

2.1. K-center problem

Imagine you have a delivery service. You want to place your k delivery hubsat locations that minimize the maximum distance between customers and their nearest hubs. This is the k-center problem. Formally, given a complete graph on n vertices with nonnegative (but possibly infinite) edge costs, and a positive integer k, the k-center problem is to find a set of k vertices, called centers, minimizing the maximum distance to any vertex and from its nearest center. In k-center problem given n cities and distances between all pairs of cities, the aim is to choose k cities called centers so that the largest distance of any city to its nearest center is minimal. The k-center problem is NP-hard. Without the triangle inequality the problem is NP-hard to approximate within any factor. Hence the edge costs assumed to satisfy the triangle inequality [6].

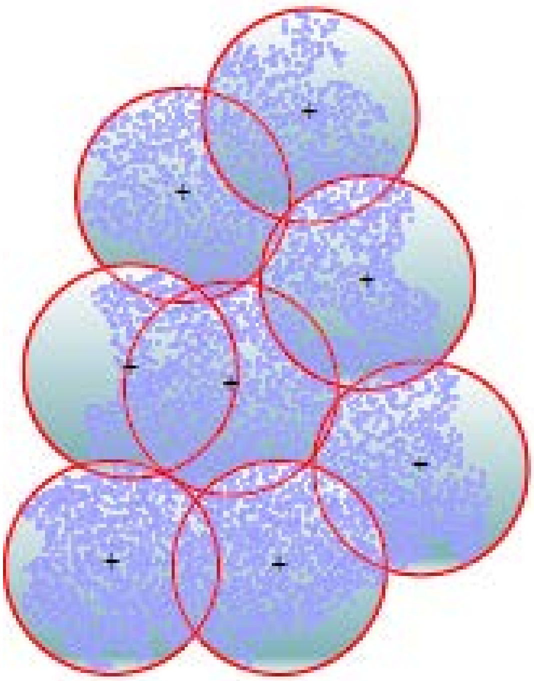

Figure 2.1-1 : K-center diagram

2.2. Hierarchical clustering

Hierarchical clustering creates a hierarchy of clusters, which may represent in a tree structure called a dendrogram. The root of the tree consists of a single cluster containing all observations, and the leaves correspond to individual observations. Algorithms for hierarchical clustering are generally either agglomerative, in which one starts at the leaves and successively merges clusters together; or divisive, in which one starts at the root and recursively splits the clusters. Basic idea is closer nodes lies on same cluster. In this thesis Agglomerative clustering is used. Agglomerative clustering starts with N clusters, each of which includes exactly one data point. A series of merge operations then followed, that eventually forces all objects into the same group [7]. In this thesis for agglomerative hierarchical clustering complete linkage is taken as the cost or measure of similarity. Incomplete linkage, the similarity between two clusters is the minimum similarity between points in different clusters thus complete linkage is explicitly trying to minimize the diameter.

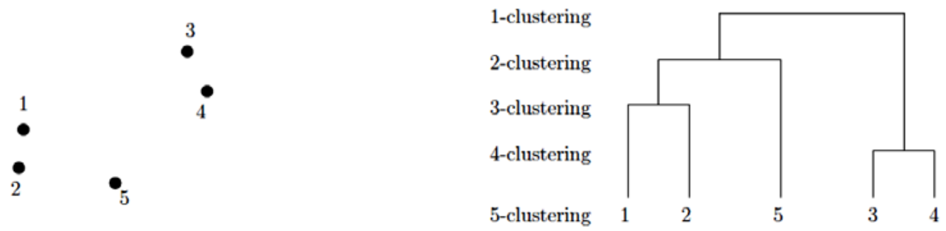

Figure 2.2-1 : 5 node hierarchical clustering with dendrogram

2.3. 1-center problem

In n-dimensional space, the 1-center problem is to compute the smallest n-sphere that contains all of a given set of points. It is also called closing circle problem. Given a set of n demand points, a space of feasible locations of a facility and a function to calculate the transportation cost between a facility and any demand point, find a location of the facility which minimizes the maximum facility-demand point transportation cost.1-center problem is defined in the plane as follows. Given are n points (x1, y1), . . . , (xn, yn) so as to minimize [8]

max{((x-xi)2+(y-yi)2)1/2}

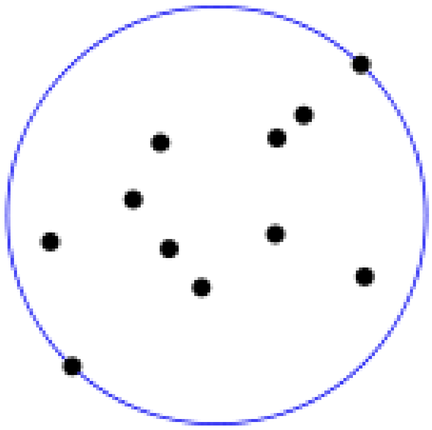

Figure 2.3-1 : 1-center problem diagram

2.4. Related works

Several approximation algorithm have been proposed in past. One of the simple and good algorithm due to Gonzalez called farthest first traversal it has approximation factor of 2. It starts with any point, and then iteratively adds in the point furthest from the ones chosen so far. Farthest first traversal takes time O(kn), which is fairly efficient. Its solution might not be perfect, but is always close to optimal [9]. The decision version of k-center problem is described by Shmoys [10] with approximation factor 2. The problem is defined as that we have given an area of radius r. The algorithm for the optimization version of problem runs the algorithm for the decision version several times with increasing value of r. Time complexity of this algorithm is $O(kn^3)$. Hochbaum and Shmoys present a 2-approximation algorithm for graphs with edge weights obeying triangle inequality [11]. This algorithm works in different fashion. The aim is to find a minimum dominating set in the pruned graph, i.e. the smallest set S of vertices such that every vertex not in S is adjacent to one of the vertices in S. Others are algorithms for facility location problems with outliers which omits the distance nodes to make facility more efficient and distance more optimized [12]. To make k center clustering robust and less erroneous to noise, Robust Hierarchical k-center clustering algorithm has been recently proposed by google researchers [3].

Chapter 3
RESEARCH METHODOLOGY

3. Research Methodology

3.1. Concept

The k-center problem is to find a set of k vertices, called centers, minimizing the maximum distance to any vertex and from its nearest center. So the first concept here is to isolate k number of groups of nodes which are closed to each other. Clustering of nodes can be used to isolate the groups of nodes. In this Thesis Agglomerative Hierarchical Clustering with complete linkage as cost factor is used. In complete linkage, the distances between clusters are determined by the greatest distance between any two objects in the different clusters (i.e., by the "furthest neighbors"). Thus this clustering results the clusters with minimum greatest distance between nodes. Then For each clusters finding center and radius of bounding circle is another step which is known as one center problem. So the second concept is bounding circle can be estimated using only four extreme nodes in 2D plane. Among which three or two nodes can be taken using condition described in algorithm. The bounding circle for that cluster will be the circumscribing circle of triangle in case of three nodes or the circle with center at midpoint in case of two nodes. The center should be one of the node so node nearer to calculated center will be considered as center and radius is maximum distance from center to nodes within the cluster.

3.2. Algorithm

Algorithm includes two steps.

First step is clustering of nodes (vertices in graph). Clustering is done using Agglomerative Hierarchical Clustering with complete linkage as cost factor.

1. Start with n clusters containing single node.
2. Find and merge most similar clusters i.e. clusters with maximum distance between nodes with each other is least or minimum.
3. Update clusters i.e. number of cluster reduced by one in each step.
4. Repeat step 2 and 3 until number of cluster is equal to k.

From this step clusters nodes are separated but k center has to be calculated explicitly.

Next step is solving 1-center problem (smallest circle problem). For each clusters, the center of circle with minimum radius enclosing all nodes within it has to be calculated.

Solving 1-center problem method is based on extreme nodes.

