

# Design and Analysis of a Quantum Circuit to Cluster a Set of Data Points

Sanjay Chakraborty\*, Sambit Halder, Shamik Kundu

Department of Computer Science and Engineering, Institute of Engineering and Management, India

Copyright©2016 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** Clustering is a powerful tool which has been used in several applications (such as, fraud detection, pattern matching etc.). It is known to all that quantum counterpart is more superior to classical one on some specific areas. Quantum computation can solve certain problems much faster than classical computation. This paper mainly proposes a quantum circuit to cluster a given set of points with respect to pre-defined points. It uses binary quantum logic and superpositions principle of quantum mechanics. This proposed approach is a collaboration of data mining techniques with quantum computation. Very few works have been done in this area. In this paper, some of the previous works are also analyzed on the idea of applying quantum concept to the classical approach of clustering algorithms.

**Keywords** Quantum Circuit, Reversible Logic, Qubits, Quantum Fourier Transform, Quantum Clustering

## 1. Introduction

In modern world, power dissipation is one of the critical issues while measuring the performance of an integrated circuit. From Landauer's statement, each bit of information loses  $kT \ln 2$  joules of energy during logical computations (where  $k$  is the Boltzmann's constant and  $T$  is the operating temperature) [5]. It has a major impact on various technologies like VLSI, Embedded systems. However, the use of reversible/quantum logic circuit helps to reduce the temperature by evacuating the radiated heat. In reversible gate the number of inputs is equal to the number of outputs and no fan-out is permitted. In this paper we present a circuit to cluster a set of given points with respect to pre-defined points. We use quantum gates and circuitry such as qFT circuit, discrete phase gate, quantum comparator circuit. We also use quantum registers to store the qubits. Some important reversible logic (necessary for the proposed work) are described below.

### 1.1. The Qubit

The basic unit of information in quantum computing is called the qubit, which is short form quantum bit. Like a bit, a qubit can also be in one of two states. In the case of a qubit, we label these two states by  $|0\rangle$  and  $|1\rangle$ . In quantum theory an object enclosed using the notation  $|\psi\rangle$  can be called a state, a vector, or a ket. While a bit in an ordinary computer can be in the state 0 or in the state 1, a qubit can exist in the state  $|0\rangle$  or the state  $|1\rangle$ , but it can also exist in a superposition state i.e. a state that is a linear combination of the states  $|0\rangle$  and  $|1\rangle$ . If we label this state  $|\psi\rangle$ , a superposition state is written as  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  Here  $\alpha, \beta$  are complex numbers of the form  $z = x + iy$  [1].

The advantages of quantum computers over classical ones are: In digital computer circuit an  $n$  bit representation of bits can be one of  $2^n$  states but only one at a time whereas in quantum computers  $n$  qubits can be in any superposition of up to  $2^n$  states simultaneously. The main advantage of quantum computing is a logarithmic reduction in the number of separate quantum systems needed to span the quantum memory. Quantum algorithms are much faster when compared to digital ones for searching, calculations and storing. Quantum computers are much more secure than digital ones. Quantum technology is inherently reversible and one of the most potential technologies for future computing systems. Heat generation in quantum circuits is less than that in digital circuits [11][12].

### 1.2. Various Quantum Gates

#### 1.2.1. Hadamard gate

The Hadamard gate acts on a single qubit. It maps the basis state  $|0\rangle$  to  $(|0\rangle + |1\rangle) / \sqrt{2}$  and  $|1\rangle$  to  $(|0\rangle - |1\rangle) / \sqrt{2}$  and represents a rotation of  $\pi$  about the axis  $(x+z) / \sqrt{2}$ . Eventually, if the combination of two rotations,  $\pi/2$  about the Y-axis followed by  $\pi$  about the X axis, it is represented by Hadamard Matrix.

$$H_0 = + 1$$
$$H_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$H_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

1.2.2. Discrete Phase Gate

Another useful gate that can be applied in the development of quantum algorithms is a variation of the phase gate is called the *discrete phase gate*. We denote the discrete phase gate by  $R_k$ , where

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{(2\pi i/2^k)} \end{pmatrix}$$

