Print ISSN 0974-0678, Online : 2320-9593

BIONANO FRONTIER

Vol. 10 Issue - 2

July - December 2017

UGC, APPROVED JOURNAL

NAAS RATING 4.17
IMPACT FACTOR 4.856

A Peer Reviewed / Listed by UGC / Indexed Journal
of Science & Technology

Published From : Mumbai
www.bionanofrontier.org

Print ISSN 0974-0678, Online : 2320-9593

BIONANO FRONTIER

A Peer Reviewed / Listed by UGC / Indexed Journal of Science & Technology

Journal of Bionano Frontier Index / Abstracted / Cited / Listed With : Indian Science Abstract , Medicinal and Aromatic Plant Abstract (MAPA / CSIR), NAAS, Medline, Google Search Engine, Genamics- Journal Seek Newzealand, E-ISRJC, Phillipines International Scientific Citation Agency, Indian Journal Seeker, Revistas Cientificas (Barcelona, Spain), & Connect e-Journals.)

OBJECTIVE AND SCOPE:

BIONANO FRONTIER : The main objective of this multi-disciplinary peer reviewed journal is to encourage the academicians and researchers those who are actively engaged in research specially in Bio-Science, Applied Biology, Microbiology, Botany, Zoology, Biochemistry, Biotechnology, Cell Biology, Molecular Biology, Genetics, Immunology, Pathology, Neurobiology, Proteomics & Genomics, Agricultural Sciences, Seed Technology, Plant Pathology & Entomology, Plant & Animal Physiology, Cell and Tissue Cultures, Nutrition, Fermentation technology, Enzymology, Agri-Biotechnology, Cytology, Hydrology, Pharmacology, Pharmaceutics, Pharmacognosy, Pharmacogenomics, Pharmacy Practice, Natural Chemistry, Analytical Chemistry, Synthetic Chemistry, Physical Chemistry, Organic & Inorganic Chemistry, Environmental Chemistry, Phyto Chemistry, Soil Chemistry, Industrial Chemistry, Medicinal Chemistry, Food Chemistry, Biopharmaceutics, Bioanalytical Chemistry, Polymer Chemistry, Drug discovery, Novel Drug delivery systems, Nuclear Pharmacy, Chemoinformatics, Nanotechnology, Bioinformatics, Environmental Sciences, Agriculture and allied branches of clinical research in Medical and Pharmaceutical field. The Journal would publish original articles, review articles. All articles published in "**BIONANO FRONTIER**" are evaluated by the experts, however ; the Editors and Publisher are not always in agreement with the opinions views expressed by the contributors/authors in their articles and take no responsibility for inaccurate, misleading data.

Subscription Rates :

Individual:	1 year Rs.800/-
Institution	1 year Rs.800/-
Life Membership Individual	Rs.5000/-

Address For Correspondence:
Dr. P. M. Dongre - Editor
C/O. Department of Bio-Physics, University of Mumbai,
Vidyanagari, Santacruz (E), Mumbai - 400098.
09969051198, Email: jbionano@gmail.com
Dr. V M Jamdhade - Executive Editor
C. /O. Department of Botany
B N Bandodkar College of Science, Bunder Road,
Chendani, Thane (W) 400 601
09869992925 / 09987653712, Email: jbionano@gmail.com
www.bionanofrontier.org

Editor : Dr. P. M. Dongre, C/o. Department of Bio - Physics, University of Mumbai, Kalina, Vidyanagari, Mumbai - 400 098 (M.S.) **Printed & Published by** Dr. V. M. Jamdhade and Printed at Laxmi Printing Press, B-28, Ganesh Tower, Near Railway Station, Thane (W) - 400 602.

SOFTWARE (REMOTE SENSING (RS) AND GEOGRAPHICAL INFORMATION SYSTEM (GIS) TECHNOLOGIES) ORIENTED MORPHOMETRIC ANALYSIS OF KR-34 WATERSHED IN WESTERN MAHARASHTRA, INDIA.

UJWALA KOLTE*
*Assistant Professor
Indira College of Commerce and Science

Abstract : To study scope for watershed development and management morphometric features of Krishna River (KR) 34 watershed have been selected. Geographical Information System (GIS) based morphometric analysishas been carried out to evaluate linear, relief and aerial morphometric parameters of the KR 25 watershed usingsoftware's Geomedia Professional 5.0, ERDAS Imagine 9.1, and ArcGIS 9.2,to evaluate linear, relief and aerial morphometric parameters of the KR 25 watershed.

Analysis reveals that, KR 34 watershed area of Basin (Km2) is 266.753 and have 5th stream order. Total length of basin (Long Axis) (Km) is 21.147 Km. Mean stream length of the watershed is (Lsm) 0.64. Bifurcation ratio (Rb) value is stand for 6.71. The circulatory ratio value 0.52 reveals that, KR 34 watershed is elongated in shape where infiltration rate is very high. Elongation ratio is 0.87have wide chances to percolate water in entire watershed. The form factor value is 0.60, which is higher reveals that, watershed intervention will definitely improve water level. It is also noted that, watershed development in such area gave effective results on the occurrence of the ground water.

Keywords : Geographical Information System (GIS), Morphometric analysis, watershed, Krishna River Basin.

INTRODUCTION:

Morphometric of watershed is defined as the measurement andmathematical analysis of the configuration, the earth's surface, shape and dimension of its landforms.Most of the morphometric studies are became integral part of watershed program. GIS based studies important in soil loss assessment and planning the watershed development program, fixing up priority areas for conservation by designing the watershed activities. Morphometric analysis were carried out in a number of Indian watersheds and subsequently used for water resources development and management projects as well as for watershed characterization and prioritization (Chalam et al. 1996, Singh and Singh, 1997, Chaudhary and Sharma 1998).

Scientific planning of the watersheds using Geographical Information System (GIS) is became an integral part of the watershed development program at government level but not efficiently used. Spatial information technology (SIT) i.e. remote sensing (RS), Geographical information system (GIS) has proved to be efficient tools in the delineation of drainage pattern and water resource management. The Global Positioning System (GPS) is also proved to be efficient satellite-based navigation system made up of a network of 24 satellites placed into orbit by the U.S. Department of Defense. GPS was originally intended for military applications, but in the 1980s, the government made the system available for civilian use. GPS works in any weather conditions, anywhere in the world, 24 hours a day. GPS technology is also proved to be innovative in the morphometric analysis. (Bharat Kakade,2009)

Systematic description of the geometry of a drainage basin and its stream-channel system requires measurement of linear aspects of the drainage network, aerial aspects of the drainage basin, and relief aspects of channel network and contributing ground slopes.(Nageswara Rao ,2010)Drainageanalysis is also known as fluvial morphometric which provides the information regarding the factors which control the development of the drainage. Morphometric analysis measures of the shape of the basin, area of the basin and the length of the stream. On the basis of projection of the system to horizontal plane, the linear properties such as length, area, arrangement etc. are calculated. This type of study is 'planimetric' which means measures in a single plane. (Nageswara Rao, 2010). To understand the morphometric features Geographical Information System (GIS) based morphometric analysis has been carried out to evaluate linear, relief and aerial morphometric parameters of the KR 34watershed usingsoftware's Geomedia Professional 5.0, ERDAS Imagine 9.1, and ArcGIS 9.2,to evaluate linear, relief and aerial morphometric parameters of the KR 34watershed. The objectives of present research are as follows

1.1 Objective of Research

I To analyze linear, relief and aerial morphometric parameters of the KR 34watershed

ii. To study co relevance of morphometric parameters and watershed development in KR 34 watershed

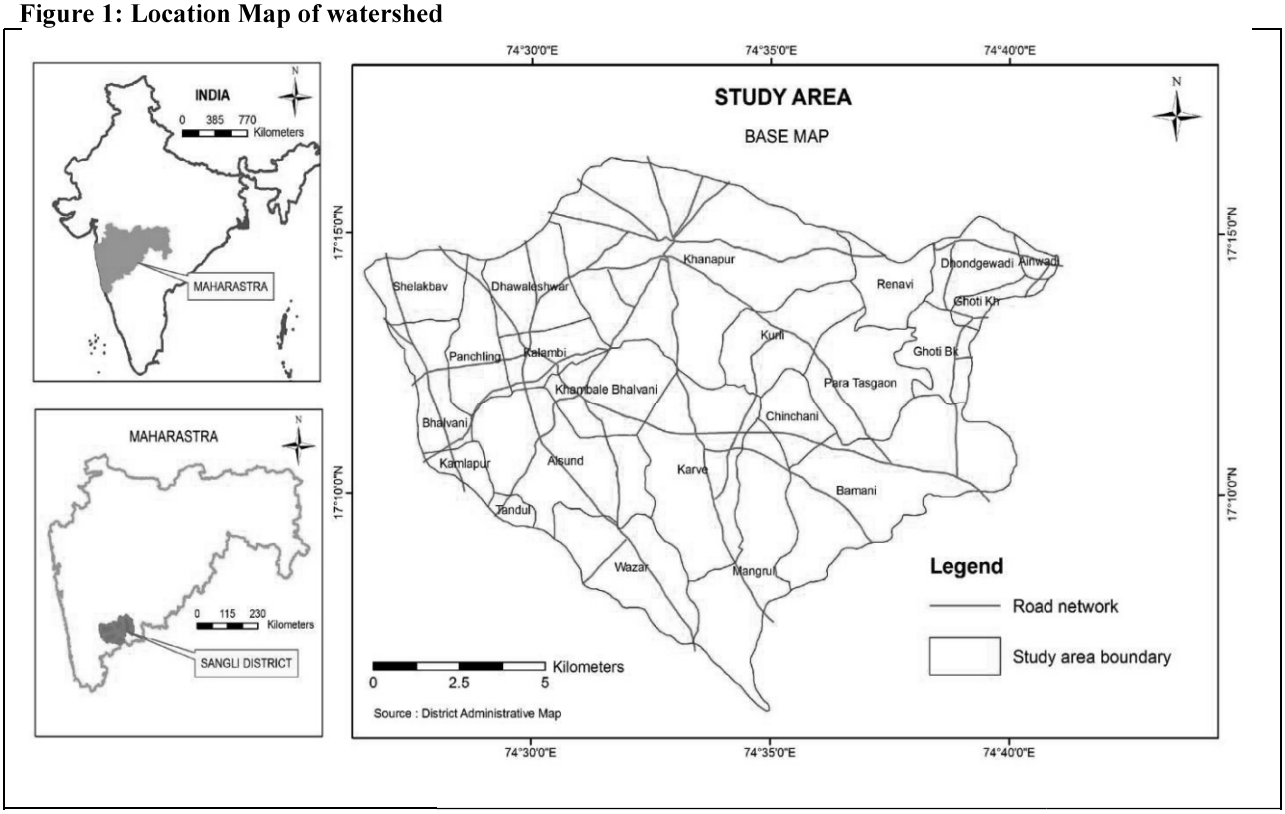
iii.To analyze potential of Geographical Information System (GIS) in planning of watershed development and management.

GIS based morphometric analysis is need to carry forward while planning watershed development activities. Considering same aspect morphometric analysis have been carried to understand nature of KR 34 watershed.

1.2. Location of the study area

The Sangli district is one of the southern-most district of Maharashtra state. It is stretches between the latitudes of 16°45' N and 17°33' N and longitudes of 73°41' E and 75°41' E. Sangli district is made up of 10 Talukas including 730 villages and 8 cities. Total geographical area of district is 8, 65,404 hectares. Out of this 6, 25,000 hectors area is under cultivation. The area under forest cover is 47,000 hectares.The Krishna basin and Bhima (Tributary of Krishna) occupies the major portion of the district. Krishna basin occupies the area of 69.03 lakh hectare which is 22% of the total area of the Maharashtra state. The KR 34 watershed fall under Krishna basin have area of 266.753 km^2.**(**Figure 1**)**

Figure 1: Location Map of watershed



Software used-

In this study the morphometric analysis of KR 34watershed were done by following a systematic method, comprising the software packages of Geomedia Professional 5.0 and ArcGIS 9.2. For this study, different types of data have been collected from the concern departments. All these data come under two categories such as

1) Primary Data 2) Secondary data - As described below

2.1 Primary Data:-Satellite Digital data acquired by Landsat ETMP (Enhanced Thematic Mapper Plus) sensor acquired for January 2006 has been used as primary data source. From this 30m spatial resolution imagery all types of spatial information such as land use/land cover, water bodies, morphological features, drainage network etc. are collected by using digital image processing and image classification module of ERDAS Imagine 9.1 software.

2.3 Secondary Data:- Survey of India (SOI) Toposheets of KR 34 watershed of Sangli district which stretches betweenlatitudes of 17 , 05'- 0' N and 74°,25'- ' Ewith the scale of 1:50000 published during 1973 to 1979 are used to extract contours, drainages and base maps. The administrative maps are prepared from the data collected from Collecterate of Sangli District in Maharashtra. Geology, geomorphology and soil thematic maps are collected from the Ground Water Survey and Development Agency, Pune, GIS cell in Central Building, Pune, and Agricultural Engineering Department of Sangli District.

3.Physiography and Climate of study area

In Maharashtra, there is a wide spatial and temporal variation in the distribution of rainfall. The State receives more than 90% of the rainfall from the SW Monsoon between the months of June to September. Based on the distribution of rainfall, the State can be divided into three distinct zones. These are: High rainfall zone (> 3,000 mm), Low Rainfall zone (400 – 750 mm) and Assured rainfall zone (1,000 – 1,250 mm). The high rainfall zone is limited to the Konkan coastal belt, adjacent Sahyadri hill ranges and to the western parts of the Kolhapur, Satara, Nashik and Pune districts. Towards the west of the arid region the rain spell gap ranges between 14 to 20 days and as such occupies a very narrow belt. The belt therefore does not experience severe droughts and scarcity. Towards the east, the rainfall decreases rapidly and falls under the rain shadow zone that experiences frequent droughts and scarcity. This region is therefore designated as the drought prone region of the state. Studies related to the rainfall duration and gap between consecutive rain spells have revealed that the axis of the arid region (Dhokarikar, 1991) is located in areas with the largest gap (25-30 days). The KR 34 watershed fall in the scanty rainfall region.Drought proneness of KR 34 (Khanapur and Atpadi block in Sangli district) is calculated percentage of drought years to total years considered for the analysis itis Khanapur and Atpadi (45.8%).

4. Geo-hydrology

Ground water occurs in unconfined conditions in near surface weathered or unweathered vesicular /fractured zone within a depth of 20 meter. Major portion of KR 34 fall in the Deccan proper, the DTW of > 30 m bgl are encountered in Umarga and Paranda taluka of Osmanabad district, Mohol and North Solapur talukas of Solapur district, Jat, Kavathe-Mahankal, Kadegaon and Khanapur (KR 34 watershed occupies major area of Khanapur and Atpadi talukas) taluka from Sangli district and Udgir taluka of Latur district.Similarly, isolated pockets of groundwater depletion from Bhor, Purandar, Junnar and Shirur talukas (Pune district),

Niphad and Chandwad from Nashik district, Kopargaon, Rahuri and Nagar (Ahmednagar district) where DTW > 25 m bgl were also noticed.The water levels in these wells showed seasonal fluctuations of up to 5.5 m. Even moderate intensity droughts affect sharp decline in the water levels in these dug wells. Consequently, depletion in water levels between 2 to 4 m and even more as compared to the average post monsoon levels of previous five years is common. In the KR 34watershed average fluctuation is 3.40 m in the command area and 5.41 meter in the non-command area. The stage of ground water development revealed that, it is up-to 57.52%, therefore there is no significant decline in ground water level and it is recorded as rising watershed is fall in the safe category.

5. Topography and Slope

The KR 34watersheds and its environs are located in the eastern part of Sangli district, having a matured topography reflected by undulating terrain with hills, knolls and inselbergs of Western Ghats and composed of relatively weathered and un-weathered rocks.The maximum and minimum elevations of watershed are 884and 69meter from Mean Sea Level (msl) respectively. The relative relief of the watershed is 3 . Slopes are one of the most important terrain characteristics and play a vital role in geomorphologic and run-off processes, soil erosion and land use planning. The general slope of KR 34 watershed in the direction of North- South side.

6. Geology of KR 34 watershed

Geology of the KR 34 watershed is made up of vesicular jointed basalt (V.J.B). KR 34 watershed fall in the Non- Command area. Aquifer in soft rock areas and depth of weathered zone and or maximum depth of fractures under unconfined zone is located on 13.10 meter(GSDA, Sangli 2008-2009).

RESULTS AND DISCUSSION :

7.1 Morphometric analysis of KR 34 watershed

Morphometric analysis is carried out for the shape of the basin, area of the basin and the length of the stream. On the basis of projection of the system to horizontal plane, the linear properties such as length, area, arrangement etc. are calculated. The plan metric' measurements were carried out in a single plane.

7.2 Liner Aspects

The linear aspects Stream order, stream length, mean stream length, stream length ratio and bifurcation ratio etc. are linear aspects that were determined and results have been given in tabulated formTable 1.

Table 1: Linear aspects of the KR-34, watershed is as follows.

Sr.No	Morphometric parameters	Watershed No. KR-34
1.	Stream order	5th
2.	Total Stream length (Lu)	610.759
3.	Average of Mean stream length (Lsm)	0.64
4.	Mean Stream length ratio (RL)	0.49
5.	Bifurcation ratio (Rb)	6.71
5.	Relief Ratio (Rh)	14.90
6.	Drainage density (D)	2.28
7.	Stream frequency (Fs)	3.55
8.	Drainage Texture (Rt)	11.81
9.	Form Factor (RF)	0.60
10.	Circulatory ratio (Rc)	0.52
11	Elongation ratio (Re)	0.87
12.	Length of overland flow (Lg)	0.88

7.3 Stream Order (Nu)

The first step of drainage basin analysis to draw the drainage divide and trace all the streams occurring within it. The smallest fingertip tributaries are designated as order 1. Where two first order channels join, a channel segment of order 2 is formed; where two of order 2 join, a segment of order 3 is formed; and so on. The stream through which all discharge of water and sediments passes is the stream of the highest order by using Strahler's Method. (Strahler, A.N.1964).

7.4 Drainage pattern of the study area –

In the following table stream orders and length of streams of the study watershed KR34, have been given asTable 2.

Table 2: Stream order of KR 34 watershed.

SN	Stream order	No of streams	Length of Stream (Km)
1	I st order	707	306.679
2	II nd order	184	156.739
3	III rd order	42	84.057
4	IV th order	16	46.511
5	V th order	1	16.260
	Total	950	610.299

1.Total no of stream orders – 950 2.Total stream length of all order- 610 .299 Km

It is observed in the present study, the maximum number of streams is in lower order i.e. 5th order for watershed KR 34.(Table 3)

Table 3: Basin Characteristics of KR 34 watershed.

Sr.No	Basin Characteristics	KR 34
1	Total stream length Km	610 .299
2	Total no of stream orders (Km)	950
3	Area of Basin (Km^2)	266.753
4	Length of Basin (Long Axis) (Km)	21.147
5	Relief ratio- Meter	
6	Maximum elevation	884
	Minimum elevation	569
	Relative relief-	315
7	Basin Perimeter (Km)	80.411
8	Square of Basin Length	447.195
9	Square of Basin perimeter	6465.928

7.5 Stream length (Lu):

The numbers of streams of various orders in a watershed are counted and their lengths from mouth to drainage divide are measured with the help of GIS software's. The stream length (Lu) has been computed based on the law proposed by Horton (1945)forthe all study watersheds. Generally, the total length of stream segments is maximum in first order streams and decreases as the stream order increases. Total length of all streams of all order (I to V) is calculated for this watershed. Total stream length of watershed KR 34 is 610.759 km comprised of 950streams of Ist to Vthorder. Sometimes there may be changes in the computation of streams.This change may indicate the streams of high altitude, lithological variation and moderately steep slopes (Singh and

Singh, 1997).

7.6 Mean stream length (Lsm):

According to Strahler (1964), the mean stream length is a characteristic property related to the drainage network and its associated surfaces. The mean stream length (Lsm) has been calculated by dividing the total stream length of order 'u' and number of streams of segment of order 'u', given in Table No.1. Mean stream length (Lsm) of any given order is greater than that of the lower order and less than that of its next higher order in both the sub-watersheds which might be due variations in slope and topography. Average of Mean stream length (Lsm) of the KR 34watershed is 0.64 km.

7.7 Stream Length Ratio(RL):

Stream length ratio (RL) may be defined as the ratio of the mean length of the one order to the next lower order of stream segment.

Stream Length Ratio RL = Lu / Lu – 1

Where, RL = Stream Length Ratio, Lu = Total stream length of the order 'u'

Lu – 1 = Total stream length of its next lower order.

Horton's law (1945) of stream length stated that mean stream length segments of each of the successive orders of a basin tends to approximate a direct geometric series with stream length increasing towards higher order of streams. The RL values of the KR 34 watershed is0.49 km,The RL values differ from watersheds to watersheds. Generally, a variation in slope and topography affects the stream length ratio.(Horton,1945)

7.8 Relief Ratio (Rh):

Relief ratio, (Rh) is ratio of maximum relief to horizontal distance along the longest dimension of the basin parallel to the principal drainage line (Schumm, 1956).

Relief Ratio Rh = H / Lb

Where, Rh = Relief Ratio, H = Total relief (Relative relief) of the basin in Kilometers

Lb = Basin length

The Rh normally increases with decreasing drainage area and size of sub-watersheds of a given drainage basin (Gottschalk, 1964).

7.9 Digital Elevation Model(DEM) —

In this study the values of Rh for the watersheds KR 34watershed is 14.90.It is noticed that, the high values of Rh indicate steep slope and high relief (m). These values may indicate the presence of basement rocks that are exposed in the form of small ridges and mounds with lower degree of slope.(Gottschalk,1964)

7.10 Bifurcation Ratio (Rb)

After the drainage network elements have been assigned their order numbers, the segments of each order are counted to yield the numbers N_u of segments of the given order u it is obvious that the number of stream segments of any given order will be fewer for the next lower order but more for the next higher order. (http://www.geog.ouc.bc.ca/thysgeog) The ratio of number of segments of a given order N_u to the number of segments of the higher order. (N_u + 1) is termed the bifurcation ratio R_b. The term bifurcation ratio (Rb) is the ratio of number of the stream segments of given order to the number of segments of the next higher order (Schumm, 1956).

Bifurcation Ratio Rb = Nu / Nu + 1

Where, Rb = Bifurcation Ratio, Nu = Total no. of stream segments of order 'u'

Nu + 1 = Number of segments of the next higher order

Strahler (1957)demonstrated that, bifurcation ratio shows a small range of variation for different regions or for different environment except where the powerful geological control dominates. In this study the mean bifurcation ratio of study watersheds are KR 34.Bifurcation ratio characteristically ranges between 3 and 5 for drainage in which the geologic structure does not distort the drainage pattern. This pattern is observed in the KR 34watershed. Bifurcation ratio above 5 indicates structural control of drainage. In such cases, the development of higher order streams is normally facilitated by head ward erosion and guided by linear zones of structural weakness. Such streams are significant because they enhance the recharge and thereby the potential of groundwater and this type of stream pattern is recorded in the KR 34 watershed. Abnormally high bifurcation ratios might be expected towards dip direction. Thus, bifurcation ratio is a significant parameter throwing light on groundwater regime.

7.11 Drainage Density (D)

Horton (1932) has introduced drainage density (D) into American hydrologic literature as an expression to indicate the closeness of spacing of streams. It is the total length of streams of all orders per drainage area. In other words drainage density (Dd) is the total length (L) of the stream in the basin divided by the area (A) of the whole basin, or Lu / A. It is thus average length of streams for each unit area.

Drainage Density D = Lu / A

Where, D = Drainage Density, Lu = Total stream length of all orders

A = Area of the Basin (Sq.Km.)

The drainage density value of KR 34watersheds is 2. 8km/ Sq.Km.The low drainage density indicated that the region has highly permeable subsoil and dense vegetative cover whereas high drainage density is attributed to impermeable subsurface materials and mountainous relief.

7.12 Stream Frequency (Fs)

Stream frequency is the ratio of the number of streams of all orders within a watershed. This helps to measure the topographic texture.Horton (1932) introduced stream frequency (Fs) which is the total number of stream segments of all orders per unit area.

Stream Frequency (Fs) = Nu / A

Where, Fs = Stream Frequency, Nu = Total no. of streams of all orders

A = Area of the Basin (Sq.Km.)

Stream frequency value of watersheds KR34is 3.5 . It is noted that, the Fs exhibits positive correlation with the drainage density value of the watershed indicating the increase in stream population with respect to increase in drainage density, thus runoff is also high with increasing stream population.

7.13 Drainage Texture Ratio (Rt)

Drainage texture ratio (Rt) is one of the important concepts of geomorphology which means the relative spacing of drainage lines. Drainage lines are numerous over

impermeable areas than permeable areas. According to Horton (1945), Rt is the total number of stream segments of all orders per perimeter of that area.

Drainage Texture (Rt) = Nu / P

Where, Rt = Drainage Texture, Nu = Total no. of streams of all orders

P = Perimeter (km)

In the study the drainage textures value of watershedKR 34 watersheds is11.81

7.14 Form Factor (Rf)

According to Horton (1932), form factor (Rf) may be defined, as the ratio of basin area to square of the basin length.

Form Factor (Rf) = A / Lb^2

Where, Rf = Form Factor, A = Area of the Basin (Sq.Km.)

Lb^2 = Square of Basin length

In the study areas values of RF for study watershed KR34 is . 0, given in (Table No.1) Thus, a result indicates that all the watersheds are elongated in shape.

7.15 Circularity Ratio (Rc)

It is the ratio of the area of the basin to the area of a circle having the same circumference as the perimeter of the basin (Miller 1953)

Circularity Ratio (Rc) = 4 x Pi x A / P^2

Where, Rc = Circularity Ratio, Pi = 'Pi' value i.e., 3.14

A = Area of the Basin (Sq.Km.), P^2 = Square of the Perimeter (Km)

In the study, the Rc (Table 1) values of KR 34 is 0.52.Rc values indicated that, all the watersheds were not much circular.

7.16 Elongation Ratio (Re)

According to Schumm (1956) elongation ratio (Re) is the ratio between the diameter of the circle of the same area as the drainage basin and the maximum length of the basin.

Elongation Ratio (Re) = 2 v (A / Pi) / Lb

Where, Re = Elongation Ratio, A = Area of the Basin (Sq.Km.)

Pi = 'Pi' value i.e., 3.14, Lb = Basin length

In the study areas the Re values of KR34 is0.8 given in Table No 1.The highest values of Re indicates high elongated watersheds with less relief and steep slope.

7.17 Length of Overland flows (Lg)

The length of overland flow (Lg) approximately equals to half of the reciprocal of drainage density (Horton, 1945).

Length of Overland flow Lg = 1 / D x 2

Where, Lg = Length of Overland flow, D = Drainage Density

The computed values of watersheds KR 34 is 0.88given in (Table No.1)

CONCLUSION :

Remote sensing and GIS have been used in the delineation and updation of drainage in the present study and these updated drainage have been used for the morphometric analysis. The morphometric analysis of the drainage networks of the KR 34 watershed exhibits the somewhat circular, elongated drainage pattern. The variation in stream length ratio might be due to changes in slope and topography. The variation in values of bifurcation ratio among the KR 34 watershed is ascribed to the difference in topography and geometric development. The stream frequencies for both sub-watersheds of the study exhibits positive correlation with the drainage density values, indicating the increase in stream population with respect to increase in drainage density.The Remote Sensing and GIS techniques have proved to become indispensable day-to-day management tools for efficient management of natural resources even at micro-watershed level.A holistic approach of this resource analysis further provides in optimizing the use of available natural resources to meet the growing demands of food, fodder and fuel wood on sustainable basis without affecting the geo-ecosystems in the area.

Morphometric analysis were carried out in a number of Indian watersheds and subsequently used for water resource development and management projects as well as for watershed characterizationand prioritization. Morphometric parameters of thefour major watersheds KR 34Table 1. The important property, bifurcation ratio (R_b) reflecting geologic and tectonic characteristics of the watershed area are calculated for sub-watersheds (10,000 to 50,000 hectors). Higher value of Bifurcation ratio (R_b) for a sub-watershed indicates high runoff, low recharge and mature topography and is expected in the region of steeply dipping rock strata.

The values of R_b also indicated that the basin has suffered less structural disturbances. As per Rb valuesKR 34is 6.71. In general, the shape of a basin affects stream flow hydrographs and peak flows. According to Circulatory Ratio (Rc) of KR 34watersheds have an elongated shape. The values of circulatory ratiofor these watersheds are 0. 2. Higher value of Rcindicates mature to old stage topography. On the basis of this information it is hypothetically considered that historically, this area had surrounded with leaf falling dense forest, hence less stream population is observed.

Various other parameters such as topography, slope, forest cover are important in the morphometric analysis. The watershed KR 34have high potential to generate the natural resources through watershed development.

Values of Form Factor (Rf) for KR 34, are 0.60. More value of form factor Rf indicates that the basin will have a flatter peak of flow for a longer duration. Flood flows of such elongated basins are easier to manage than from the circular basin. The sub-watershed with high value of time of concentration (tc) will produce less runoff rate and vice-versa. Sub-watersheds with higher elongation ratio (Re) and relief Ratio (Rh) are considered critical from erosion point of view and should be provided with suitable soil and water conservation measures. Drainage density (Dd) and Stream frequency (Fs) are computed for all the sub-watersheds and are given in Table No 1.It was observed that the sub-watersheds having large area under dense forest have low drainage frequency and high Fs values having more agricultural land. The value of Fs for KR 34 is 3.55 demarcates the area under agriculture. High value of Df in the sub-watershed KR 34 produces more runoff compared to others. A very few literature and studies have been reported in India regarding GIS applications in watershed planning.

The morphometric analysis of different KR 34watersheds showed their relative characteristics with respect to hydrologic response of the watershed. Morphometric parameters coupled with integrated thematic

maps, viz. land use/land cover, soil and drainage density and soil information can help in decision making process for water resources management. In poorly managed land, contour bunds can be constructed to increase the groundwater recharge which would eventually help in cultivation of kharif and rabbi crop instead of cash crop.These measures are expected to bring down the soil erosion rates as well as improvement in water resources regime. Land development activity, area treatments and percolation tanks are recommended based on the land use and drainage pattern to increase the irrigated area and recharge of the study area.

