An Efficient Cancer Disease Prediction System through Quantum Computing Technique

Milan Jain
M. Tech. Research scholar
Dept. of Computer Science and Engineering
Technocrats Institute of Technology, Bhopal

Setu Kumar Chaturvedi, Ph.D
Professor & Head
Dept. of Computer Science and Engineering
Technocrats Institute of Technology, Bhopal

ABSTRACT
Diagnosis and exact recommendation for a cancer patient is very much sensitive and of greater influencing in cure of this diseases. Traditionally the diagnosis procedure was very time consuming and not so accurate, so here in this respect there is a need of more accurate, fast and easily to understand method / diagnosis tool which gives reliable results. Here we use the shor algorithm, Hce tool and we get accurate and frequent result, when we adopt the approach of quantum computing with the traditional tools of diagnosis.

Index Terms
Cancer, Cancer classification, Quantum Computing, Quantum Algorithm, Hierarchical Clustering Explorer.

1 INTRODUCTION
A Quantum computing is one of the fastest growing areas of research and technical field, which revolves two area of century computer science and quantum mechanics and both of the field having its own importance. It is used for the calculation for searching anything on internet, modelling the national economy, forecasting the weather etc. A simple computer operate any operation at a time due to this it takes too time for solve the problem and this is overcome by quantum computer [26]. Quantum computing is an emerging technology. The key to improving computer performance has been the reduction of size in the transistors used in modern processors [24]. The quantum computer is one proposal that may have merit in dealing with the problems presented.

The performance of a computer is improved, according to Moore's law, if the performance keeps improving by means of technological innovations, which have occurred over the last few decades, the number of transistors per chip may be doubled every 18 months. Furthermore, processor clock frequency could reach as much as 40 GHz within 10 years [6]. In this way, quantum computers can be used to solve certain computationally intense problems where classical computers require large amounts of processing time. Notwithstanding, further improvements will be necessary to ensure quantum computers' proper performance in future, but such improvements seem obtainable. Currently, there exist some algorithms utilizing the advantage of quantum computers. For instance, the polynomial-time algorithm for factoring a large integer with O (n3) time was proposed by Peter Shor [21]. This algorithm performs factoring exponentially faster than classical computers. This algorithm could factor a 512-bit product in about 3.5 hours with 1 GHz clock rate [20], whereas the number field sieve could factor the same product in 8400 MIPS years [4]. (One MIPS year is the number of instructions that a processor can execute in a year, at the rate of millions of instructions per second.) First, the main characteristics of quantum computers, superposition states, and interference are introduced. In 1982, the Nobel prize-winning physicist Richard Feynman thought up the idea of a 'quantum computers', a device which uses the effects of quantum mechanics to its advantage [9]

In the classical model of computing, the most common building block is bit, it can only lie in two distinct states, 0 or 1. In quantum computers the rules are changed [8],[9],[25]. Not only can a 'quantum bit', usually referred to as a 'qubit', exist in the classical 0 and 1 states, this also be in a coherent superposition of both states. In 1982, Feynman states that quantum computers would be able to simulate quantum mechanics systems with a highly greater degree of accuracy than is possible with classical computers [22].

1.1 Cancer
It stands for malignant neoplasm, which is class of diseases involving out of control cell growth. This can harm the body due to spread of damage cells in different part of body by blood stream. All the tumours are not a cancer tumours, the starting tumour is not entering into neighbouring tissue and not transfer over body. Humans are affected by 200 different known cancers [5].

The cancer is undefined as a result the cells grow uncontrollable. The cancer causes are diverse, tough, and only half understood. There are many reasons which are responsible for cancer like use of tobacco, user diet, alcohol, some infections, radiation, poor in physical activity, and environmental pollutants [3]. These source can injured or combine genes with present genes due to which cells to cause cancerous mutations [18]. Near about 5–10% cancers find directly for inherited genetic defects [2]. The detection of Cancer can be done by number of ways like presence of some signs and symptoms, screen testing, microscopic examination of a tissue sample. There are many possibilities of surviving this disease by its type and where the cancer is situated. Cancer can harm all age's person, and some types of cancer are famous in children and possibility of increasing cancer with age. The survey in 2007 the 13% of all human deaths worldwide (7.9 million) by causing cancer. In the developing world the Rates are increase due to people live to an old age and the change in lifestyle [17].