We notice that,

$$R_k|0\rangle = |0\rangle, R_k|1\rangle = e^{(2\pi i/2^k)}|1\rangle$$

1.2.3. Quantum Fourier Transform

A quantum computer can be used to compute a quantum analog of the Fourier transform. Consider two states of  $d=n$  qubits

$$|x\rangle = |x_{n-1} x_{n-2} \dots x_0\rangle$$

$$|y\rangle = |y_{n-1} y_{n-2} \dots y_0\rangle$$

where each  $x_i, y_i \in \{0,1\}$ . Then

$$x \cdot y = x_0y_0 + x_1y_1 + \dots + x_{n-1}y_{n-1}$$

The quantum Fourier transforms maps the state

$$|\psi\rangle = |x_{n-1}x_{n-2}\dots x_0\rangle \text{ into the state}$$

$$\frac{1}{2^{n/2}} (|0\rangle + e^{2\pi i[0.x_{n-1}]1}|1\rangle) \otimes (|0\rangle + e^{2\pi i[0.x_{n-2}x_{n-1}]1}|1\rangle) \otimes \dots \otimes (|0\rangle + e^{2\pi i[0.x_0\dots x_{n-2}x_{n-1}]1}|1\rangle)$$

Where,

$$\frac{x_0}{2} \rightarrow 0.x_0$$

$$\frac{x_0}{2^2} + \frac{x_1}{2} \rightarrow 0.x_0x_1$$

$$\frac{x_0}{2^3} + \frac{x_1}{2^2} + \frac{x_2}{2} \rightarrow 0.x_0x_1x_2$$

and so on, to represent a binary fraction. There are two important things to notice here—superposition states and the introduction of phases. We build a circuit with Hadamard gates to introduce superposition states and for the phases, we will use the discrete phase gate. The Quantum Fourier transform circuit applied to a three qubit state is given below [1]:

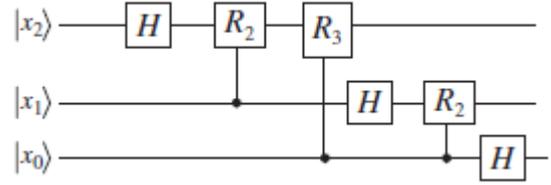


Figure 1. Quantum Fourier Transform

1.2.4. Controlled Gate:

Controlled gate will allow us to implement an *if-else* type of construct with a quantum gate. In a controlled classical gate, we include a control bit  $C$ . If  $C=0$ , then the gate does nothing, but if  $C=1$ , then the gate performs some specified action. Controlled quantum (or *controlled unitary*) gates work in a similar fashion, using a control qubit to determine whether or not a specified unitary action is applied to a target qubit. The specific gate we will use in our circuit is given below,

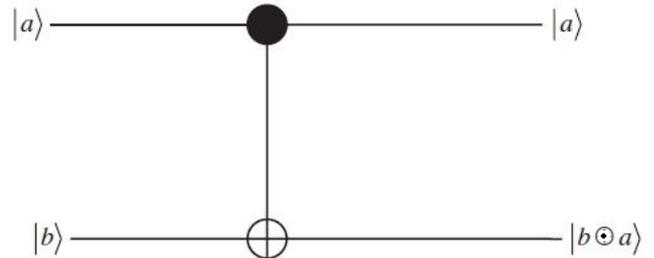


Figure 2. Controlled Quantum Gate

If the control qubit is  $|0\rangle$ , then nothing happens. If the control qubit is  $|1\rangle$ , then whatever is in  $|a\rangle$ , it is passed through the target bit.

The possible input states to the CNOT gate are  $|00\rangle, |01\rangle, |10\rangle$  and  $|11\rangle$  and the action of the Controlled gate on these states is as follows:

$$|00\rangle \rightarrow |01\rangle$$

$$|01\rangle \rightarrow |00\rangle$$

$$|10\rangle \rightarrow |10\rangle$$

$$|11\rangle \rightarrow |11\rangle$$

In our case, we will encounter only the last two cases.