REFERENCES :

1. Bharat Kakade- Watershed Manual – A practical Guide for Watershed Development Practitioners and Trainers.
2. *Dhokarikar*, B.G. (1991) Groundwater Resource Development in Basaltic Rock Terrain of Maharashtra. Water Industry Publication, Pune.
3. Gottschalk, L.C. (1964) Reservoir Sedimentation. In: V. T. Chow (ed), Handbook of Applied Hydrology. McGraw Hill Book Company, New York, Section 7-I.)
4. Ground Water Survey and Development Agency, (GSDA, Sangli 2008-2009).
5. Horton, R.E. (1932) Drainage Basin Characteristics. Trans. Am. Geophysical. Union, Vol.13, pp.350-361.
6. Horton,R.E. (1945) Erosional Development of Streams and their Drainage Basins; Hydro physical Approach to Quantitative Morphology. Geol. Soc. Am. Bull., vol.56, pp.275-370.
7. http://www.geog.ouc.bc.ca/thysgeog.
8. Miller, V.C. (1953) A Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain area, Virginia and Tennessee, Proj. NR 389-402, Tech Rep 3, Columbia University, Department of Geology, ONR, New York.)
9. Nageswara Rao - 2010- "Morphometric analysis of Gostani river basin in Andhra Pradesh state, India using spatial information technology", International Journal of Geomatics and Geosciences.Vol.1.No.2.
10. Sachumm,S.A. (1956) Evolution of Drainage Systems and Slopes in Badlands at Perth Amboy, New Jersey. Geol. Soc. Am. Bull., vol.67, pp.597-646.)
11. Singh, S. and Singh,M.C. (1997) Morphometric Analysis of Kanhar River Basin. National Geographical. J. of India, vol. 43(1), pp. 31-43.
12. Strahler, A.N.(1964) Quantitative geomorphology of drainage basins and channel networks In: V. T. Chow (ed), Handbook of Applied Hydrology. McGraw Hill Book Company, New York, section 4-II.
13. Strahler,A.N. (1957) Quantitative analysis of watershed geomorphology. Trans. Am. Geophys. Union, vol.38, pp.913-920).

ANDREV : REVERSE ENGINEERING TOOL FOR EXTRACTING PERMISSION OF ANDROID MOBILE APPLICATION.

MANISHA PATIL*
Asst. Professor
Indira College of Commerce and Science
Pune. 411033, Maharashtra, India.
E-mail: manisha.patil@iccs.ac.in

Abstract : *One of the leading and the most popular operating system for smart phones and tablets is an Android. Being an open source platform has also become a prime target for the attackers as growing users. This paper focuses on the work done on the Android platform by performing static analysis on the permission-based framework and permission extraction tool is designed. Extracted a number of permission-based features by reverse engineering the Android application (apk) files using the batch scripted tool. AndRev tool is used to decompile apks in batch mode. Features have been stored in feature vectors. The analysis is done using feature vectors in order to study the pattern of permissions in applications as per the category. Two categories of apks, namely, General and Entertainment apps are studied with an initial dataset of 50 applications each.*

Keywords: Android Mobile Apk, User Privacy, Permission Extraction, Decompile.

INTRODUCTION:

Google created a format for Android application package (APK) and used to install an application onto the Android OS. Official Google Play Store used as the data source for apps. There were more than millions Apps that have frequently been downloaded, but there is no way to validate them all or to get absolute purely benign Apps, so our best approach is to trust Google Play Store's review system and only download highly popular Apps or Apps from trusted and well-known sources. There are also different categories of Apps, which might affect the distribution of permissions. For example, an App from entertainment Category would be regarded as normal when requesting for access to contacts, but it should be malicious for an App such as basic calculators from general category. In this study downloaded a certain amount of top Apps in each category accordingly. As a result, got 100 valid Apps, which ought to be a good sampling of Apps. The next step will be decompiling and extracting permissions.

RELATED WORK

Schmidt [2] et al. proposed a solution based on monitoring events occurring on Linux-kernel level. They use kernel system calls, network activity events and file system logs to detect anomalies in the system. At that time, there were no real Android devices available, so they failed to test their system properly.

Shabtai [1] et al. proposed Adnromaly — a framework for anomaly detection on Android smartphones. The framework continuously monitored the information obtained from the Smartphone. Then, it applied machine learning to classify the collected data as benign or malicious. Yet they could not find real malware to test their proposal. Enck et al [8] used de-compilation and static analysis techniques to study 1100 free applications from the official Android Market to understand a broad range of security-related metrics associated with these applications. They discovered that sensitive information is widely leaked in applications. For instance, more than half of the applications include at least one advertisement libraries that collect and send private information, e.g. the location of the phone.

Pridgen & Wallach [7] examined a sample of 114,000 apps and found that the number of permissions required by apps is increasing, and consequently, posing a privacy risk to Android users.

Felt et al. [2] and Kelley et al. [4] suggested that many users have a low comprehension of the Android permissions system – that is the permissions system may be insufficient in providing adequate user privacy in the hands of a novice user.

Kern & Sametinger [6] took a different approach and recommended the use of fine-grained individual permissions control on a per app basis. This means that each Android app would have each of their permissions explicitly listed and the user would either deny or allow the permission request.

Zhou et al. [11] designed a system that could control an app's access to sensitive permissions Berthome et al. [3] proposed a set of two apps, comprising (1) the Security Monitor, a third party app installed onto the device, and (2) the Security Reporter, which would be injected into a decompiled target app. The injected app is able to monitor the targeted app and can then report to the Security Monitor with details such as resource requests.

Juanru, Dawu & Yuhao [5] used a similar technique of decompiling Android apps to aid with their Android malware research.

Xu, Saïdi & Anderson [10] developed a solution called Aurasium that automatically repackages Android apps to have sandboxing and policy enforcement abilities in order to enhance user privacy.

Kirin [9], an application certification for Android. Kirin performs a permission check on the application during installation. When a user installs an application, Kirin extracts its security configurations and checks them against the security policy rule that it already has. If an application fails to pass all the security policy rules, Kirin can either delete it or alert the user.

```
please wait apks are decompiling...
C:\xampp\htdocs\andrev>for /F %f in ('dir /b apk\') do java -jar apktool.jar d -s "apk\%f"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\ADM.apk"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\am.apk"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Android"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Archery"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Cheat"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Chrome"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\CM"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Crime"

C:\xampp\htdocs\andrev>java -jar apktool.jar d -s "apk\Flashlight_5.2.4_170.apk"
I: Using Apktool 2.0.0-RC4 on Flashlight_5.2.4_170.apk
I: Loading resource table...
I: Decoding AndroidManifest.xml with resources...
I: Loading resource table from file: C:\Users\Manish\apktool\framework\1.apk
I: Regular manifest package...
I: Decoding file-resources...
I: Decoding values */* XMLs...
I: Copying raw classes.dex file...
I: Copying assets and libs...
I: Copying unknown files...
I: Copying original files...
```

Fig 1 : Decompiling Apks in a batch mode using AndRev Tool

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Applicatio	Category	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p	android.p a
2	ADM.xml	games	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
3	am.xml	games	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
4	Flashlight	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Addlogoo	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	April-Layc	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
7	EyeEm-Ca	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
8	FastLite-S	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
9	FastMesse	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
10	FolioforFa	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
11	HDCamera	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12	HDRMax-F	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
13	InboxMes	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
14	InfinitePa	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
15	InstaQuot	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	InstaRepo	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
17	MagicPain	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
18	PlantNan	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
19	Repostan	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
20	VoiceRec	general	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Fig 2 : Outcome of AndRev Tool an Excel Sheet
(Permission declared : 1 and not Declared : 0 value is set)

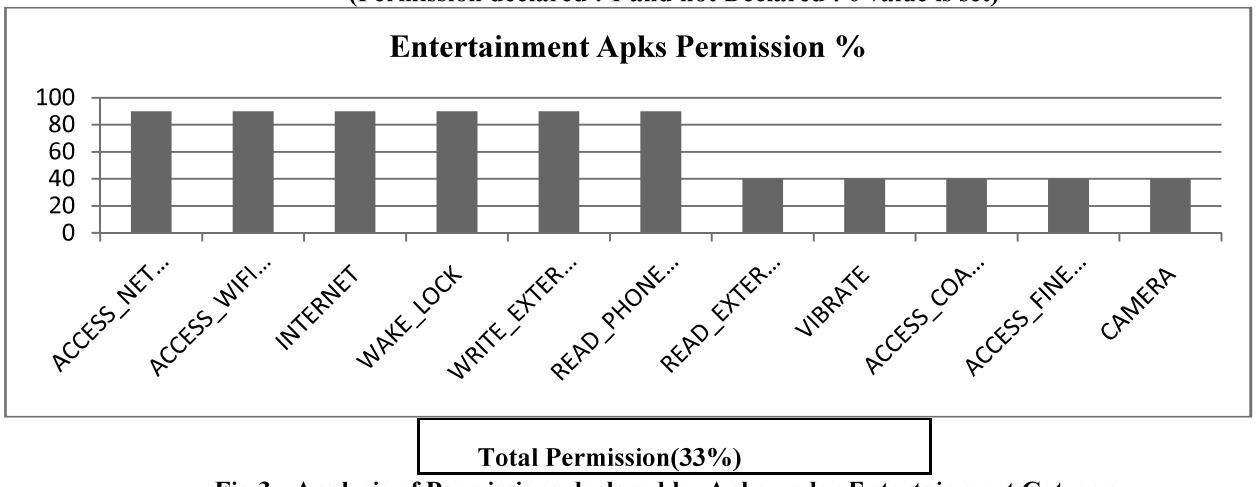


Fig 3 : Analysis of Permissions declared by Apks under Entertainment Category.

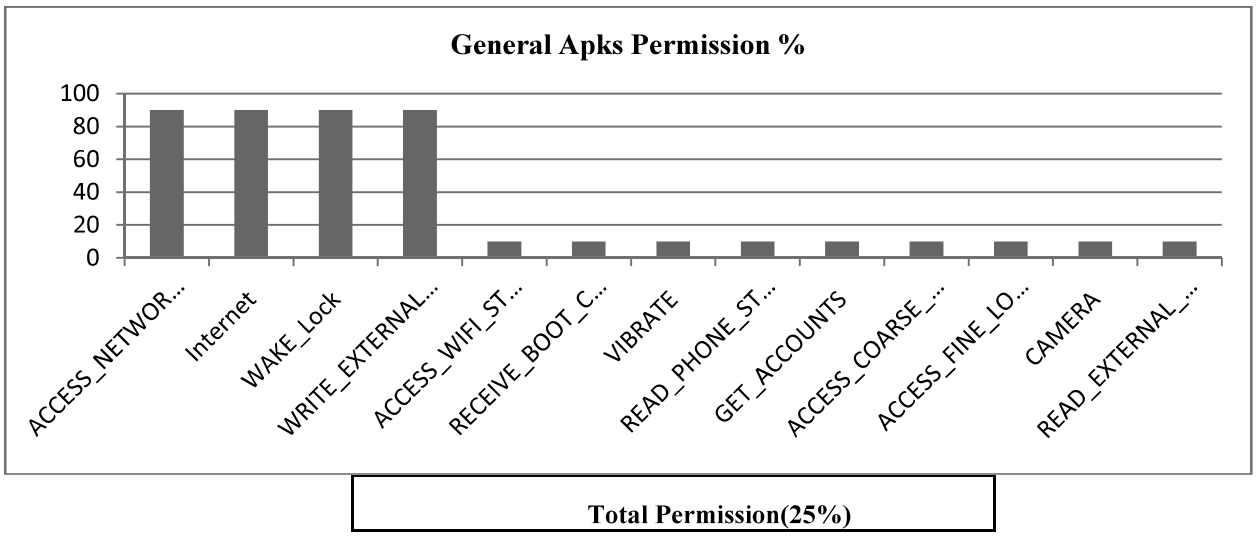


Fig 4 : Analysis of Permissions declared by Apks under General Category.

CONCLUSION :

The study shows, the approaches proved valuable in protecting smart phones but they have restrictions. In particular, the Android system has been in a dominant position in the market of Smartphone operating system. Malware for the Android system is also growing significantly. Therefore, it is necessary to develop a security suite for the Android phones, such as signature-based anti-virus technology, Smartphone firewall, access control mechanisms and lightweight Intrusion Detection Technology. AndRev tool used for reverse engineering mobile apps and to extract permissions . Resultant excel sheet (Feature vector of 100 apks) is further analyses. Out of the two categories, entertainment app apks are using 33% of permissions while general app apks are using 25% of permissions. This study concludes general apps are more safe for user privacy than entertainment apps due to unexpected permissions are declared less in a number.

REFERENCES :

[1] Asaf Shabtai, Uri Kanonov, Yuval Elovici, Chanan Glezer, and Yael Weiss. "Andromaly: a behavioral malware detection framework for android devices". Journal of Intelligent Information Systems, pages 1–30, 2011. 10.1007/s10844-010-0148-x

[2] Aubrey-Derrick Schmidt, Hans-Gunther Schmidt, Jan Clausen, Kamer Ali Y¨uksel, Osman Kiraz, Ahmet Camtepe, and Sahin Albayrak. "Enhancing security of linux-based android devices". In in Proceedings of 15th International Linux Kongress. Lehmann, October.2008.

[3] P. Berthome, T. Fecherolle, N. Guilloteau & JF. Lalande, "Repackaging Android Applications for Auditing Access to Private Data", ARES 2012, pp. 388-396.

[4] A.P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin & D. Wagner, "Android permissions: User attention, comprehension, and behavior", SOUPS 2012, p. 3.

[5] L. Juanru, G. Dawu & L. Yuhao, "Android Malware Forensics: Reconstruction of Malicious Events", ICDCSW 2012, pp. 552-558.

[6] M. Kern, & J. Sametinger, "Permission Tracking in Android", UBICOMM 2012, pp. 148-155.

[7] T. Book, A. Pridgen & DS. Wallach, "Longitudinal Analysis of Android Ad Library Permissions", arXiv preprint arXiv: 1303.0857, 2013.

[8] W. Enck, D. Octeau, P. McDaniel and S. Chaudhuri, "A study of Android application security," 20th Usenix Security Symposium, 2011.

[9]W. Enck, M. Ongtang, and P. McDaniel, "On Lightweight Mobile Phone Application Certification," Proc. 16th ACM Conf. Computer and Communications Security (CCS 09), ACM, 2009, pp. 235-245.

[10] Xu R., H. Saïdi & R. Anderso, 'Aurasium: Practical policy enforcement for android applications', 21st USENIX conference on Security symposium, 2012, pp. 27-27.

[11] Y. Zhou, X. Zhang, X. Jiang & V. Freeh, "Taming information-stealing smartphone applications (on Android)", TRUST 2011, pp. 93-107

"SP CRYPTOLOGY ALGORITHM"

PROF. DANI SUHAS PRABHAKAR AND PROF. BHALSING PRAJWAL VITTHAL
Assistant Professor, MSc (Computer Science),Asst.Professor,MSc.Comp,MBA,SET,DIT
New Arts, Commerce &Science College,
Indira College of commerce and science,
Ahmednagar , Maharashtra.Pune, Maharashtra. .
Email-suhas.dani17@gmail.com & prajwalmotog@gmail.com

Abstract:-To provide security to data(mostly during exchange), different encryption methods/algorithms are used. Encryption is the encoding process to convert the readable data (Plain Text) into non-readable format (Cipher Text) to provide the security to avoid the cyber-crime. Thus, to provide the secure services to network and data, we can use the proposed technique.

This paper presents, symmetric (Private) key encryption technique with the combination of DES, IDEA and MD5 algorithms to provide the security to data against different attacks,and in this technique we are not exchanging the key.

Keywords:- Encryption, Decryption, Key, Plain Text, Cipher Text, DES, IDEA and MD5.

INTRODUCTION:

We have heard the adage "information is power" and that is certainly true when it comes to cyber crime. Access to your personal information is what gives hackers the power to tap into your account and steal your information. But the right information can also empower you to protect yourself from being caught up in the thriving industry that is cyber crime. To avoid cyber crime strong encryption is the basic mechanism. Encryption can prevent cyber criminals from accessing and exploiting data.

Encryption is the encoding method to convert the readable data(Plain Text) into non readable format(Cipher Text). The theme of encryption is nothing but the concept of cryptography. Every encryption and decryption process has two aspects: 1) the algorithm2) key used to encrypt and decrypt the data. In general, the algorithm used for encryption and decryption processes is usually known to everybody. However, it is the 'key' used for encryption and decryption that makes the process of cryptography secure. Broadly there are two cryptographic mechanism.

1) Symmetric key cryptography involves usage of the same key is used for encryption and decryption.

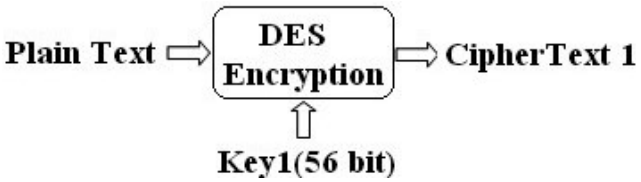
2) Asymmetric key cryptography involves the usage of one key for encryption and another, different key is used for decryption.

In symmetric type, we need to exchange the key because without key receiver cannot decrypt the data. But this symmetric key cryptography suffers through key exchange problem. Without key no one can decrypt the data, so if attacker gets the key during transit then he can be able to decrypt the data. So we must secure the key from unauthorized access. Every time the communicators will not communicate using same key, so key should be changed after specific amount of time orfor every new communication.

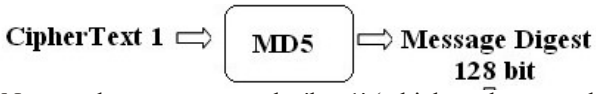
When we send the key (to secure that key), we can follow this proposed method.

SENDER SIDE ENCRYPTION PROCESS:-

1) Consider, the 64bit key. Take the plain text, encrypt it by using key and DES (DataEncryptonStandard) algorithm.



2) Calculate the message digest by using MD5 algorithm.
Input for MD5 = Cipher Text 1(output of step1)
Output of MD5= 128 bit Message Digest(say MD)



3) Now we have to encrypt the 'key1' (which we have used in step1's encryption process) by using IDEA algorithm this process is called key wrapping. For IDEA encryption process, 128 bit key is required. So for this IDEA encryption, we are providing the message digest(step2's output) as a key, say it is key2.

Input for IDEA = Key1(which we have used in step1's encryption)
Output of IDEA = Cipher Text 2(i.e. encrypted key1)
Key to IDEA encryption = Message Digest (step2's output)

4) Send the Cipher Text 1 and Cipher Text 2 to the receiver.

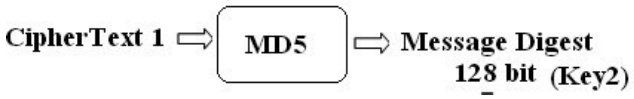
RECEIVER SIDE DECRYPTION PROCESS:-

1) Receiver receives Cipher Text 1 and Cipher Text 2.

2) Take the Cipher Text 1 and calculate the Message Digest of it, using MD5.

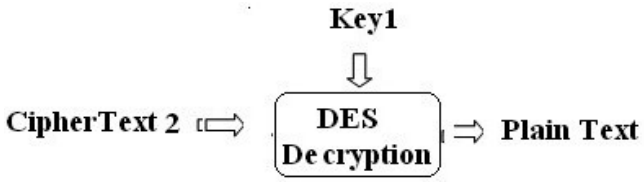
For a given message, the message digest must always be the same.

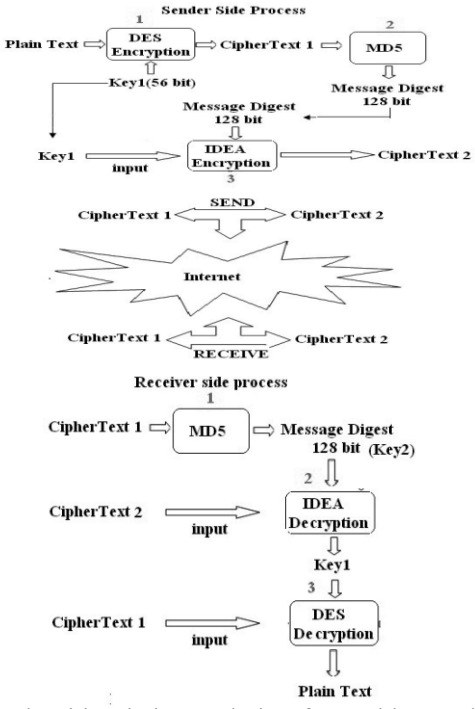
----property of message digest.



3) Now the IDEA decryption is perform
Input for IDEA = Cipher Text 2 . Output of IDEA = Key1 (its nothing but DES's key)
Key to IDEA encryption = Message Digest(step1's output)

4) Now theDES decryption is perform. Input for DES = Cipher Text 1
Output of DES = Original Plain Text Key = Key1(step2's output)





Our algorithm is best solution for problems which created from symmetric key cryptography and asymmetric key cryptography,here is the comparison.

Feature	Symmetric Key. Crypt.	Asymmetric Key.Crypt.	SP Crypt. Algo
Encryption/Decryption key	Same key used for both	Diff. key used for both	Same key used for both
Speed for operation	Very fast	slow	Very fast
Size of output text	Usually same or less than original text	More than original text	Less than original text
Key exchange	Required and create problem	No problem	No problem
Scalability	An issue here	Issue handle quite well	No issue
usability	Mainly for encryption and decryption	Used for encryption ,decryption as well as digital signature	Mainly for encryption and decryption

CONCLUSION :

1) Strong Encryption because through this method, we are providing the security to the data, by using two level encryption process(DES and IDEA).
2) We are also securing the key by encrypting it with the help of message digest.
3) Solution to key exchange problem.

REFERENCES :

1] Cryptography and Network Security By AtulKahate (Tata Mcgraw-hill Publishing Company Limited)
[2] Cryptography and Information Security By V.K. Pachghare (PHI Learning Private Limited)
[3] EliBiham and Adli Shamir, "Differential Cryptanalysis of full DES".
[4] Dan Boneh and Glenn Durfee "Cryptanalysis of low exponent RSA"
[5] Introduction to Computer Security By Matt Bishop and Sathyanarayana (PEARSON EDUCATION)
[6] Applied Cryptography Protocols, Algorithms, and Source Code in C By Bruice Schneier (Wiley India)
[7] http://www.cryptool.org

FORWARD FREQUENCY BASED LEAST FREQUENTLY USED (FFBLFU) PAGE REPLACEMENT ALGORITHM

MADHAVI AVHANKAR AND BHAKTI SHINDE
Indira College of Commerce and Science Indira College of Commerce and Science
madhavi.avhankar@icc.ac.in bhakti.shinde@iccs.ac.in

Abstract:- Page replacement algorithms are used in Virtual Memory Management (VMM) system to decide which page to evict from memory whenever page fault occurs and free page is not available for allocation. Over the years the number of page replacement algorithm have been proposed. Each algorithm aims to decrease overall page fault rate that is increase the hit ratio of pages. In this paper a new page replacement algorithm has been introduced – Forward Frequency Based Least Frequently Used(FFBLFU), which gives better performance than traditional LFU. This paper remonstrate the advanced version of LFU refer as FFBLFU. The general idea behind this new algorithm(FFBLFU) is, while applying LFU if two or more pages found with the same smaller count, then instead of applying the FIFO to evict the page choose the page to be removed which has less frequency count in future for next five pages(except the page being replaced)

Keywords: - LFU algorithm, FFBLFU algorithm, Hit ratio, Virtual Memory Management, Page Replacement algorithms.

INTRODUCTION:

In computer operating system page replacement algorithms for paging decides which page in Virtual Memory Management (VMM) to swap out when a page of memory needs to be allocated and free page is not available for allocation. In paging OS divides the process address space or virtual address space into number of fixed size units called pages same as main memory is also divided into fixed size units called frames. While executing the process the virtual address that is the CPU generated address or logical address must be translated into the physical address, which is the memory management unit generated address. This translation has been done by special hardware unit – Memory Management Unit. This address translation has been done for every memory unit and MMU uses page table which contains the address mapping information to make the translation. If the required logical address is not found (mapped) to Main Memory , MMU trap the OS. This trap is called page fault, which gives an opportunity to the OS to bring the desired page from secondary memory to Main Memory and then update the page table correspond [2]. Every running process has its own virtual address space and OS and MMU must keep track of each page and location of each page of the process. Whenever the processor needs the requested page it first find the corresponding entry in the cache hit has occurred and if it is not present in the cache, cache miss has occurred and when cache miss has occurs the OS applies page replacement algorithm to choose a page from cache for replacement to make a place for the requested page. Hence one of the important factor to measure the performance of page replacement is hit ratio . Hit Ratio = Total number of hit Counts/ Total number of Reference Counts To represent hit ratio as a percentage: Hit Ratio % =Hit Ratio *100 Pages are brought into main memory only when the executing process demands them, this is known as demand paging. A page fault typically occurs when a process references to a page that is not marked present in main memory and needs to be brought from secondary memory. In this existing page needs to be discarded. The selection of a such page is performed by page replacement algorithms which try to minimize the page fault rate at the least overhead[1]

Demand Paging : Demand Paging is quite similar to paging and only copies a disk page into physical memory when as attempt is made to access it. Rather than copying the entire process into main memory, it uses lazy swapper known as pager. When a process is to be brought in instead of copying the whole process the pager brings only necessary pages into memory. Thus it decreases the swap time and amount of physical memory needed as it avoids reading into memory pages that will not be used anyway. Valid – invalid bit scheme is used to distinguish between those pages that are in memory and those that are on disk and access to a page marked invalid causes a page fault trap; this page fault trap is known as page fault and is handled by as follows :

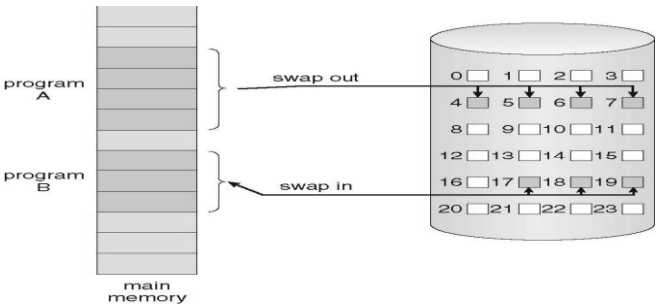


Figure 1.Process executes and accesses pages that are memory resident

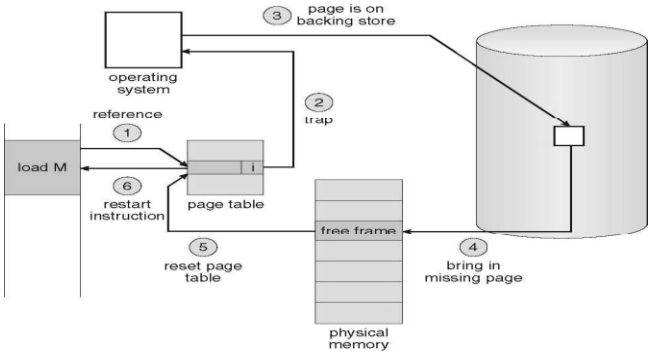


Fig.2 Steps in Handling a Page Fault

1. Operating system looks at table to decide if it is invalid reference or valid reference, if the reference is invalid abort the process
2. If it is valid get the empty frame
3. Swap page into frame
4. Reset tables
5. Set validation bit = v
6. Restart the instruction that caused the page fault
For performance of demand paging :-
Let a : memory access time(usually 50 - 500ns)
p : probability that a memory reference cause a page fault
s: time to serve a page fault

Then Effective Access Time(EAT) =Hit Rate x Hit Time + Miss Rate x Miss Time EAT = (1-p) x a + p x s

There are various page replacement algorithms available , and each algorithm is designed to reduce the page faults.
2. Page Replacement Algorithms: Page Replacement Algorithms uses the techniques in which if no frame is available to replace the page, find one which is not currently being used and free it, frame can be free by writing the contents to swap space and change the page table to indicate that page is no longer in memory and freed frame can be used to hold the page for which the page fault occurred [4]. There are number of page replacement algorithms are present , some of them are as follows :
I.First In First Out (FIFO) :- This is the simplest page replacement algorithm, as the name suggest if the frame is not available then according to FIFO it selects the oldest page for replacement. Queue data structure is used by this algorithm to hold the references of all pages in memory, when the page is brought into memory it get inserted at the tail of the queue and while page replacement, the page at the head of the queue get replace. This algorithm suffers from Belady's anomaly i.e Page fault rate may increase as the number of allocated frames increased [5]
II.Optimal Page Replacement :-This is the best possible page replacement algorithm is easy to describe but impossible to implement, it replaces the page which has not been used for the longest period of time in future. As it requires the knowledge of future string it is difficult to implement. This algorithm never suffers from Belady's Anomaly and gives lowest page fault rate among all.
III. Least Recently Used (LRU) :- Least Recently Used algorithm works on the idea that the page that has not been referenced for longest period of time will get replace that is LRU keeps track of page usage over a short period of time. LRU replacement associates with each page the time of that page's last use. Though the LRU can provide near optimal performance in theory it is expensive to implement in practice.
IV. Least Frequently Used (LFU) :- Least Frequently Used algorithm is the Counting Based Page Replacement algorithm,this algorithm replaces the page with the smallest frequency count in the past. The idea behind this selection is that an actively used page should have large reference count.
V. Second Chance:- Second chance algorithm is the next version of FIFO page replacement algorithm, in this algorithm every page is associated with reference bit, when page has been selected for replacement reference bit is checked, if it is zero then replace the page but if it is one give the second chance to the page and move on to select the next page in FIFO order.