1.2 History of Cancer
The earliest cancer record from 3000 BC in the Egyptian Edwin Smith Papyrus and the breast cancer [12]. Hippocrates (ca. 460 BC – ca. 370 BC) define different type of cancer and showing to them with the Greek word *karkinos* (crab or crayfish).[152] the doctors of 15th, 16th and 17th centuries accept to dissect bodies to know about the death reason[13].

The physician John Hill state that the cause of nose cancer is tobacco in 1761[13]. This followed in 1775 by surgeon Percivall Pott scrotum of cancer was same disease among chim-

ney sweeps [14]. In the 18th century, it discovered 'cancer poison' pass from the primary stage by lymph nodes to other stage ("metastasis") and first formulated by the English surgeon Campbell De Morgan between 1871 – 1874[11].

2 LITERATURE SURVEY

This section reviews the current literature and related work in the areas of cancer with different methods and technology through examination of various research papers, journals and online resources.

A. E. Kaplan, P. Meystre. In this paper, a method for makes possible the analysis of the most fundamental aspects of light-matter interaction, and promises uprecedented accuracy in the determinant. [1]

Craig S. Lent and P. Douglas Tougaw. This paper provides a paradigm for computing with interacting quantum dots, quantum-dot cellular automata (QCA). Arrays of quantum-dot cells could be used to perform useful computations. [7]

Fortunato Bianconi, Valerio Brunori, Paolo Valigi, Francesco La Rosa, and Fabrizio Stracci. In this paper, it present a modern web-based management system that allows to integrate different sourcing, validation and data elaboration thus providing a new evaluation system for the oncology network based on cancer registries.[10]

Hualong Yu1, Jun Ni2, Yuanyuan Dan3, Sen Xu4. This paper presents a skewed gene selection algorithm that introduces a weighted metric into the gene procedure. [15]

Huseyin Seker, Student Member, Michael O. Odetayo, Dobrila Petrovic, and Raouf N. G. Naguib. The aim of this paper is to investigate the fuzzy –nearest neighbor (FK-NN) classifier as a fuzzy logic method that provides a certainty degree for prognostic decision and its markering. [16]

Lipo Wang, Feng Chu, and Wei Xie. This paper is an attempt to finding the smallest set of genes that can ensure highly accurate classification of cancers from microarray data by using SVM algorithms. [19]

Shi Kuifan , Chen Yuehui , Dong Jiwen . This paper briefly reviews some neural networks and discusses aboutdisadvantages. [23]

3 PROPOSED METHODOLOGY

Here user try to collaborate these two the science of quantum computing and the techniques of cancer detection under the same hut, for accuracy of medicine technique of cancer detection and for faster the process which in the mean time benefit the different folks who fight the cancer for welfare of mankind.

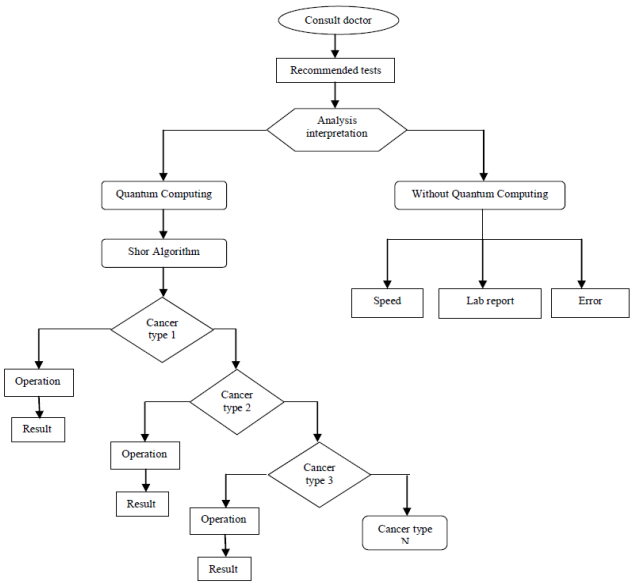


Figure 3.1 Proposed model for cancer by Shor Technique

3.1 Proposed Algorithm

Basically seven steps are used which are given below

Step1 Apply Cancer tumour size a.

Step2 Apply the number N is the number we wish to factorize.

Step3 Apply random number X, where 1<X<N-1.

Step4 Result by Shor technique.

Step5 Comparing cancer value and proposed values.

Step6 Result.