1.2.5. Quantum Comparator:

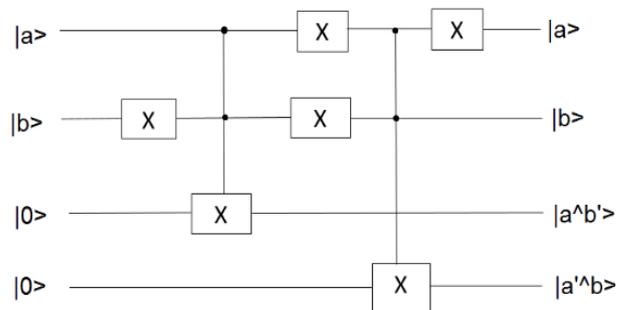


Figure 3. Quantum Comparator Circuit

The above circuit represents a quantum comparator

circuit. The circuit takes a total of four qubits as inputs out of which the last two  $|0\rangle$  are given as garbage bits. We can determine which one of  $|a\rangle$  and  $|b\rangle$  is greater than the other or which one is smaller than the other by monitoring third and fourth output line.

#### 1.2.6. Classical Concept of Clustering:

In classical concept, clustering can be done in various ways such as K-means method or K-medians method.

The procedure for performing K-means method is as follows:

1. Define the number of clusters  $k$ .
2. Initialize the  $k$  cluster centers (randomly, if necessary).
3. Decide the class memberships of the  $N$  objects by assigning them to the nearest cluster center.
4. Re-estimate the  $k$  cluster centers, by assuming the memberships found above are correct.
5. If none of the  $N$  objects changed membership in the last iteration, exit. Otherwise go to 3.

In the following diagram, there are three clusters based on three pre-defined fixed points. All the given input data points are grouped into these three clusters according to the above procedure.

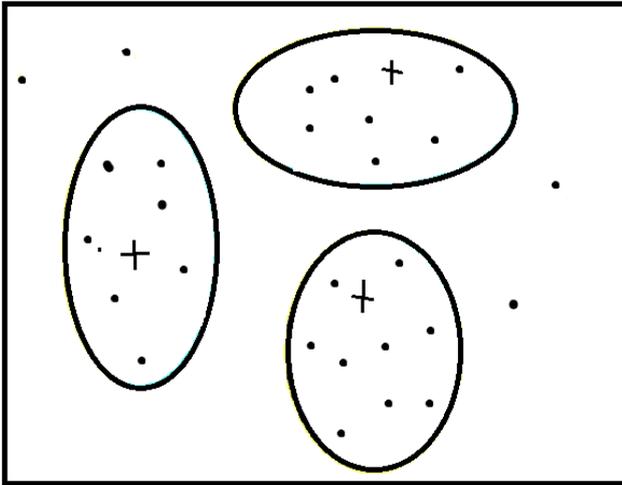


Figure 4. Clusters in K-means method

The outline of the paper is as follows. In Section 1 and 2, we review basics of quantum information processing and some existing quantum clustering approaches. In Section 3, we introduce our proposed work using the basic quantum subtractor, quantum fourier transforms and comparator logic circuits. In this section, the overall block diagram of our proposed approach is also explained. In section 4, a list of comparisons of our approach with other existing approaches is given. Finally, we conclude in Section 5 with a discussion of the issues that we have raised and also address some future works associated with it.

## 2. Literature Survey

In the paper named "Quantum Clustering Algorithms" by Esma Aïmeur, Gilles Brassard and Sébastien Gambs, they explained a way to implement K-medians algorithm using a black box implementation of the following three quantum algorithms: The `quant_find_max` algorithm, `quant_cluster_median` and an algorithm to find  $c$  smallest number. Unsupervised learning is the part of machine learning whose purpose is to give to machines the ability to find some structure hidden within data. Given a training data set, the goal of a clustering algorithm is to group similar datapoints in the same cluster while putting dissimilar datapoints in different clusters. Some possible applications of clustering algorithms include: discovering sociological groups existing within a population, grouping automatically molecules according to their structures, clustering stars according to their galaxies, and gathering news or papers according to their topic. In this part, it seeks to speed-up some classical clustering algorithms by drawing on QIP techniques. Here the concept of Grover's black box model (Oracle) is used for concerning the distance between points of the training dataset. In the classical black-box setting, a query corresponds to asking for the distance between two points  $x_i$  and  $x_j$  by providing indexes  $i$  and  $j$  to the black box. But here in quantum black box ("O"), it is possible to query the quantum black box in superposition of entries. For superposition of entries, Walsh-Hadamard gate is applied to all the input qubits initially set to  $|0\rangle$ . In that case, the resulting output is a superposition of all the triples  $|i, j, \text{Dist}(x_i, x_j)\rangle$ . The distance Oracle is illustrated below,