P contains n nodes of cluster

1. Find two nodes (let's say A and B) from P with maximum distance between them add to T.
2. while(|T|<=4)

   add node from p with maximum square root of, sum of square of distances with nodes of T. (let's say C and D nodes are added)
3. Calculate angles $Angle_{ACB}$ and $Angle_{ADB}$

   If ($Angle_{ACB}$>=90 and $Angle_{ACB}$>=90)

   Center C=midpoint of AB

   Else if ($Angle_{ACB}$< 90)

   Center C = circumscribing center(A,C,B)

   Else if ($Angle_{ADB}$< 90)

   Center C = circumscribing center(A,D,B)

   Else

   Center C = circumscribing center(C,D,X)

   Where X=A if A is farther from both D and C (here sum distance is used)
   = B else
4. node nearest to C will be k center for that cluster.
5. R=max(distance(C,p_i))

Repeat for all the clusters.

One of the existing greedy algorithm called farthest first method is implemented in this thesis to compare the result obtained for k-center problem. This is the existing greedy algorithm for computing k centers in k center problem. This algorithm includes following steps.

1. Read n nodes.
2. Pick C = {x}, for an arbitrary point x
3. Repeat until C has k centers:
4. Let y maximize d(y, C), where

d(y, C) = minx in C d(x, y)

C = C U {y}

For comparing result obtained for 1-center problem two existing methods are implemented in this thesis. One of which is bouncing bubble method, this algorithm includes following steps:

1. Read n nodes
2. Set arbitrary center=node(x) and radius=0.00001f
3. For each nodes

   Difference= node(x)-center

   Distance =distance(node(x),center)

   If(dist>radius)

      New radius=(old radius+distance)/2

      New center =old center+((distance-radius)/distance*difference)

Another existing method used for 1-center problem is Jack Ritter's method. This algorithm is simple and most commonly used one. This algorithm includes following steps.

1. Pick a point x from P, search a point y in P, which has the largest distance from x;

2. Search a point z in P, which has the largest distance from y. set up an initial ball B, with its center as the midpoint of y and z, the radius as half of the distance between y and z;

3. if all points in P are within ball B, then we get a bounding sphere. Otherwise, let p be the point outside the ball, construct a new ball covering both point p and previous ball. Repeat this step until all points are covered.

3.3. Development tools and environment

For the development of this thesis Netbeans IDE 8.0.2 is used as a platform. I have used Java as programming language for the coding in this thesis. Any external ApI is not used in this thesis.

Chapter 4
RESULT, ANALYSIS AND COMPARISONS

4. Result, Analysis and Comparisons

From the result obtained, as shown in Table 4.1, it was observed radius for k centers using hierarchical clustering and solving 1-center problem is minimum than that of farthest first method. It was seen that value of radius from output is dependent on input vertices and value of k or number of centers. It was observed that for less value of k, radius using hierarchical clustering is minimum than that of farthest first method. But for higher value of k result is not better and in some case value of radius is higher than that of farthest first method. These are due to local optimum in hierarchical clustering. But better results are obtained in many cases. Time required for higher number of nodes is seems quite slow for hierarchical clustering method than that of farthest first method. It is because time complexity is quite high for hierarchical clustering. It will be discussed in time complexity section. But the overall performance of algorithm is relatively good and yields somehow optimum result.

Table 4.1 : Output for different number of point and k

S no.	Number of Points	K=3		K=7		K=9		K=10	
		Radius using hierarchical clustering method	Radius using farthest first method	Radius using hierarchical clustering method	Radius using farthest first method	Radius using hierarchical clustering method	Radius using farthest first method	Radius using hierarchical clustering method	Radius using farthest first method
1.	43	124	125	76	76	63	67	63	52
2.	154	180	238	115	126	115	112	99	92
3.	207	207	258	145	179	122	140	120	125
4.	877	391	425	256	321	238	267	229	260

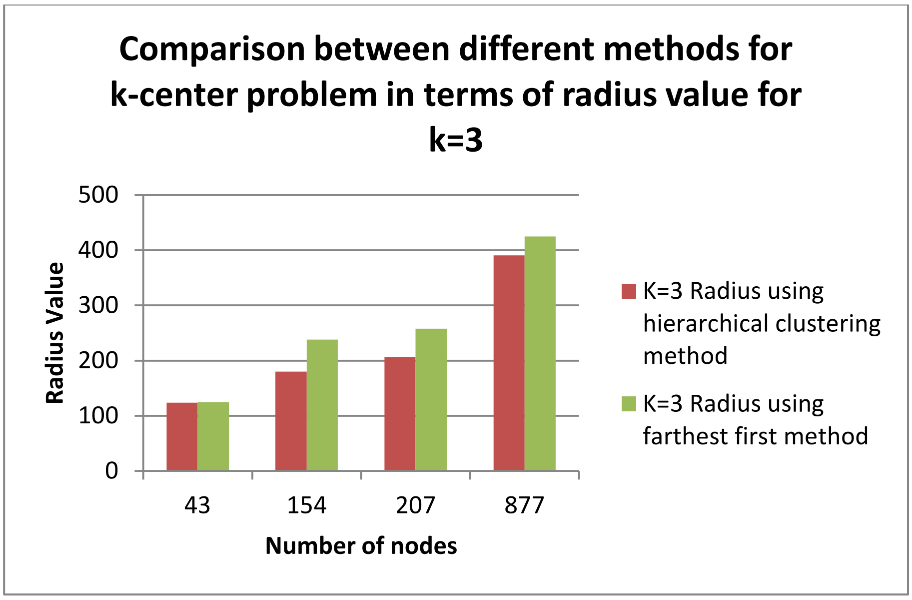

Figure 3.3-1 : Comparison between different algorithms in terms of radius for k=3

Also Table 4.2 compares three algorithms for solving 1-center problem, Extreme point method is one I proposed in this thesis and other two bouncing bubble method and Jack Ritter's method are existing methods. It compares these three algorithms for different number of points. It was observed, the bounding radius is minimum for extreme point method in each case. Based on input condition the bounding radius obtained using extreme point method is slightly less or equal to Jack Ritter's method and Bouncing Bubble Method.

Table 4.2 : Output for 1-center problem using different methods

S no.	Number of points	Radius using Extreme Point method	Radius using Bouncing Bubble method	Radius using Jack Ritter's method
1.	12	194	210	230
2.	54	358	363	358
3.	67	351	359	351
4.	154	297	317	301
5.	207	270	269	272
6.	482	608	694	627
7.	877	667	670	667

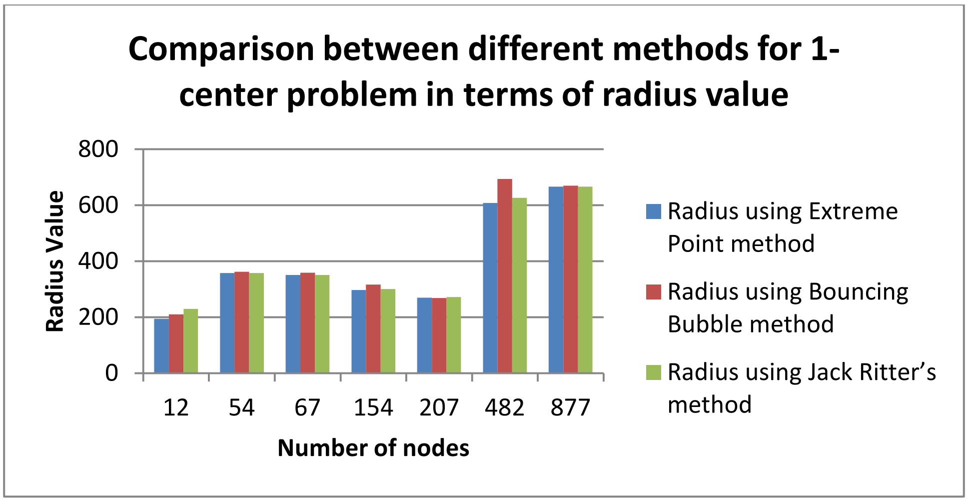

Figure 3.3-2 : Comparison between different methods for 1-center problem

Table 4.3, compares cost factors used in agglomerative hierarchical clustering, three different cost factors are compared complete linkage, average linkage and single linkage. It was observed that with cost factor as complete linkage is better and minimum radius value result was obtained. For different value of k comparison is carried out and result obtained is radius which is least for complete linkage, then for average linkage and largest radius is obtained using single linkage.

