Contemporary Engineering Sciences, Vol. 7, 2014, no. 7, 315 - 322
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2014.429

Analysis of Clustering Techniques in VLSI Cell Partitioning

V. Leela

M.Tech VLSI Design
School of Computing
SASTRA University,Thanjavur-613401,India

R. Manikandan

School of Computing
SASTRA University, Thanjavur-613401,India

Abstract

Circuit partitioning plays a dominant role in VLSI physical design of chips. In this paper the newly proposed rank based k-medoid clustering algorithm is discussed, in order to partition the combinational circuit based on their interconnection distance among cell groups. Clustering analysis of the given circuit ,partition the set of objects into non overlapping subsets. The proposed ranked k-medoid clustering algorithm method reduces the size of the interconnections distance and also speed up the large scale partitioning problems without any reduction of partitioning accuracy. The ranked K-medoid algorithm has overcome the local optimum problem and has showed the improved result when it is compared with K-mean and K-medoid clustering algorithm.

Keywords: Clustering, Ranked k-medoid, K-mean

1. Introduction

A cluster analysis group the data for the given information of the user based on the objects and its relationship.The goal of clustering is to cluster the objects which are similar within the group and different from the objects present in other group. Clustering analysis has been used in various applications like physical geography, slippy map optimization, market research and security.The clustering algorithm are of various types like Exclusive Clustering (K-Mean), Overlapping Clustering (Fuzzy-CMean), Hierarchical clustering, and Probabilistic Clustering. In this paper our main approach is partitioning without overlapping and mainly exclusive type algorithms such as K-Mean and K-medoid, since the partitioning of the circuit based on distance measure should not overlap and should minimize the total distance measure improving the speed and accuracy even at large data sets also.

One of the simplest and most commonly used algorithm for partitioning approach is K-mean which re computes each of the centers with average of the object distances within the previous centre. Even though K-mean is efficient ,it undergoes various demerits like sensitive to outliers and local optimum problems[1].In this paper we have introduced the Rank-based clustering technique in K-medoids where the initialization of K does not leads to local optimum and it helps to find the best K using silhouette function. Since the rank-based approach helps to find the similarity near most medoids faster, reducing the computational time. The combinational circuit is taken as example in VLSI based application instead of datasets and its performance is compared with K-Mean and K-medoid along with proposed algorithm for the better efficient results.

2. Existing methods

2.1*K-Mean clustering:*

K-mean is one of the popular unsupervised learning algorithm in clustering process.[2]The algorithm is fast and convergent in finite number of steps.K-Mean clustering is the highly used method for various application like cluster analysis,features learning and vector quantization.K-mean clustering mainly partition the n values into k clusters in which each metrics belongs to each cluster with the nearest mean with the given iterations.

The algorithm contains the following steps,

(1) Initiate the process with k points into the space, (ai the initial point of centroids) representing the objects that has to be clustered.

(2) Assign each of the object with a group by choosing the closest neighbour.

(3) After the assignment of centre k to each groups of objects, calculate the distance of the next k centroids.

(4) Repeat the steps until the centroids no longer move and its converge else return step 2.

2.2 K-medoid clustering

The k-medoid algorithm is similar to k-mean algorithm ,where both the techniques reduces the distance between the points in a cluster with the assigned center.The k-medoid chooses the data points as medoids. K-Medoid clustering algorithm is more sensitive to noise and outliers [4] but it minimizes the sum of

pairwise dissimilarities.The algorithm has following steps
Step 1 : (Intialization)
1-1. Calculate the distance based on the dissimilarity measure, between every objects present in the data using the Euclidean equation (1) follows:

$$D(p,q)=\sqrt{(q_1-p_1)^2+(q_2-p_2)^2+....+(q_n-p_n)} \qquad -(1)$$

1-2. Evaluate the medoids among the objects and fix the centers of the clusters as K.
1-3. Determine the position for each objects and then sort all the objects based on the smallest distance order. Choose the objects having the lowest value as initial group medoids.
1-4. Locate the object of the data with respect to the nearest neighbour medoid using similarity measure and calculate the current value of the medoids and the distance between the medoids and objects.
Step 2 : (Updation and labeling)
Recalculate and change the previous medoid of the cluster to reduce the over-al distance between the objects in the cluster.
Step 3 : (Assignment)
3-1. Allocate the object of the data to the nearest neighbor medoid
3-2.Evaluate the distance of all the objects with respect to new medoids .If the result is same as the old value, then stop the iterations. Else, return back to the Step 2.