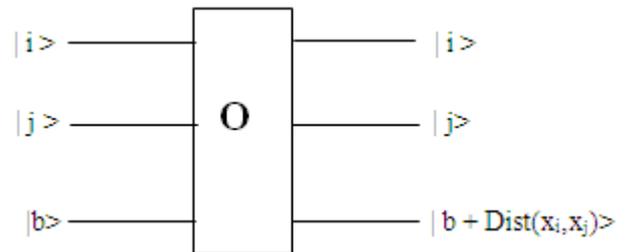


Figure 5. Block diagram of distance Oracle

In the above figure,  $\text{Dist}(x_i, x_j)$  represents the distance between them. The addition  $|b + \text{Dist}(x_i, x_j)\rangle$  is performed in an appropriate finite group between the ancillary register  $b$  and the distance  $\text{Dist}(x_i, x_j)$ . One classical clustering algorithm and its corresponding quantum oracle representation are given in table 3. The algorithm is used to find the pair of points that are farthest apart in the data set i.e. diameter of the dataset. The algorithm proposed by them is  $(n/k)^{.5}$  times faster than the classical approach. In their paper, they have shown how to speed up a selection of classical clustering algorithms by quantizing some of their parts [2] [12].

In the paper named "K-means v/s K-medoids: A Comparative Study" by Shalini S Singh and N C Chauhan, they gave a detailed discussion of two of the various partitioning based clustering methods namely K-means and K-medoids. They also cited various strengths and

weaknesses of both the algorithms and came into a conclusion that the advantage of k-means is its low computation cost, while drawback is sensitivity to noisy data and outliers. On the other hand, k-medoid is not sensitive to noisy data and outliers, but it has high computation cost.

In the paper —Multi-Valued Logic Gates for Quantum Computation by Ashok Muthukrishnan and C. R. Stroud, Jr, he authors have developed a new approach for multi-level quantum system by decomposing any arbitrary unitary operations of  $d$  levels ( $d > 2$ ) into logic gates of two systems by which the number of ions required for computation is reduced by a factor  $\log_2 d$ . In this paper, the authors have efficiently reduced the complexity of the multi valued simulation by a factor of  $\log_2 d$ . But the use of non-logarithmic scaling brings inefficiency in the construction part. A lot of elementary gates are required which makes the circuit a little complex [9].

To enhance the clustering ability of self-organization network, a paper introduces a quantum inspired self-organization clustering algorithm. First, the clustering samples and the weight values in the competitive layer are mapped to the qubits on the Bloch sphere, and then, the winning node is obtained by computing the spherical distance between sample and weight value. Finally, the weight values of the winning nodes and its neighborhood are updated by rotating them to the sample on the Bloch sphere until the convergence [10].

One very new approach is introduced in the paper [7], which is referred to as ‘Dynamic Quantum Clustering’. In this approach, a full-fledged dynamical scheme is introduced which uses a time-dependent Schrodinger equation. Moreover, Hamiltonian formalism is approximated by a truncated calculation within a set of Gaussian wave functions (coherent states) centered on the original points. This allows for analytic evaluation of the time evolution of all such states, opening up the possibility of exploration of relationships among data-points through observation of varying dynamical-distances among points and convergence of points into clusters.

### 3. Proposed Work

The circuit for subtracting two qubits is shown in Fig. 6. The overall circuit for the clustering algorithm is given in Fig. 5.