2.1 Proposed Algorithm: Forward Frequency Based LFU page replacement algorithm (FFBLFU)
The concept behind the proposed algorithm is while applying the traditional Least Frequently Used algorithm if more than one pages are having same smallest count the instead of applying First In First Out (FIFO) , to evict the page, choose the page to be removed which has less frequency count in future for next five pages(except the page being replaced). The reason behind this selection is that the page which occurs minimum number of times in future will get replace because it has less chances to require in future.
The Procedure for the proposed algorithm is as follows:

```
#define max 20
/*Structure used to store reference string, page table and all related values
*/
struct PageTbl
{
int RefString[100]; int PT[10]; int nof; int nor;
int PFL; int PFH; } P[100];
int f;
/*Procedure to initialize all the data structure&/
void initialize()
{
int i;
for(i=0;i<P[f].nof;i++)
        P[f].PT[i]=-1;
}
```

```
/*Procedure to get input values for Reference string*/
void Accept()
{
        int i;
        for(i=0;i<100;i++)
                P[f].RefString[i]=0;
        for(i=0;i<P[f].nor;i++)
                P[f].RefString[i]=rand()%10;

        for(i=0;i<P[f].nor;i++)
        {
                printf("[%d]=",i);
                printf("%d\t",P[f].RefString[i]);
        }
}
```

```
/*Procedure of Forward Frequency Based Least Frequently Used */
void FFBLFU()
{
        int i,j,k,Faults=0;
        initialize();
        for(k=0,i=0; k<P[f].nof && i<P[f].nor; i++)
        {
                printf("\n Next Page  : ");
                printf("mm%4d \n",P[f].RefString[i]);
                if((Search(P[f].RefString[i]))==-2)
                {
                printf("PF =%d\n",P[f].RefString[i]);
                P[f].PT[k]=P[f].RefString[i];
                Faults++;
                k++;
                }
        }
        k=0;
        while(i<P[f].nor)
        {
                if(Search(P[f].RefString[i])==-2)
                {

                k = GetFFBLFU (i,k);
                P[f].PT[k]=P[f].RefString[i];
                k=(k+1)%P[f].nof;
                Faults++;
                }
                i++;
        }
        P[f].PFH=Faults;
        printf("\nTotal Page Faults: %d %d",Faults,P[f].PFH);
        f++;
 }
int Search(int s)
{
        int i;
        for(i=0;i<P[f].nof; i++)
                if(P[f].PT[i]==s)
                        return(i);
        return(-2);
}

/* Procedure to get the victim page according to FFBLFU concept*/
int GetFFBLFU(int e, int s)
{       int i,j,cnt1=999,cnt2,posi=0,Tfrm[10],w,c=0,flag=0;
        for(i=0;i<P[f].nof;i++)
        Tfrm[i]=-1;
        i=s;
        posi=s;
        do
        {
                cnt2=0;
                for(j=e-1;j>=0;j--)
                {
                        if(P[f].PT[i]==P[f].RefString[j])
                        {
                                cnt2++;

                        }
                }
        if(cnt2<=cnt1)
                {
                        Tfrm[i]=cnt2;
                        for(w=0;w<P[f].nof;w++)
                        {
                                if(Tfrm[w]!=-1)
                                {
                                if(Tfrm[w]>cnt2)
                                        Tfrm[w]=-1;
                                }
                        }
                }
                if(cnt2<cnt1)
                {
                        cnt1=cnt2;
                        posi = i;

                }

                i=(i+1)%P[f].nof;
        }while(i!=s);

        for(i=0;i<P[f].nof;i++)
        {
                if(Tfrm[i]!=-1)c++;
        }
        cnt1=999;
        if(c>=1)
        {
                i=s;
                do
                {
                cnt2=0;
                for(j=e+1;j<e+6 && j<P[f].nor ;j++)
                {
                        if(Tfrm[i]!=-1)
                        {
                        if(P[f].PT[i]==P[f].RefString[j])
                        {

                                cnt2++;

                        }
                        flag=1;
                        }
                }
                if(flag==1)
                {
                if(cnt2<cnt1)
                {
                        cnt1=cnt2;
                        posi = i;
                }
                flag=0;
                }
                        i=(i+1)%P[f].nof;
        }while(i!=s);
        }
        return(posi);
}
```

2.3 Experimental result analysis : For the implementation of the proposed Page replacement policy, we used a 'C' language, in which number for the reference string get generated randomly which is used to run the simulation program. Following tables and Figures shows comparative result of two algorithm that is Traditional LFU and Proposed Forward Frequency Based LFU for a 20 and 30 random number as memory reference which is used for the evaluation respectively.

Table 1. Comparative analysis among Traditional LFU and FFBLFU for 20 random reference string.

No. of Frames	Avg. Number of Page Fault for 20 reference string	
	LFU	FFBLFU
1	19	19
2	17	15
3	16	13
4	14	12
5	13	10

Figure 1. Comparative analysis among Traditional LFU and FFBLFU for 20 random reference string

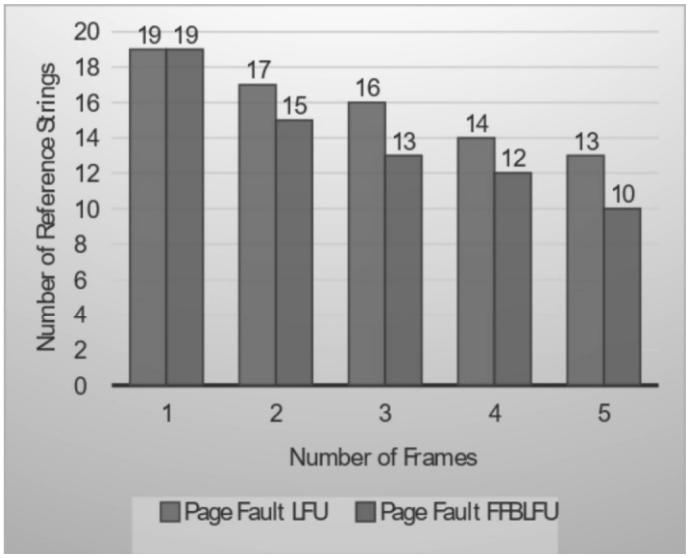
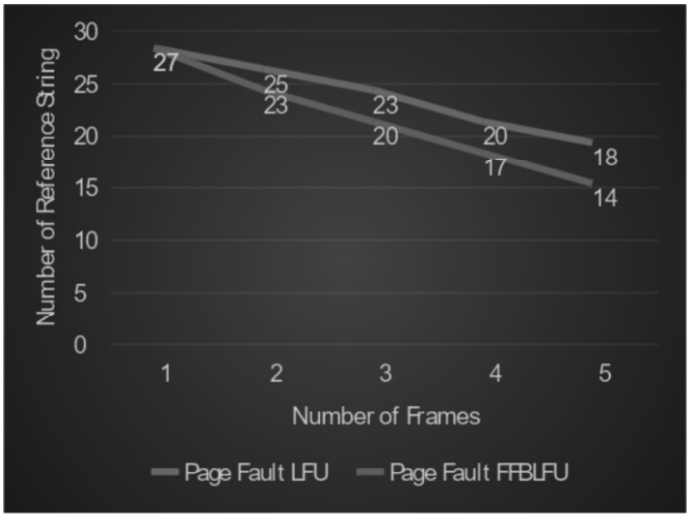


Table 1. Comparative analysis among Traditional LFU and FFBLFU for 30 random reference string

No. of Frames	Avg. Number of Page Fault for 30 reference string	
	LFU	FFBLFU
1	27	27
2	25	23
3	23	20
4	20	17
5	18	14

Figure 1. Comparative analysis among Traditional LFU and FFBLFU for 30 random reference string



CONCLUSION :

In virtual memory, it is assigned several frames to be implemented in each process. So, the main issue about virtual memory is to use an algorithm for page replacement. This paper proposes a new page replacement algorithm – Forward Frequency Based Least Frequently Used (FFBLFU) to improve traditional Least Frequently Used (LFU) algorithm, also this paper discussed various page replacement algorithm like First in First out (FIFO), Optimal Page Replacement (OPT), Least Frequently Used (LFU), Second Chance, Least Recently Used (LRU). Above tables and Figures shows comparative result of two algorithm that is Traditional LFU and Proposed Forward Frequency Based LFU which conclude that the proposed algorithm – Forward Frequency Based Least Frequently Used (FFBLFU) is more efficient than traditional LFU as it gives less page faults compared to traditional LFU algorithm and the upcoming researches have focused on this algorithm, to increase the performance of system and to decrease the page fault rate.

REFERENCES :

[1] A study of page replacement algorithm by Mr. Mohit Saini , Ms. Manju Verma
[2] Analysis of page replacement algorithm with variable number of frames by Manisha Kerge
[3] Abraham Silberschatz, Peter Baer Galvin, Greg Gagne, "Operating System Concepts", Eight Edition
[4] Yeskendir Sultanov, Kazakh, British Technical University, Almaty, Kazakhstan. "Design and Implementation of Page Replacement Algorithms based on an Inverted Page Table" Advanced Engineering Technology and Application An International Journal
[5] K. Ravindra Babu, G. Pavan Kumar, "Enhanced Hybrid LRU Page Replacement Algorithm" International Journal of Advanced Research in Computer Science and Software Engineering
[6] Foundation of Operating Systems By Dp Sharma
[7] William Stallings, Operating Systems Internal and Design Principles, 5th Edition , 2006.

CANCER CLASSIFICATION BASED ON DIFFERENT CLASSIFICATION TECHNIQUES

UJWALA KOLTE, NINAD THORAT
ujwala.kolte@iccs.ac.in, ninad.thorat@iccs.ac.in
Indira College of Commerce and Science

Abstract : *The datasets of bio-informatics are very large in size. They do not convey any useful information directly. One need to generate knowledge and extract information from the data sets in order for them to be useful. Today data mining techniques are very handy and powerful and can be used with very large datasets. Medical Data mining can be used to generate knowledge for automated diagnosis of diseases and prediction. Medical mining can be used to extract hidden and important connection between features and diseases. In this paper we are proposing a comparative study of different classification algorithms for identification of breast cancer. We will be comparing our results using a software tool WEKA based on algorithms namely Random forest ,Naive Bayes, Multilayer perceptron and ZeroR.*

Keywords: Classification, Random forest ,Naive Bayes, Multilayer Perceptron , ZeroR. **1. 1.**

INTRODUCTION:

Automated disease diagnosis is gaining popularity as it can be hugely helpful in early and accurate diagnosis of diseases. This can be of great advantage to the patients who are diagnosed with diseases in early stages. Also early diagnosis can be helpful for accurate treatment of patients helping the medically as well as financially as they will be given medicines which they need exactly. Breast cancer is a deadly disease and is a most common cancer amongst woman. This year (2017) an estimated 252710 new cases will be registered among woman worldwide.

In order to beat such deadly diseases computer technology can be used which include data mining techniques which can generate knowledge out of huge amount of data-sets. Medical data mining is data mining on medical data in which one analyses the data predicts useful information out of it. It can be used to generate predictive model to increase the accuracy of diagnosis in any specific disease. Some work has been done in the field of data mining and that is as follows:

(Parneet Kaur et. al 2015), focused on forecasting the failure in the core subjects. They applied dataset of 788 students Random Forest, Support Vector Machine, Decision Tree and Neural Network. They got highest predictive accuracy 93% using decision tree algorithm[2] .(Vili Podgorelec et. al. 2005) applied various data mining techniques on medical data to predict various diseases [3]. (Doron Shalvi et. al.1998) has suggested that how unsupervised neural networks are useful for performing medical data mining [4]. (Cindy et. al. 2006) have developed a mechanism about association rules based on parameters collected from breast cancer patients[5]. (Ranjit et. al. 2006) proposed a methodology for improving classification accuracy of Naive Bayes classifier algorithm related to medical data [6]. (Yasuyuki Tomota et. al. 2004) proposed a prediction model for Childhood Allergic Asthma (CAA).[7] (Jorng-Tzong Horng et. al. 2004) has presented a system for genetic factor identification responsible for cervical cancer.[8] (Xio Wang et. al. 2014) and (Lin Hua et. al. 2014) have proposed machine learning approaches for determination of disease susceptibility [9][10].

2. Tools and Techniques

Weka is a powerful data mining tool . WEKA contains 49 data pre-processing tools, 76 classification/regression algorithms, 8 clustering algorithms, 3 algorithms for finding association rules, 15 attribute/subset evaluators plus 10 search algorithms for feature selection[1]. Following are few algorithms supported be WEKA and used by us in this paper-

Random Forest : It is a data mining algorithm that is used for classification , regression and other tasks. It operates by constructing multitude of decision trees at training time and giving the class as output which is mean or mode of individual trees.

Naive Bayes : Naive Bayes is a kind of classifier which uses the Bayes Theorem. It predicts membership probabilities for each class. Naive Bayes classifier assumes that all the features are unrelated to each other. Presence or absence of a feature does not influence the presence or absence of any other feature.

Multilayer Perceptron : An MLP is a logistic regression classifier where the input is first transformed using a learnt non-linear transformation . This transformation projects the input data into a space where it becomes linearly separable. This intermediate layer is referred to as hidden layer. A single hidden layer is sufficient to make MLPs a universal approximator.

ZeroR : ZeroR is the simplest classification method which relies on the target and ignores all predictors. ZeroR classifier simply predicts the majority category (class). Although there is no predictability power in ZeroR , it is useful for determining a baseline performance as a benchmark for other classification methods.

3. Proposed work

We are using the available breast cancer dataset which we will use to train our neural network. We with the help of WEKA will create a neural network and use classifier algorithms such as Random Forest, ZeroR, Multilayer Perceptron, Naive Bayes. We will use some part of dataset to first train the network and then with the other part we will test

the accuracy of neural prediction. We will be splitting our dataset into train and test data. We will be using training data : test data ratio as 90:10,70:30 and 50:50.We will analyse performance of prediction of neural network for different algorithms (classifiers) .

In the dataset we will be training with the help of features such as Class, age, menopause:, tumor-size, inv-nodes, node-caps, deg-malig, breast, breast-quad, irradiat. Based on the above features of the dataset we will be comparing performance of classification algorithms on the basis of parameters as correctly classified instances , incorrectly classified instances , classification accuracy , kappa statistics, time taken. We will be applying following work flow for our work.

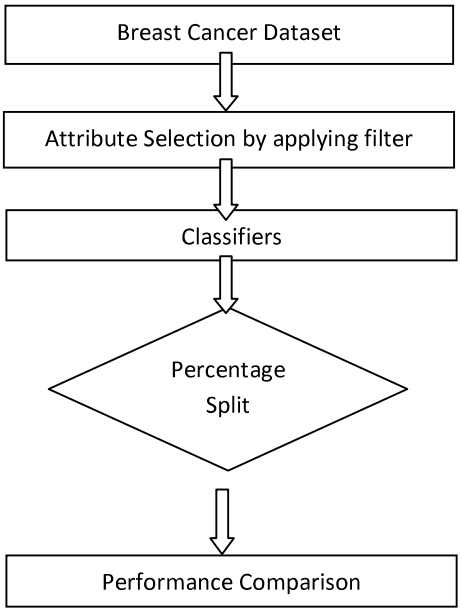


Table I : performance comparison using 90% percentage split (Total instances :286, Total trained : 258, Total tested : 29)

Classification Algorithms	Correctly classified	Incorrectly classified	Classification Accuracy (%)	Kappa Statistics	Time (Sec.)
Random Forest	21	8	72.4138	0.3256	0.02
Naive Bayes	19	10	68.9655	0.2606	0.02
Multilayer Perceptron	22	7	75.8621	0.394	0.01
ZeroR	19	10	65.5172	0	0

Table II : performance comparison using 70% percentage split (Total instances :286, Total trained : 200, Total tested : 86)

Classification Algorithms	Correctly classified	Incorrectly classified	Classification Accuracy (%)	Kappa Statistics	Time (Sec.)
Random Forest	60	26	69.7674	0.2897	0.06
Naive Bayes	59	27	68.6047	0.306	0.03
Multilayer Perceptron	60	26	69.7674	0.2995	0
ZeroR	54	32	62.7907	0	0

Table III : performance comparison using 50% percentage split (Total instances :286, Total trained : 143, Total tested : 143)

Classification Algorithms	Correctly classified	Incorrectly Classified	Classification Accuracy (%)	Kappa Statistics	Time (Sec.)
Random Forest	99	44	69.2308	0.2082	0.02
Naive Bayes	104	39	72.7273	0.335	0.02
Multilayer Perceptron	105	38	73.4266	0.3559	0
ZeroR	96	47	67.1329	0	0.01

Figure 1: Comparison of classifiers with percentage split 90%

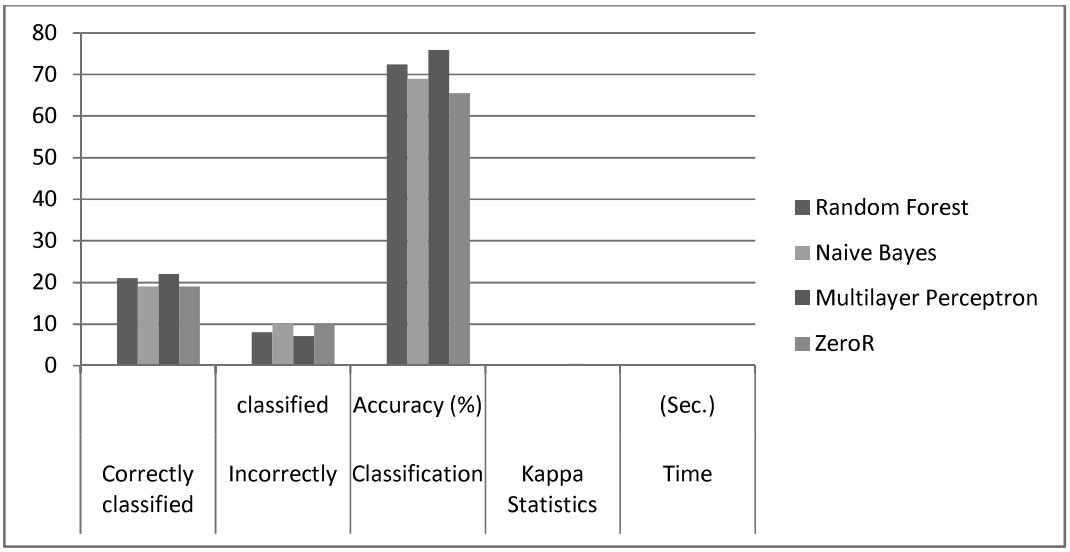


Figure 2: Comparison of classifiers with percentage split 70%

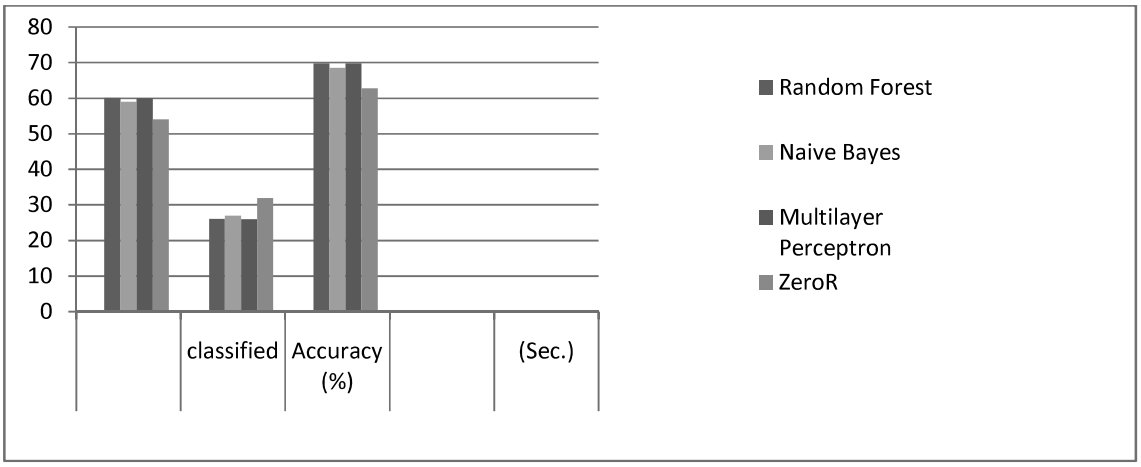
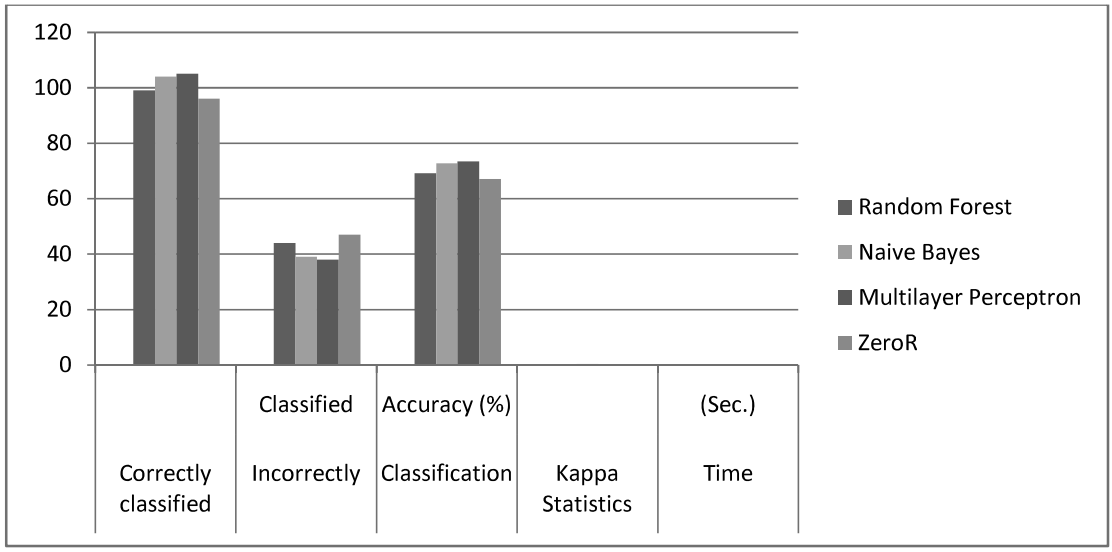


Figure 3: Comparison of classifiers with percentage split 50%



CONCLUSION :

The simulation results show that for 90% percentage split , classification accuracy of multilayer perceptron is highest (75.8621%). For 70% percentage split , Random Forest classifier and Multilayer Perceptron are performing identically having classification accuracy 69.7674%. And for 50% percentage split, multilayer perceptron is performing better with classification accuracy 73.4266%. Hence, we can conclude that performance of Multilayer Perceptron is better than Random Forest, Naive Bayes, and ZeroR classifiers. Thus, breast cancer can be classified more accurately using Multilayer Perceptron classifier than other three classification algorithms. Future scope of this work lies in comparing these classifiers using more number of parameters.

REFERENCES :

1. Anand Kishor Pande, Dharmveer Singh Rajpoot(2016) A comparative study of classification techniques by utilizing WEKA , 978-1-5090-2684-5/16/$31.00 ©2016 IEEE
2. Parneet Kaur, Manpreet Singh and Gurpreet Singh Josan, " Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," 3rd International Conference on Recent Trends in Computing, ProcediaComputer Science 57, Elsevier,pp.500-508, 2015.
3. Vili Podgorelec, Marjan Heriko, I. Rozman, "Improving Mining of Medical Data by Outliers Prediction" In IEEE Symposium on Computer-Based Medical Systems (CBMS'05), pp.91-96, ISBN: 0- 7695-2355-2, 2005
4. Doron Shalvi and Nicholas DeClaris, "An Unsupervised Neural Network Approach to Medical Data Mining Techniques", In IEEE proc of International Joint Conference on Neural Networks, pp. 171-176, Vol.1, ISBN: 0-7803-4859-1, 1998
5. Cindy L. Bethel et al., "Mining for Implications in Medical Data", In IEEE proc of 18th International Conference on Pattern Recognition (ICPR'06), pp. 1212 – 1215, vol. 1, 2006, ISBN: 0-7695-2521-0
6. Ranjit Abraham, Jay B.Simha, Iyengar, "A comparative analysis of discretization methods for Medical Datamining with Naïve Bayesian classifier", In proc of IEEE international conference on information Technology, pp. 235 - 236, 2006, ISBN: 0-7695-2635-7
7. Yasuyuki Tomita , Shuta Tomida, Yuko Hosegawa, Uoichi Suzuzki, Taro Shirakawa , Takeshi Kobayashi and Hiroyuki Honda (2004), Artificial neural network approach for selection of susceptible single nucleotide polymorphisms and construction of prediction model on Childhood Allergic Asthma , BMC Bioinformatics BioMed Central
8. Jorng-Tzong Horng, Kai-Chih Hu, Li-Cheng Wu, Hsien-Da Huang, Horn-Cheng Lai, and Ton-Yuen Chu (2004), Identifying the Combination of Genetic Factors that Determine Susceptibility to Cervical Cancer, IEEE Transactions on Information Technology in Biomedicine (Volume: 8, Issue: 1, March 2004)
9. Xiao Wang, Qinke Peng and Tao Zhong (2014), Analysis of Disease Association and Susceptibility for SNP Data Using Emotional Neural Networks , 2014 International Joint Conference on Neural Networks (IJCNN) July 6-11, 2014, Beijing, China
10. Lin Hua, Hong Xia, Ping Zhou, Li An(2014), Apply Support Vector Regression to Extract the Potential Susceptibility Genes of Chronic Obstructive Pulmonary Disease, 2014 7th International Conference on BioMedical Engineering and Informatics (BMEI 2014)

IMPACT OF NOMADIC MOBILITY MODEL AND MANHATTAN GRID MOBILITY MODEL ON PERFORMANCE OF AODV ROUTING PROTOCOL

VAISHALI DHOKCHAWLE AND DR. JANARDAN PAWAR
Asst.Pof ICCS and Vice-Principal , ICCS
Vaishali.dhokchawle@gmail.com, janardanp@iccs.ac.in

Abstract: *A Mobile Ad-hoc Network (MANET) is a collection of wireless mobile nodes forming a network without using any infrastructure. All mobile nodes function as mobile routers that discover and maintain routes to other mobile nodes of the network and therefore, can be connected dynamically in an arbitrary manner. The mobility attribute of MANETs is a very significant one. The mobile nodes may follow different mobility patterns that may affect connectivity, and in turn protocol mechanisms and performance. There are two types mobility models ,one is whose movements are independent of each other's named as entity mobility models and other is whose movements are dependent of each other's called as group mobility models. Finally performance behavior of Ad-hoc On-demand Distance Vector (AODV) routing protocol with Manhattan mobility model and Nomadic Mobility model with respect to throughput, end to end delay and routing load is analyzed on NS-2.34 simulator.*

Keywords : *Mobile ad hoc network (MANET), Nomadic Mobility Model, Manhattan Mobility Model, AODV, mobility.*

INTRODUCTION:

In the 1990s, with the advent of note-book computer the concept of commercial ad-hoc network arrived which give birth to mobile nodes. Within the IETF, the MANET working group was born, and made effort to standardize routing protocols for ad hoc networks. There are two kinds of mobile wireless networks. The first is known as infra-structured networks with fixed and wired gateways that is "one-hop" wireless network include wireless local area networks (WLANs). The second type of mobile wireless network is the infrastructure less mobile network, known as the MANET. MANET is usually a self-organizing and self-configuring "multi-hop" network which does not require any fixed infrastructure. In MANET, all nodes are dynamically and arbitrarily located, and data delivered across network with the help of other nodes in the network. Since the topology of the network is constantly changing at unpredictable times, the issue of routing packets between any pair of nodes is crucial task. Thus, mobility of nodes in the network plays indispensable role in evaluating performance of routing protocol in MANET. Various mobility model are advert in [4].Studies shows that most of simulation is done using Random Waypoint Mobility model [5,6].

There are various mobility models which are proposed for different applications such as military battlefield, civilian environments like taxicab, sports stadium, boat and small aircraft. In commercial sector such as emergency/rescue operations for disaster relief efforts, e.g. in fire, flood, or earth-quake ,ship-to-ship ad hoc mobile communication, law enforcement, Wireless LAN (WLAN), GPRS, and UMTS. Personal area network, as short-range MANET can simplify the intercommunication between various mobile devices (such as a PDA, a laptop, and a cellular phone).Each of these application required different mobility patterns.

In this paper performance of AODV routing protocol is analyzed with respect to application such as battle field mobility (Nomadic Mobility area) and city road mobility (Manhattan Grid Mobility model) using simulator NS-2.34.

2. Literature review:

Three different components of the problem have been exposed in [4]:

1) Performance measures (throughput, overhead, end to end delay) of protocol can vary widely with mobility model. e.g. - If wrong mobility model is used for simulation, simulation result may show that very low overhead (5 %) but when deployed significant large overhead may be used (50%).

2) Simulation may show that when transmission range increases throughput increases but on deployment it may decrease.

3) AODV throughput is better than DSDV in RWP, however with RPGM, DSDV is better.

In [14], Valentina comes to conclusion that DSDV in more stable in Reference Point Group Mobility (RPGM),Random Waypoint (RW), Gauss-Markov (GM) and Manhattan Grid (MG) and best in GM and MG.AODV is best in RPGM, group model and in entity models, AODV have the highest routing overhead with the increase of node speed, but has acceptable average delays. DSR have lowest routing overhead, on the count of higher average delays, with MG and GM models, at higher node speeds. DSR best with the RW model.

In [10], DSR is best for Packet delivery ratio out of DSDV, DSR and AODV for three mobility models Random Waypoint, Gauss Markov and Manhattan grid when compared for performance analysis under varying number of nodes.

Work of Manjusha S.in [7], revels that in Random Waypoint, DSR is best because it has high PDR ratio and less routing load. For RPGM mobility model, PDR for AODV and DSR is nearly same. For Manhattan Grid, DSR perform better as it has high PDR and less routing load.

Narinder Pal in [9], concludes that OLSR give the best performance in terms of throughput, and end-to-end delay and AODV, OLSR and GRP routing protocol in Vector Mobility Model outperforms Random Way Point Model.

3. Problem Statement

If MANET for real life applications (campus area, city street way, express way, shopping mall, cafeteria area, museum) are deployed based on simulation results of wrongly

selected mobility model, resulting behavior of network after deployment will be not as expected and there will be time as well as money wastage. To avoid such problem proper mobility model for concern application has to be studied thoroughly and selected properly.

Thus, behavior of various MANET routing protocol for specific application need to be studied to avoid further complication after MANET deployment.

4. Mobility Model

1) Nomadic Community Mobility Model

It is group mobility model, where movement of nodes depends on group leader. Movement happens in group. Best examples of this model are ancient nomadic societies moved from location to location or consider a class of students touring an art museum. Within each community or group of mobile nodes maintains their own space, where they move randomly using Random Walk mobility model [5,6].

2) Manhattan Grid Mobility Model

It was proposed by Bai at al.[10] in 2003.This model is categorized as entity mobility model. Manhattan mobility models the mobility of nodes on street defined by maps. Each street has bidirectional moment. The mobile nodes are is free to move along horizontal and vertical lines in grid and at intersection of grid, the mobile node can turn left or right or go straight. The Manhattan model employs a probabilistic approach in the selection of nodes movements, since, at each intersection, a vehicle chooses to keep moving in the same direction. The probability of going straight is 0.5 and taking a left or right is 0.25 each. From the above discussion, it is very clear that this model is not suitable for highway systems. Although this model provides flexibility for the nodes to change the direction, it imposes geographic restrictions on node mobility [8].