4 EXPERIMENTAL SETUP

The Hierarchical Clustering Explorer (HCE) is a visualization tool for interactive exploration and analyzing multiple attribute datasets. The tool was developed by the Human-Computer Interaction Lab at the University of Maryland. The tool allows users to load their own datasets and explore them through different visualization and statistical methods. When users first load their dataset, they are presented with an overview the data is shown in the form of colour mosaic that is able to get more details of the data with manipulating different views when we needed. For further help of users it evaluates their different hierarchical clustering dataset, comparison and performing dynamic queries. The colour mosaic consists as its main window, a graphical representation of a table which maps each element's attributes to a colour value. Users can apply a hierarchical clustering algorithm by choosing their own parameters. The clustering results are displayed in a dendrogram form in the color mosaic view.

4.1 Data Filtering

4.1.1 Present call filtering

There are two outputs from the noise calculations; one is the continuous p value assignment, and the other is a simple present/absent threshold. When the probe set detection p value reaches a certain level of significance, then the probe set is assigned a current call and other whose probe sets having less noise ratios are assigned as absent call.

4.1.2 Standard deviation filtering

Users can filter out rows based on the standard deviation. The default threshold is 1.

4.2 Data Transformation

4.2.1 Log transformation (Natural log)

Users sometimes want to transform the variable to get a good output.

4.2.2 Normalization

Users can normalize the input data either **row-to-row** or **column-to-column**.

$\frac{x-m}{\sigma}$	Values will be standardized, means calculate mean from deviation and then divide the deviation by the standard deviation. After standardization, each row (, or column) will have the same mean (0) and the same standard deviation (1).
$\frac{x}{control}$	Simply divide values by the value which is from first column or row.
$\frac{x}{median}$	Simply divide values by the median.
rescale to a new range	Linearly transform each row or each column to get the values of new range.

Table 4.1 Normalization of input data

4.3 Linkage method

When hierarchical clustering algorithm merges two clusters to generate a new bigger cluster, it should calculate the distances of the previous and next clusters. We have 5 different linkage methods. Let C_n be a new cluster, a merge of C_i and C_j. Let C_k be a left. Average Linkage (UPGMA : Unweighted Pair Group Method with Arithmetic Mean)

$$DIST(C_n, C_k) = \frac{|C_i|}{|C_i|+|C_j|} DIST(C_i, C_k) + \frac{|C_j|}{|C_i|+|C_j|} DIST(C_j, C_k) \quad (1)$$

- Average Group Linkage (Centroid Linkage) :

$$DIST(C_n, C_k) = DIST(Mean(C_n), Mean(C_k)) \quad (2)$$

- Complete Linkage :

$$DIST(C_n, C_k) = Max(DIST(C_i, C_k), DIST(C_j, C_k)) \quad (3)$$

- Single Linkage :

$$DIST(C_n, C_k) = Min(DIST(C_i, C_k), DIST(C_j, C_k)) \quad (4)$$

- Shneiderman's 1by1 Linkage :

Let C_{n-1} be the newly merged cluster in the previous iteration. Let C_m be the closest cluster to C_{n-1}, and C_p be the closest cluster to C_m.

If $|DIST(C_{n-1},C_m) - DIST(C_m,C_p)|<THRESHOLD$, merge C_{n-1} and C_m instead of searching two new closest clusters globally.

4.4 Choose a distance/similarity measure

The formula for Pearson's correlation is as follows.

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{(\sum X)^2}{N})(\sum Y^2 - \frac{(\sum Y)^2}{N})}}$$

It can be used if the numbers are converted into z scores:

$$r = \frac{\sum z_x z_y}{N}$$

where z_x is the variable X converted into z scores and z_y is the variable Y converted into z scores.

4.5 Experiment Results and Analysis

4.5.1 Experiments Performed on Original Dataset of Cancer

ID	TRT	AGE	WEIGHIN	STAGE	TOTAL-CIN	TOTALCW2	TOTALCW4	TOTALCW6
1	0	52	124	2	6	6	6	7
11	0	59	175.8	2	6	7	16	13
15	0	69	167.6	1	6	6	6	11
26	0	56	158	3	6	11	15	15
39	0	46	149	4	7	8	11	11
14	1	42	162.6	1	4	6	8	7
16	1	44	261.4	2	6	11	11	14
24	1	68	226	4	12	11	12	9
42	1	73	181.5	0	8	11	16	.
44	1	67	187	1	5	7	7	7
50	1	60	164	2	6	8	16	.
58	1	54	172.8	4	7	8	10	8