3. Proposed Algorithm

3.1 *Ranked K-medoid Algorithm*

The Ranked k-medoid clustering is similar to k-medoid algorithm aims at partitioning a set of objects into k clusters in which each objects belongs to the cluster with the nearest medoids based on the rank index.The k-medoid algorithm has certain disadvantages like whenever there is a different intial sets of medoids as input it leads to different final clustering.The proposed algorithm has overcome the local optimum problem and the accuracy ,speed of clustering remains stable even at large data sets .[6]

Let us consider the n objects having s number of variables ,where each are grouped into k(k<n) clusters,where the k is assigned randomly.The Euclidean distance are used as a similarity measure between the objects i and objects j is given by the proposed algorithm has the following steps as follows:
Step 1:(initialization)
1.1Calculate the distance of every pairs of the objects based on their dissimilarity (Eg.Euclidean correlation) in equation(1).
1.2 Arrange the calculated distance using the sorting (ascending) method.
1.3 Select the initial cluster k by allocating each objects to near most objects and calculate the overall distance of the objects with respect to Medoids.
1.4 Choose the points k from the data using the sorted index points as reference and set the values of medoids.

Step 2(Update)
2.1 Find the new medoid form the initial medoids in each cluster,where it minimizes the total distance of the objects by selecting the near most neighbors.
2.2 Select the medoid assignment that has the lowest cost and update the medoids in each groups of the cluster by replacing with new medoids.
Step 3(Assignment and labeling)
3.1Calculate the overall distance between the clusters.and the sum is equal to the previous calculated value , then stop the iterations process.
3.2 Repeat step 2 until maximum iteration is achieved where the medoids does not change the position and its converge.

4. Experimental Results:

The K-mean and K-medoid algorithm are used for partition the datasets into a group in order to minimize the distance between the points with a center or medoids.The proposed algorithm is taken up with combinational circuit gate matrix as example to generate data with the given mean and variance and are compared with K-mean and K-medoid using Matlab R2010a.

4.1 Circuit representation:

Let us consider the combinational CMOS circuit given below in fig(1) for the data generation through its interconnection distance .The Cmos circuits has the interconnection with gates through the wire and it is given in the form of a matrix K along with the input matrix I .The input matrix I gives the details of wire connection with the previous and next gate and the K matrix represents the type of gates used in the circuits.The K matrix is generated with a column based on the connection of the wires or gates .[1]The total number of column in K matrix is similar to number of gates present in the Cmos circuits fig(1).The matrix represents the specific type of gate,a wire and no gate.The gates are listed in Table 1.Each number in the matrix I represents the wire or output of a specific gate from the respective column of K matrix.

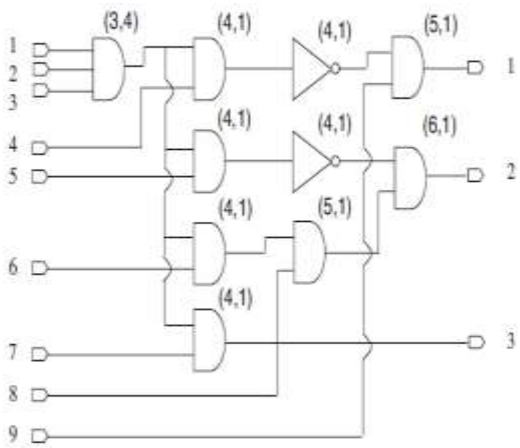


Fig. 1. Combinational Cmos circuit with input and fan out.