4. Routing Protocols

Routing protocol is used to find route between source node and destination node. Various type of routing protocols are proposed for MANET. Which routing protocol is best for particular scenario in needed to evaluate as each routing protocol has unique feature. Routing protocol is classified into two-proactive or table drive and reactive or on demand routing.

Ad-hoc On Demand Routing Protocol (AODV)-

It is on demand or reactive routing protocol. If any source wants to send packets to a destination, it broadcast a route request packet (RREQ).The neighboring nodes in turn broadcast packets to its neighbor till it reaches its destination. During this forwarding process, intermediate nodes are recorded from which first copy of packet is reached. This record is stored in their route tables, which helps for establishing a reverse path. If additional copies of the same RREQ are received later on, they are discarded. The replay is send using reverse path. For route maintenance, when a source nodes moves, it can re-initiate a route discovery process. In case of link failure in the path, source node is informed and source node may decide to re-discover route.

5. Simulation Setup

Here simulation is performed using mobility models Nomadic Mobility Model and Manhattan Grid Mobility Model on AODV MANET routing protocol using parameters given in Table-1.Experiment is performed using NS 2.34 on fedora 21 operating system with Intel core i3 processor. Set of simulation scripts for NS2 simulation environment merged with BonnMotion scenario generation tools.

Parameters	**Values**
Routing Protocol	AODV
Simulation Area	1000m X 1000 m
Simulation Time	600 sec
Number of Nodes	50,100,150
Queue Length	100
Queue Type	Priority Queue
Mobility Models	Nomadic Mobility Model, Manhattan Grid Mobility Model
CBR Traffic	Agent -UDP, Packet Size-512 , Max packet Size-1000, Send rate-8 mbps, Maximum Connection-9
Mac_Protocol	802_11

Table-1: Simulation Parameter for Experiment in NS2.34 simulator.

5.1 Mobility Generation

In Nomadic mobility model, maximum speed is 5meter/sec and minimum speed is 0.25 meter/sec with pause time of 15 sec and group size of 20 nodes.

./bm -f 50nodesNomadic Nomadic -i 360 -x 1000 -y 1000 -d 600 -a 20 -s 2.0 -r 2.5 -c 60 -h 5.0 -l 0.25 -p 15.0 _-R 1

The above command creates Nomadic scenario with initial phase of (-i) 360 seconds in 1000X1000 area(-x,-y) for duration (-d) 600 seconds ,average mobile nodes per group is 20 nodes (-a),group size deviation 2.0 (-s), maximum distance from center 2.5 (-r) ,maximum pause time 60 seconds (-c), maximum speed of 5.0 meter/second (-h),minimum speed of 0.25 meter/second (-l), maximum pause time 15 seconds (-p) and random seed of 1 (-R).

For Manhattan Grid mobility model, following BonnMotion command is used.

./bm -f 50nodesManhattanGrid ManhattanGrid -i 360 -R 1 -x 1000 -y 1000 -d 600 -n 50 -u 3 -v 5 -q 5.0 -t 0.4 -c 0.2 -e 0.5 -m 1.0 -s 0.2 -p 0.40 -o 120

The above command creates Manhattan Grid scenario with initial phase of (-i) 360 seconds ,random seed (-R) 1 in 1000X1000 area(-x,-y) for duration (-d) 600 seconds , number of mobile nodes 50 (-n),number of X-axis blocks is 3 (-u),number of Y-axis blocks is 5 (-v),update distance is 5.0 (-q),turn probability is 0.4 (-t),speed change probability (-c) 0.2, minimum speed of 0.5 m/s (-e) ,mean speed of 1.0 (-m),speed standard deviation of 0.2 (-s),pause probability of 0.40 (-p),

maximum pause time 120 seconds.

5.2 Traffic generation at application layer

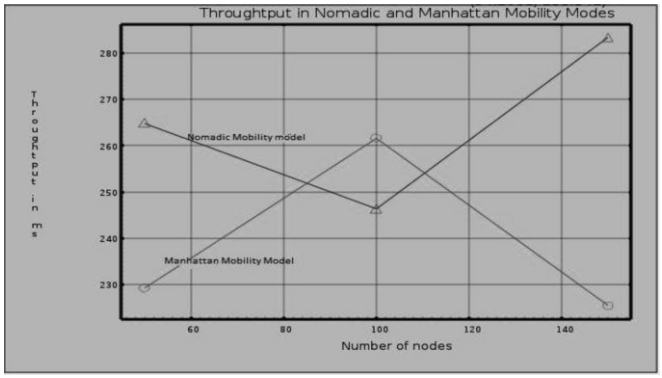
Constant bit rate (CBR) traffic is generated using "cbrgen.tcl" file in present at "~ /ns-2.34/indep-utils/cmu-scen-gen/cbrgen.tcl" of NS2.34 simulator. It can generate traffic using beneath tcp or udp transport layer protocol.

#ns cbrgen.tcl -type udp -traffic cbr -nn 10 -seed 1.0 -mc 9 -rate 8.0 >10cbr

This command will generate 9 CBR (-mc) connections randomly chosen 10 (-nn) nodes, default packet size is 512 bytes with rate of 8.0 mbps (-rate) and seed value of 1.0(-seed) which is redirected to user defined file "10cbr".

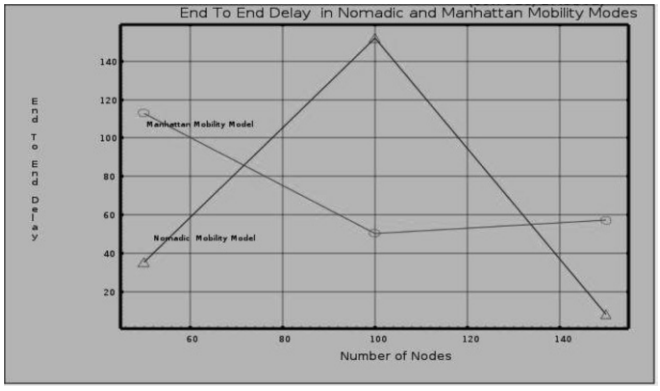
6. Simulation Result and Analysis

Performance of AODV routing protocol is analyzed in Manhattan Grid and Nomadic mobility models environment for metrics such as for throughput, end-to-end delay and routing load.



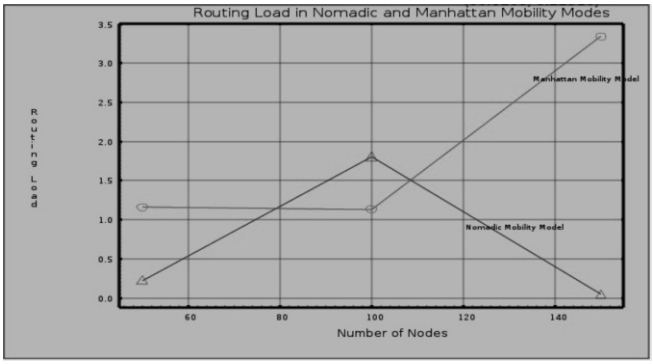
Graph-1: Throughput Comparison in Nomadic Model and Manhattan Grid Model for AODV Routing Protocol

From Graph-1, it is observered that AODV routing protocol throughput is up by 19.58% in Nomadic mobility model than Manhattan model for 50 mobile nodes and for 100 nodes AODV does well by 74.03% in Nomadic model again than in Manhattan mobility model.



Graph-2: End To End Delay in Nomadic and Manhattan Grid Mobility Model for AODV Routing Protocol

From graph-2, it is observed that end-to-end delay for 50 numbers of mobile nodes, AODV performances is 59.63% better in Manhattan Model than Nomadic Model. AODV performance well for 100 numbers of mobile nodes by 50.28% in Nomadic Model with drop by 6.54% for 150 numbers of mobile nodes.



Graph-3: Routing Load comparison in Nomadic and Manhattan Grid mobility Model for AODV Routing protocol

Graph-3 revels that AODV routing load is 73.67% higher in Manhattan Model than in Nomadic model for 50 mobile nodes. For 100 mobile nodes it is 32.88% greater in Nomadic Model than in Manhattan Model. For 150 mobile nodes routing load in Nomadic model suddenly drop by 97.99% compared to Manhattan Model.

From Graph-3, it is observed that AODV routing load is better in Nomadic Model than in Manhattan Grid model. AODV is better choice students touring art museum. So choosing appropriate mobility model according to situation plays vital role for analysis purpose.

CONCLUSION :

From this empirical analysis, AODV routing protocol behave differently in Manhattan Grid and Nomadic Mobility Model for 50,100,150 number of nodes.

1. The performance of MANET routing protocols can vary significantly with different mobility models. As concluded from above experiment, performance of AODV routing protocol is affected by Mobility models in which it is applied.
2. AODV routing protocol performs well in Nomadic mobility model as for 50 numbers of mobile nodes having high throughput, low end to end delay and low routing load. For 100 numbers of mobile nodes, throughput aberrantly goes down with high end to end delay and high routing load. For 150 numbers of mobile nodes AODV routing protocol gives good throughput with minimum end to end delay and minimum routing load.
3. While AODV routing protocol in Manhattan Grid mobility model, for 50 numbers of mobile nodes throughput is low, end to end delay is high and routing load is high. For 100 numbers of mobile nodes throughput is high, end to end delay is low and routing load is low. For 150 numbers of mobile nodes

throughput is low, end to end delay is high and routing load is high.

4. Real life example of Nomadic mobility model is people moving in group for instance group of students/visitors/tourist visiting museum or historical monuments. For 50 and 100 number of people moving in group AODV routing protocol is worthy choice.
5. Real life example of Manhattan Grid is movement of people along roadways in cities. For 100 numbers of mobile nodes AODV routing protocol is superior choice to get high throughput.

REFERENCES :

[1] B. Liang and Z. Haas. Predictive distance-based mobility management for PCS networks. In Proceedings of the Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), March 1999.

[2] F. Bai,Narayanan Sadagopan,A. Helmy,"IMPORTANT: a framework to systematically analyze the Impact of Mobility on Performance of Routing Protocols for Adhoc Networks", in Proceedings of IEEE Information Communications Conference (INFOCOM 2003),San Francisco,Apr.2003

[3] Fan Bai and Ahmed Helmy, "A SURVEY OF MOBILITY MODELS in Wireless Adhoc Networks".

[4] F. Bai, N. Sadagopan, and A. Helmy, "Important: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks," in Proc. of IEEE INFOCOM 2003, (San Francisco, CA), March/April 2003.

[5] M. Sanchez. Mobility models. http://www.disca.upv.es/misan/mobmodel.htm. Page accessed on May 30th, 2002.

[6] M. Sanchez and P. Manzoni. Anejos: A java based simulator for ad-hoc networks. Future Generation Computer Systems, 17(5):573–583, 2001.

[7] Manjusha S. Patil Bokade, Prof. M.N.Thakare , Prof. B. J. Chilke,"Simulation of MANET Routing Protocols DSDV, DSR and AODV for different Mobility Models", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 6, November-December 2014 ISSN 2278-6856

[8] Mobility Models, Broadcasting Methods and Factors Contributing Towards the Efficiency of the MANET Routing Protocols: Overview, Shafinaz Buruhanudeen, MohamedOthman, Mazliza Othman, Borhanuddin Mohd Ali. Paper ID: 123

[9] Narinder Pal, Renu Dhir, "Analyze the Impact of Mobility on Performance of Routing Protocols in MANET Using OPNET Modeller",International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013 ISSN: 2277 128X

[10] Renu Bahuguna, Hardwari Lal Mandoria,"Simulation Based Performance Comparison of MANET Routing Protocols Using Different Mobility Models and Node Density", International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 5, Issue 7, July 2015.ISSN: 2277 128X

[11] S. Sesay et al., "Simulation Comparison of Four Wireless Ad Hoc Routing Protocols", Information Technology Journal 3(3), 2004, pp: 219-226.

[12] S. Shah et al, "Performance Evaluation of Ad Hoc Routing Protocols Using NS2 Simulation", Conf. of Mobile and Pervasive Computing, 2008.

[13] S. Basagni, I. Chlamtac, V.R. Syrotiuk, and B.A. Woodward. A distance routing effect algorithm for mobility (DREAM). In Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM), pages 76–84, 1998.

[14] Valentina timcenko, "MANET Routing Protocols vs. Mobility Models: Performance Analysis and Comparison", proceedings of the 9th WSEAS International Conference on applied informatics and communications (AIC '09)

TO DESIGN A STRUCTURED MODEL FOR COMPARISON AND SUMMARIZATION OF EVENTS OF ANY DISCUSSION FORUM

MRS. VISHRANTI GHARGE AND MRS.KAVITA DHAKAD
Indira College of Commerce and Science, Pune
vishranti.gharge@gmail.com,, kavita.dhakad1@gmail.com

Abstract : *This structured model is used to help and determine some problems and findings at discussion forums. Model shows the process to analyze discussion forums from television news channel broadcasting and other websites and generate the comments based on the events discussed by Business Intelligence tools. By the use of analytical techniques and tools measure trustworthiness of contents based on similarities and contradictions generated by speech or audio-video segments from media and merging the data together to show comparison and summarization in the form of report.*

Keywords : Discussion forum, Business Intelligence tools (BI), Data Analytics, Text Mining and Data Visualization.

INTRODUCTION:

Now days everybody is getting information via television news channels and by different ways like Web Medias. There are so many regulating issues in India like Education System, complicated taxes and licensing systems and so on. Also, monopoly by government controlled institutions on certain goods and services delivery and the lack of transparent laws and processes.

So, this proposed structured model will help to show the detail process, reports or contents based on similarities and contradictions generated by speech or audio-video segments from media (Television news channels, Websites etc.) broadcasting relate to any subject. It will be a good initiative to develop such a useful model which will be helpful to all the people all the way, when it is implemented in the form of application.

This model is applied for any business, where all discussion forums is generated. Here we designed an integrated structured system which will give all the detail reports in the form of summarized comments or charts from audio or video discussion forum.

Now after recording the audio or video files, audio/video to text transcription services is used , It is incredibly accurate and reliable, boasting a hybrid transcription process, while using the transcription process, speech recognition technology is first used and then expert transcriptionist come to ensure the perfect text file. It is more efficient and effective way of converting audio/video to text by the use of powerful transcription services and can benefit to the research.

Apply Data Mining procedure for extracting information from sets of data. Data mining is mining knowledge from data so by the use of this method generates the knowledge base. Data mining is used for knowledge discovery, query language, classification and prediction, decision tree induction, cluster analysis.

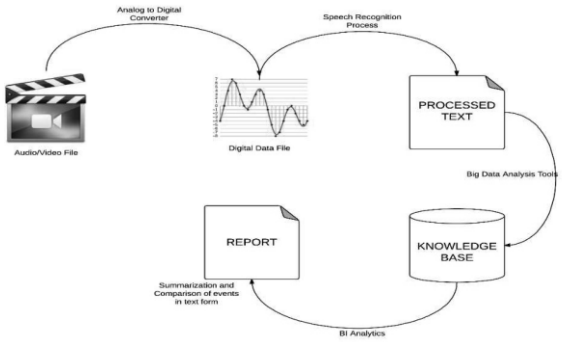
Business intelligence influences software and services to transform data into actionable intelligence that informs businesses to take tactical decisions.BI tools access and analyze knowledge base and present analytical findings in reports summaries, dashboards, graphs, charts and maps to provide users with detail intelligence about the state of business.

Objectives

1. To transform speech or audio-video segments to digital form, apply speech recognition process and acquire organized event knowledge base.
2. To generate the reports or charts by applying BI Analytics and text mining techniques, produce structured Evidences from knowledge base.

Methodology

1. To analyze multi-model contents from television news channel broadcasting and other websites, there is first process of voice recognition and convert voice into electrical signals and these signals transformed into text. Through the use of voice to text transcription service converting input voice recording into text file. The software first identifies the audio segments containing speech, and then it recognizes the language being spoken if it is not known a priori, and finally it converts the speech segments to text and time-codes. Speech engine loads the list of words to be recognized and compare with its own acoustics model. It then determines which words in the grammar the audio most closely matches and returns the results. Speech recognition followed by four steps like Signal Processing, Grammar Design Phoneme recognition, Word recognition and result generation. The result is a fully annotated XML document including speech and non-speech segments, speaker labels, words with time codes, high quality confidence scores and punctuations. This XML file can be directly converted into plain text.
2. This text file is processed by text mining technique and generates high quality information is typically derived through the devising of patterns and trends through such as statistical pattern learning into knowledge base. Using BI analytics this knowledge base converts it to actionable information which can be used for data visualization.
3. Visualized data measures the trustworthiness of contents based on similarities and contradictions generated by speech or audio-video segments from media and in this way, we get appropriate text data in minimum time to compare.



CONCLUSION :

Proposed system structure will be helpful for analyzing the discussion held between two or more persons on media based on any events. It will be useful for data visualization based on his or her views shared. This model used analytical techniques and tools to measure reliability of contents based on similarities and contradictions generated by speech or audio-video segments from media and merging the data together .This produces comparison and summarization reports. Summarized reports and comments are used to detect fraud or corruption held at different places, decision making. The data is collected from different sources (Audio or video segments) to carry out analysis and associations that will be helpful to come to a tangible conclusion.

REFERENCES :

1. Mary K. Patt(2017),Article on Business Intelligence https://www.cio.com/article/2439504/business-intelligence/business-intelligence-definition-and-solutions.html
2. Thomas H. Davenport, Jeanne G. Harris(2007),"Competing on Analytics: The New Science of Winning",
Harvard Business School Press.
3. Rick Sherman(2014), "Business Intelligence Guidebook: From Data Integration to Analytics", MK Publisher,
4. Jiawei Han, Jian Pei, Micheline Kamber (2011),"Data Mining: Concepts and Techniques",MK Publisher ,Third Edition.
5. https://transcribeme.com/blog/the-affordable-easy-way-to-convert-audio-to-text
6.https://www.tutorialspoint.com/data_mining

RESEARCH REVIEW ON THE EFFECT OF SMARTPHONE EM RADIATION ON HUMAN BODY

VIJAY DAULATRAO MORE[1] AND AVINASH SHINGTE[2]
[1]Assistant Professor, Indira College of Commerce And Science, Tathwade, Pune, 411033
E-mail: vijay.more@iccs.ac.in
[2]Assistant Professor, Indira College of Commerce And Science, Tathwade, Pune, 411033
E-mail: avinash.shingte@iccs.ac.in

Abstract : *In recent years, there has been a tremendous growth of smart phones with over several billions of users worldwide. Smart phone has made a great impact with our society. It is hard to imagine lives without mobile phones. But the use of mobile phones and it's interaction with the human body raised the question of whether the electromagnetic waves from mobile phone is harmful or not? This paper gives a brief review of all possible effects of mobile phone radiation with human tissues. This review paper will provide answers for public health concerns about the risk of using mobile phones. The growth in the use of mobile phone has raised the concerns about the possible interaction between the electromagnetic fields (EMF) radiation and the biological effects on human tissues, particularly the brain and the human immune system. The Wireless communication is experiencing a dynamic growth in the global scale and the mobile phones are becoming a vital device in the global modern society. Mobile phones radiate radio frequency waves, a form of non-ionizing radiation, which can be absorbed by tissues closest to where the phone is kept.*

Keywords : Smartphone, mobile phone EM wave, human body

INTRODUCTION:

The smartphones, being a very new invention of humanity, became an inherent part of human's life. Smartphones contains advanced computing capability, such as internet communication, information retrieval, video, e-commerce and other features, that makes the device is one of the necessities for many people. “Mass cell phone mobilization” covered humanity probably ten or fifteen years ago. The growing number of smartphones and smartphone owners raises a concern about phones' effect on human health and life.

A world-wide popularization of smartphones and a little knowledge about their side effects triggered to start research on effects of smartphones on human health.

Merriam-Webster dictionary defines the smartphone as “a cell phone that includes additional software functions (as e-mail or an Internet browser)”. The highlights of this research include recent scientific

facts and research analysis of the smartphones on human health. The main key points discussed in the research paper are the effect of electromagnetic waves on human brain.

One glance on mobile phone usage in worldwide [7]

Cell Phone Usage

		*On your cell phone, do you regularly...**			
	% Saying they own a cell phone	Make phone calls	Send text messages	Take pictures or video	Use the internet
	%	%	%	%	%
U.S.	85	96	67	57	43
Spain	96	98	70	59	21
Britain	89	87	79	54	38
Germany	88	82	56	27	18
France	85	95	77	51	28
Lithuania	91	99	79	47	24
Russia	86	99	75	50	27
Ukraine	84	100	72	48	19
Poland	78	99	85	56	30
Turkey	84	97	64	44	22
Jordan	95	94	63	43	23
Israel	95	99	73	57	47
Lebanon	79	100	87	33	19
Egypt	71	98	72	58	15
China	93	99	80	54	37
Japan	86	98	81	72	47
Indonesia	55	96	96	38	22
India	53	98	49	26	10
Pakistan	48	97	44	9	6
Mexico	57	89	82	61	18
Kenya	74	100	89	31	29
MEDIAN	**85**	**98**	**75**	**50**	**23**

* Asked only of those who say they own a cell phone.

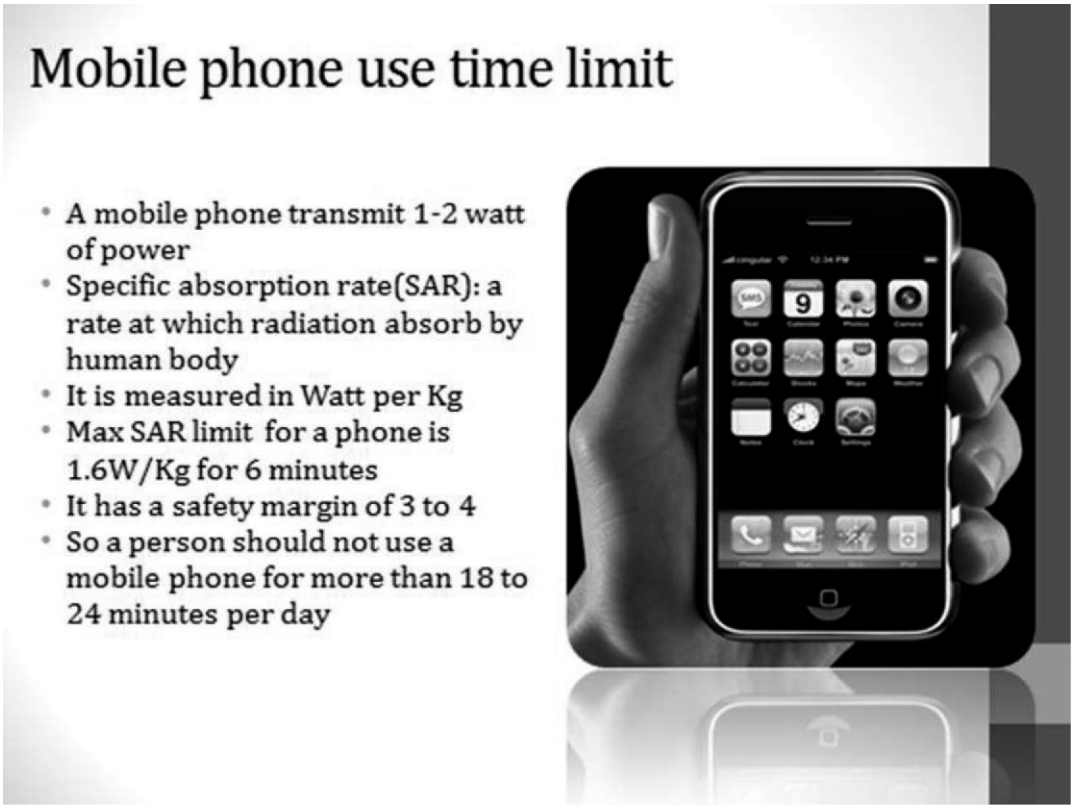
PEW RESEARCH CENTER Q68 & Q70a-d.

2. Mobile phone Effects

Several effects have been reported due to exposure to long term EM radiation from cell phones. These effects leaded us to divide the reported results into three major categories. These categories are as follows:

2.1. EM Waves

Due to the increased number of users using the mobile phone, the concern is now focused towards electromagnetic radiations emitted by the mobile phones itself. Electromagnetic radiation can be classified into ionizing and non-ionizing radiation. Ionizing radiation is the radiation with high energy which is able to remove tight bonds between electrons and atoms resulting in tissue damage while non-ionizing radiation is the radiation that has enough energy to vibrate the atoms and molecule but do not remove the electrons in the molecule [2]. This radiation mainly occurs at low frequency range. Mobile phone is designed with low power transceiver to transmit voice and data to base station is located at few kilometers. These radiations cause problems like headaches, severe pain in ear, blurring of vision, memory loss, itching, burning sensations, feeling asleep, hypersensitivity exhaustion [1] have been observed when using mobile phone. Researchers have found that these symptoms are more common in people with higher exposure to radiation of mobile phone.



Experimental research on the effects of radio-frequency radiation is very broad and heterogeneous. It includes both studies of cell cultures and tissues (in vitro) and of laboratory animals, as well as of people (volunteers). On one hand, these studies focus on functional changes in the brain and the resulting effects on cognition, and (to some extent) well-being – that is, the influence of exposure to radiation on the head. On the other hand, these studies focus on the possibility of a relationship between mobile phone use and carcinogenic processes, reproduction and development, the cardiovascular system and longevity – that is, exposure of the whole body. These studies found very small and reversible biological and physiological effects that do not necessary lead to diseases or injuries. Also, the research findings on the changes at the molecular level associated with the development of cancer are inconsistent and contradictory. Experimental research on the biological effects of RF and microwave fields is very broad and includes studies of volunteers, animals and in vitro, cell-based techniques. In vitro experiments that show abnormal cell proliferation, changes in cell membranes, and movement of ions and substances across membranes are difficult to extrapolate to people [13]. Finally, there is no evidence of non-thermal effects on human health. It is important to distinguish between biological (or physiological) effects and psychological and health effects. The demonstration of an RF or microwave radiation effect in experimental research does not necessarily mean that such exposure will lead to harmful effects on human health [14]. Human bodies, with the aid of their immune, nervous or endocrine systems, can effectively resist some external pressures, adapt to them and maintain the stability (homeostasis) disrupted by those changes.

2.2. Effect of electromagnetic waves on human brains

The smartphone is a source of the eminence of electromagnetic waves. Numerous studies have been conducted in the past years to identify the effect of electromagnetic waves emitted from the cell phones on human health. The topic has been studied for a long time, but in past, it touched on a rather narrow circle of people, mostly staff of broadcast and specialized radio stations. Even at that time, measures taken to protect people from radiation apply only on those who work near powerful sources of radiation. And, despite the revolutionary changes in the field of telecommunications, as well as many discoveries and emissions, the impact of electromagnetic waves of different frequencies hotly debated ever since. As soon as mobile phones more and more part of our lives, the world is continuing research to proof whether cell phones are harmful to human health? Today there is no official statement announced by laboratory or medical center to answer this question. The complexity of the analysis of the statistical data makes the task more difficult for researchers. The impact of harmful radiation emitted from cell phones waves is still being studied [15].

Nevertheless, researchers comment “that more work is needed to interpret the results, which some called "puzzling." The statement that cell phones can cause cancer has been not confirmed. The studies failed to proof that cellphones make a major risk develop cancer among frequent users. The main issues while conducting studies are some people may not accurately report the usage as they don't exactly remember how often they use the cell phone (excluding speaker phone or ear buds) and it is still difficult to measure the impact of other factors that may accelerate the cancer development for excessive cell phone users[3][4].

The WHO/International Agency for Research on Cancer (IARC) has classified radiofrequency electromagnetic fields as "possibly carcinogenic to humans" and associates it with wireless phone use. In May 2011, a Working Group of 31 scientists from 14 countries met in France to assess the potential carcinogenic hazard from exposure to radiofrequency electromagnetic fields emitted by wireless communication devices, microwaves, radio, and television signals. The Working Group made a conclusion that "the evidence, while still accumulating, is strong enough to support a conclusion and the 2B classification (carcinogenic to human)". It means that a risk of hazard exposure emitted from the cell phones exists to cause cancer, and therefore, additional observations are required [11][12].

2.3. Genotoxic, Blood Brain Barrier (BBB), and immune system effects.

Some research studies show association between cell phone radiation and cell changes including damage to chromosomes, alterations in the activity of certain genes and a boosted rate of cell division. They have studied the effects of EMF radiation on the rat brain. He found that the exposure to MRI induced leakage of Evans Blue labeled proteins normally not passing the BBB of rats. The study exposed male and female Fischer of 344 rats in a transverse electromagnetic transmission line chamber to microwaves of 915 MHz as continuous wave and SAR varied between 0.016 and 5 W/kg. All rates were sacrificed by perfusion-fixation of the brains under chloral hydrate anesthesia about 1 hour after the exposure. The brains were per fused with saline for 3-4 minutes and thereafter fixed in 4% formaldehyde for 5- 6 minutes. The central coronal sections of the brains were dehydrated and embedded in paraffin and sectioned at 5 microns [10].

2.4. DNA and cancer effects.

Can microwaves disrupt the covalent bonds of DNA? The fundamentals of thermodynamics and physics indicate this is impossible. Numerous studies have concluded that there is no evidence to support the existence of the 'Microwave Effect', and yet, some recent studies have demonstrated that microwaves are capable of breaking the covalent bonds of DNA. The exact nature of this phenomenon is not well understood, and no theory currently exists to explain it.

Most of the studies mentioned above concluded that the microwave effect, if it existed, was indistinguishable from the effects of external heating. However, it was recently demonstrated that the microwave effect is distinguishable from external heating by the fact that it is capable of extensively fragmenting viral DNA, something that heating to the same temperature did not accomplish. This experiment consisted of irradiating a bacteriophage PL-1 culture at 2450 MHz and comparing this with a separate culture heated to the same temperature. The DNA was mostly destroyed, a result that does not occur from heating alone.

Demonstrating that radiation causes adverse effects on health would signal a widespread public health problem. Mobile phones have been in extensive use for a relatively short period of time, and their technology has progressively changed, from analogue to digital systems. Mobile phones and base stations emit radio frequency or microwave radiation. Exposure to such a radiation could affect health directly. The use of mobile phones also results in indirect effects, such as car accidents and interference with health equipment.