The various components used in the circuit in Fig. 5 are as follows:

1. Quantum Fourier Transformation Circuitry
2. Quantum Subtractor Circuitry
3. Quantum Inverse Fourier Transform Circuitry
4. Quantum Comparator Circuitry
5. Controlled Quantum Gate
6. Quantum registers for storing Qubits

A set of qubits is given. We need to cluster them with respect to some pre-defined points. Here for simplicity, the circuit diagram is shown for two pre-defined points. At a specific moment, the given stream of input qubits is denoted by  $|a\rangle$ .  $K_1$  and  $K_2$  are the two pre-defined points.  $|a\rangle$  is sent as one of the inputs in both of the Subtractor Circuitry.  $K_1$  and  $K_2$  are passed through the Quantum Fourier Circuitry first and the output of Quantum Fourier Circuitry is sent as the second input of the Subtractor Circuitry. The Subtractor Circuitry calculates the difference between input points and the pre-defined points with respect to which clusters are to be made. The differences from both the Subtractor Circuitry are sent to the Quantum Comparator. Quantum Comparator determines which pre-defined point, the input qubit is nearer to. According to the output of Comparator, the controlled Quantum gates determine in which Cluster, the input Qubit belongs.

To explain how the Subtractor Circuit works, we show a general example below.

Let us consider two integer numbers  $a$  and  $b$  stored in distinct registers, each constituted by  $n$  qubits; if the circuit for qFT is applied to the ket  $|a\rangle = |a_1 a_2 a_3 \dots a_n\rangle$ , the representative state of number  $a$ , its most significant qubit will be in a superposition of the form,

$$|\varphi_1(a)\rangle = \frac{(|0\rangle + e^{2\pi i 0.a_1 a_2 \dots a_n} |1\rangle)}{2^{1/2}}$$

At this point, the -ve phase shift gate is applied.

For simplification, only the state of the most significant qubit during the calculation is shown [3],

$$\begin{aligned} |\varphi_1(a)\rangle &\rightarrow \frac{1}{2^{1/2}} (|0\rangle + e^{2\pi i 0.a_1 a_2 \dots a_n + 0.b_1} |1\rangle) \quad \text{after } R_1 \text{ from } b_1 \\ &\rightarrow \frac{1}{2^{1/2}} (|0\rangle + e^{2\pi i 0.a_1 a_2 \dots a_n + 0.b_1 b_2} |1\rangle) \quad \text{after } R_2 \text{ from } b_2 \\ &\dots\dots\dots \\ &\rightarrow \frac{1}{2^{1/2}} (|0\rangle + e^{2\pi i 0.a_1 a_2 \dots a_n + 0.b_1 b_2 \dots b_n} |1\rangle) \quad \text{after } R_n \text{ from } b_n \\ &= |\varphi_1(a+b)\rangle \end{aligned}$$

A specific example will greatly clarify the above mentioned procedure. Let us consider the integer 3 which, in a binary form is expressed by 11. So it can be coded in state  $|11\rangle$ ; we want to subtract integer 1 which is coded in state  $|01\rangle$ .

$$|11\rangle \xrightarrow{\text{qFT}} \left\{ \begin{array}{l} \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i (\frac{1}{2} + \frac{1}{4})} |1\rangle) \\ \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i (\frac{1}{2})} |1\rangle) \end{array} \right\} \xrightarrow{R}$$

$$\begin{aligned} & \left\{ \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i(\frac{1}{2} + \frac{1}{4} - \frac{1}{4})} |1\rangle) \right. \\ & \left. = \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i(\frac{1}{2} + \frac{1}{2})} |1\rangle) \right. \\ & = \left\{ \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i(\frac{1}{2})} |1\rangle) \right. \xrightarrow{\text{qFT}^{-1}} |10\rangle \\ & \left. \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i(1)} |1\rangle) \right\} \end{aligned}$$

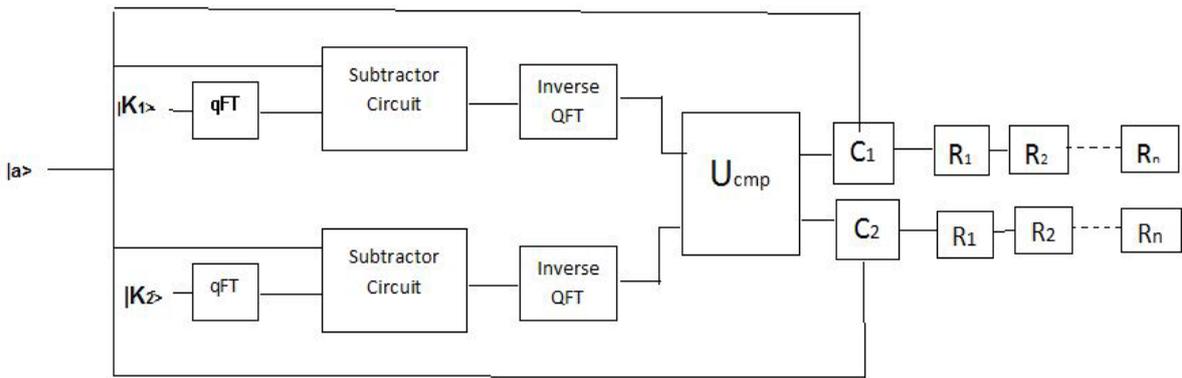
1) Thus state  $|10\rangle$  is obtained that represents integer 2.[3]