Table 4.2 Dataset of Traditional Cancer Technique

4.5.2 Experiments Performed on Proposed Dataset of cancer

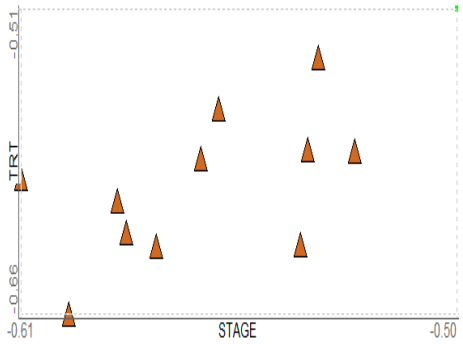
NAME	ID	AGE	a	X	N	SHOR
Ankya	22	20	2	2	3	1
Sujit	102	81	2	4	7	2
Dinesh	46	41	3	5	8	5
Ruchi	68	55	7	6	7	6
Kiran	88	33	1	7	9	0
Simple	70	58	3	8	10	2
Rohan	114	75	4	9	11	5
Pradeep	52	41	4	10	13	3
Servesh	125	34	1	4	9	0
Vinay	75	43	6	6	11	5
Asish	190	46	4	5	9	4
Shati	76	56	1	7	10	0
Upendra	108	62	3	7	10	3
Kapil	77	45	3	7	11	2
Sanjeev	16	47	4	8	11	4
Sanchita	110	62	2	9	13	3
Shobhit	79	50	3	2	5	3
Nirupma	40	50	2	11	15	1
Sheetal	41	87	4	12	17	3
Rakesh	21	52	2	4	2	0

Table 4.3 Dataset by Shor Technique

5 RESULTS

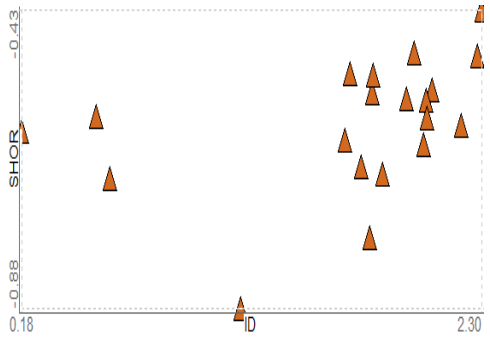
The comparative results analysis of Traditional Cancer Technique and shor technique.

5.1 Dendrogram Display by Traditional Cancer Technique



Graph 5.1 Results for Traditional Cancer Technique

5.2 Dendrogram Display by Shor Technique



Graph 5.2 Results for Shor Technique

5.3 Table View by Traditional Cancer Technique

ID	NAME	TOTALCIN	TOTALCW2	TOTALCW6	TOTALCW4	TRT	STAGE	AGE	WEIGHT
24.000000	24.000000	-0.43	-0.44	-0.47	-0.43	-0.58	-0.54	0.35	2.54
58.000000	58.000000	-0.47	-0.46	-0.46	-0.42	-0.58	-0.53	0.38	2.53
39.000000	39.000000	-0.48	-0.46	-0.39	-0.39	-0.63	-0.54	0.35	2.54
14.000000	14.000000	-0.48	-0.44	-0.42	-0.40	-0.54	-0.54	0.25	2.57
11.000000	11.000000	-0.51	-0.50	-0.39	-0.34	-0.62	-0.59	0.43	2.51
26.000000	26.000000	-0.54	-0.44	-0.36	-0.36	-0.66	-0.60	0.46	2.50
1.000000	1.000000	-0.48	-0.48	-0.45	-0.48	-0.63	-0.58	0.66	2.44
15.000000	15.000000	-0.50	-0.50	-0.41	-0.50	-0.61	-0.59	0.65	2.44
44.000000	44.000000	-0.50	-0.46	-0.46	-0.46	-0.56	-0.56	0.52	2.49
42.000000	42.000000	-0.48	-0.43	-0.61	-0.34	-0.60	-0.61	0.62	2.45
50.000000	50.000000	-0.49	-0.45	-0.60	-0.30	-0.59	-0.57	0.52	2.48
16.000000	16.000000	-0.45	-0.39	-0.36	-0.39	-0.51	-0.50	0.00	2.61