The studies cover the effects of RF and microwave radiation between 100 MHz and 60 GHz and focus both on the functional changes in the brain (influence of exposure to RF and microwave fields on the head) and on carcinogenic processes, reproduction and development, the cardiovascular system and longevity (as a result of whole body exposure to RF and microwave fields)[9]. Since the radiation from mobile phones and signal stations does not have enough energy to break chemical or molecular bonds directly, there is no basis in theory to suggest that they can damage DNA. Moreover, a biological mechanism that explains any possible carcinogenic effect from RF or microwave fields has yet to be identified. Because of the difficulties in interpreting findings from laboratory studies, the hypothesis that RF or microwave radiation is harmful and could have effects on health that have not yet been recognized cannot be rejected. Indirect experimental results are difficult to extrapolate.

Clinical effects

Within human population studies, epidemiological studies provide the most direct information on the long-term effects on health of any potential harmful agent. To assess the adverse effects on health that may result from the use of mobile phones, research with a specific focus on cancer has been carried out. By the end of the 1990s, the number of studies was small and the works presented major methodological limitations, the most outstanding one being the lack of enough people with an exposure time long enough to accurately assess the potential adverse late effects on health of mobile phone use.

The majority of those studies suggested the need for additional, high-quality research. As a result of these recommendations, a series of multinational case-control studies, coordinated by the International Agency for Research on Cancer (IARC), were set up after a detailed feasibility study was carried out in 1998 and 1999[5][6].

Safety Measures:

A. There should be some law in every country for hearing on public health threat by exposure to transmitted radio frequency radiation.

B. Radiation emitting devices should require some health testing prior to approval.

C. Continuously transmitted devices should be banned.

D. There should be some warning label on the cell phones and cordless phones.

E. While talking on cellphone, keep phone at speaker mode, with phone a hand's away or use a wired handset. Moreover, use a Bluetooth emitter, as it will decrease levels of microwave radiations.

F. Turn off handset when not in use.

G. Try to use phone when having full signal strength, when signal strength is poor it emits high amount of radiation.

H. Don't use Cell Phones in elevators, cars and in planes because it emits more radiation in enclosed metal spaces.

I. When at home, use wired landline as it emits less amount of radiation.

Future

It is desired to work on the Effect of mobile phone on Human Health employing the available two mobile phone technologies in the Country viz. GSM 3G and 4G. The work on this is being carried out in the laboratory with a number of

subjects and using the above mentioned mobile communication technologies using different type of modulation, emitted power and operating frequency. Response of human body is being studied under three conditions: when no phone is used, when a GSM 3G phone is used and while using a GSM 4G phone [8]. The characteristics of body signals so obtained would help to determine which type of communication technology is more suitable for human being in concern with human health. Results obtained will help designing a class of communication devices which have minimal effect on human health

CONCLUSION :

In this paper we reviewed and summarized some of the crucial research done to study the effect of cell phone radiation. The work was motivated by the fact that the public is concern about the danger of using cell phone. Long-term experimental follow-up is much needed. Cancer for instance, needs time to develop the reason why only studies took over 10 years were able to indicate link between cell phone and human tissues, particularly the brain and the human immune system. It's hard to reach a conclusion whether cell phone is harmful because most of the existing studies have not shown the same findings. And the effects may be very different depending on the type of electromagnetic radiation. GSM and GPRS (3G) phones for instance use pulsed radiation. In this case the levels rise and fall rapidly. 4G phones on the other hand use continuous levels. Since the 4G phones are becoming very common, we are planning to make more deep studies focuses on the effect of 4G cell phones on human[8].

The influence of cell phones and their effects on human health are still being tested and studied. However, addiction and huge reliance on cell phones carry some risks on human development and health. The risks are psychological, social, physical and emotional.

REFERENCES :

[1] Cheung, W. S. and F. H. Levien (1985). Microwaves made simple, principles and applications. Artech House, Inc. Denham, MA.

[2] Chipley, J. R. (1980). "Effects of microwave radiation on microorganisms." Adv. Appl. Microbiol. 26:129-145.

[3] Cope, F. W. (1976). "Superconductivity – a possible mechanism for non-thermal biological effects of microwaves." J. of Microwave Power 11: 267-270.

[4] National Radiological Protection Board. Mobile phones and health 2004: report by the Board of NRPB. Chilton, Didot, Oxfordshire, National Radiological Protection Board, 2004 (Documents of the NRPB, volume 15, no. 5;

[5] http://www.hpa.org.uk/radiation/publications/documents_of_nrpb/pdfs/doc_15_5.pdf, accessed 15 September 2006).

[6] WHO. The International EMF Project: health effects of static and time varying electric and magnetic fields: progress report 1998-1999. Geneva, World Health Organization, 1999 (document number: WHO/SDE/OEH/99.9;

[7] Pew research center: Smartphone Ownership 2013. Retrieved http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013

[8] Aruna Tyagi, Manoj Duhan, Dinesh Bhatia (2011) Effect of Mobile Phone Radiation on Brain Activity GSM Vs CDMA. IJSTM.

[9] Robert Leeb, Felix Lee, Claudia Keinrath, Reinhold Scherer, Horst Bisch,et al. (2007) Brain-Compute Communication: Motivation, Aim, and Impact of Exploring a Virtual Apartment Ieee Transactions On Neural Systems And Rehabilitation Engineering 15(4): 473-82.

[10] Scarella O Clatz, Lanteri S, Beaume G, Oudot S, Piperno S, et al. (2006) Realistic numerical modeling of human head tissue exposure to electromagnetic waves from cellular phones.

[11] WHO. WHO Workshop: sensitivity of children to EMF exposure. Geneva, World Health Organization (http://www.who.int/peh-emf/meetings/children_turkey_june2004/en/, accessed 15 September 2006).

[12] IARC. Studies of cancer in humans. Preamble to the IARC Monographs on Evaluation of Carcinogenic Risks to Humans. Lyon, International Agency for Research on Cancer. (http://monographs.iarc.fr/ENG/Preamble/currentb2studies humans0706.php, accessed 6 October 2006).

[13] Hardell L et al. Cellular and cordless telephones and the risk for brain tumors. European Journal of Cancer Prevention, 2002, 11:377–386.

[14] Auvinen A et al. Brain tumors and salivary gland cancers among cellular telephone users. Epidemiology, 2002, 13:356–359.

[15] Hardell L, Mild KH, Carlberg M. Further aspects on cellular and cordless telephones and brain tumours. International Journal of Oncology, 2003, 22:399-407.

APPLICATION OF FUZZY RATING SCALE IN INDUSTRIAL ROBOT SELECTION

P. J. PAWAR[1], POONAM PAWAR[2], SHITAL PASHANKAR[2]
[1]*K. K. Wagh Institute of Engineering Education and Research, Nasik. Maharashtra, India.*
[2]*Indira College of Commerce and Science, Pune*

Abstract: *The selection of robots to suit a particular application and production environment from among the large number available in the market has become a difficult task. Various aspects such as product design, production system, and economics, need to be considered before a suitable robot can be selected. Some of the attributes in robot selection such as man machine interface, programming flexibility etc. are expressed qualitatively. This qualitative data is usually converted into values using Likert scale. However, Likert scale has an important limitation of loss of significant amount of information. Hence in this paper an attempt is made to apply fuzzy rating scale for quantification of qualitative attributes of robot selection.*

Keywords: Industrial Robots, Likert scale, Fuzzy rating scale

INTRODUCTION:

Recent developments in information technology and engineering sciences have been the main reason for the increased utilization of robots in a variety of advanced manufacturing facilities. Robots with vastly different capabilities and specifications are available for a wide range of applications. The selection of robots to suit a particular application and production environment from among the large number available in the market has become a difficult task. Various aspects such as product design, production system, and economics, need to be considered before a suitable robot can be selected. The selection problem is particularly relevant in view of the likely lack of experience of prospective users in employing a robot. Indeed, robots are still a new concept in industry as a whole, and so it is not unusual for an industry to be a first-time robot purchaser. Many precision-based methods for robot selection have been developed to date.

The objective of a robot selection procedure is to identify the robot selection attributes, and obtain the most appropriate combination of the attributes in conjunction with the real requirements of the industrial application. A robot selection attribute is defined as a factor that influences the selection of a robot for a given industrial application. These attributes include: cost, configuration , load capacity, weight and size of the robot, type and number of end effectors, type of control, velocity of movements, type of programming, programming flexibility, reliability, repeatability, positioning accuracy, resolution, number of degrees of freedom, number of joints, their sequence and orientation, motion transformation characteristics, ease of operation, work volume, drive system, man-machine interface, vendor's service contract, training, delivery period, maintainability, ease of assembly, ease of disassembly, types and number of sensors used, availability or assured supply, management constraints, etc.[1]

Efforts need to be extended to determine attributes that influence robot selection for a given industrial application, using a logical approach to eliminate unsuitable robots, and for selection of a proper robot to strengthen the existing robot selection procedure. Pertinent attributes and the alternative robots involved are to be identified. Values of the attributes and their relative importance are to be obtained. An objective or subjective value, or its range, may be assigned to each identified attribute as a limiting value, or threshold value, for its acceptance for the considered robot selection problem. An alternative robot with each of its selection attributes, meeting the acceptance value, may be short-listed. After short-listing the alternative robots, the main task to choose the alternative robot is to see how it serves the attributes considered. As most of the attributes in robot selection such as programming flexibility, reliability, ease of operation, man-machine interface, vendor's service maintainability, ease of assembly, ease of disassembly, availability or assured supply, management constraints, etc. are expressed qualitatively, they need to convert into suitable rating for application of multi-attribute decision making method for robot selection. Likert scale is commonly employed for the rating of this qualitative data [2].

II. LIKERT SCALE V/S FUZZY RATING SCALE

Likert scaling, originally introduced by Rensis Likert [3], is the most widely used psychometric scale in survey research. It asks respondents to indicate their levels of agreement with a declarative statement. For a 5-point Likert scale, for example, each scale point could be labeled according to its agreement level: 1 = strongly disagree (SD), 2 = disagree (D), 3 = neither disagree nor agree (NN), 4 = agree (A), and 5 = strongly agree (SA). Depending on what is being measured, the scale labels may be worded differently. When measuring frequency, for instance, labels like "never-always" can be used; when measuring attitude, belief, or characteristic of the respondent, labels like "not very much-very much" are suitable. A well designed Likert scale should state the opinion, attitude, or belief being measured in clear terms and use the appropriate wording for scale points. Likert scales have been widely used to measure observable attributes in various social science measurement areas. However, Likert scale suffers from following serious drawbacks:

a) Likert scales fail to approximate intervals of ordinal data

b) The respondents are forced to make a choice from the given options that may not match their exact responses. They have to either select an answer from an insufficient range of responses or respond to an "acceptable" answer in the closed format. This miss-matching further worsens the information distortion problem.

c) Significant amount of information is lost and/or distorted due to the built-in limitations of the Likert method.

Whereas Likert scales or associated codings discretize concerned attributes into a small number of potential values, the use of the free response format would allow attributes to take either a large finite or infinite number of potential values. The spirit of Statistics, as the science of variation, randomness and chance, would be better captured by using this free response format than a Likert-like (or a coded Likert-like) one. Furthermore, the fuzzy scale is rich and expressive enough to find a value in it fitting appropriately the valuation/opinion/rating involving subjective perceptions in most of real-life situations, even if we constrain ourselves to find it in some operational classes of fuzzy sets, like trapezoidal, S- and Π-curves (Eshragh and Mamdani [4]). The fuzzy scale is a flexible method to obtain a numerical value from ordinal variables. Fuzzy Scale design could easily incorporate the knowledge of an expert in the fuzzification of input and in the building-up of control block-rules, as they allow for the best adaptability. These are strengths, but also weaknesses as they involve extremely subjective, ambiguous, and private decisions, when scientific procedures should be objective, unambiguous, and public. Actually, the architecture of the Fuzzy Scale could be public, but, for example, the number of rules grows exponentially and their control is arduous. Furthermore, the aggregation of variables proceeds through a tree, varying nonlinearly their impact on the output, as their influence depends on the levels of the knot where they enter. The responses are a sort of weighted average with unknowable weights. However, in spite of these difficulties, the Fuzzy Scale could be a valid and reliable tool to represent situations described by qualitative ordinal variables comparable with others. The Fuzzy Scale does not perform single-item analysis well, as it works on an aggregation of a small number of variables (as few as two or three). Many popular scales, indexes, and measures may be conceived as fuzzy sets denoting graded concepts. The combination of computational intelligence with Fuzzy Scale could represent a useful route for the analysis of individual behavior or judgments. Human action is characterized by great complexity as it involves a lot of variables, spatial, individual and/or group heterogeneity and contamination between the actors. Soft computing techniques could be an interesting strategy to model social actions because they allow for a representation of reality in its wholeness, without introducing restrictions or reductions, even if they still require further refinement and adjustments to compete with traditional methods.

III DEVELOPMENT OF FUZZY LIKERT SCALE

Chen & Hwang [5] proposed an approach to solve MADM problems in a fuzzy environment. The approach is of two steps. In the first step, fuzzy data is converted into crisp scores. Then, as the next step, this data in the form of decision matrix is used to rank the alternatives by using MADM methods. In the following section, this method has been explained. This method logically converts linguistic terms into their corresponding fuzzy numbers.

Fuzzy numbers 'W' are converted into the crisp score as:

$$\mu_{\max}(x) = \begin{Bmatrix} x, 0 \le x \le 1 \\ 0, otherwise \end{Bmatrix} \qquad \mu_{\min}(x) = \begin{Bmatrix} 1-x, 0 \le x \le 1 \\ 0, otherwise \end{Bmatrix}$$

The maximum and minimum fuzzy numbers should be selected in a manner that they can be fitted automatically into comparison scale. The left score of the fuzzy number is calculated as below:

$$\mu_L(W_i) = Sup_x \left[\mu_{\min}(x)^{W_i(x)} \right]$$

By using the above formula, the left score is converted into crisp number between 0 & 1. It is the maximum value of the intersection of fuzzy number *W*j and the minimum fuzzy number. Similarly, the right score can be converted into crisp number by using the following formula.

$$\mu_R(W_i) = Sup_x \left[\mu_{\max}(x)^{W_i(x)} \right]$$

The total score is calculates as:

$$\mu_T(W_i) = \frac{\left[\mu_R(W_i) + 1 - \mu_L(W_i) \right]}{2}$$

IV. APPLICATION OF FUZZY LIKERT SCALE IN ROBOT SELECTION

This example problem considers five robot selection attributes, and three alternative robots. The objective and subjective information of the attributes is given in Table 1. Man–machine interface (MI) and programming flexibility (PF) are expressed subjectively in linguistic terms, and these attributes are assigned objective values with the help of Table 1. LC, MI, and PF are beneficial attributes, and higher values are desirable. PC and R are non-beneficial attributes, and lower values are desirable.

Table 1. Robot selection attributes information [6]

Robot	PC ($1,000)	LC (kg)	R (mm)	MI	PF
Robot 1	73	48	0.15	A	H
Robot 2	71	46	0.18	AA	VH
Robot 3	75	51	0.14	BA	H

PC: Purchasing cost; LC: Load carrying capacity; R: Repeatability error; MI: Man-machine interface; PF: Programming flexibility; A: Average; AA: Above average; BA: Below average; H: High; VH: Very high.

The total score using the fuzzy rating scale are obtained by methodology discussed in section III and is presented in Table 2 for the data considered in this example. The objective data of the attributes are given in Table 2. The decision maker can appropriately make use of this 5 point scale The conversion of the linguistic terms into corresponding fuzzy crisp score is given in Table 2.

Table 3. Crisp value is calculated using right score(μ_R) and left score (μ_L)

Qualitative measure of intangible factors	Right (μ_R)	Score	Left Score (μ_L)	Crisp Score
Below Average	0.2		1	0.1
Average	0.385		0.762	0.311
Above average	0.596		0.537	0.529
High	0.752		0.409	0.671
Very High	1		0.2	0.9

Table 2. Robot Membership function to calculate Fuzzy numbers

Qualitative measure of intangible factors	Fuzzy numbers	Membership function (μ)
Below Average	(0,0,0.25 n)	$\mu_w(x) = \begin{Bmatrix} 1, x = 0 \\ (0.25 - x)/0.25, 0 \le x \le 0.3 \end{Bmatrix}$
Average	(0.1, 0.28, 0.45)	$\mu_w(x) = \begin{Bmatrix} (x - 0.1)/0.18, 0.1 \le x \le 0.28 \\ (0.45 - x)/0.17, 0 \le x \le 0.3 \end{Bmatrix}$
Above Average	(0.35, 0.56, 0.65)	$\mu_w(x) = \begin{Bmatrix} (x - 0.35)/0.21, 0.35 \le x \le 0.56 \\ (0.65 - x)/0.09, 0.56 \le x \le 0.65 \end{Bmatrix}$
High	(0.5, 0.72, 0.85)	$\mu_w(x) = \begin{Bmatrix} (x - 0.5)/0.22, 0.5 \le x \le 0.72 \\ (0.85 - x)/0.13, 0.72 \le x \le 0.85 \end{Bmatrix}$
Very High	(0.7, 1, 1)	$\mu_w(x) = \begin{Bmatrix} (x - 0.7)/0.3, 0.7 \le x \le 1 \\ 1, x = 1 \end{Bmatrix}$

The crisp scores assigned to various linguistic tems in Table 3 corresponds to the conversion scale developed for this given application as shown in Figure 1.

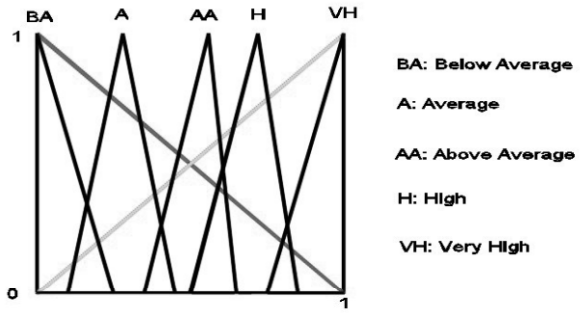


Fig. 1. Fuzzy conversion scale for given application.

Once the crisp score to the linguistic variables using conversion scale as shown in Fig. 1, the ranking can be done by any of the multiple attribute decision making method such as TOPSIS, Weighted Product method (WPM), Simple additive weights method (SAW), VIKOR, etc, For demonstration purpose simple additive method is used for the present example considering equal weights to all attributes. This is a two-step process. In the first step the attribute data is normalised as shown in Table 4. In second step, the attribute added for each robot is added to get the robot selection index (RSI) as shown in Table 5.

Table 4. Robot selection attributes with crisp value

Robot	PC ($1,000)	LC (kg)	R (mm)	MI	PF	
Robot 1	73		48	0.15	0.311	0.671
Robot 2	71		46	0.18	0.529	0.9
Robot 3	75		51	0.14	0.1	0.671

Table 5. Robot selection Index values

Robot	PC ($1,000)	LC (kg)	R (mm)	MI	PF	RSI
Robot 1	0.972	0.941	0.933	0.587	0.745	4.178
Robot 2	1	0.901	0.777	1	1	**4.678**
Robot 3	0.946	1	1	0.189	0.745	3.88

RSI: Selection Index;

It is observed from the Table 5, that the Robot 2 is best choice with respect to the given attributes as indicated by highest value of RSI.

Now, if the Likert scale is used, for which the conversion scale is BA=1, A=2, AA=3, H=4, VH=5 for this application, the robot selection attributes in quantitative form are as shown in Table 6.

Table 6. Robot selection attributes using Likert scale

Robot	PC ($1,000)	LC (kg)	R (mm)	MI	PF	
Robot 1	73		48	0.15	2	4
Robot 2	71		46	0.18	3	5
Robot 3	75		51	0.14	1	4

The normalized data and RSI then would be as shown in Table 7.

Table 7. Robot selection Index values of example 1

Robot	PC ($1,000)	LC (kg)	R (mm)	MI	PF	RSI
Robot 1	0.972	0.941	0.933	0.666	0.8	4.312
Robot 2	1	0.901	0.777	1	1	**4.678**
Robot 3	0.946	1	1	0.333	0.8	4.079

RSI: Selection Index;

It is observed from the Table 7, that the Robot 2 is best choice with respect to the given attributes as indicated by highest value of RSI.

CONCLUSION :

This paper provides the methodology for application of fuzzy rating to robot selection. For demonstration purpose three robots are considered from which the best one is to be selected with respect to five criteria namely cost, load carrying capacity, repeatability error , man machine interface and programming flexibility. Out of which man machine interface and programming flexibility are to be expressed in qualitative manner only. For quantification of this linguistic data two scales (a) Fuzzy rating scale and (b) Likert scale are employed, the result of comparison of these two methods reveals that there is no significant effect on ranking of the robot and the ranking remains same with both the methods and will usually remain same for any other application also. However, it can be seen from Table 5 and Table 7 that the RSI for Robot 1 and Robot 3 are different with these two methods. Although it does not have much impact on ranking but these methods offers different percent effect (or contribution) due to different RSI, of the attributes for various alternatives if needs further analysis. When such an analysis is essential, fuzzy rating scale will be a much better option as it is based on general perception which is converted into crisp sores as against specific perception in case of Likert scale. This is due to the fact that fuzzy rating method able to capture near exact information without loss of much data.

REFERENCES :

[1] R. V. Rao. Decision making in the manufacturing environment. Springer-Verlag, London. 2007.

[2] J. C. Chan. Response-order Effects in Likert-type Scales. Educational Psychology Measurements. 1991. Vol. 51, p. 531-540.

[3] Likert Rensis. A Technique for the Measurement of Attitudes. Archives of Psychology. 1932. Vol. 140, p. 1–55.

[4]F. Eshragh, E.H.Mamdani,: A general approach to linguistic approximation. International Journal of Man-Machine Studies. 1979. Vol. 11, p. 501–519.

[5]S.J.Chen, C.L.Hwang. Fuzzy multiple attribute decision making methods and applications. Lecture notes in Economics and Mathematical Systems. 1992. Springer, New York.

[6]P.P.Bhangale, V.P.Agrawal, S.K.Saha. Attribute based specification, comparison and selection of a robot. Mechanism and Machine Theory. 2004. Vol. 39, p.1345–1366.

AN ANALYSIS OF WEAK BASE CLASS PROBLEM IN C++ AND WAYS TO AVOID WEAK BASE PROBLEM

PROF. VISHAL VERMA
Indira College of Commerce & Science
vishal.verma@iccs.ac.in
PROF. GANESH BHONDVE
Indira college of commerce & science
ganesh.bhondve@iccs.ac.in

Abstract : *Abstract: C++ strongly supports the concept of Reusability. The C++ classes can be reused in several ways. In this context virtual functions are used. Once a class has been written and tested, it can be adapted by another programmer to suit their requirements. This is basically done by creating new classes, reusing the properties of the existing ones. When the function in base class is virtual and when it is overridden at that time the function call depends upon the objects to which base class pointer is pointing. Another situation is when we are calling a base class virtual function from base class non-virtual function at that time we except the base class function should get called but here instead of base class the derived class function gets called which is overridden. Here the base class is called as Weak base class. In this paper we have situation where we face weak base class problem, Internal implementation of weak base class and how to avoid weak base class problem.*

Keywords: *Inheritance, Virtual Function, Infinite Recursion, Virtual Table, Abstract Class, Pure Virtual Function, Virtual Pointer, Runtime Polymorphism, Stack Overflow.*

INTRODUCTION:

There are many scenarios where we use Runtime Polymorphism. When we want to use runtime polymorphism automatically comes concept of inheritance for reusability. While using inheritance Base class and Derived class comes in picture. For achieving runtime polymorphism we use virtual function concept. Generally derived class can access base class functions. Weak base class is the class which access derived class function which is against the concept of inheritance. In this paper we have represented the concept of weak base class when it comes while programming in C++, and different ways to avoid weak base class problem in C++ programming language.

A Program of function overriding to demonstrate weak base class program :

```
#include<iostream.h>
#include<conio.h>
class Base
{
        public :

        virtual void Addition()

        {
                int a, b;
                cout<<"\nEnter values for 2 numbers  a &
b :";
                cin>>a>>b;
                cout<<"Addition in base class : a + b =
"<<a+b<<endl;

        }

        void Subtraction()

        {
                int a, b;
                cout<<"\n\nEnter value of a & b :";
                cin>>a>>b;
                cout<<"Subtraction in base class a - b =:
"<<a-b<<endl;

        }


        virtual void Multiplication()
        {

        int a, b;
                cout<<"\nEnter values for 2 numbers  a &
b :";
                cin>>a>>b;
                cout<<" Multiplication in base class : a *
b =      "<<a*b<<endl;

        }
};

        class Derived : public Base
        {

                public:
```

```
                        void Addition()
                        {
                        int a, b, c;
                        cout<<"\nEnter values for 3
number of a,b & c :";
                        cin>>a>>b>>c;
                        cout<<"Addition in Derived class
: a + b + c = "<<a+b+c<<endl;
                        }
};

int main()
{
        clrscr();

        Base *bPtr;
        Derived dObj;

        bPtr = &dObj;

        bPtr->Addition();
        bPtr->Subtraction();
        bPtr->Multiplication();
        getch();
        return 0;
}
```

Expected output:

Enter values for 3 number of a,b & c: 10 20 30
Addition in Derived class: a + b + c = 60

Enter value of a & b : 50 20
Subtraction in base class a - b = 30

Enter value of a & b : 5 2
Multiplication in base class a * b = 10

Explanation of above output:

The reason for this output is here in base class Addition() function is virtual and it has been overridden in derived class. In the base class Subtraction () function is normal function.

Let us see the code in main() function. In the main() function we have code in which first we have created the base pointer bPtr and we have created object of derived class dObj. By using Base class pointer we have given call to bPtr->Addition () and bPtr->Subtraction() respectively. Here Addition () method is virtual so first derived class method is called. Then bPtr->Subtraction () due to this call base class function will call as it's a normal function defined only in base class. Now at last due to bPtr->Multiplication () this call a function from base class will call as Multiplication() has defined virtual in base class but it has not been again redefined in derived class.
We will see above concept by using virtual table and virtual pointer

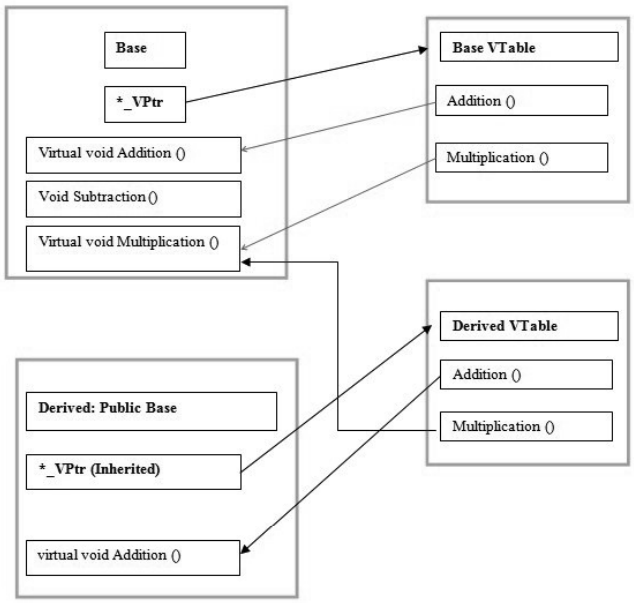


Fig 1: Virtual table and Virtual pointer

To implement virtual functions, C++ uses concept of virtual table. The **virtual table** is a lookup table of functions used to resolve function calls in a dynamic binding manner. The virtual table also called as “VTable” , “virtual function table”, “virtual method table”, or “dispatch table”.There is separate VTable for every single class used in C++ program. The compiler also adds separate hidden pointer to a base class *_Vptr. This “*_Vptr” is a real pointer . Because of “*_Vptr” pointer size of class gets incremented by one pointer size. The “*_Vptr” is inherited in derived class as shown in diagram.
Now if we slightly modify the code in base class Subtraction function as follow

```
void Subtraction()
        {
           Addition();
                int a, b;
                cout<<"\n\nEnter value of a & b :";
                cin>>a>>b;
                cout<<"Subtraction in base class a - b =:
"<<a-b<<endl;

        }
```

From the above base class function code we want to call base class Addition() function.When we execute the code again the Addition() function of derived class gets called instead of base class and we get the following output.

Output:

Enter values for 3 number of a,b & c: 10 20 30
Addition in Derived class : a + b + c = 60

Enter value of a & b : 50 20
Subtraction in base class a - b = 30

```
Enter value of a & b : 5 2
Multiplication in base class a * b = 10
```

The base class which gives call to its derived class function called as **Weak base class.** In the above the base class becomes weak base class and hence the problem is called weak base class problem. Actually the base class should not class any derived class function but here base class is calling derived class function this is problem in C++ programming language.

Infinite Recursion Problem (Stack Overflow) causes due to weak Base class

In the following example we have two classes one is Base class and another is Derived class. We have written two virtual functions increment1() and increment2() in base class. In derived class we have redefined increment2() function. When we execute the following code the code will be in infinite loop and hence there will be stack overflow. This **infinite recursion** (stack overflow) problem causes due to weak base class.

```
#include<iostream.h>
#include<conio.h>
class Base
{
            int counter;
public:
           Base()
           {
           counter = 0;
          }
 virtual  void increment1()
{
   increment2();
  }
 virtual void increment2()
{
   counter++;
}
};
class Derived : public Base
{
 void increment2()
{
   increment1();
  }
};

int main()
{
           clrscr();
           Base *bPtr;
           Derived dObj;
           bPtr = &dObj;
           bPtr->increment2();
           getch();
           return 0;
}
```

Avoiding Weak base class Problem

- We can avoid weak base class problem by declaring pure virtual functions in C++ and mandatory redefining that pure virtual function in derived class.
- Weak base class can also be avoided using the access specifiers which are used in C++ like public, private and protected when we are dealing with runtime polymorphism using virtual function. These changes prevent derived class from relying on implementation details of Base class and allow derived class to expose only those Base class functions that are applicable to themselves.
- There is another solution to avoid weak base class problem in C++ we can use Abstract class concept to avoid such problem.
- To avoid weak base class problem we can also use the concept of singleton class used in C++ programming language

CONCLUSION :

While development in C++ programming language there are many cases where we use function overriding. When we want to use virtual function for function overriding automatically comes concept of inheritance for reusability. While using inheritance Base class and Derived class comes in picture where exactly the virtual function is used. While redefining virtual functions we may come different problems like weak base class, derive class malfunctioning due to weak base class and infinite recursion(stack overflow) due to which harm to our development possible. so while dealing with runtime polymorphism in C++ use of avoiding weak base should be considered as mentioned above.