As from definition, in complete linkage the distance between two clusters is the maximum distance between and observation in one cluster and an observation in the other cluster. This method ensures that all observation in a cluster are within a maximum distance and tends to produce clusters with similar diameters. Also for average linkage, the distance between two clusters is the mean distance between an observation in one cluster and an observation in the other cluster. Similarly for single linkage, the distance between two clusters is the minimum distance between and observation in one cluster and an observation in the other cluster. The single linkage method tends to identify long chain-like clusters. Thus from the definition we can say that complete linkage will end up with clusters with minimum inter cluster distance clusters. So the result obtained as shown in table 4.3 is as expected and single linkage will have maximum radius among there for higher number of clusters.

Table 4.3 : Output using different cost factors in hierarchical clustering

S no.	Number of centers (n=54)	Radius using Complete Linkage	Radius using Average Linkage	Radius using Single Linkage
1.	K=2	140	169	143
2.	K=3	124	107	143
3.	K=4	102	102	143
4.	K=5	102	101	122
5.	K=6	76	76	122
6.	K=7	76	76	107
7.	K=8	76	76	107
8.	K=9	63	63	107
9.	K=10	63	63	99

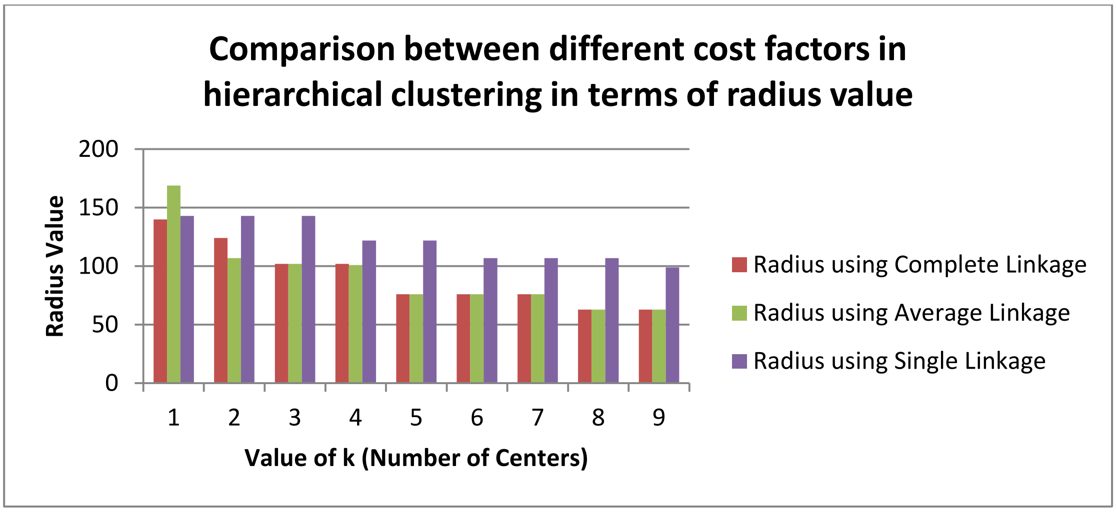

Figure 3.3-3 : Comparison between different cost factors in hierarchical clustering in terms of radius value

The result obtained in graphical form can be presented as following.

Agglomerative hierarchical Clustering using complete linkage as cost factor

This example output is of random input points in 2D metric space which is after performing agglomerative hierarchical clustering using complete linkage as cost factor. Numbers of nodes is 207 and clustered into 5 clusters.

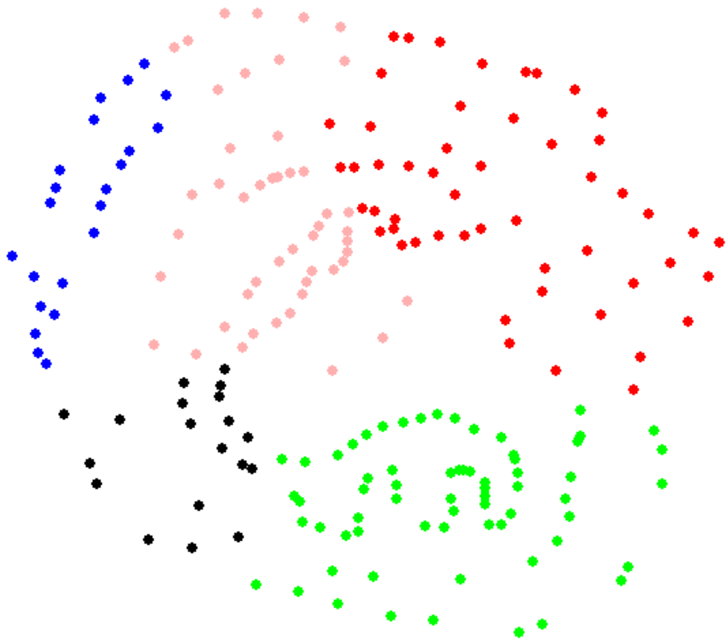

Figure 3.3-4 : output of clustered nodes for k=5 using complete linkage as cost factor

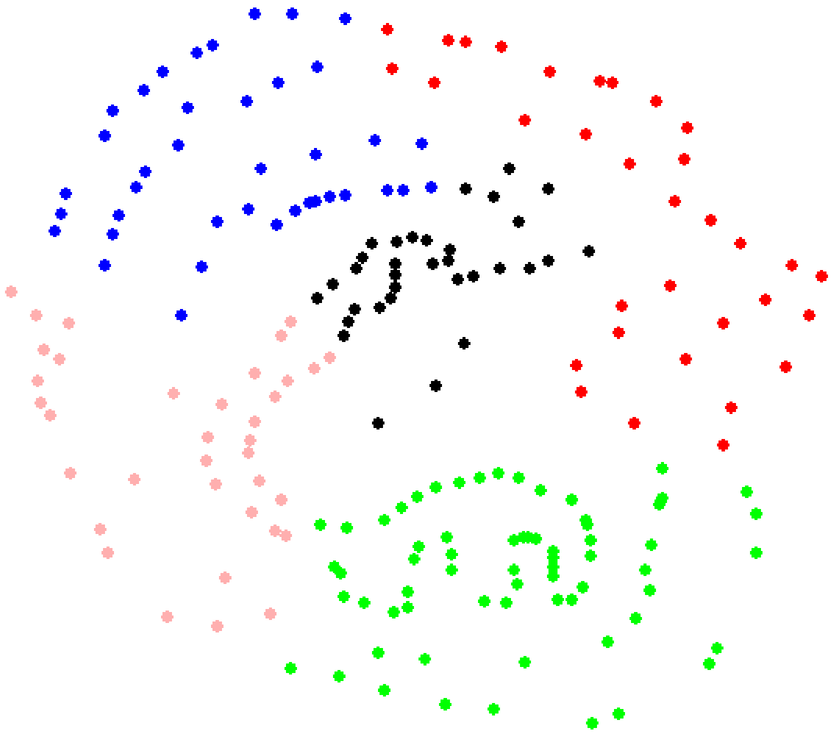

Figure 3.3-5 : output of clustered nodes for k=5 using average linkage as cost factor