Now an overall working of the proposed circuit is shown by taking a specific example.

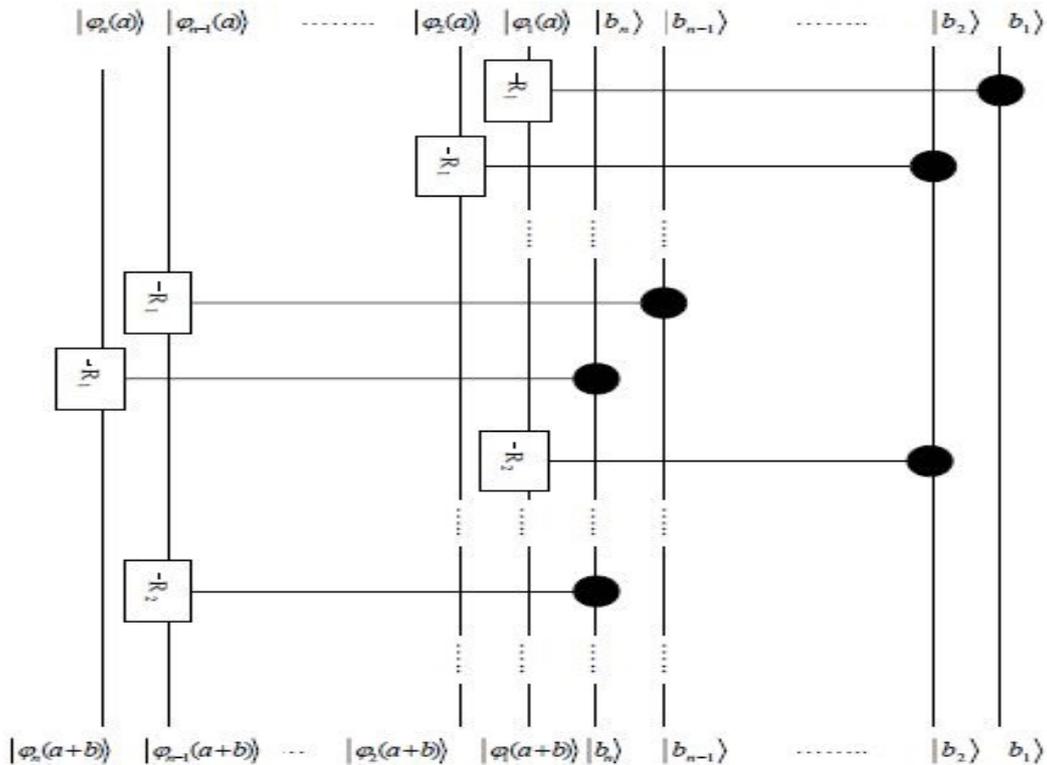
Let the two pre-defined points be  $|011\rangle$  and  $|110\rangle$  and the set of given qubits are  $\{|001\rangle, |010\rangle, |110\rangle$  and  $|111\rangle\}$ ,

**Table 1.** Classification of the input qubits

I/P Qubits	Subtractor with inverse qFT	Ucmp Active O/P line	K <sub>1</sub>  011>	K <sub>2</sub>  110>
001>	011>- 001>= 010>  110>- 001>= 101>	C <sub>1</sub>	✓	
010>	011>- 010>= 001>  110>- 010>= 100>	C <sub>1</sub>	✓	
100>	100>- 011>= 001>  110>- 100>= 010>	C <sub>1</sub>	✓	
111>	111>- 011>= 100>  111>- 110>= 001>	C <sub>2</sub>		✓