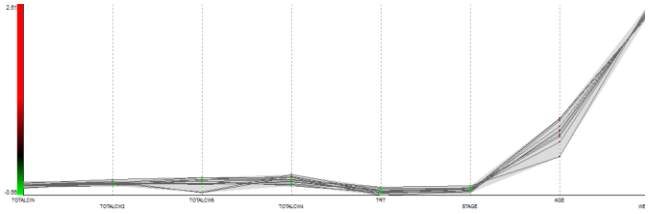
5.4 Table view by Shor

NAME	X	SHOR	A	N	ID	AGE
Sheetal	-0.66	-0.69	-0.39	-0.22	0.59	2.13
Rakesh	-0.49	-0.59	-0.38	-0.49	0.53	2.18
Nirupma	-0.83	-0.88	-0.35	-0.14	1.19	1.72
Shobhit	-0.50	-0.50	-0.53	-0.43	1.99	1.04
Sanchita	-0.62	-0.60	-0.45	-0.35	2.05	0.86
Sanjeev	-0.62	-0.62	-0.35	-0.15	0.18	2.25
Kapil	-0.60	-0.64	-0.46	-0.32	2.03	0.89
Upendra	-0.57	-0.57	-0.47	-0.39	2.04	0.90
Shati	-0.65	-0.68	-0.45	-0.35	1.84	1.18
Asish	-0.43	-0.43	-0.42	-0.36	2.30	0.18
Vinay	-0.52	-0.55	-0.52	-0.33	2.07	0.87
Servesh	-0.48	-0.50	-0.41	-0.30	2.28	0.25
Pradeep	-0.72	-0.78	-0.41	-0.25	1.78	1.21
Rohan	-0.59	-0.57	-0.47	-0.43	1.95	1.05
Simple	-0.64	-0.67	-0.46	-0.39	1.75	1.32
Kiran	-0.57	-0.61	-0.38	-0.32	2.20	0.45
Ruchi	-0.49	-0.53	-0.53	-0.49	1.80	1.31
Dinesh	-0.64	-0.53	-0.53	-0.37	1.70	1.42

Sujit	-0.56	-0.56	-0.51	-0.44	1.80	1.30
Ankya	-0.52	-0.63	-0.52	-0.41	1.67	1.45

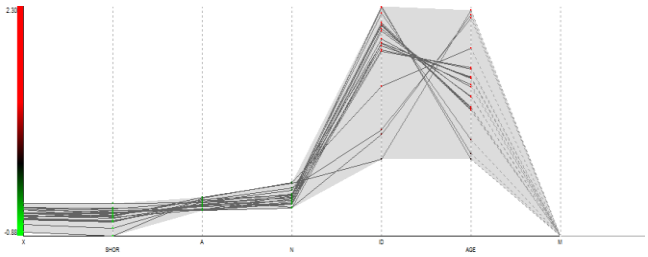
Table 5.3 shows the result of Traditional Cancer Technique which shows Minimum and maximum value of dataset, displayed using bar graph in following graph 5.1. Table 5.4 shows the result of Shor Technique which shows Minimum and maximum value of dataset, displayed using bar graph in following graph 5.2.

5.5 Profile Search by Traditional Cancer Diagnosis



Graph 5.3 Result for parallel patient by Traditional Cancer Diagnosis

5.6 Profile Search by Shor Technique



Graph 5.4 Result for parallel patient by Shor Cancer Diagnosis

Figure 5.3 shows the result of Profile search by Traditional Cancer Technique and this graph shows the report of all the Cancer Patient Diagnosis by Cancer Technique and Figure 5.4 shows the result of Profile search by Shor Technique and this graph shows the report of all the Cancer Patient Diagnosis by Shor Technique.

6 CONCLUSIONS

With classical computers gradually approaching their limit, the quantum computing promises to deliver a new level of computational power. A quantum computer have power to perform calculations across the multitude of parallel universes gives it the ability to quickly perform tasks that classical computer will never be able to practically achieve.

A quantum computing user start work after diagnoses of cancer, and by quantum analysis we get exact dataset and with the help of this we can increase the accessing speed .It provide an interrelation and make a more appropriate approach for accurate disease detection.

7 REFERENCES

[1] A. E. Kaplan, P. Meystre "Quantum Electronics on the Quantum and Nonlinear Optics of Single Atoms, Ions, and Electrons". In IEEE Journal of quantum electronics, vol. qe-23, no. 1, January1987.

[2] American Cancer Society. Retrieved July 22, 2013.

[3] Anand P, Kunnumakkara AB, Kunnumakara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, Sung B, Aggarwal BB ,"Cancer is a preventable disease that requires major lifestyle changes". *Pharm. Res.* **25** (9): 2097–116. doi:10.1007/s11095-008-9661-9. PMC 2515569. PMID 18626751, September 2008.

[4] B. Preneel (Ed.), Factorization of a 512-bit RSA modules, Lecture Notes in Computer Science, Vol. 1807 (Berlin: Springer- Verlag, 2000).