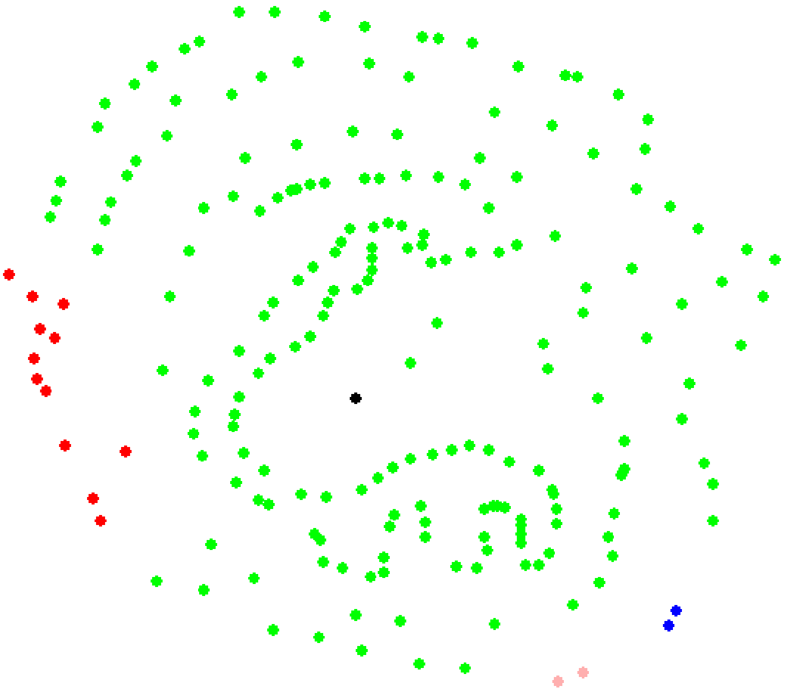

Figure 3.3-6 : output of clustered nodes for k=5 using single linkage as cost factor

One Center Problem

This output is for solving 1-center problem. Number of nodes is 207. Each cluster after clustering from above process is solved using extreme point method to get center and bounding radius. Other methods, bouncing bubble and jack ritter's method is also used to solve 1-center problem. Extreme point method gives optimum result than that of bouncing bubble method and Jack Ritter's method

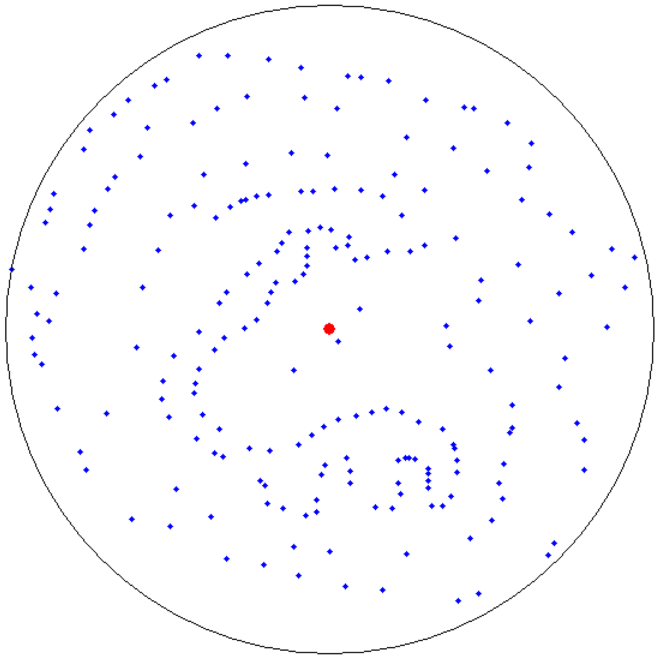

Figure 3.3-7 output of 1-center problem using extreme point method r=270

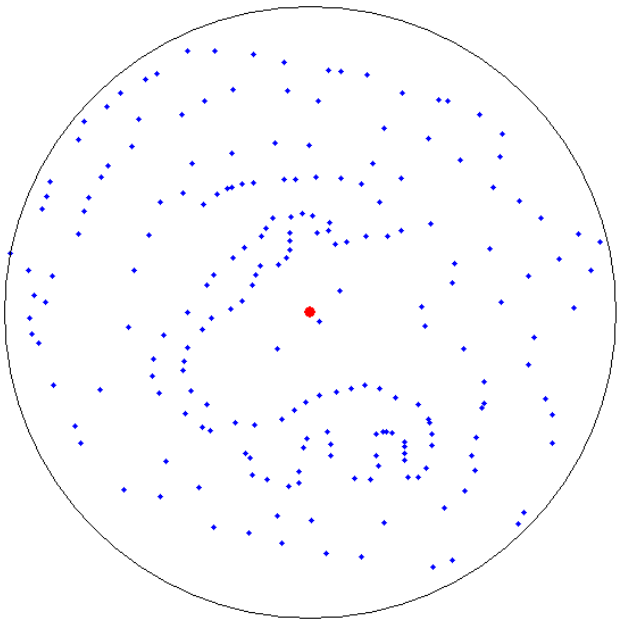

Figure 3.3-8 : output of 1-center problem using bouncing bubble method r=270

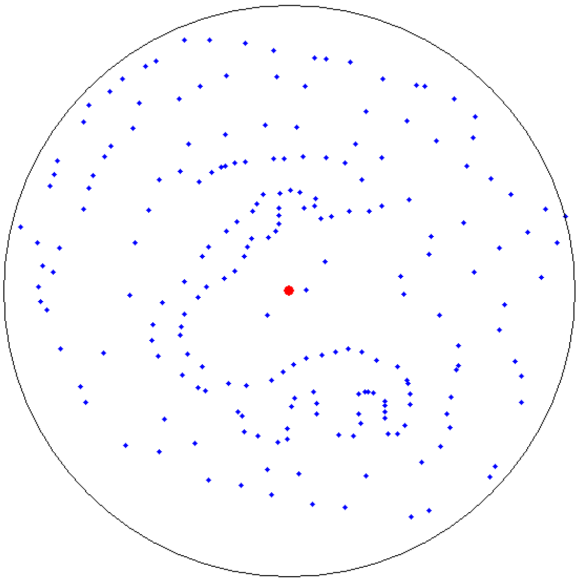

Figure 3.3-9 output of 1-center problem using Jack Ritter's method r=272

K-center Problem

After Clustering and solving one center problem for each clusters following output is obtained. Number of nodes is 63.

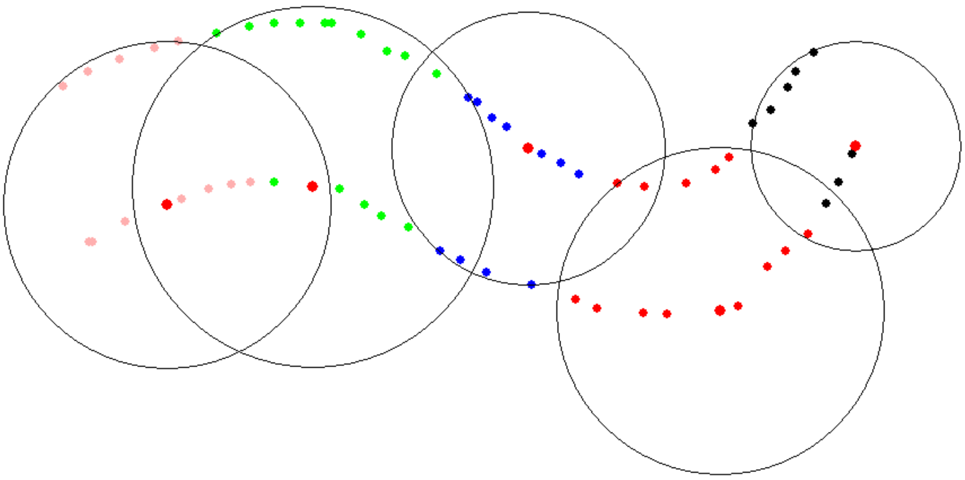

Figure 3.3-10 Output for k-center problem using agglomerative hierarchical clustering method