Gate number	Type
No gate	0
1	AND
2	XOR
3	OR
4	INV
5	WIRE
6	XNOR
7	BUFFER
8	NOR
9	NAND

TABLE 1:Gate matrix representation

$$k = \begin{bmatrix} 1141 \\ 5141 \\ 5115 \\ 5150 \\ 5500 \\ 5000 \\ 0000 \\ 0000 \end{bmatrix}$$

$$i=\begin{bmatrix} 1 & 10 & 19 & 28 \\ 2 & 11 & 0 & 32 \\ 3 & 0 & 0 & 0 \\ 4 & 10 & 20 & 29 \\ 0 & 12 & 0 & 30 \\ 0 & 0 & 0 & 0 \\ 5 & 10 & 21 & 31 \\ 0 & 13 & 23 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 6 & 10 & 22 & 2 \\ 0 & 14 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 7 & 15 & 24 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 8 & 16 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

4.2 Partitioning :

The partitioning of the matrix is carried out through clustering process.[7] The matrix of the circuits are taken and the covariance of a matrix is evaluated in order to find the near-most value of the function and to get the inter relationship present in the interconnection representation of the matrix I.Given a matrix consisting of *n* independent observations x_1, ..., X_n of a n-dimensional random vector With E as expected values, following X as random values in data.

$$\mathrm{COV}(X)=E[(X-E[X])(X-E[X]^T] \quad -(2)$$

The matrix from the large data sets has more similar sets of value which increases the space and processing time.so this can be reduced by finding the unique value in the matrix Ai the similarity matrix values which are common are reduced in this step. By applying Cramer's rule ,the solution of a system of linear equations with as many values are unknowns, it helps whenever the system has a unique solution. Therefore unique matrix can be found by considering a system having matrix of n by n as A

$$\text{Unique matrix}=\frac{\det A(i)}{\det A} \qquad \text{where } i=1,...n \quad -(3)$$

where Ai is the matrix formed by the distance between the gates

The dissimilarity among the objects are found using various functions like Euclidean distance, cosine etc.In this paper we are using the Euclidean distance of a gate matrix in-order to find the dissimilarity based on distance of gates and its interconnection present over the circuit. The Euclidean distance of a matrix is calculated by

$$\|X-Y\|=\sqrt{(x-y)(x-y)} \qquad -(4)$$

Where the Euclidean distance between X and Y is just interconnection length of the gates in the matrix.The matrix sets of the interconnection matrix is now allowed calculated using the equations in order to find out the Mean in equation and Variance of the matrix in equation for the data generations.

$$\mu = \sum xC(x) \qquad -(5)$$

Where x is the random values and C(x) gives the probability of the value of x.

$$Var(X) = K[(X-\mu)^2] \ -(6)$$

Where x is the random values in matrix with K as expected value of the matrix. Generates data with as many clusters and n points in each of them.

4.2.1 Silhouette

The Silhouette is a new graphical display method used for the partitioning of data based on the relationship between object lies within its cluster. The silhouette is formed based on comparison between separation and its tightness of cluster.In this paper the data of the cmos circuit gate matrix is clustered by k-means, k-medoid, ranked K-medoid with the help of silhouette function.

Let us consider the points i in x(i) is the average of all the dissimilarity inside the same group of cluster and y(i) be the lowest average dissimilarity of i to any other cluster where i is not a member of nearest points.It is given as

$$S(i)=\frac{x(i)-y(i)}{\max\{x(i),y(i)\}} \qquad \text{------(7)}$$

Which can be written as:

$$s(i)=\begin{cases}1-a(i)/b(i),\\0,\\b(i)/a(i)-1,\end{cases} \qquad \text{--(8)}$$

The above equation (8) the value of S(i) can be inbetween in equation (9)

$$-1\le S(i)\le 1 \quad \text{--(9)}$$

Calculate to find whether s(i) to be close to the value of 1 ,if so a(i)<<b(i) as a(i) is the dissimilar distance of own cluster and b(i)is the unmatched to its neighboring cluster.Thus an s(i) close to one means that gates interconnection distance value is appropriately clustered. If s(i) value is negative one then leaves the gate connection is so afar leaving in next group.And s(i) value is zero then the objects is present on the border of two clusters.

The clustering algorithm are given with circuit matrix mean and variance for data generation and Silhouette outputs for clustering using k-1, k, k+1 with the best K value in fig 2,3 and the proposed K-medoid showed best result in the delay optimization on several iteration of given data shown in table 2.The fig 4 shows that proposed algorithm is compatible with large data sets and execution is better with less processing time.