REFERENCES :

a. Programming Language Pragmatics(Michael L. Scott)
b. C++ Primer by Stanley B. Lippman, Jossee Lajoie, and Barbara E. Moo
c. Jumping into C++ Kindle Edition By Alex Allain
d. Accelerated C++: Practical Programming(By **Andrew Koenig, Barbara E. Moo)**
e. E Balagurusamy object oriented programming with C++
f. Effective Modern C++ : 42 Specific ways to Improve Your Use of C++11 and C++14(1st Edition) by Scott Meyers

STUDY OF TCP INCAST PROBLEM IN DATA CENTER NETWORKS

MAHENDRA N. SURYAVANSHI[1], DR. AJAY KUMAR[2]
1 Asst. Professor, Indira College of Commerce and Science, Pune-33;
Research Student, Computer Science Department, Savitribai Phule Pune University, Pune-07, INDIA
2 Director, JSPM's JTC Jayawant School of Computer Applications, Pune-33, INDIA

Abstract : *In data center networks the overall throughput collapse when multiple storage servers involved in synchronized request workload, simultaneously send data to a single client. This throughput collapse is called as TCP Incast problem. This paper analyses the Incast problem and discusses application level, transport level and data-link level solutions for TCP Incast problem.*

Keywords: TCP, data center network, synchronous read, goodput, Incast.

INTRODUCTION:

Data center is a pool of computing and storage resources clustered together using communication networks to host Internet-based applications (e.g. search engines, video data hosting, social networking, large-scale computing) and data storage[9][6]. Applications hosted by data center are either data intensive or communication intensive. The thousands of servers may be harnessed to fulfill a simple web search request or database query [6][5]. Building data center networks using commodity TCP/IP and Ethernet networks is attractive because of the low cost, ease-of-use and desire to share the bandwidth over multiple compute resources [2]. In TCP/IP protocol suite TCP is the most popular transport protocol and considered as the backbone of the internet. It provides reliable, byte-stream, connection-oriented services and operates over heterogeneous network topologies. TCP offers flow control and congestion control. Modern implementation of TCP uses slow start, fast-retransmit, fast-recovery and congestion avoidance algorithms. TCP used in data center networks to provide reliable communication between various clients and servers located at data center network on time [3].

Synchronized read/write operations are commonly performed in data center networks for multiple servers to one client. HDFS, Lustre, Panasas, pNFS, Cassandra, MapReduce are the network file systems used in data center networks to facilitate synchronized read/write operations [8]. Synchronized request workload requires that many-to-one communication pattern to take place between multiple servers and single client in reliable and on time manner.

There are applications need many-to-one communication patterns in data center networks some of the examples are:

1)Social Networking sites: User logs in to the social networking site. If user's complete profile is stripped across multiple storage servers, request for fetching complete profile can be sent to number of storage servers within the data center. These servers send their part of requested data to client simultaneously.

2)Web search applications(search engine): Client submits search query to web search application. There could be hundreds of thousands of storage servers that contain requested data for search query. All storage servers respond with their part of result to the client simultaneously.

3)Data warehousing applications and applications used to maintain big organizations data, banks data, government data, hospitals data etc. are all follows many-to-one communication pattern where multiple servers send data to the single client.

Many-to-one communication pattern in data center network environment is shown figure-2 where multiple servers simultaneously sending data to a single client.

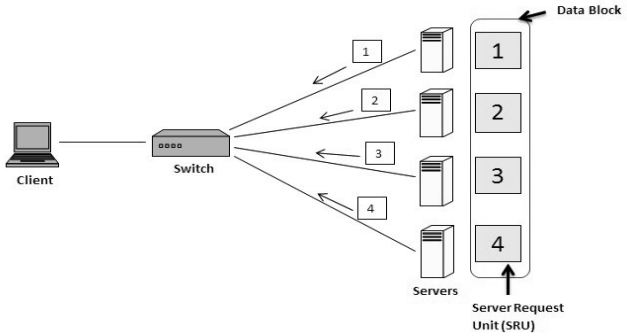


Figure-1: one client requesting data from multiple servers through Synchronized reads in data center network

We refer to Amar Phanishayee et al. [7] for the TCP Incast problem. Using TCP for synchronized request workloads in data center networks, When multiple storage servers simultaneously sends data to a single client beyond the storing capacity of the client switch buffer, the throughput collapse which exhibits TCP Incast problem.

There is one or more packet loss at client switch buffer when it overflows. Server involved in a synchronized request experiences a timeout, other servers can finish sending their responses, but the client must wait minimum of 200ms before receiving the remaining parts of the response. During this time period the client's link may be completely idle. Furthermore client issue the next data block request only when all servers have responded with their portion of data. Hence the overall throughput degrades significantly [4].

The performance of data center network in terms of goodput vs number of servers involved in synchronized workload in figure-2 [7] shows that as number of servers increases goodput decreases. The goodput of network is 700 Mbps when 2 servers involved in synchronized workload. Goodput further drops to less than 50% for 5 servers. Eventually for 22 servers and onwards, goodput drops to less than 100 Mbps.

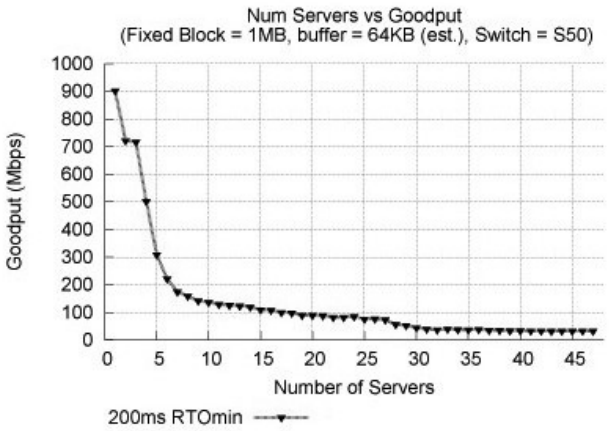


Figure-2: Goodput decreases as the number of servers involved in synchronized Workload increases in data center network [7]

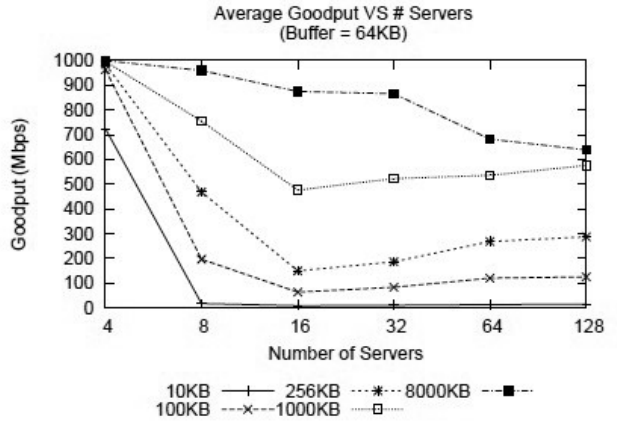


Figure-3: Effect of varying SRU size [1]

II - Solutions for Incast problem

Solutions for Incast problem in data center network can be proposed at various layers i.e. application layer, transport layer and data-link layer.

1) Application Layer Solution

i) Limiting number of storage servers: According to figure-2, as the number of servers increases the goodput decreases. To reduce the impact of Incast problem one of the solution at application layer is to limit the number of servers participating in synchronized data transfer. To achieve this, one need to identify at what range, servers are experiencing degraded throughput, and then limit the acceptable number of servers that can participate in synchronized data transfer. According to [2] Elie Krevat et al. described that Panasas network file system uses the idea of limiting the number of servers participating in any one data transfer. It divides large pool of servers into small RAID group of a limited range. Furthermore, limit any communication to one RAID group at a time [2].

ii) Increasing SRU (server request unit) size: The amount of idle link time can be reduced by increasing the SRU size at each server. One implicit benefit of increasing SRU size is data can be stored at small number of servers. Figure-3 Panishayee et al. [1] is the simulation result at different SRU size. The large SRU size improves goodput. Average goodput for SRU size 8MB is 10 times more than average goodput for SRU size 10KB. Generally SRU size of 8MB is quite impractical and applications process the data in the range of 1-256 KB SRU size [1].

iii)Throttling Data Transfers: A client can advertise a small TCP receive buffer and can throttle data transfer rates of servers which are participated in synchronized data transfer. This technique helps to limit TCP window size and to improve the scalability. This ensures that more requests can be made to large number of servers. Throttling data transfer can be used to mitigate the impact of Incast, but it has side effect of underutilization of client's link capacity since TCP windows are not allowed to increase to the proper size. To address this, client can adjust the throttle rate whenever necessary. Due to complexity, throttling technique is in rare use [2].

iv) Staggering Data Transfers: A client can request only a subset of the total data block at a time to stagger the data transfer. This limits the amount of data being transferred synchronously as only a subset of servers are sending data at any one time. This staggering can either be made by the client or at the servers. Client can make staggering by requesting less data to limited number of servers at once. To affect staggering, servers can also deterministically or randomly delay their response to a request. This technique limits the number of servers that are involved in synchronized data transfer. This releases the pressure on client switch buffer. Switch will not drop packets and no timeout will occur. Eventually Incast problem will be avoided [2].

v) Global Scheduling of Data Transfers: In multiple workload situations where client makes multiple requests to different subset of servers, global scheduling of data transfers is required. A server can only transfer data to a client if it has clients SRU token. When server receives client request, server may fetch the data in advance and waits for the appropriate SRU token. A client can send request packets to all servers containing data for its multiple requests, but only the K servers that have been allocated that clients SRU tokens can transfer the data. This will restrict the total number of synchronized servers sending data to any given client. Limited number of servers are participating in synchronized data transfer hence it may not overflow the clients switch buffer. There may not be any packet loss and hence no throughput collapses [2].

1) Transport Layer Solutions

Reducing TCP minimum RTO value: TCP timeouts are unavoidable, but time spent on waiting for a timeout can be reduced. TCP implementations use an RTO_{min} value of 200ms. This value is generally greater than round-trip times. By reducing the TCP's RTO value, TCP Incast problem can be avoided. Figure-4 [8] shows that reducing RTO_{min} from 200ms to 200µs improves the goodput for the network up to 47 servers.

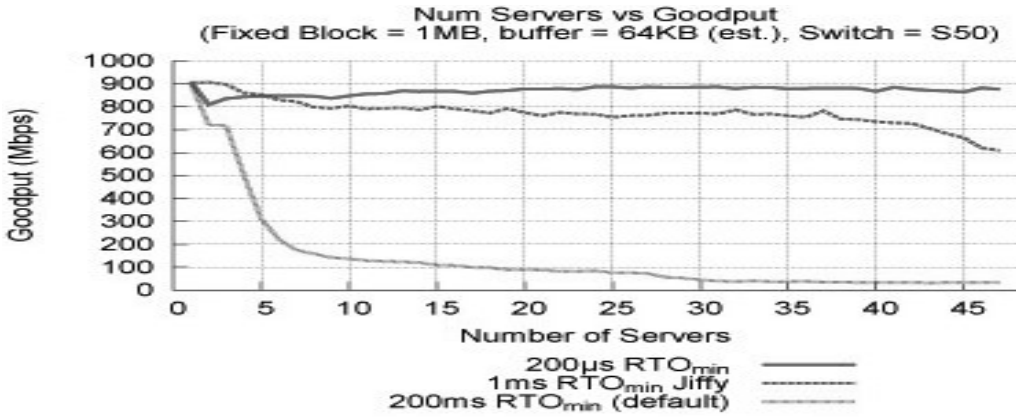


Figure-4: Goodput of network for different RTO_{min} values [8]

To achieve timely response to packet losses and along with avoiding premature timeouts, there is need to determine right *RTO* value. A premature timeout can lead to spurious retransmission.

1) Data Link Layer Solutions

Provisioning Larger Switch Buffers: Timeouts are the primary cause of Incast, and the root cause of timeouts is packet losses. Use of Larger switch buffers can reduce the significant packet losses at client side. Increasing the switch buffer size at client side doubles the number of servers that can transmit before the system experiences Incast. The effect of increasing switch buffer size can be seen on goodput as shown in figure-5 [1].

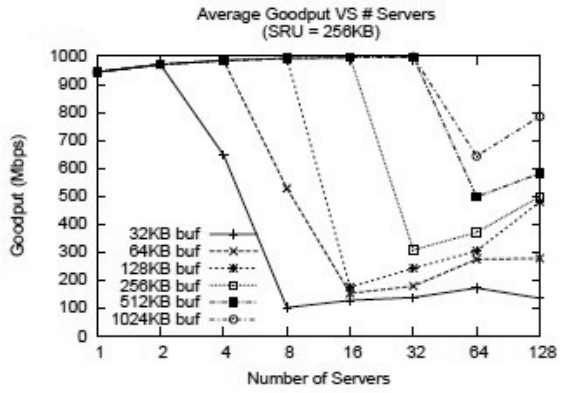


Figure-5: Effect of increasing switch buffer on goodput of network [1]

For 128 number of synchronously communicating servers, there is continuous increase in network goodput from 15%, 29%, 47%, 50%, 60% to 80% of network link capacity, as switch buffer size is increasing in order 32KB, 64KB, 128KB, 256KB, 512KB to 1024KB. Unfortunately switches with larger buffers costs more. Hence this cannot be the optimal solution for TCP Incast problem.

CONCLUSION :

Many-to-one traffic pattern is common in data center networks, where data is stripped across multiple storage servers. TCP Incast occurs when client request a data block stripped across multiple synchronously communicating servers. The simultaneous data transfer from multiple servers to a single client overloads the clients switch buffer. This results into one or more timeouts and retransmissions. This paper contains application level, transport level and data-link level techniques to avoid Incast problem, but still research is needed to solve the TCP Incast problem completely for better throughput, decreasing latency and to overcome the congestion in data center networks.

REFERENCES :

1) Amar Phanishayee, Elie Krevat, Vijay Vasudevan, David G. Andersen, Gregory R. Ganger, Garth A. Gibson, Srinivasan Seshan, "Measurement and Analysis of TCP Throughput Collapse in Cluster-based Storage Systems", 6th USENIX Conference on File and Storage Technologies (FAST '08) Feb. 26-29, 2008.

2) Elie Krevat, Vijay Vasudevan, Amar Phanishayee, David G. Anderson, Gregory R. Ganger, Ganger A. Gibso, Srinivasan Seshan, "On Application-level Approaches to Avoiding TCP Throughput Collapse in Cluster-based Storage Systems", proceedings of the 2nd international Petascale Data Storage Workshop (PDSW '17) November 11, 2007.

3) Irfan Riaz Shohab, Muhammad Younas, Ramzan Talib, Umer Sarwar, "Application Layer's Approaches for TCP Incast problem in data center Networks", International Journal of Computer Science and Mobile

Computing, Vol.3 Issue 4, April 2014, pg 459-474.
4) Jiao Zhang, Fengyuan Ren, Li Tang, Chuang Lin, "Modeling and Solving TCP Incast Problem in Data Center Networks", IEEE Transactions on Parallel and Distributed Systems, VOL. 26, No. 2, FEB 2015.
5) Juha Salo,"Data Center Network Architectures", http://www.cse.tkk.fi/en/publications/B/10/papers/Salo_final.pdf
6) Kashif Bilal, Samee U. Khan, Joanna Kolodziej, Limin Zhang, Khizar Hayat, Sajjad A Madani, Nasro Min- Allah, Lizhe Wang, Dan Chen, "A Comparative Study of Data Center Network Architectures", Proceedings 26th European Conference on Modelling And Simulation.
7) Vijay Vasudevan, Amar Phanishayee, Hiral Shah, Elie Krevat, David G. Andersen, Gregory R. Ganger, Garth A. Gibson, Brian Mueller, "Safe and Effective Fine-grained TCP Retransmissions for Datacenter Communication", SIGCOMM '09, August 17-21, 2009.
8) Vijay Vasudevan, Hiral Shah, Amar Phanishayee, Elie Krevat, David Anderson, Greg Ganger, Garth Gibson, "Solving TCP Incast in Cluster Storage Systems", http://www.pdl.cmu.edu/PDL-FTP/Storage/vasudevan_fast09wip.pdf
9) Yang Liu, Jogesh K. Muppala, MalathiVeeraraghavan, "A Survey of Data Center Network Architectures",http://www.ece.virginia.edu/mv/pubs/recent-samples/Data-Center-Survey.pdf.

TUMOR DETECTION AND ANALYSIS USING CT-SCAN IMAGES

DR. SUMEDH GAIKWAD*, JYOTI SHROTE AND RUPALI SURASE
Indira College of Commerce & Science,Tathwade, Pune (M.S.) India
Dept. of CSIT, Dr. B.A.M. University, Aurangabad, (M.S.)
*Corresponding author: sumedh25@yahoo.com

Abstract : *In the field of computer science, medical image study is used as a system of detection and access control. Tumor detection is a type of biomedical examination. We have chosen the Agile Software Model for our system modeling. In this study, we have gone through all four phases of this model Agile methodologies and its proposed model. The analysis comprises DFD (Data Flow Diagram), Flowchart, & E-R Diagram. Pre-processing is done using the binarization, image contrast, histogram equalization, Feature detection (SURF Method), Plot ROC Curve on the database.*

According to the consequence we have proved that the tumor is detected precisely with various types into brain. By using the result of the ROC, curve is plotted by calculating positive and negative result and studied out the specificity and sensitivity onto X-axis and Y-axis. The overall detection gain is 99%, depending on various tumor images.

Keywords : Tumor, Brain, CT-Scan

INTRODUCTION:

Basic approaches to display one-, two-, and three-dimensional (3D) biomedical data are introduced. As a determination, segmentation, image enhancement systems, texture study and their application in diagnostic imaging will be discussed. To widespread this outline, storage, communication of medical images and retrieval, are also introduced. In addition to this theoretical background, an overview of useful software tools is given. In particular, MATLAB (2012) for medical image enhancement and visualization (3D rendering) will be prudently demonstrated. It is developed by the National Institutes of Health, MATLAB (2012) is open source and freely available in the public domain [1].

Medical imaging is the method and process used to produce images of the human body or parts and function for clinical purposes (medical procedures seeking to examine, diagnose, or disease) or medical science (including the study of usual anatomy and physiology). Even though imaging of detached organs and tissues can be attained for medical reasons. Such procedures are not usually referred to as medical imaging, but to a certain extent are a part of pathology.

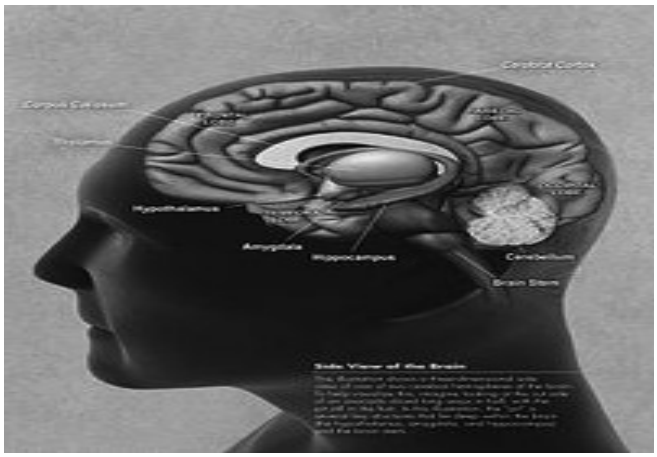
As a discipline and in its broadest sense, it is part of biological imaging and integrates nuclear medicine, radiology, endoscopy, analytical radiological sciences, medical photography, (medical) thermographs, and microscopy.

As a field of scientific investigation, medical imaging institutes a sub-discipline of medicine, medical physics or biomedical engineering, depending on the context: Research and expansion in the area of instrumentation, modelling, image acquisition and quantification are usually the purview of medical physics, computer science, and biomedical engineering. Research into the application and analysis of medical images is usually the domain of radiology and the medical sub-discipline significant to medical ailment or region of medical science (cardiology, neuroscience, psychology, & psychiatry etc.) under investigation. Many of the techniques established for medical imaging also have scientific and industrial applications [2-3].

Databases

Various types of medical image analysis are as follows,

Cancer Detection : Various types of tumors are detected like brain tumor, breast tumor, chest tumor, bone tumor. All of these detected using various techniques such as CT-Scan, MRI, Ultrasound, Biopsy, etc.



TB Diagnosis:

Tuberculosis is detected by finding Mycobacterium tuberculosis bacteria in a clinical sample taken from the patient. While other examinations may strongly suggest tuberculosis as the diagnosis, they cannot confirm it.

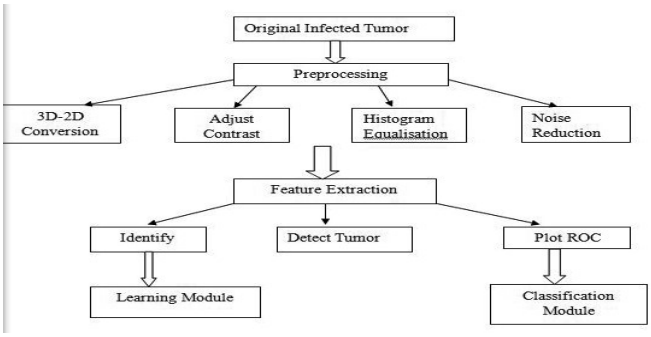
A comprehensive medical evaluation for tuberculosis (TB) must comprise a physical examination, medical history, and chest X-ray and microbiological analysis (of sputum or some other suitable sample). It may also consist of a tuberculin skin test, other scans and X-rays, surgical biopsy etc.

Diabetic Retinography

Diabetic eye disease refers to a group of eye problems that people with diabetes may face as a complication of diabetes. All can cause severe vision loss or even blindness.

Diabetic retinopathy is harmful to the blood vessels in the retina. Diabetic retinopathy is the utmost common diabetic eye disease and a prominent cause of blindness in adults. It is initiated by changes in the blood vessels of the retina. In some individuals with diabetic retinopathy, blood vessels may swell and leak fluid. In other persons, abnormal new blood vessels grow on the surface of the retina. The retina is the light-sensitive tissue at the rear of the eye. A healthy retina is obligatory for good vision. If you have diabetic retinopathy, at first you may not notice alterations to your vision. But over time, diabetic retinopathy can get of poorer quality and cause vision loss. Diabetic retinopathy generally affects both eyes [4].

Proposed Methodology



The architecture of system consists of various phases which are required for the maximum accuracy of the recognition and verification of the individual.

It consists of phases such as noise removal, background elimination, and detection.

Original Infected Tumor

Phases of the System

There are following phases of the system as follows:

1) Acquisition: The aim of CT scanner hardware is to obtain a large number of transmission measurements through the patient at altered positions. Single CT image may comprise approximately 800 rays taken at 1,000 different projection angles. Before the acquisition of the next slide, the table that the patient lies on, is moved slightly in the cranial-caudal direction (the "z-axis" of the scanner)

2) Pre-processing: This is second phase of the system. In this phase various operations like 3D to 2D conversion of image, contrast image, histogram equalization. pre-processing will be helpful for image enhancement. Preprocessing plays an important role to improve quality of an image.

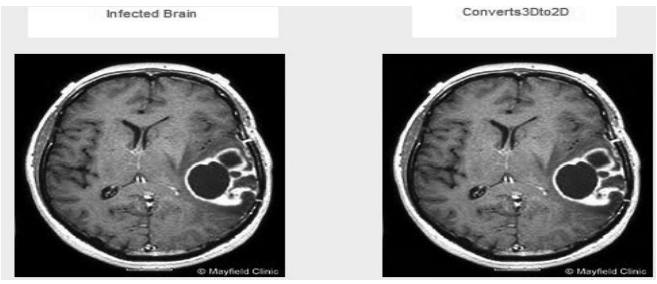


Figure VII.3D to 2D Conservation

3) Noise Reduction: It is one of the most important processes in image processing. It is also called smoothing or noise filtering. Images are often corrupted due to noisy channels. Median filter is widely used for smoothing and restoring images corrupted by noise. Median filter has attractive properties for suppressing impulse noise while preserving edges.

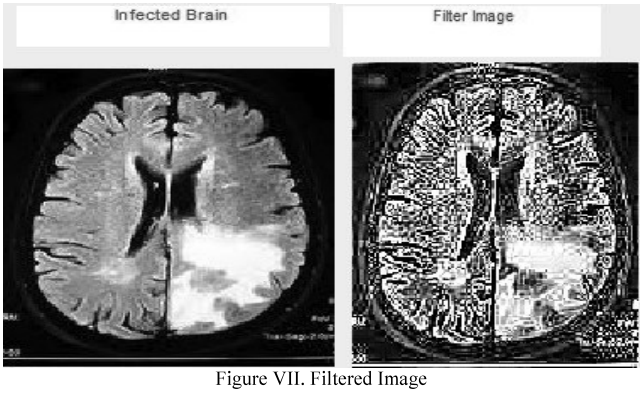
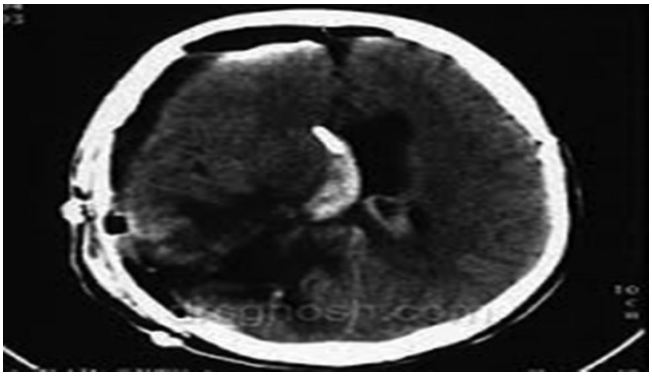


Figure VII. Filtered Image

4) Tumor Normalization: Tumor normalization is one of the important part of getting correct result. Tumor sizes may vary due to conformations of various tumors. Height and width of tumor varies & depends onto stages of tumor. Even the same person has various tumors with different size. We need to detect these all tumors and try to destroy it.

5) Feature Detection: It is the number of tumors which belongs to the same image. This feature provides information about exact location of the tumors.



There are two types of tumor detection

a) Manual Detection:

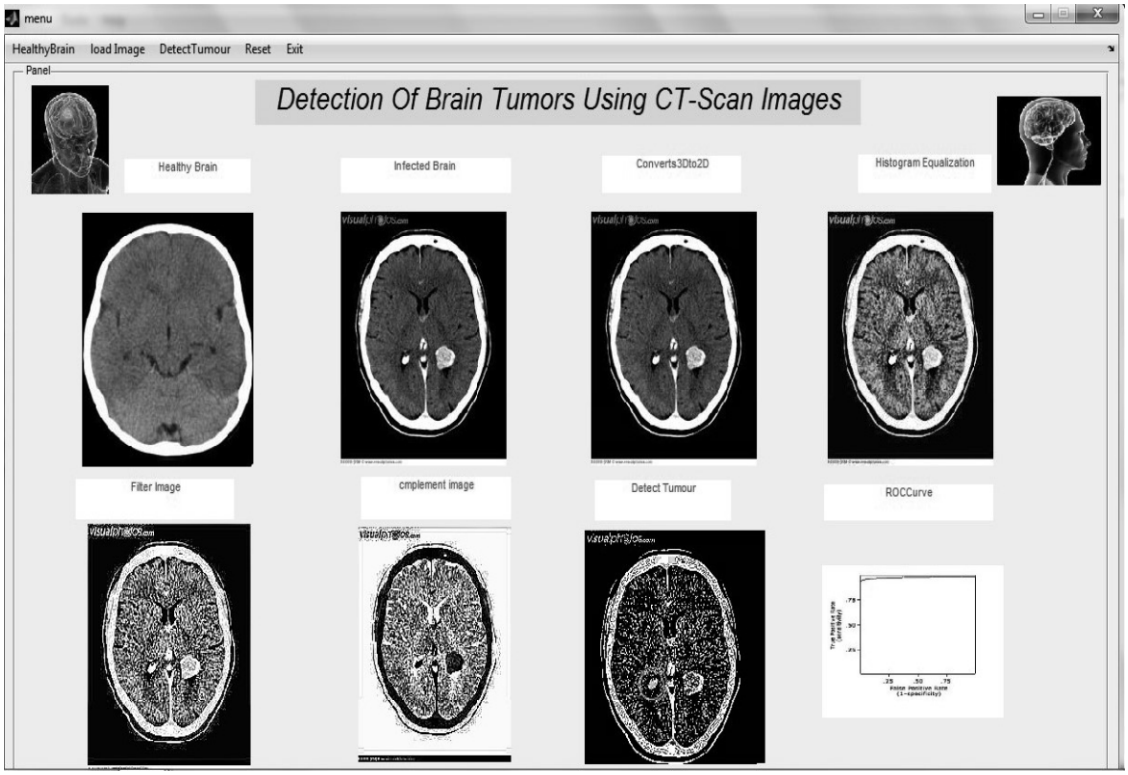
This type of detection includes detection which is identified by human eyes for checking the result of system which recognise it correct or not.

b) System Detection:

This type of detection includes detection of brain tumors with varying sizes. These include true positive and true negative type of result. Based on that ROC (Resistance Optimance Classification) is plotted.

6) Decision Making:

Once a similarity is obtained, the decision implies that the computation of decision threshold. If tumor is identified with correct location then it comes under TRUE POSITIVE type and if not detected correctly then it comes under TRUE NEGATIVE type of classification.



Sr. No.	Image Name	Manual Detection	System Detection			
			TP	TN	FP	FN
1	download(1)	One Tumor in Left Parital Lobe	✓			
2	download(2)	One Tumor in Right Parital Lobe	✓			
3	images(2)	One Tumor at Bottom of Left Parital Lobe	✓			
4	images(3)	One Tumor at Top of Left Parital Lobe	✓			
5	images (12)	One Tumor at Bottom of Right Parital Lobe	✓			
6	images(14)	One Tumor in Left Parital Lobe	✓			
7	images(5)	One Tumor At Center Of Brain	✓			
8	images(19)	One Tumor At Middle of Center	✓			
9	images(10)	One Tumor At Bottom Of Left Parital Lobe	✓			
10	images(18)	Two Tumor At Top Of Left Parital Lobe	✓			

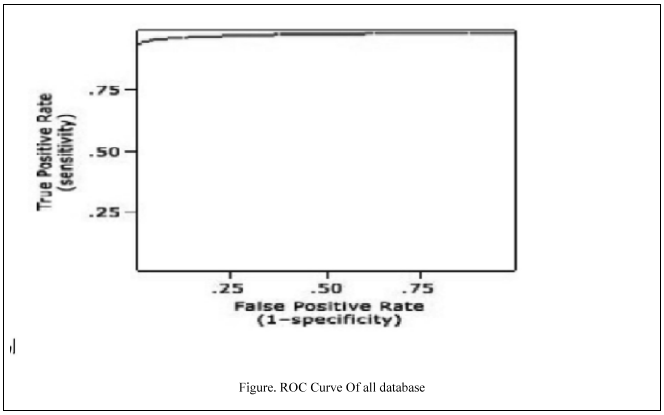


Figure. ROC Curve Of all database

CONCLUSION :

In computer science, medical image analysis is used as a form of detection and access control. Tumor detection is a type of biomedical analysis.