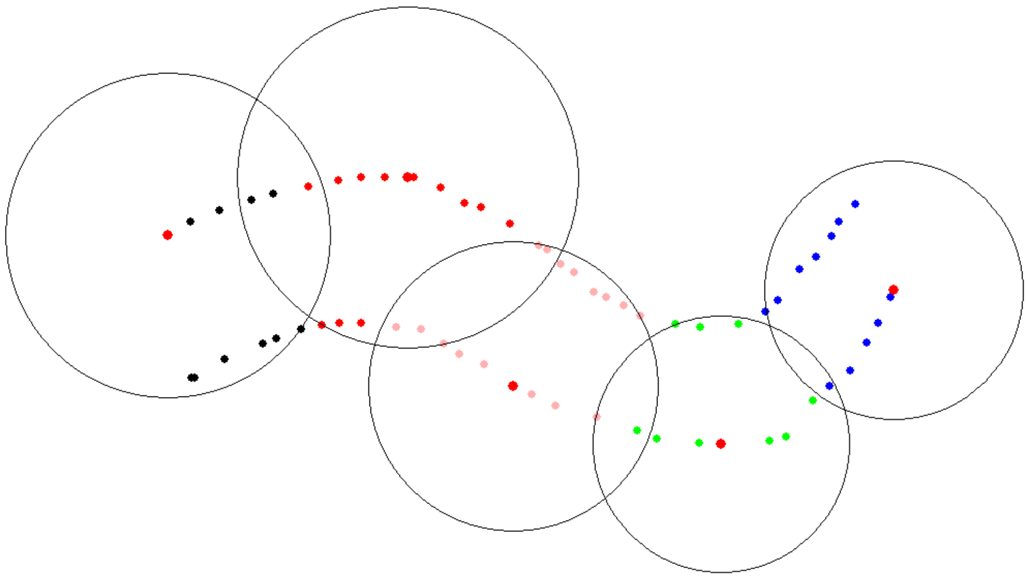

Figure 3.3-11 Output for k-center problem using agglomerative farthest first method

4.1. Time complexity

The time complexity of agglomerative Hierarchical clustering algorithm is calculated as $O((n-k)n^2)$ but value of k can be 1 so time complexity of algorithm will be $O(n^3)$. Similarly solving 1-center problem using extreme point method has time complexity of $O(km^2)$. For k=1, m=n so time complexity will be $O(n^2)$.Thus overall time complexity of k center problem for the method used in this thesis is $O(n^3) + O(n^2)$ so it is $O(n^3)$.

The farthest first method has time complexity of O(kn) it is relatively faster. Similarly the algorithm for 1-center problem bouncing bubble has time complexity of O(nd) and Jack Ritter's method with time complexity O(n).

Table 4.4 compares the method proposed in this thesis with farthest first method to compute k-center problem in terms of calculation time. The machine used for obtaining the results had Intel® core™ i5-3210M CPU @2.50GHz 2.50GHz and 4 GB ram. Figure 4.12 shows the result of table 4 in graphical form.

Table 4.4 : Comparison of calculation time among different algorithms

S. No.	number of points	hierarchical clustering method	farthest first method
1	100	353 ms	250ms
2	500	1 sec 282ms	611ms
3	1000	6sec 41ms	824ms
4	1500	20sec 17ms	904ms

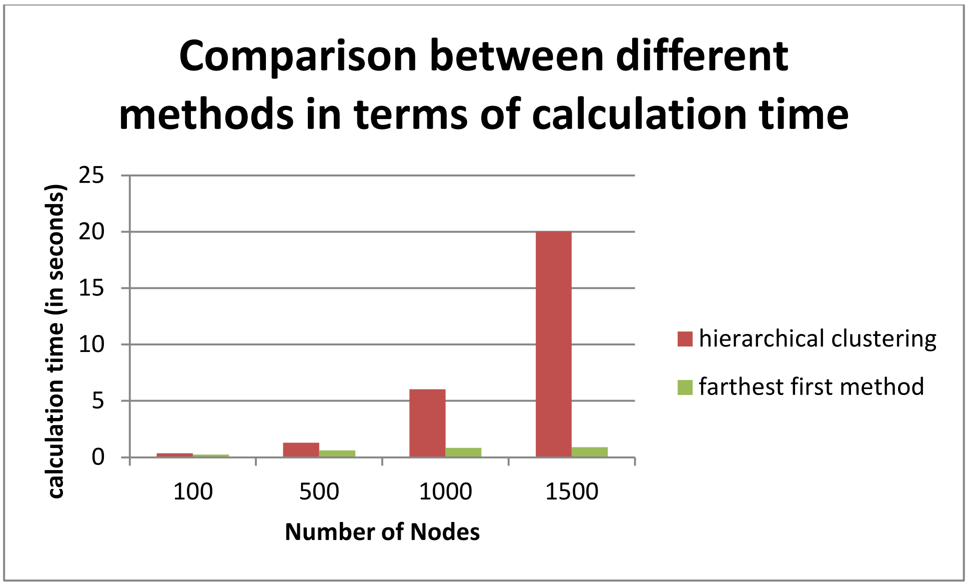

Figure 4.1-1 : Comparison between different methods in terms of time complexity

Chapter 5
CONCLUSION

5. Conclusion

In this thesis k-center problem is solved in metric space graph first by clustering nodes of graph and then finding k center within each cluster. This is one new approach, the method discussed here implements agglomerative hierarchical clustering with complete linkage as cost factor to compute clusters and for each clusters 1-center problem is solved to get overall result for k-center problem. Agglomerative hierarchical clustering is existing greedy method for clustering and the method used to solve 1-center problem, extreme points method is implemented in this thesis. This method is new approach which would solve the problem to get optimum value. Other method used to solve 1-center problem is bouncing bubble method and Jack Ritter's method. From the results, it was observed that extreme point method gives minimum value result than bouncing bubble method and Jack Ritter's method with little higher time complexity.

In this thesis the proposed method is also compared with existing algorithm for k-center problem called farthest first method, this is greedy method which is simple and efficient and widely used to solve k-center problem. For small number of centers, the proposed method i.e. agglomerative hierarchical clustering and solving 1-center problem yields better and minimum results than that of farthest first method but for higher number of centers it may not get minimum result and depends on input conditions.

Chapter 6
LIMITATION AND FUTURE ENHANCEMENT

6. Limitation and Future Enhancement

The proposed method in this thesis has many limitations. Since it is an heuristic approach the output may stuck in local optimum. The agglomerative hierarchical clustering is not perfect to calculate minimum inter cluster distances so the final result may not get optimum. Also it has time complexity of $O(n^3)$ which is not suitable for higher number of nodes. The algorithm for 1-center problem, extreme point method proposed in this thesis is quite complex and has higher time complexity than other available algorithm though it gives better results than most of existing algorithms.

Further researches can be done to improve time complexity of the algorithm. Better clustering algorithms can be used to get better clusters with less inter cluster distances, which was required in k-center problem. The algorithm for 1-center problem, extreme point method can be further improved to get less error results, so further researches can be done on it.

REFERENCES

References

[1] N. G. R. R. Shiva Chaudhuri, "The p-neighbour k-center problem," Vols. 65,3,131-134, 1997.

[2] R. R. &. D. Garg, "THE ANALYTICAL STUDY OF K-CENTER PROBLEM SOLVING TECHNIQUES," *International Journal of Information Technology and Knowledge Management,* vol. 1, pp. 527-535, July-December 2008.

[3] S. L. V. M. l. R. Silvio Lattanzi, "Robust Hierarchical k-center clustering," january 2015.

[4] P. M. L. Sanjoy Dasgupta, "Performance guarantees for hierarchical clustering," 24 july 2010.

[5] A. foul, "A 1-center problem on the plane with uniformly distributed demand points," in *Operations Research letters*, may 2006.

[6] I. L. Gortz, "Topics in Algorithms Data Structures on trees and approximation algorithms on graphs," no. Department of Theoretical Computer Science, may 2005.

[7] J. M. Z. A. E. Rahmat Widia Sembiring, "A Comparative Agglomerative Hierarchical Clustering Method to Cluster Implementaed Course," *Journal of computing,* Dec 2010.

[8] N. Megiddo, "The Weighted Euclidean 1-center problem," in *Mathematics of Operations Research*, Tel Aviv University, nov 1983, p. vol 8 no. 4.

[9] T. F. gonzalez, "Clustering to minimize the maximum intercluster distance," *Theoretical Computer Science,* pp. 293-306, 1985.

[10] D. B. Shmoys., "Computing Near-optimal Solutions to Combinatorial Optimization Poblems," Technical Report, Ithaca, NY 14853, 1995.

[11] D. S. H. a. D. B. Shmoys, ""A best possible heuristic for the k-center problem"," in *Math of Operational Research Vol 10. No. 2.*, 1985, pp. 180-185.

[12] S. K. D. M. M. G. n. Moses Charikar, "Algorithms for facility loacation problems with outliers," *Discrete Algorithms,* pp. 642-651, 2001.