**Figure 6.** Block diagram of the Proposed circuit



**Figure 7.** Subtractor Circuit used in the proposed work

**Table 2.** Comparisons of the proposed quantum clustering approach with other existing approaches

Existing Quantum Clustering Approaches	Proposed Quantum Clustering Approach
'Quantum Clustering Algorithms' approach is not necessarily realistic because it requires the availability of a quantum black box that can be used to query the distance between pairs of points in superposition [2].	In proposed approach, we are not using any black box or quantum oracle concept. However, we have built our proposed quantum model using some basic quantum circuits (like, qFT, quantum comparator, quantum subtractor etc.)
Circuit complexity is more in exiting approaches [7, 10].	Circuit complexity is less because it is a combination of small quantum circuits.
Quantum self-organization network clustering algorithm is more powerful than our proposed approach but it is a costly approach in terms of computing efficiency for enhancing clustering ability.	This approach is less costly in terms of computing efficiency.

## 4. Comparisons

The proposed quantum approach has several benefits over the existing approaches to some certain extents. This approach is specially developed to cluster a set of data points which can be applied in various applications where clustering plays a major role. The comparisons details are given in the table 2.

## 5. Conclusion & Future Work

The proposed circuit in this paper only work for one iteration. There is a huge scope of improvement that can be made if the circuit is made to work for multiple iterations. That way we can refine the clusters to greater accuracy. The proposed circuit can be used in complex computational units in future research. In the beginning of 21<sup>st</sup> century, we see a sudden outburst of noisy data which are very difficult to organize by finding correct patterns with the help of traditional data processing tools. Quantization of traditional clustering algorithms can greatly help us in organizing huge amount of data. Though the research done in this field is very limited but the results are promising and there is room for a lot of works to be done.

## REFERENCES

- 
- [1] David McMahon, 'Quantum Computing Explained', ISBN: 78-0-470-09699-4 (cloth), 2008.
- [2] Esma Aïmeur, Gilles Brassard and Sébastien Gambs. 'Quantum Clustering Algorithms', DOI: 10.1145/1273496.1273497 Conference: Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007.
- [3] G. Florio and D. Picca, 'Quantum implementation of elementary arithmetic operations'. INFN (Sezione di Bari) and Dipartimento di Fisica, Università di Bari, Via Orabona, Vol. 4, Bari, Italy, 5 Mar 2004.
- [4] Shalini S Singh N C Chauhan, 'K-means v/s K-medoids: A Comparative Study', B.V.M. Engineering College, V. V. Nagar, Gujarat, India, National Conference on Recent Trends in Engineering & Technology, 13-14 May 2011.
- [5] A. T. Monfared and M. Haghparast, "Design of New Quantum/Reversible Ternary Subtractor Circuits" , Journal of Circuits, Systems, and Computers, World Scientific, Vol. 25, No. 2, 2016.
- [6] Williams, Colin P, 'Explorations in Quantum Computing', ISBN 978-1-84628-886-9, 2011.
- [7] Marvin Weinstein and David Horn, "Dynamic quantum clustering: a method for visual exploration of structures in data", Stanford Linear Accelerator Center, Stanford, CA, USA, 18.08.2009.
- [8] Jiawei Han and Kamber, "Data Mining: Concepts and Techniques", Elsevier, ISBN- 978-1-55860-901-3.
- [9] Ashok Muthukrishnan and C. R. Stroud, Jr., 'Multi-Valued Logic Gates for Quantum Computation', The Institute of Optics, University of Rochester, Rochester, New York 14627, February 1, 2008.
- [10] Ziyang Li1, Panchi Li, "Clustering Algorithm of Quantum Self-Organization Network", 16 June 2015, DOI: 10.4236/ojapps.2015.56028.
- [11] S. Chakraborty, S. B. Mandal, S. H. Shaikh, "Ternary Quantum Circuit for Color Image Representation", 3<sup>rd</sup> International Doctoral Symposium on Applied Computation and Security Systems (ACSS-2016), Springer-AISC, 2016 (accepted).
- [12] S. Chakraborty and L. Dey, "Image Representation, Filtering, and Natural Computing in a Multivalued Quantum System", Handbook of Research on Natural Computing for Optimization Problems, IGI-Global, 2016.