We have chosen the Agile Software Model for our system modeling. In this study we have gone through all four phases of this model, Agile methodologies and its proposed model. We have seen the analysis which includes DFD (Data Flow Diagram), Flowchart, E-R Diagram.

We have applied the binarization, image contrast, histogram equalization, Feature detection (SURF Method), Plot ROC Curve on the database. We have used the varieties into database.

According to the result we have proved that the tumor is detected accurately with various types into brain. By using the result of the ROC curve is plotted by calculating positive and negative result finding out the specificity and sensitivity onto X-axis and Y-axis.

The overall detection gain is 99%, depending on various tumor images. In future we will try to implement the structured approach and fuzzy logic will be applied.

REFERENCES :

[1] S. U. R. F. (SURF), "Herbert bay, and Reas Ess, Tinne Tuytelaars, Luc Van Gool," Elsevier preprint, 2008.
[2]. "Blob detection – the free encyclopedia, "2011, [Online ; accessed 14 July2010].
3]. https://secure.wikimedia.org/wikipedia/ en/wiki/Blob detection
[4]. M. B. David Lowe, "Invariant features from interest point groups," BMVC, 2002.
[5]. www.ncbi.nlm.nih.gov/pubmed/22755708
[6]. www.bic.mni.mcgill.co
[7].Rad.usuhs.mil/medvix/parent.php3? mode=homepage
[8]. www.imageprocessingplace.com

COMPARITIVE ANALYSIS OF SCTP BANDWIDTH FOR PERFORMANCE MEASUREMENT OF WI-FI NETWORK IN ANDROID SMARTPHONES

Jyoti Shrote
Asstt.Prof., Indira College of Commerce and Science, Pune.
Jyoti.shrote@iccs.ac.in
Vijaya Kumbhar
Asstt.Prof., Indira College of Commerce and Science, Pune.
contactvijayak@iccs.ac.in

Abstract - Now days the wireless technology is playing very important role . This technology is used to transfer data efficiently over long distances. It is ever developing field, the devices can be developed to support communication with higher data rate and security. The increasing number of wireless devices and slow Internet connectivity which causes the poor Wi-Fi performance. The main reason of such performance degradation includes the channel allocation. In this paper the performance analysis of channel allocation of SCTP protocol are taken into account .The paper tried to focus on SCTP bandwidth as it one of the important attribute in order to analyze the performance.

Keywords*: Wi-Fi, Android Smartphone's, SCTP Protocol*

INTRODUCTION:

Computer networks transfer's data from one machine to another on any network using ISO/OSI layers or TCP/IP layers. This Layer has important layer transport layer. The primary role of transport layers is to provide end to end communication service between two or more applications running on different hosts. It also provides functions such as flow control, error recovery and reliable delivery. The transport layer employed one of two protocols, transmission control and user datagram protocol. The choice of transport protocol depends on the requirement of the application in the terms of quality of service. Application that requires reliability in order to delivery of the data, it uses the TCP whereas once if it can tolerate a certain degree of loss it prefers UDP because it provides faster degree of packets. To extend transport layer functionality the new protocol the stream control transmission protocol(SCTP).like TCP ,SCTP, offer a point to point connection oriented reliable delivery transport service for application communicating over an IP network. SCTP provides a number of functions that are critical for telephony signal ling transport and at the same for communication SCTP supports for multi-homing and partial ordering. it establish a session with another SCTP host over multiple interfaces identified by separate IP address. Thus SCTP can benefit applications that require reliable delivery and fast processing of multiple unrelated data stream protocol.[1,2,4,7,9,10].
Research on extending SCTP to support concurrent multipath transfer, consisting in simultaneously sending data over different size and number of request Serially and Bidirectional.[3,5,9,10]

II Wi Fi Network

WI FI is an international standard describing the characteristics of a wireless LAN It connects laptop computers, office equipment, personal assistants etc. and creates wireless local area networks IEEE 802.11 defines two topologies: the infrastructure mode and the Ad hoc mode. In infrastructure mode, each station connects to an access point (AP) via a wireless link. The set-up formed by the access point and the stations located within its coverage area are called the basic service set (BSS). It is possible to connect several access points (BSS), by a distribution system (DS), to form a large network covered by several cells. The set-up is called the Extended Service Set (ESS).When using ad hoc mode, stations are able to communicate directly, in peer to peer communication mode, an access point.[15,16]

2. Literature Review:

In communication Wired technology playing the very important role from several years but these technology have been drawbacks of using cable, it is very difficult to use for long distance communication. Even also reliability does not occur. Therefore these drawbacks can be overcome by using wireless communication. The wireless communication has been allowed for transferring data over long distance The advantage of using wireless communication is Reliability of data, greater mobility and possibilities to move devices and connect it freely without utilization of cables. Wireless communication communicates via satellite. The data can transfer with the help of wireless network such Wi-Fi.[1,4,7,18]
Wi-Fi is the wireless fidelity which allows an electronic device to communicate over wireless signal. Fidelity gives compatibility between wireless equipment from different manufacturer. Wi-Fi works on physical and data link layer It allows local area network to operate without cables and wiring. It is very much popular for the home and business network . [1,3,5,16,17]
The rapid growth of digital wireless telephony gives rise to an increasing demand for data services as well. This is achieved through the concept of transport layer connection established between different interface pairs at the two end points. During the normal operation SCTP always uses at most one path at the time for communication. [10,12]
SCTP is the fundamental member of a family of protocols designed by the SIGTRAN group to allow SS7messages to be transported over an unreliable IP infrastructure
All data transferred between the hosts is encapsulated in SCTP packets. SCTP packet contains a common header and a

sequence of structures called 'chunks'. [11,13,14]
The Stream Control Transmission Protocol (SCTP is a Transport Layer protocol, serving in a similar role as the popular protocols: TCP and UDP. Indeed, it provides some of the same service features of both, ensuring reliable, of messages with congestion control like TCP, and preserving data message boundaries similarly to UDP. However, differently to TCP and UDP, SCTP offers such advantages as multi-homing and multi-streaming capabilities. The main difference to TCP is the multihoming mechanism and the concept of several streams within a connection. Where in TCP a stream is referred to as a sequence of bytes, an SCTP stream represents a sequence of messages. Multi-streaming allows data to be partitioned into multiple streams that have the property of being independently delivered to the application at the receiver. This means that the loss of a data chunk that belongs to a certain stream will only affect the delivery within that stream, without affecting the delivery of other streams. This prevents head-of-line blocking problem that can occur in TCP, as TCP supports only a single data stream within sliding window control with adapted versions of the well known TCP slow-start, congestion avoidance, fast retransmit and fast recovery mechanisms SCTP congestion control mechanisms include two major differences with the equivalent TCP mechanisms. First, the direct dependence of SCTP on the number of bytes acknowledged, rather than the number of acknowledgements received, to increase the congestion window. Secondly, the implicit dependence of SCTP on SACK messages for acknowledging the received data chunks.[16,18]

3. Research Objective:
Following research questions were aimed in writing this paper:
- What is the performance of SCTP protocol in android 5.1 and Android 6.
- What is the performance of SCTP over TCP?
- Is SCTP better than TCP? How?

4. Experimental Work:
Various tools are available for performance measurement of Wi-Fi network such as Wi-Fi analyzer wireshark, Acrylic Wi-Fi (Windows), AirGrab WiFi Radar (Mac OS X), Cain & Abel (Windows), Homedale (Windows), LizardSystems Wi-Fi Scanner, WirelessNetView , Wireless Diagnostics (Mac OS X Lion and later), SL Speed Test, Wi-Fi Network Analyzer, Wi-Fi Spectrum Analyzer, Wireless Manager and even Wi-Fi Hotspots, InSSIDer, Xirrus Wi-Fi Inspector, Connectify, WeFi, Hotspot Shield,Plug and browse.
These tools can be installed on different operating systems depending on their compatibility. It can run on Windows, Mac OS, Linux, Android The tools which runs on android smartphones are Speedtest.net, 3G 4G WiFi Map & Speedtest, Wifi Analyzer, Network Signal Info, WiFi Expert, WiFi Manager, WiFi Connection Manager [18]. Amongst these tools,iperf tool has been selected for experimental work as this tool runs on the operating systems like windows, linux, Mac OS , Linux, reeBSD, OpenBSD, NetBSD, VxWorks,

5. Iperf tool:
iPerf is used for TCP performance tuning and it also measures throughput ,bandwidth and jitter ,data loss in case of UDP tests. There are two components of iperf tool server and client. It is an open source command line tool.[7].

6. Setup:
The parameters for the experiment are no. of requests and size of data. There are three cases in which the experiment was carried out viz. unidirectional, bidirectional (sequential and parallel). The device used for this experiment was Android 6 (marshmallow) smart phone.
Following are the different commands used for the experiment of TCP and UDP bandwidth measurements:
a)To send data of different size in one direction:
perf –c 192.168.43.144 --sctp –n 10
b)To send more no. of requests in one direction:
perf –c 192.168.43.144 --sctp -r 10
c)To send data from both directions sequentially:
perf –c 192.168.43.144 --sctp –d –n 10
d)To send data from both directions parallel:
perf –c 192.168.43.144 --sctp -p -n 10
e)To send no. of requests from both directions sequentially:
perf –c 192.168.43.144 --sctp –d –r 10
f) To send no. of requests from both directions parallel:
perf –c 192.168.43.144 --sctp –p –r 10

7. Results and Interpretation:
Case I-a) Unidirectional-Changing the size of data

Size of data (in KB)	Android 5.1		Android 6	
	Client (Mbits/Sec)	Server (Mbits/sec)	Client (Mbits/sec)	Server (Mbits/sec)
10	0.33	16.5	2.02	14.0
100	1.01	18.8	1.92	13.7
1000	1.09	19.7	1.76	22.7
10000	1.78	20.5	1.04	56

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).

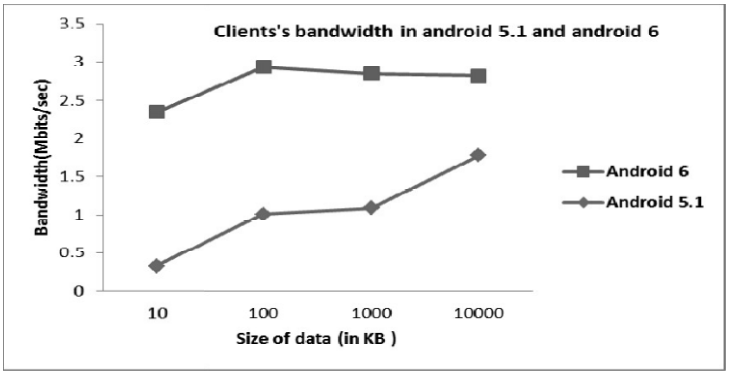


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

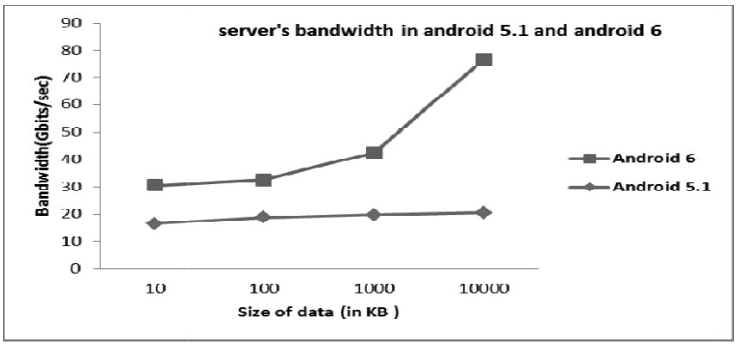
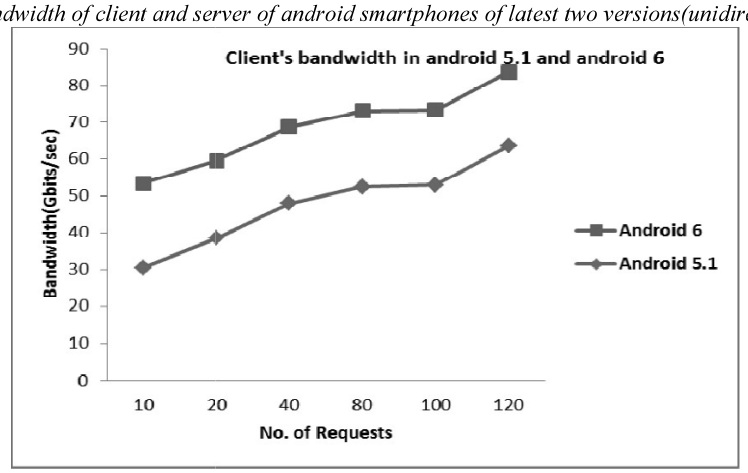


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Case I-b) Unidirectional-Changing the number of requests

No. of Requests	Android 5.1		Android 6	
	Client (Mbits/Sec)	Server (Mbits/sec)	Client (Mbits/sec)	Server (Mbits/sec)
10	30.5	22.7	22.9	1.26
20	38.5	22.6	21.0	1.25
40	48.0	22.4	20.7	1.25
80	52.5	20.3	20.5	1.25
100	52.9	20.1	20.4	1.25
120	63.5	20.0	20.1	1.00

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).



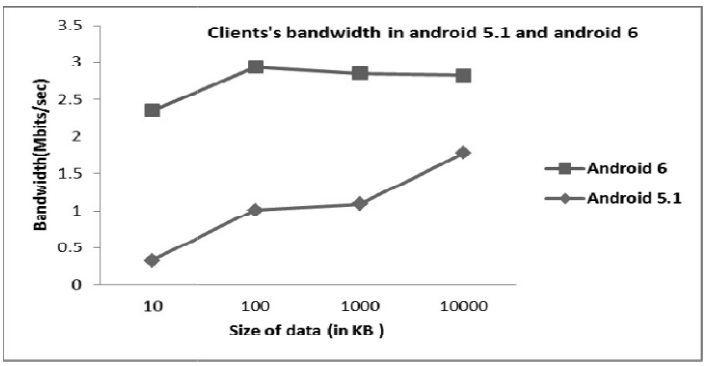


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

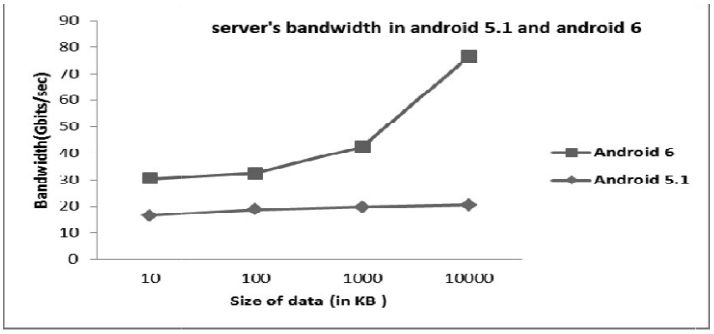


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Case I-b) Unidirectional-Changing the number of requests

No. of Requests	Android 5.1		Android 6	
	Client (Mbits/Sec)	Server (Mbits/sec)	Client (Mbits/sec)	Server (Mbits/sec)
10	30.5	22.7	22.9	1.26
20	38.5	22.6	21.0	1.25
40	48.0	22.4	20.7	1.25
80	52.5	20.3	20.5	1.25
100	52.9	20.1	20.4	1.25
120	63.5	20.0	20.1	1.00

Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

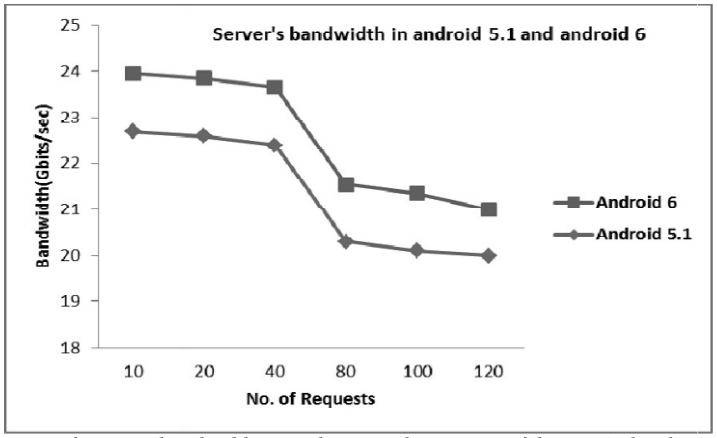


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Interpretation:

From the graphs above, it can be observed that Android 5.1 is utilizing less bandwidth and android 6 is utilizing more bandwidth.

Case II-a) Bidirectional (Sequential) -Changing the size of data

Size of data (in KB)	Android 5.1		Android 6	
	Client (Mbits/Sec)	Server (Mbits/sec)	Client (Mbits/sec)	Server (Mbits/sec)
10	1.42	27.8	18.9	6.61
100	1.20	27.8	22.9	6.62
1000	1.544	20.9	20.9	6.68
10000	1.920	4.08	21.09	6.92
100000	1.75	23.4	21.13	6.70

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).

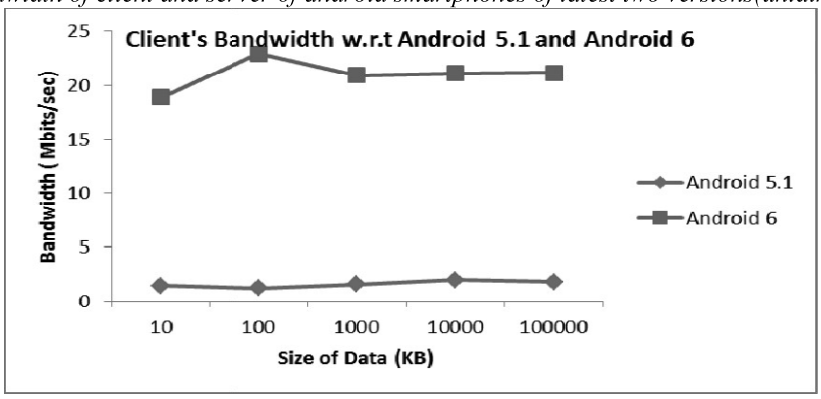


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

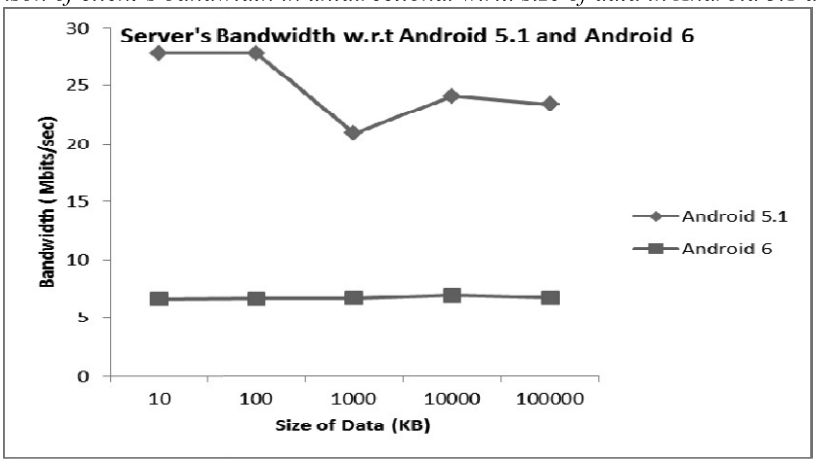


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Interpretation:

From the graphs above, it can be observed that Android 5.1 is utilizing less bandwidth and android 6 is utilizing more bandwidth, and there is a vast different between the bandwidths of android 5.1 and android 6.

Case II-b) Bidirectional (Sequential)-Changing the number of requests

No. of Requests	Android 5.1		Android 6	
	Client (Mbits Sec)	Server (Mbits /sec)	Client (Mbits /sec)	Server (Mbits/ sec)
2	10.8	26.2	12.03	28.2
20	27.0	25.2	16.08	25.2
40	31.7	24.3	12.06	26.3
80	34.5	22.85	18.03	25.85
100	33.7	26.2	17.09	27.2
120	46.7	26.2	24.06	26.2

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).

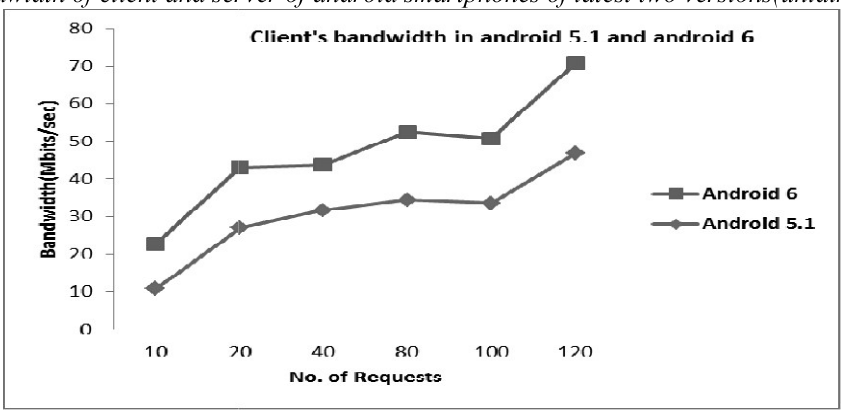


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

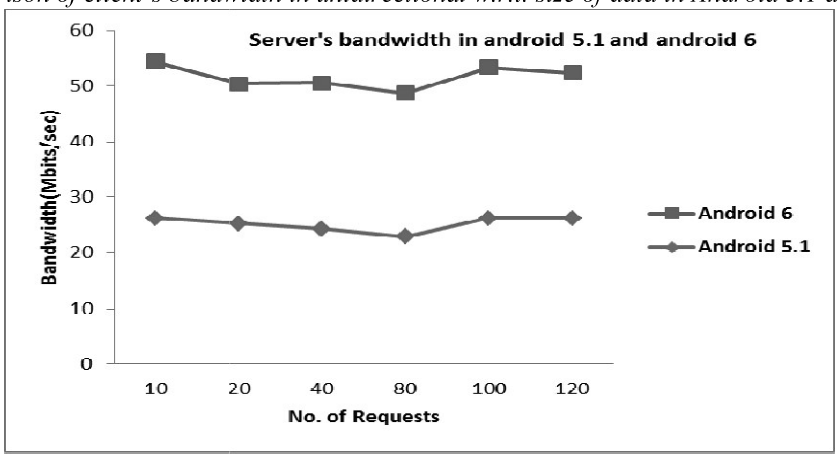


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Interpretation:

From the graphs above, it can be observed that Android 5.1 is utilizing less bandwidth and android 6 is utilizing more bandwidth.

Case III-a) Bidirectional (Parallel) -Changing the size of data

Size of data (in KB)	SCTP		SCTP	
	Client (Mbits /Sec)	Server (Mbits /sec)	Client (Mbits /sec)	Server (Mbits/ sec)
10	3.40	21.0	20.10	3.61
100	3.29	21.0	20.10	3.62
1000	3.54	19.4	16.6	3.68
10000	3.95	21.5	20.7	3.92
100000	3.40	20.6	20.7	3.87

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).

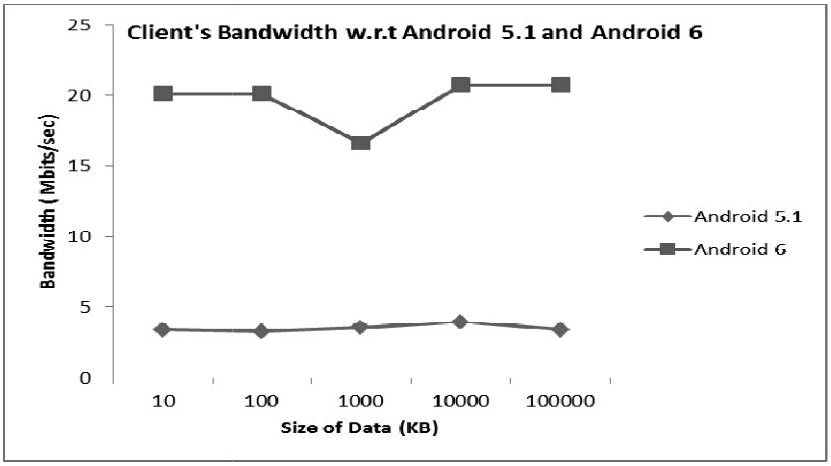


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

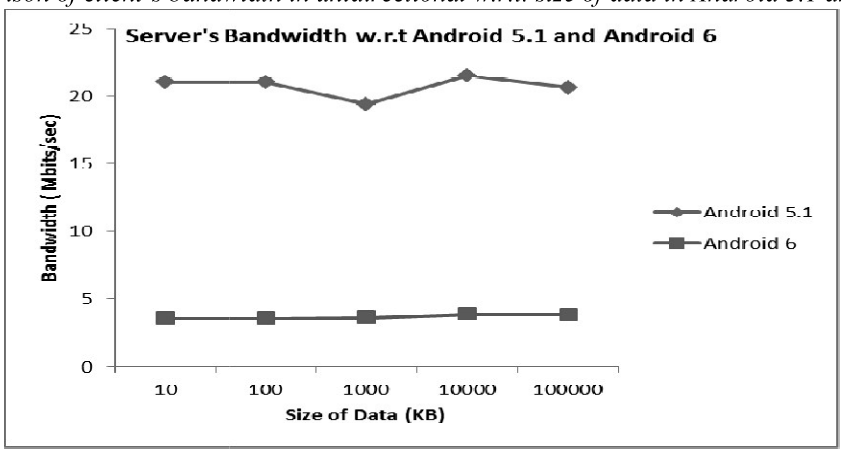


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Interpretation:

From the graphs above, it can be observed that Android 5.1 is utilizing less bandwidth and android 6 is utilizing more bandwidth, and there is a vast different between the bandwidths of android 5.1 and android 6.

Case III-b) Bidirectional (Parallel)-Changing the number of requests

No. of Requests	Android 5.1		Android 6	
	Client (Mbits Sec)	Server (Mbits /sec)	Client (Mbits /sec)	Server (Mbits/ sec)
10	19.6	18.8	1.06	2.03
20	28.3	19.7	1.04	2.05
40	36.6	20.4	1.03	2.03
80	41.6	24.8	1.05	2.05
100	43.2	25.8	1.06	2.07
120	45.4	28.2	1.05	2.03

Table 1 : Bandwidth of client and server of android smartphones of latest two versions(unidirectional).

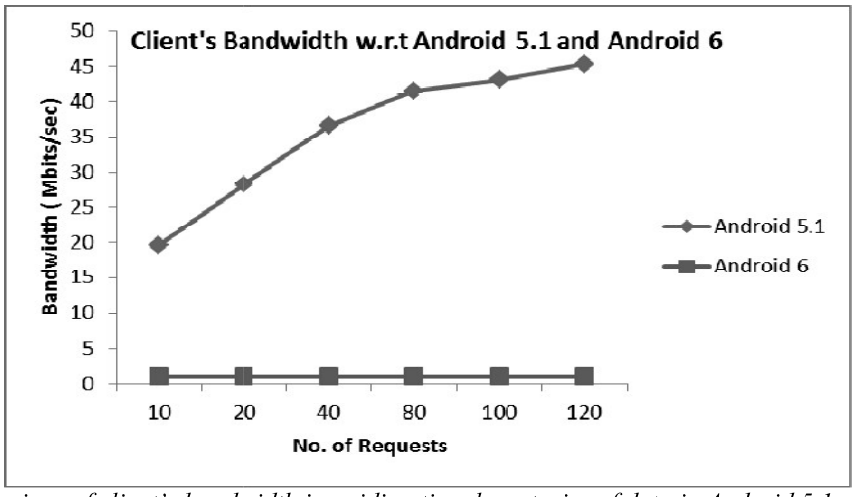


Fig.1: Comparison of client's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

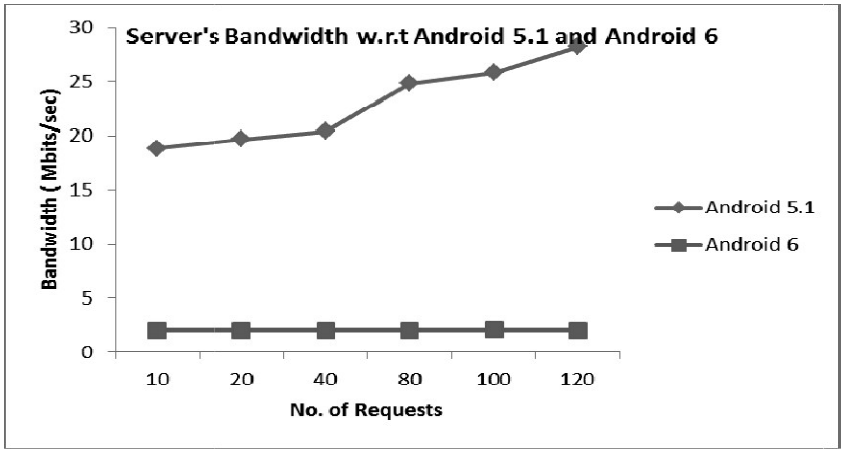


Fig.2: Comparison of server's bandwidth in unidirectional w.r.t. size of data in Android 5.1 and Android 6

Interpretation:

From the graphs above, it can be observed that Android 5.1 is utilizing less bandwidth and android 6 is utilizing more bandwidth, and there is a vast different between the bandwidths of android 5.1 and android 6.

CONCLUSION :

From all the graphs above, it can be concluded that in both versions Server device is taking constant bandwidth whereas client device is either utilizing less or more bandwidth. Compared to TCP, SCTP is a better protocol as for some of the experiments, TCP protocol failed at some level but SCTP didn't get failed at all.