[13] B. Tian, "Bouncing Bubble: A fast algorithm for minimal Enclosing Ball problem," 2012.

[14] J. Ritter, "An efficient bounding sphere.," *In Andrew S. Glassner, editor, Graphics Gems. Academic Press,* Boston, MA, 1990.

APPENDICES

APPENDICE

Appendix A : Snapshots of the User Interface

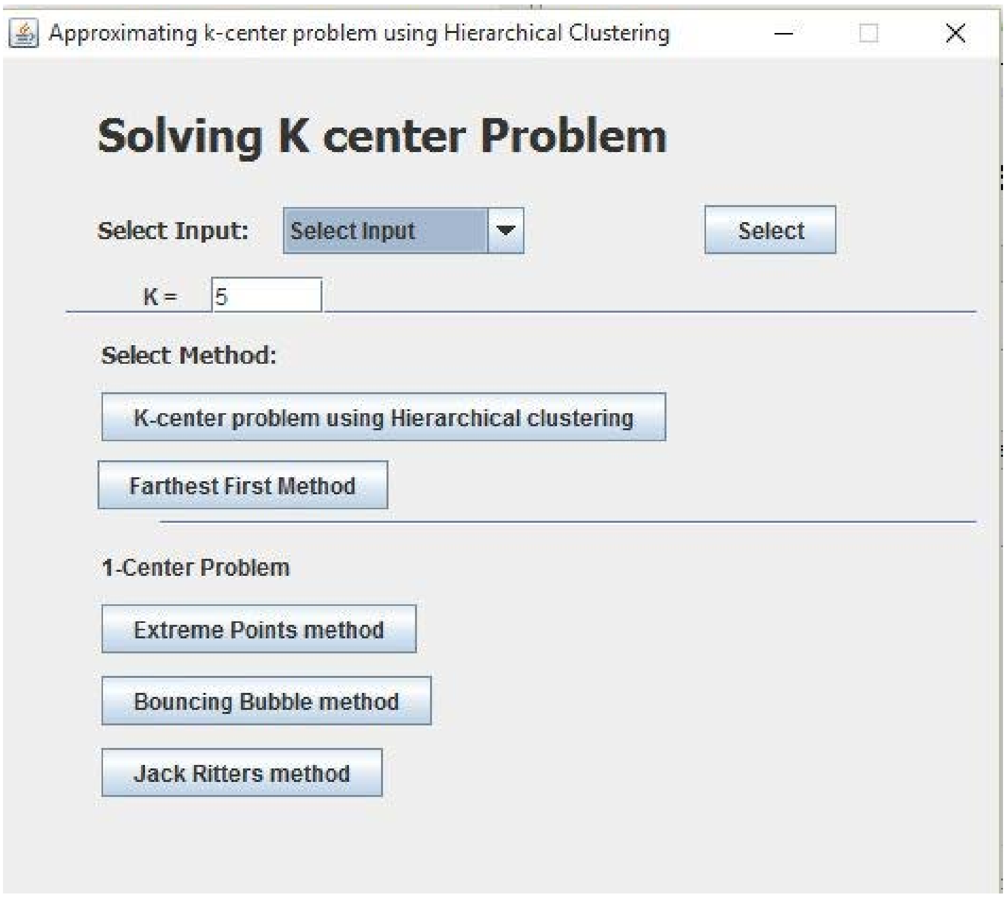

Figure A-1 User Interface for Approximating k-center problem using Hierarchical Clustering

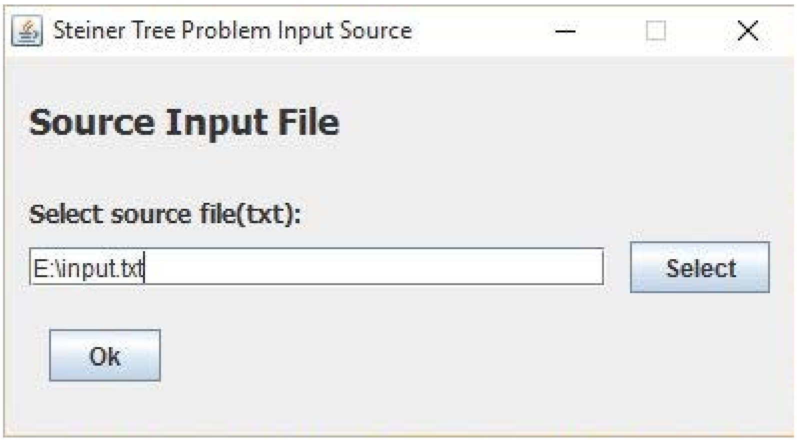

Figure A-2 Selecting Input file with inputs nodes in it.

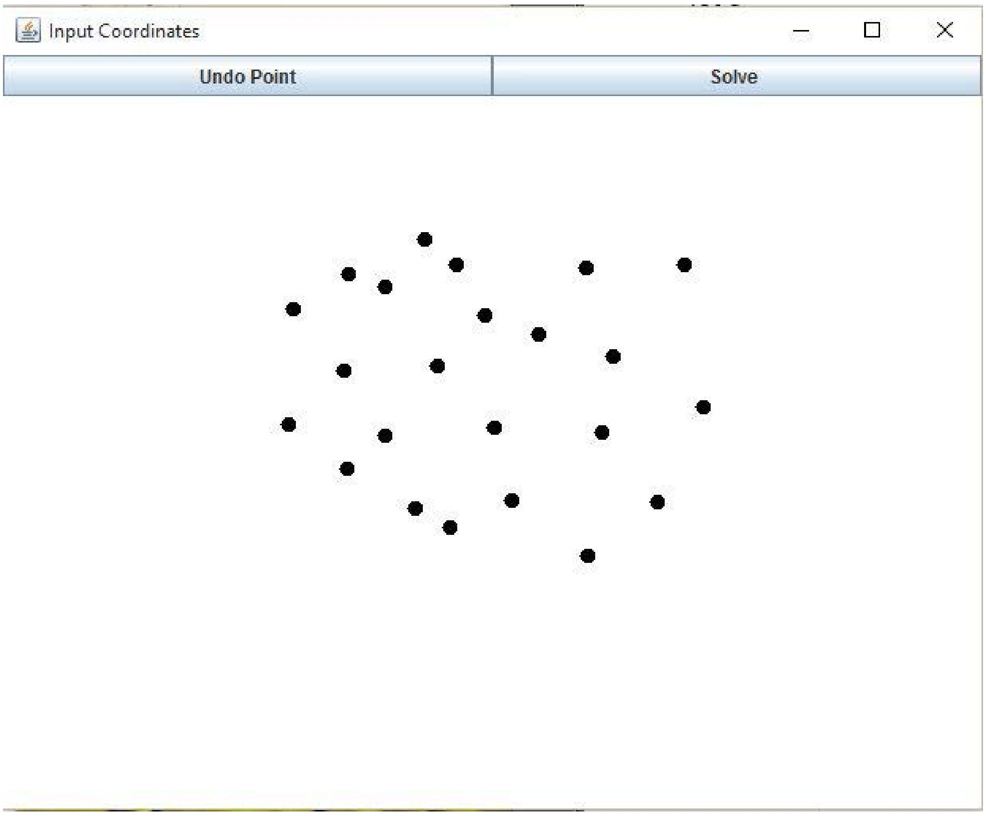

Figure A-3 Taking input by drawing vertex points.

Appendix B Snapshot of output for various methods.