No of cluster	Iterations	Delay(s)	Ranked k-medoid	K-medoid	K-Mean
3	100	36	1	2	1.8
4	50	45	1.2	2.5	2.6
5	150	52	1.3	3.4	2.5
3	170	55	1,8	3.5	3

Table 2:The Performance of algorithm

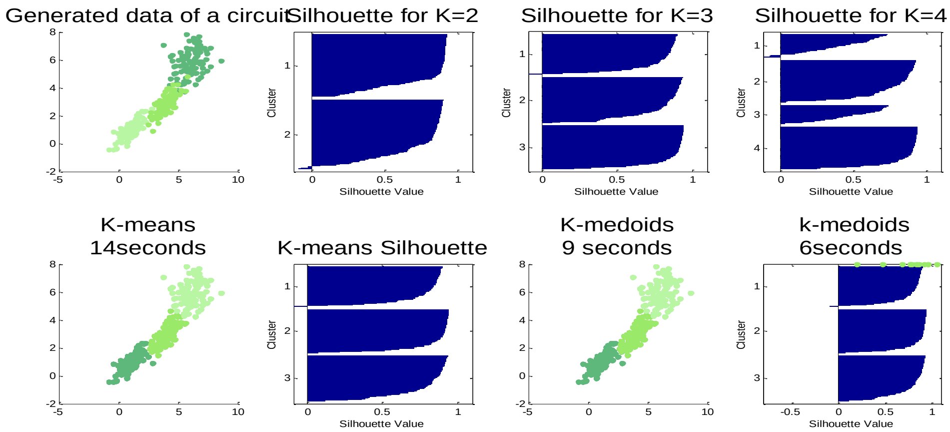
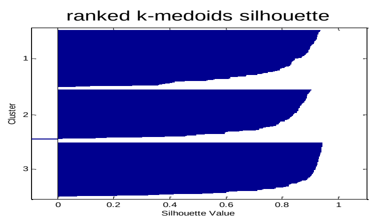


FIGURE 2:K-Mean and K-Medoid using silhouette function



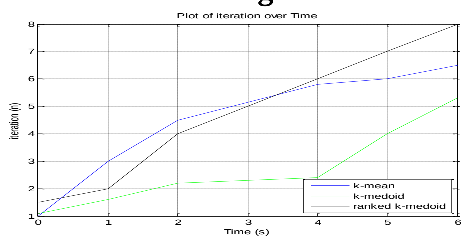


Figure 3:Ranked K-Medoid using silhouette function
Figure 4:Comparison result between cluster algorithm

5. CONCLUSION

In this paper we proposed a novel ranked based K-medoids,for optimum solution to VLSI based Circuit partitioning concepts.The combinational circuit is taken as example instead of dataset,where the gate matrix are simplified as mean and variance to generate the datasets.The Local optimum problem also minimized along with improvement in accuracy and speed when compared with K-Mean and K-Medoid algorithm.our algorithm founds that the clustering can be done even in circuits without causing any problem in partitioning accuracy and accepts even large circuit matrix as data sets with stable performance and speed.

References

[1] Ganesh K. Venayagamoorthy, Scott C. Smith, Gaurav Singhal,Particle swarm-based optimal partitioning algorithm for combinational CMOS circuits,Engineering Applications of Artificial Intelligence (2006), doi:10.1016/j.engappai.2006.06.011.

[2] V.Leela, K.Sakthi priya, R.Manikandan , A comparative analysis between k-mean and y-means algorithms in fisher's iris data sets, International Journal of Engineering and Technology (IJET), 5(1):2013, 245-249.
[3] Monica Sood , Shilpi Bansal ,K-Medoids Clustering Technique using Bat Algorithm, *International* Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 , 5(8): 2013,20-22.
[4] H.S. Park , C.H. Jun, A simple and fast algorithm for K-medoids clustering, Expert Systems with Applications, 36(2):2009, 3336–3341.
[5] Peter J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis Journal of Computational and Applied Mathematics 20 (1987) 53-65.
[6] Seyed Mohammad Razavi Zadegan, Mehdi Mirzaie, Farahnaz Sadoughi, Ranked k-medoids: A fast and accurate rank-based partitioning algorithm for clustering large datasets ,Knowledge-Based Systems 39 (2013), 133–143.
[7] Dr.Sumithra devi K.A. and Vijayalakshmi.M.N,Classification Algorithms in AchievingPartitioning Optimization for VLSI Applications International Journal of Computer Theory and Engineering, 2(6): 2010 ,1793-8201.

Received: February 1, 2014