[5] Cancer Research UK "How many different types of cancer are there, CancerHelp UK", Retrieved 11 May 2012.

[6] C.P. Williams & S.H. Clearwater, Exploration in quantum computing (New York: Springer-Verlag, 1997).

[7] Craig S. Lent and P. Douglas Tougaw "A Device Architecture for Computing with Quantum Dots" .In Proceeding of the IEEE, Vol. 85, No. 4, April 1997.

[8]Cue the qubits: Quantum computing - How to make a quantum computer. The Economist Volume 342 Issue 8005

[9] David Deutsch ,Quantum computation., Physics World, 1/6/92

[10] Fortunato Bianconi, Valerio Brunori, Paolo Valigi ,Francesco La Rosa, and Fabrizio Stracci, Information Technology as Tools for Cancer Registry and Regional Cancer Network Integration, IEEE Transaction on system, man and cybernetics-part A:systems and humans, VOL. 42, NO. 6, November 2012.

[11] Grange JM, Stanford JL, Stanford CA ,"Campbell De Morgan's 'Observations on cancer', and their relevance today". Journal of the Royal Society of Medicine **95** (6): 296–9. doi:10.1258/jrsm.95.6.296. PMC 1279913. PMID 12042378, (2002).

[12] Hajdu SI, Thun, MJ, Hannan, LM, Jemal, "A note from history: landmarks in history of cancer, part 1, *Cancer* **117** (5): 1097–102. doi:10.1002/cncr.25553. PMID 20960499, (March 2011).

[13] Hajdu SI, Thun, MJ, Hannan, LM, Jemal, "A note from history: landmarks in history of cancer, part 2.". *Cancer* **117** (12): 2811–20. doi:10.1002/cncr.25825. PMID 21656759 ,(June 2011).

[14] Hajdu SI, Thun, MJ, Hannan, LM, Jemal, "A note from history: Landmarks in history of cancer, part 3.". *Cancer* **118** (4): 1155–68. doi:10.1002/cncr.26320. PMID 21751192 , (July 2011).

[15] Hualong Yu1, Jun Ni2, Yuanyuan Dan3, Sen Xu4, Mining and Integrating Reliable Decision Rules for Imbalanced Cancer Gene Expression Data Sets, Tsinghua Sci-

ence and technology ISSNll1007-0214ll07/10llpp666-673 Volume 17, Number 6, December 2012

[16] Huseyin Seker, Michael O. Odetayo, Dobrila Petrovic, and Raouf N. G. Naguib, A Fuzzy Logic Based-Method for Prognostic Decision Making in Breast and Prostate Cancers, IEEE transaction on information technology in biomedicine , VOL. 7, NO. 2, JUNE 2003.

[17] Jemal A, Bray, F, Center, MM, Ferlay, J, Ward, E, Forman, D "Global cancer statistics", *CA: a cancer journal for clinicians* **61** (2): 69–90. doi:10.3322/caac.20107. PMID 21296855, (February 2011).

[18] Kinzler, Kenneth W.; Vogelstein, Bert, *The genetic basis of human cancer* , New York:McGraw-Hill, Medical Pub. Division. p. 5. ISBN 978-0-07-137050-9,(2002).

[19] Lipo Wang, Feng Chu, and Wei Xie "Accurate Cancer Classification Using Expressions of Very Few Genes". In IEEE/ACM Transaction On Computational biology and bioinformatics, Vol. 4, No. 1, January-March 2007.

[20]M. Oskin, F.T. Chong, & I. Chuang, A practical architecture for reliable quantum computers, IEEE Computer, 79–87, January2002.

[21] P.W. Shor, Algorithm for quantum computation: Discrete logarithm and factoring, Proc. 35th IEEE Annual Symp. On Foundations of Computer Science, Santa Fe, NM, 24–134, November 1994.

[22] Seth Lloyd ,Universal Quantum Simulators, Science 23/8/96

[23] Shi Kuifan , Chen Yuehui , Dong Jiwen, Classification of Cancer Cells Using Neural Networks in Combination with Expert Experience, Tsinghua Science and technology,ISSN 1007-0214 22/ 23 pp853-855 Volume2, Numberi December 1997

[24] Tom Thompson ,When silicon hits its limits, Byte 1/4/96

[25]Two-bit heroes - Computing with quanta. The Economist Volume 338 Issue 7948

[26]Vishal Sahni, Quantum Computing, Tata McGraw Hill Education Private Limited New Delhi (2010)