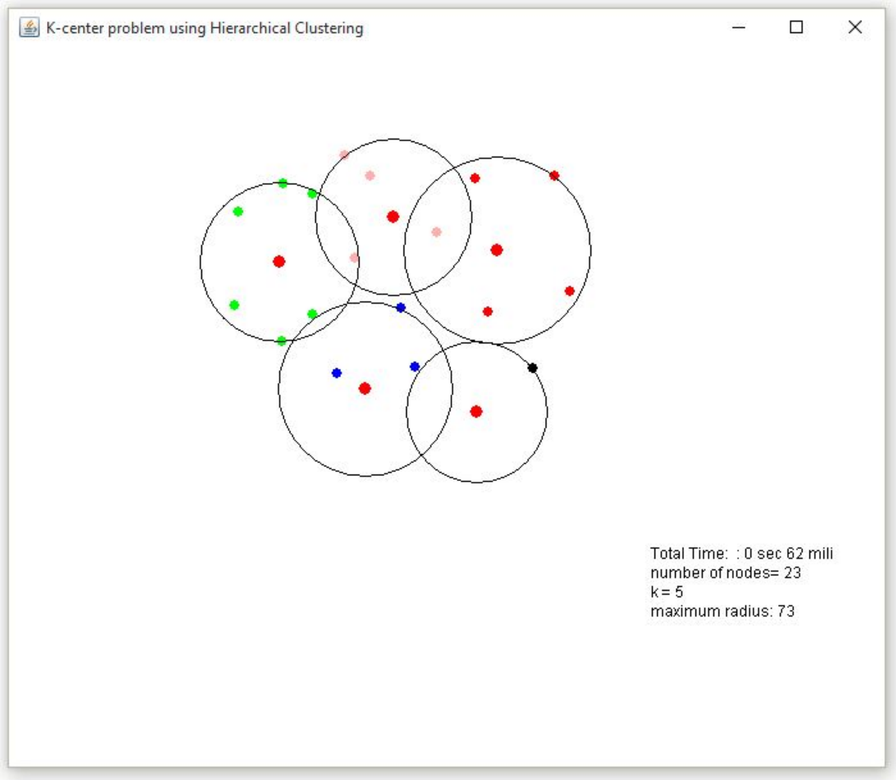

Figure B-1 output for k center problem using hierarchical clustering.

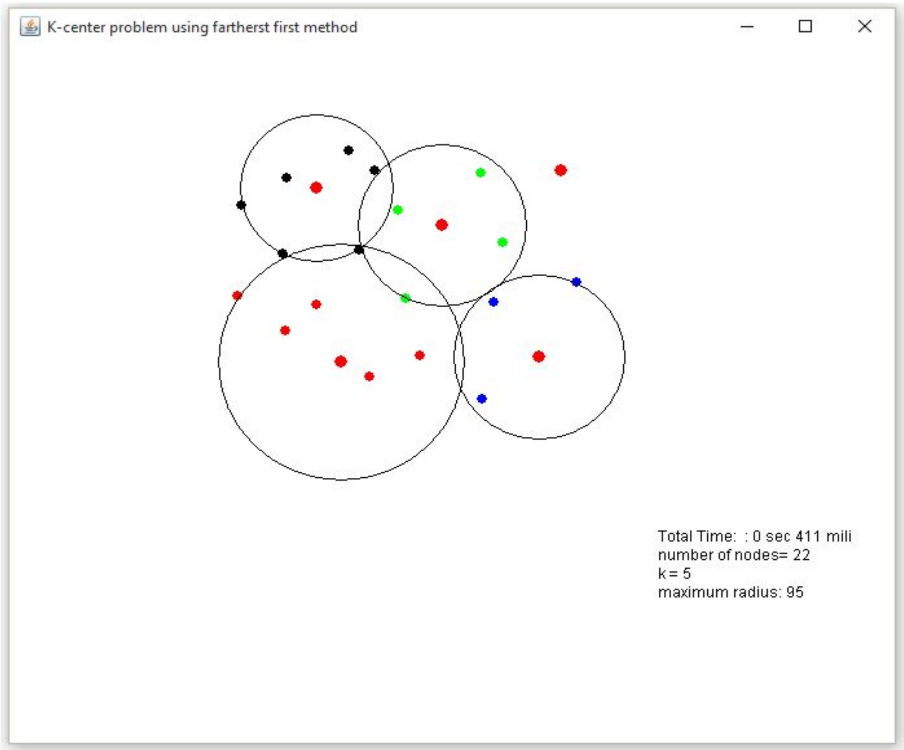

Figure B-2 Output for k-center problem using farthest first method.

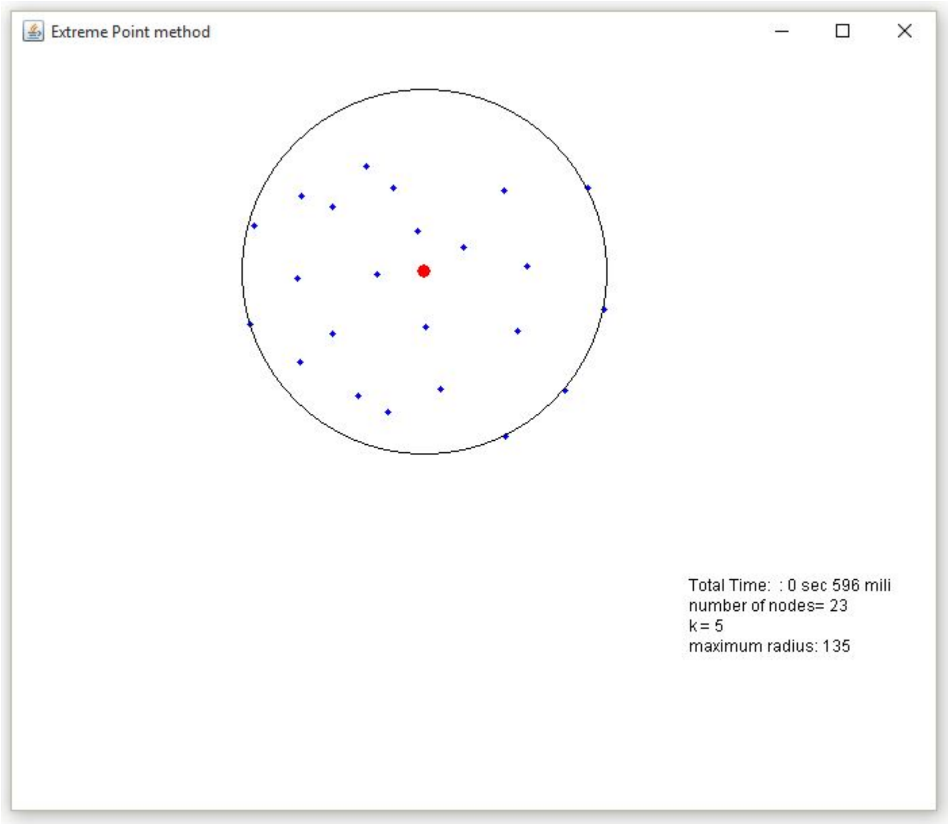

Figure B-3 Output for 1-center problem using extreme point method.

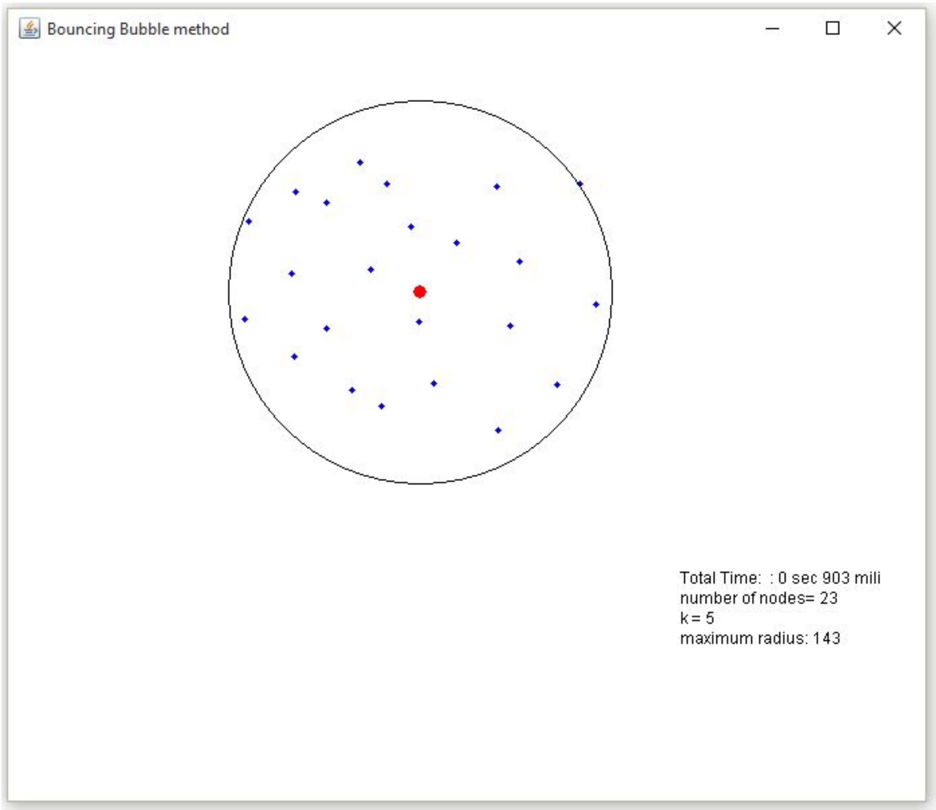

Figure B-4 Output for 1-center problem using Bubble method.

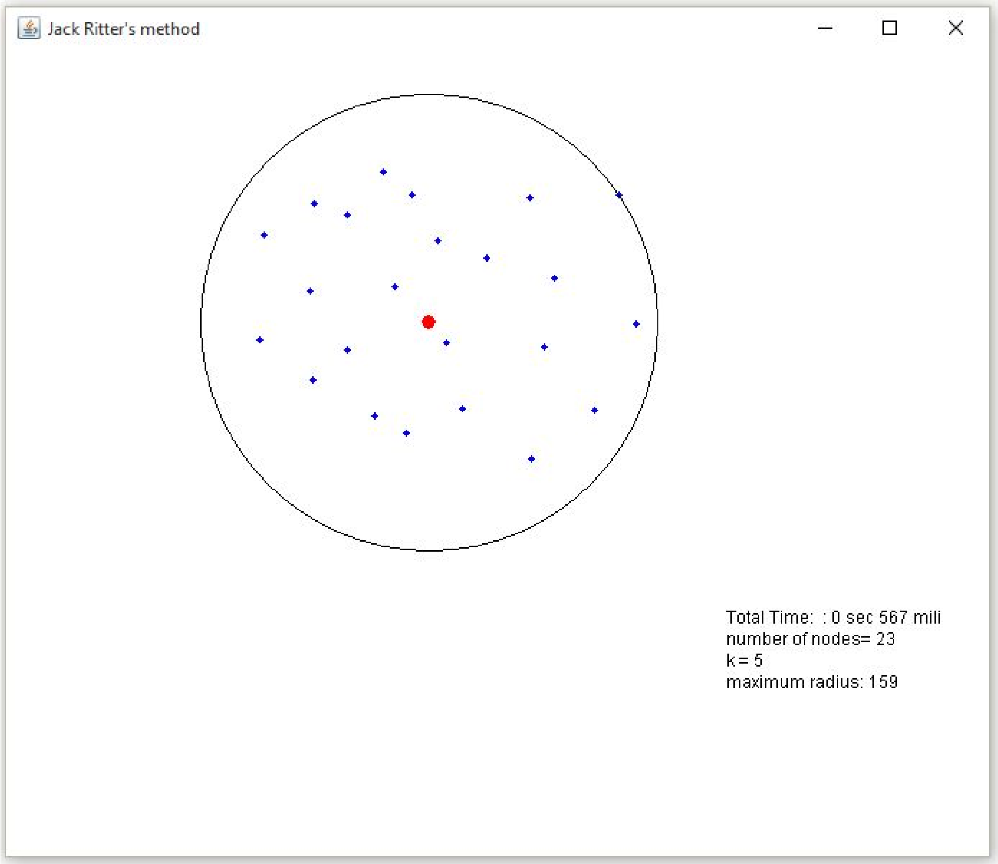

Figure B-5 Output for 1-ceter problem using Jack Ritter's method.

Appendix C Snapshot of the output for large points.

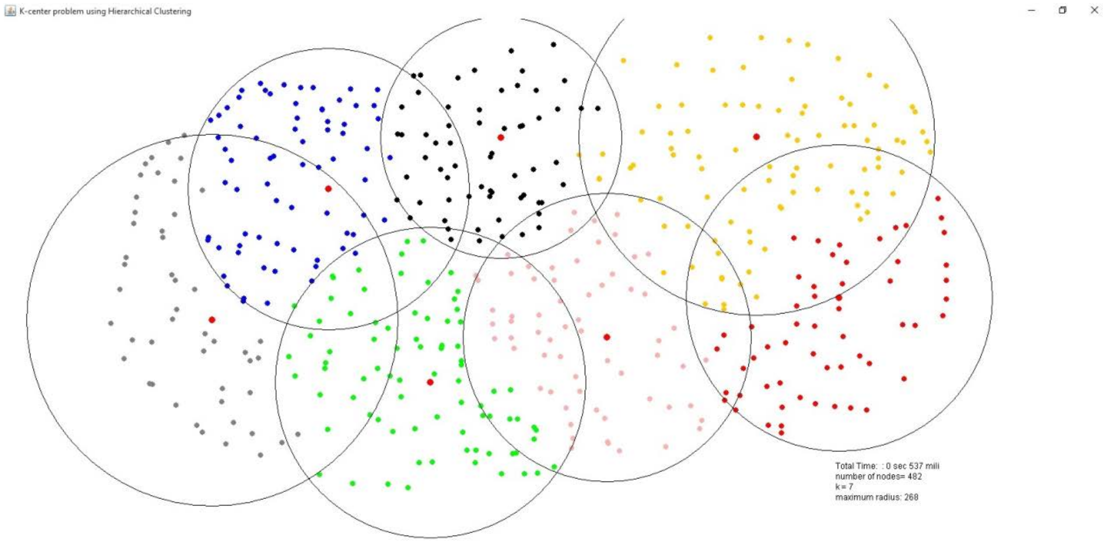

Figure C-1 Output for k-center problem for 482 points using hierarchical clustering method.

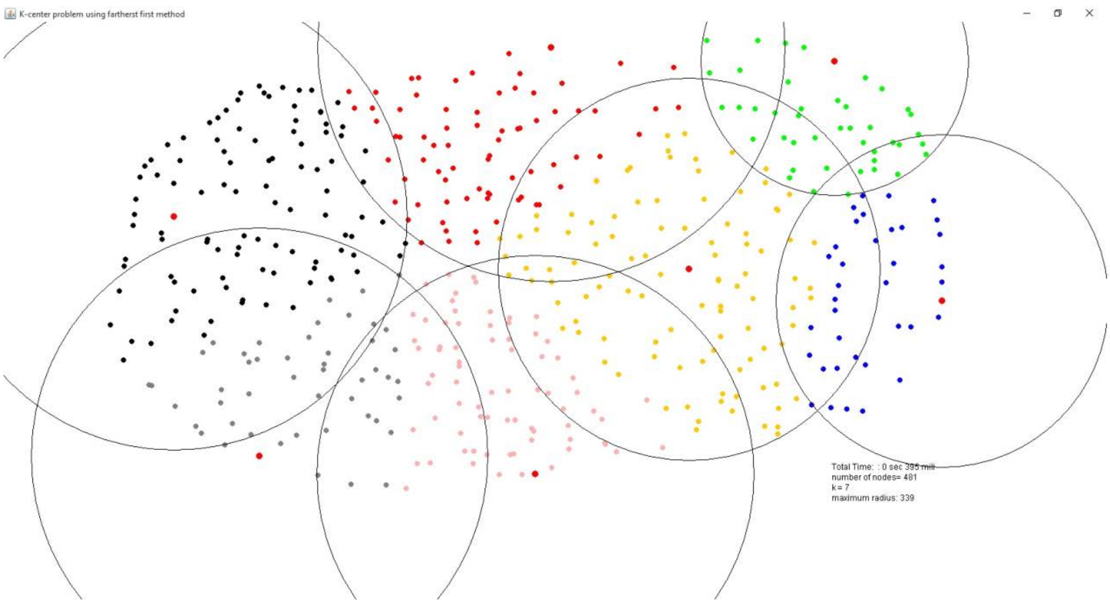

Figure C-2 Output for k-center problem for 482 points using Farthest First Method.