REFERENCES :

1. 802.11a: A Very-High-Speed, Highly Scalable Wireless LAN Standard, White Paper, 2002, www.proxim.com
2. Bhardwaj, D., Kataoka, K., Vikram, V. A., & Hirani, V. (2015, December). Hybrid approach to distributed Wi-Fi performance assessment for multi-floor structures. In *Software, Knowledge, Information Management and Applications (SKIMA), 2015 9th International Conference on* (pp. 1-8). IEEE.
3. Boussen, S., Tabbane, N., & Tabbane, S. (2009, November). Performance analysis of SCTP protocol in WiFi network. In *Computer Sciences and Convergence Information Technology, 2009. ICCIT'09. Fourth International Conference on* (pp. 178-182). IEEE.
4. da Silva, R. F., Chen, W., Juve, G., Vahi, K., & Deelman, E. (2014, October). Community resources for enabling research in distributed scientific workflows. In *e-Science (e-Science), 2014 IEEE 10th International Conference on* (Vol. 1, pp. 177-184). IEEE.
5. Forouzan, A. B. (2006). *Data communications & networking (sie)*. Tata McGraw-Hill Education.
6. Hou, X., Deshpande, P., & Das, S. R. (2011, October). Moving bits from 3G to metro-scale WiFi for vehicular network access: An integrated transport layer solution. In *Network Protocols (ICNP), 2011 19th IEEE International Conference on* (pp. 353-362). IEEE.

http://mashtips.com/android-apps-for-wifi-network/

7.

http://searchnetworking.techtarget.com/tip/Measure-

wireless-network-performance-using-testing-tool-iPerf
8.
http://www.computerworld.com/article/2515158/desktop-apps/mobile-wireless-6-useful-wi-fi-tools-for-windows.html
9.
http://www.networkcomputing.com/data-centers/8-free-wifi-analyzers/2071006633
10.
http://www.networkworld.com/article/2925081/wi-fi/7-free-wi-fi-stumbling-and-surveying-tools-for-windows-and-mac.html#slide8
11. https://iperf.fr/
12.https://www.acrylicwifi.com/en/ wlan-software/wifi-analyzer-acrylic-professional/
13.
14. Kim, K. H., Nam, H., & Schulzrinne, H. (2014, April). WiSlow: A Wi-Fi network performance troubleshooting tool for end users. In *INFOCOM, 2014 Proceedings IEEE* (pp. 862-870). IEEE.
15. Perotto, F., Casetti, C., & Galante, G. (2007, March). SCTP-based transport protocols for concurrent multipath transfer. In *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE* (pp. 2969-2974). IEEE.
16. Ramanathan, R., & Rosales-Hain, R. (2000). Topology control of multihop wireless networks using transmit power adjustment. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE* (Vol. 2, pp. 404-413). IEEE.
17. Unni, S., Raj, D., Sasidhar, K., & Rao, S. (2015, May). Performance measurement and analysis of long range wi-fi network for over-the-sea communication. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2015 13th International Symposium on* (pp. 36-41). IEEE.
18. Yassin, M., Rachid, E., & Nasrallah, R. (2014, November). Performance comparison of positioning techniques in Wi-Fi networks. In *Innovations in Information Technology (INNOVATIONS), 2014 10th International Conference on* (pp. 75-79). IEEE.

OPEN FLOW RUNOFF ESTIMATION USING "V" NOTCH METHOD - A COMPARATIVE ANALYSIS OF TWO MICRO-WATERSHEDS IN THE KRISHNA BASIN (KR-25) IN MAHARASHTRA.

DR. SATISH S. PATIL, +DR.C.B. PAWAR*, RAMDAS S. BOLAGE*, SHANTILAL T. GHALME*
+ Professor and Head, Dr. BAMU, Aurangabad, Maharashtra.
** Assistant Professor, Dept of Computer Science,*
Indira College of Commerce and Science, Tathwade, Pune.

Abstract - *India has average annual rainfall is 1170 mm. Annual runoff computation in every watershed is very much important for planning and management of available water. It is requires to be store 70% of available runoff through various water harvesting structures. The V notch method is simple and reliable method to compute runoff available for rainfall. In this research paper an attempt have been made to calculate the available runoffs in two micro watersheds namely Renavi and Wasumbhe villages in Khanapur block of Sangli district of Maharashtra in Krishna river basin.*

Key Words- Runoff, Watersheds, Watershed treatments.

INTRODUCTION:

Weirs are typically installed in open channels such as streams to determine discharge (flow rate). The basic principle is that discharge is directly related to the water depth above the crotch (bottom) of the V; this distance is called head (h). The V-notch design causes small changes in discharge to have a large change in depth allowing more accurate head measurement than with a rectangular weir[1]. Runoff computations have significance in planning and management of available water in particular catchments.

Watersheds selected for present study fall in Khanapur block, where annual rainfall is 500-600 mm. The annual runoff is computed using V notch method in two micro-watersheds namely Renavi and Wasumbhe village. The villages with and without watershed treatments have been selected to compute runoff. The village Renavi has high number of water harvesting structures, whereas village Wasumbhe has few quantity of water harvesting structures. Watershed development program have been carried out in the experimental village Renavi in the span of 2002 -2004. The discharges calculated with V notch method are helpful in planning and management of runoff, soil erosion, agriculture, land and vegetation in control village Ghoti- Wasumbhe.

Research Methodology –

The objective of present research work is as follows.

A. Objectives-

1. To analyze the runoff from the micro-watershed.
2. Identifying the importance of the watershed development and Management.
3. Suggesting the region specific watershed treatment in the soil and water conservation.
4. "V" notch methodology is useful for the measurement of open flows and discharges from both experimental and control villages. This method reveals the importance of the scientific development and management of the soil and water resources. To assess the discharge "V" notch of 120^0 have been used in the entire study period.

B. Methodology -

Experimental village Renavi - Geographical details

Renavi village have the total population-2226 (As per census of 2001). Renavi village have geographical area of 1674 hector, of which 1426 hector area is under cultivation. Barren land of the village is 118 hector and forest canopy is spread over the 148 hector. Village watershed is composed of the 02 micro watersheds. Area suitable for the watershed is 1685 hector. Total available runoff of the village is 2310 TCM, of which 178 TCM runoff have been arrested through the watershed development work.

Table No. : Geographical details of Renavi Village

Geographical Area: 1692.00 hector	Area under watershed: 1674.00 hector
Area under cultivation: 1426 hector	Water budget *(Jalsankalpa)*
Barren land: 118 Hector	A) Total available run off :2310 TCM
Area under forest cover: 148.00 hector	B) Total arrested runoff :178 TCM

(Jalja, Department of Agriculture, Sangli district, 2005)

Control Village- Ghoti- Wasumbhe

Ghoti – Wasumbhe village is located in the semi-arid tract of Khanapur block. Geographical area of the village is composed of 873.38 hector area. Of the total area about 657 hector area is under cultivation of Kharif and Rabbi crops. About 164 hector area has the forest area controlled by Social Forestry department. Village has 152 hector barren lands. The arrested runoff is only 82 TCM and there is scope for harvesting 1205 TCM runoff.

Table No. : Geographical Details of Ghoti -Wasumbhe

Geographical area: 873.38	Watershed area- hector
Area under cultivation- 657 Hector	Water budget *(Jalsankalpa)*
Barren land hector- 152	A) Total available runoff- 1205 TCM
Forest area- 164	B) Total arrested runoff- 82 TCM
Watersheds:1	

(Jalja, Department of Agriculture, Sangli district, 2005)

Assessment of open flow's using V notch of the 120^0 to study discharges in the Experimental and Control villages:

To assess the discharges from the micro-watersheds of the study villages the "V" notch has been used. Assessment of discharge has immense importance in the planning of watershed development, management, computation of runoff, conservation of soil etc. "V" notch reading in the experimental village shows the importance of area treatment in the arresting of runoff and soil conservation, whereas in case of control villages this methodology generates base for scientific planning of the watershed activity to harvest the water and conserve soil.

Following criteria's have been adopted for the selection of the villages.

i) Criteria for selection of control village:-

1. Control village should not vary far from experimental village.
2. Control village has relatively same socio-economic and somewhat geographical conditions.
3. Control villages have not implemented watershed management program.

ii) Criteria for selection of experimental village:-

1) The experimental villages have come under drought prone region.

The experimental villages have successfully implemented the watershed management project where annual rainfall is less. The equation is developed by the Kinds vaster - Carter equation, from the Bureau of Reclamation, Water Measurement Manual, United States, 2008 is used in this study to measure the flows originated from I^{st} and II^{nd} order streams of the watershed. This equation is used in the conditions, when notch angle is greater than 90^0 and flow rate is high.

The equations have been given as follows.

$$Q = 4.28\ C_e \tan(\emptyset/2)\ (h + K)^{5/2} \quad \text{---------- (3.18)}$$

Where Q = Discharge (cfs)

C_e = Discharge coefficient

$\emptyset$ = Notch angle

h = Head (ft)

k = Head correction factor (ft).

$$C_e = 0.607165052 - 0.000874466963\ \emptyset + 6.10393334 \times 10^{-6}\ \emptyset^2$$

$$C = 0.6053572$$

$$k\ (ft.) = 0.0144902648 - 0.00033955535\ \emptyset + 3.29819003 \times 10^{-6}\ \emptyset^2 - 1.06215442 \times 10^{-8}\ \emptyset^3$$

$k = 0.0135282$, where $\emptyset$ is the notch angle in degrees

RESULTS AND DISCUSSION :

On the basis of the above formula discharge variations of the open flows have been calculated to know the discharge of the water from the upper catchment. These readings scientifically proved the discharges and importance of the area treatments in the planning of watershed development.

About 07 readings with the V notch of 120^0 has been recorded in experimental village and 09 readings has been recorded in control village during the entire study period 2009 -2010.

Table No.1 - Discharge variation in study villages

Village	Reading No	Q in feet ^{3}/s	h in cm	h in feet	h+K	$(h+K)^{5/2}$	Q in meter^3 /minute	Q in m^3 hour
Renavi	1	1.03	16.5	0.54	0.55	0.23	1.75	105
Renavi	2	0.09	6.1	0.20	0.21	0.02	0.16	9.64
Renavi	3	0.00	1.5	0.05	0.06	0.00	0.01	0.45
Renavi	4	0.75	14.5	0.48	0.49	0.17	1.28	76.50
Renavi	5	0.14	7.1	0.23	0.25	0.03	0.23	13.78
Renavi	6	0.28	9.7	0.32	0.33	0.06	0.48	28.97
Renavi	7	0.00	0.6	0.02	0.03	0.00	0.00	0.09

Table No.2 Discharge variation in Control Village

Village	Reading No	Q in feet ^{3}/s	h in cm	h in feet	h+K	$(h+K)^{5/2}$	Q in meter^3 /minute	Q in m^3 hour
Ghoti-Wasumbhe	1	1.40	18.7	0.61	0.63	0.31	2.37	142.26
Ghoti-Wasumbhe	2	1.61	19.8	0.65	0.66	0.36	2.73	163.62
Ghoti-Wasumbhe	3	0.53	12.6	0.41	0.43	0.12	0.91	54.41
Ghoti-Wasumbhe	4	0.41	11.3	0.37	0.38	0.09	0.70	41.82
Ghoti-Wasumbhe	5	0.08	5.7	0.19	0.20	0.02	0.14	8.23
Ghoti-Wasumbhe	6	0.21	8.5	0.28	0.29	0.05	0.35	21.12
Ghoti-Wasumbhe	7	0.05	4.7	0.15	0.17	0.01	0.09	5.26
Ghoti-Wasumbhe	8	0.01	1.7	0.06	0.07	0.00	0.01	0.58
Ghoti-Wasumbhe	9	0.01	2.7	0.09	0.10	0.00	0.03	1.52

(Source- on the basis of field work)

To compile the results of the watershed development program (WDP), one control village adjoining to experimental villages, namely Reanvi (KR 25), the expected outcomes have been reported in discharge variation. The control village has fast withdrawn the rainwater compared to the experimental village, which earlier affecting the drying of the channels and overall affecting the availability of water. Therefore judicious utilization and management of generated natural resource is required in the both types of villages. V notch methodology generates potential yields of the runoff from study village in both control and experimental village. In experimental village arrested runoff through watershed activities improves the ground water levels.

After monsoon (post condition) the V notch method is used to assess the runoff from the control and experimental micro watersheds in Khanapur block. Village Renavi and Control village Ghoti- Wasumbhe is located in the KR 25 watershed, whereas the village Revangaon is partially located in the KR 34 watershed. All these study villages fall in ridge portion of KR 25 and KR 34 watershed have high slope than any other village. This topography is responsible for sudden discharge of rainwater. In Renavi village highest discharge

of 105 m^3 /hr have been recorded in the WS No.1 in the month of October, Whereas the discharge of 76.50 m^3 /hr have been recorded in the WS No.3 in the month of December, 2009. In the month of January 2010, the discharges of 13.78 m3 /hr have been recorded in the WS No.3. Very negligible reading has recorded after the month of January, but streams are alive.

In Ghoti- Wasumbhe village discharge of 142 m^3 /hr at watershed no. 1 have been recorded in the month of October 2010, The flow rate is 54.41 m^3 /hr have been recorded in the watershed No. 2, in the month of October 2010. We reported flow rate of 1.53 m3/ hr in the month of January 2010. Analysis revealed that, in Wasumbhe village water is fast drained to nearest water channels, sudden losses affecting the ground water availability in this village responsible to huge dry spell of water from the month of January to end of June.

CONCLUSION :

V notch is reliable and manually controlled method for assessing of open flows, especially is drought- prone zones of India. This simple technique reveals the importance of water resource management. In present study we analyzed that with and without watershed treatment villages. This study will be useful in assessment of water resource, where the region cope with similar drought conditions.

REFERENCES :

1. LMNO@LMNOeng.com, official website, (09/01/2016, when assessed).
2.Kinds vaster - Carter equation (2008), Bureau of Reclamation, Water Measurement Manual, United States, (www.goggle.com).
3.United States, Department Interior, Bureau of Reclamation, (2001): Third Edition, and Water Measurement Manual- (12th December 2011 when assessed).
4.Website- http://www.usbr.gov/ pmts/hydraulics lab/pubs/wmm/index .htm

AN ANALYSIS OF VARIOUS FACTORS AFFECTING VELOCITY OF WEB COMMERCE APPLICATION

MRS. TEJASHRI PHALKE AND MRS. SHUBHANGI CHAVAN
B.C.A. Dept
Indira college of Commerce and Science, Pune
tejashriphalle@iccs.ac.in, shubhangi.chavan@iccs.ac.in

Abstract: *Technology is essential part of our contemporary lifestyles. Since the global business is shifting towards digitization very fast, every individual wants to make use of their time in quality manner. Concerning that now a days the majority are using Web-commerce.*

At the same time as the use of the Web-commerce, customers anticipate velocity, safety, should be credible and straightforward, and so forth. To increase income and sustain in competition growing pace of E-trade utility is a needed. Different factors are liable for speed of Web-commerce which includes Complexness, Browser Compatibility, server, website content, Source code, latency, etc.

Keywords: *Web-Commerce, Velocity, factors, performance, latency, Globe.*

INTRODUCTION:

Web commerce is the buying and selling of products and offerings by means of agencies or customers over the Globe. Web-commerce is a blend of IT (facts technologies), ITT (statistics transmission technology) and IPT (facts dealing with/processing technology). It's far about constructing relationships and creating wealth! For buying online, in an extraordinarily aggressive commercial enterprise, a poorly acting website can be a point of no return. The customers are as unreliable as butterflies. Patron has many alternatives due to the fact they are now not prepared to look forward to poorly executing internet site.

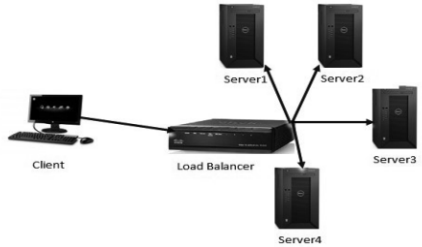
Numerous factors affect internet site pace, a number of the maximum, common are:

1. **Server**

Host server plays a decisive position on the subject of internet site's performance. Speed and reliability of web server performs an critical position in that perceptive. Server function must be to backs up its cutting-edge servers, reimburses you for downtime and is rapid. When user sends request through a browser to server, the server responds by using handing over the page through the browser. The Server will take longer time to reply or the request to go through, or to load.

Recommendations:

a. Accumulation of a load balancer will helps to improve the overall performance and security of web page. In spite of creating a center net server larger and extra powerful, we will use a load balancer to distribute traffic throughout some of servers. Load balancer is beneficial even the application is poorly written or scrambling troubles.



a. So one can improve the velocity of server, tracking actual global overall performance in real time is a great deal higher. Monitoring can seize numerous exceptional styles of problems. They encompass:

A server is down.

A server is limping, losing connections.

A server is suffering from an excessive proportion of cache misses.

A server isn't always sending correct content material.

Server activities may be monitored with tools which includes SolarWinds® Server & Application Monitor (SAM), software -as-a-service (SaaS) tracking equipment etc.

for example: Fing, and software for each Android and iOS, that is rapid, free and extraordinarily beneficial.

b. 3 Tier Deployment

3 Tier deployment layer encompass Load Balancer, net Servers, & Database servers, with Firewall at each tier layer. 3-tier structure is secured and scalable because isolation allows ensuring identity of difficulty at any layer and cleaning troubleshooting of performance issues.

1. **File Types and Size**

If the file sizes are larger it takes long time to load. Record size increases due to large photos, uncompressed facts, if javascript and css code is embedded in html code and so forth.

Recommendations: File Size may be reduced by using:

a. By means of Optimizing snap shots inside the shape of jpeg, gif, png.

b. In place of setting large photo, set the width and peak attributes of <img> attribute in HTML.

c. Compress content by HTTP compression, which enables to ship all web page statistics in a single smaller file rather than many one-of-a-kind files.

d. Optimize and compress your JavaScript and CSS files by means of combining them and minifying the source

code.

e. Place JavaScript and CSS in outside documents

f. Place Style sheet References on the top and Script References at the bottom

e.g.<html>
<link rel="stylesheet" href="mycss.css">
<body>

<!-- Content will go here -->

</body>
<script src="myscripts.js"></script>
</html>

g. Preferred web page length should not more than 2 MB.

2. **Browser Compatibility**

Browser compatibility is likewise any other tremendous component which influences internet site pace. Browsers play a function of translators and interpreter whilst its miles dispatched via servers. If the browser cannot examine the records, the response page might not carry out well.
As an admin cannot manipulate that the person is going to pick which browser, the excellent you can do is manipulate how browsers examine your website online.

Recommendations: To boom the browser compatibility distinctive equipment can be use along with Adobe Browserlab, Netrenderer, Browsera, etc.

3. **Complexity**

Complexity can be on the client side as well as the application side. Code level complexity is difficult and costly to diagnose without access to tool. At the same time as server requests are can be optimized in diverse approaches (like with parallel processing, asynchronous operations etc.) contemporary websites rely closely on client side JavaScript execution, smart caching and occasionally third parties content material.

Recommendations: The implication of all this complexity is the challenge in managing it. Paying attention to all the elements, understanding each component's contribution to page load times and quickly diagnosing the root cause of performance issues

4. **Latency**

Latency is nothing but distance from the website origin server to the user who is accessing the website. For front-end application operations, mean time to repair (MTTR) and mean time between failure (MTBF) are critical metrics, where Time to first (TTFB) which initiates reasonable user experience and search rankings. TTFB is duration of browser which receives web server's first response. According to the user expectations, servers, back-end infrastructure, front-end API and applications operates on unified monitoring, which are keys to ensure the speed and response.

Recommendations: -Organizations can get around this by either building data centers (or locating and managing servers) in many locations throughout the world or partnering with a content delivery network (CDN) that has already built a high speed network with points of presence in all of the locations, the organization needs to reach customers and employees

6. The Database

Poor website response time is not always because of server problem but because of database thathelps them. Gradual database queries or too many database queries and non-optimized databases are the problems subject with internet sites overall performance.

Recommendations: 1) While executing queries proper use of indices ought to be recommended which enhance response time?.
e.g.

- DELETE FROM Products WHERE UnitPrice = 1
- SELECT * FROM PRODUCTS
  WHERE UnitPrice BETWEEN 14 AND 16

Grouping records with a GROUP BY clause will often require sorting, so a UnitPrice index will also
help the following query to count the number of products at each price.

SELECT Count (*), UnitPrice FROM Products
GROUP BY UnitPrice

2) Sufficient availability of database resources such as CPU, memory, Disk Space.
3) Understand the load and individual response time of each service by identifying

Service's communication with the database, which queries are executed? How often are the queries executed per request? How many rows do they return? etc.
4) Check at your routers, test your cables, and take a look at your network interfaces
5) Enhance SQL queries to improve website performance

6) Rather than the usage of tables with massive quantity of information, table partitioning ought to be used.
7) If possible use of SSD storage for a database is recommended.

The following table shows the various sites with their rank, speed. Load time, server and database setup which reflect their performance.

Sr. No.	Website	Type	Ranking		Load Time	Speed Index	server	Database
			Global	Country				
1	google.co.in	Internet telecom search Engine	9	1	3.756s	2193	GWS (Google Web Server),GFE/1.3 (Google Front-End), bsfe (Blog Search Front-End), Apache and many more	Bigtable
2	facebook.com	Internet telecom social Network	2	2	5.774s	3368	HipHop,Apache HTTP Server, Scribe (log server	MySQL
3	youtube.com	Arts Entertainment TV Video	3	3	5.525s	6231	NetScalar, Apache, Python application server	MySQL
4	amazon.in	Shopping and General Merchandise	87	8	10.273 s	6810	Nginx, **Apache,SSH,postfix**	(NoSQL databases such as DynamoDB) MySQL
5	flipkart.com	Shopping and General Merchandise	170	9	5.296s	3761	Nginx	MySQL
6	yahoo.com	news and Media	6	5	5.441s	3950	POP3, IMAP,SMTP	Matrix (formerly ParAccel).
7	cricbuzz.com	Sports	367	17	7.073s	2701	Nginx	directly connect there applications to the ICC server

CONCLUSION :

The customers are as unreliable as butterflies. A bad appearing internet site outcomes in a bad user revel in, and sites with poor user stories deserve much less promotion. To be successful in today's net trade world normal overall performance of internet site is have to in which pace performs an important function. There are ranges of things which affect speed of internet site. To conquer issues caused due to these factors, possible answer are given.

REFERENCES :

1. http://www.dnnsoftware.com/wiki/working-with-large-files
https://thenextweb.com/apple/2011/01/15/how-to-use-activity-monitor-to-make-your-mac-faster/https://social.technet.microsoft.com/wiki/contents/articles/34132.sharepoint-2016-increase-file-size-limit.aspx
2.https://odetocode.com/articles/237.aspx
3.
https://bib.irb.hr/datoteka/398249.Belani_et_al_int11162d-1c_391-398.pdf
4.http://ieeexplore.ieee.org/document/7385038/
5.https://www.webpagetest.org/
7.https://googlesystem.blogspot.in/2007/09/googles-server-names.html#gsc.tab=0 b=0

WIRELESS PENETRATION TESTING

SANJEEV KUMAR, MANISH KUMAR AND SHIVENDU BHUSHAN
Network Administrator**,** Network Security Administrator, Asst. Professor, HOD-BCA
Indira College of Commerce & Science
sanjeevkumar@iccs.ac.in, manishkumar@indiraicem.ac.in, shivendu@iccs.ac.in

Abstract: Wi-Fi is developed on IEEE 802.11 standards defined by the Institute of Electrical and Electronics Engineers., and it is widely used in wireless communications. Wi-Fi sets up numerous ways to build up a connection between the transmitter and the receiver such as DSSS, FHSS, Infrared (IR) and OFDM. A Wi-Fi infrastructure generally consists of hardware components such as wireless routers and Aps, antennas, relay towers and authentication servers, and software components such as encryption algorithms, key management and distribution mechanisms.

Keywords: *WEP/WPA/WPA2, Dictionary Attack,* vulnerabilities, *Encryption*

INTRODUCTION:

Wireless networks are all around us. You want it or not, you ARE part of this system. As new wireless encryption-standards are being developed, new attack techniques are being discovered and presented at the same time. It's a real arms race. Wireless networks broadcast their packets using radio frequency or optical wavelengths. A modern laptop computer can listen in. Wprse, an attacker can manipulate packets on the fly and persuade wireless stations t accept his packets as legitimate. Wireless cracking is an information network attack similar to a direct intrusion. Two frequent types of vulnerabilities in wireless LANs are those caused by poor configuration, and those caused by weak encryption or flawed security protocols. There are two basic types of vulnerabilities associated with WLANs: those caused by poor configuration and those caused by poor encryption. Poor configuration causes many vulnerabilities. Wireless networks are often put into use with no or insufficient security settings. With no security settings – the default configuration – access is obtained simply by association. Without sufficient security settings, networks can easily be defeated by cloaking and/or MAC address filtering. Poor encryption causes the remaining vulnerabilities. Wired Equivalent Privacy (WEP) is defective and can be defeated in several ways. Wi-Fi Protected Access (WPA) are vulnerable to dictionary attacks. Different types of wireless security protocols were developed for wireless networks protection. The wireless security protocols are WEP, WPA, and WPA2, serving the same purpose but being different at the same time. No matter how protected and encrypted, wireless networks cannot keep up in safety with wired networks. The latter, at their most basic level, transmit data between two points, A and B, connected by a network cable. To send data from A to B, wireless networks broadcast it within their range in every direction to every connected device that happens to be listening.

Wireless Security Protocols

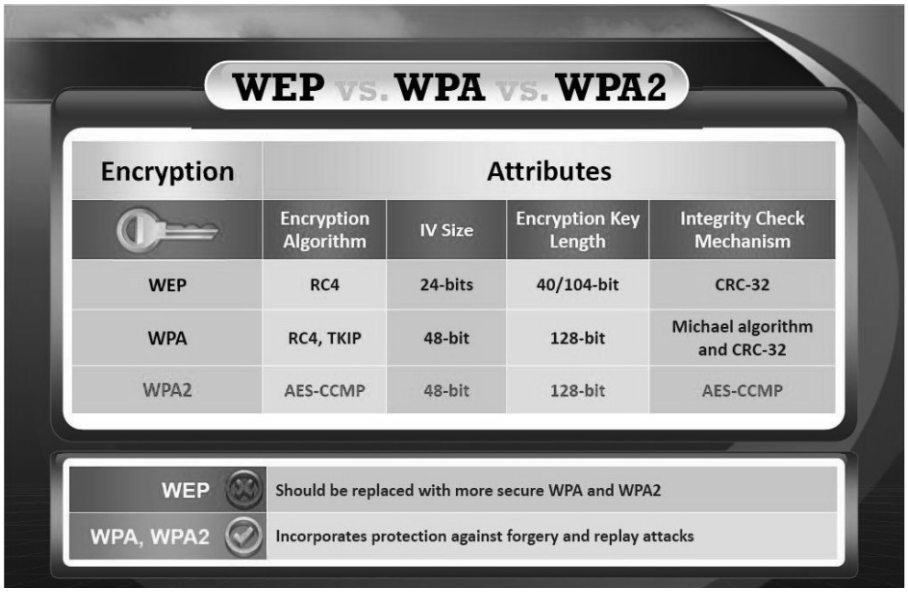
1. ***WEP*** **(Wired Equivalent Privacy) –** WEP was developed for wireless networks and approved as a Wi-Fi security standard in September, 1999. WEP was aimed to offer the same security level as wired networks, however there are a bunch of well-known security issues in WEP, which is also easy to break and hard to configure. WEP was officially abandoned by the Wi-Fi Alliance in 2004.

2. WPA (Wi-Fi Protected Access) – WPA was used as a temporary security enhancement for WEP. One year before WEP was officially abandoned, WPA was formally adopted. Most modern WPA applications use a preshared key (PSK), most often referred to as WPA Personal, and the Temporal Key Integrity Protocol or TKIP (/tiːˈkɪp/) for encryption. WPA Enterprise uses an authentication server for keys and certificates generation.
WPA, just like WEP, after being put through proof-of-concept and applied public demonstrations turned out to be pretty vulnerable to intrusion. The attacks that posed the most threat to the protocol were however not the direct ones, but those that were made on Wi-Fi Protected Setup (WPS) - auxilliary system developed to simplify the linking of devices to modern access points.

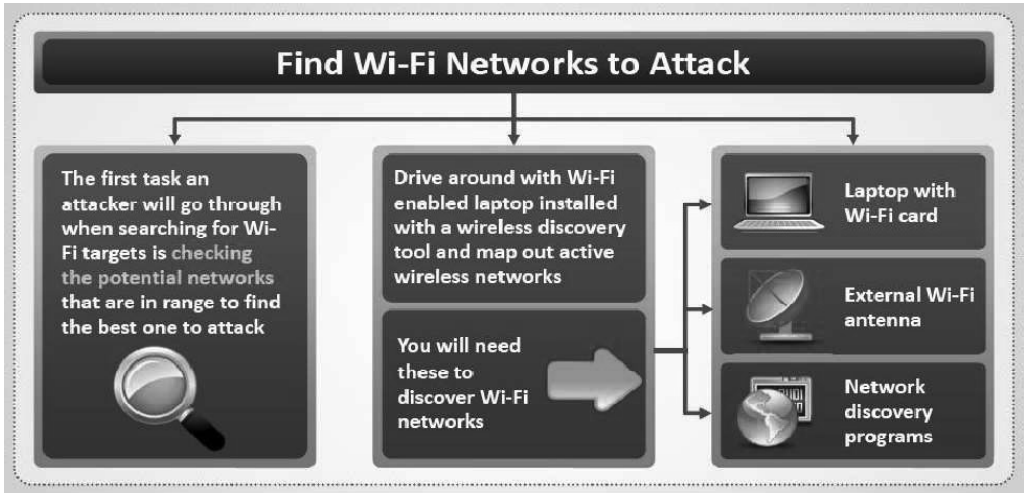
3. WPA 2 (Wi-Fi Protected Access version 2) – The most important improvement of WPA2 over WPA was the usage of the Advanced Encryption Standard (AES) for encryption. AES is approved by the U.S. government for encrypting the information classified as top secret, so it must be good enough to protect networks. At this time the main vulnerability to a WPA2 system is when the attacker already has access to a secured Wi-Fi network and can gain access to certain keys to perform an attack on other devices on the network. This being said, the security suggestions for the known WPA2 vulnerabilities are mostly significant to the networks of enterprise levels.
Unfortunately, the possibility of attacks via the Wi-Fi Protected Setup (WPS), is still high in the current WPA2-capable access points, which is the issue with WPA too. And even though breaking into a WPA/WPA2 secured network through this hole will take anywhere around 2 to 14 hours it is still a real security issue and WPS should be disabled

Wireless Standards:

802.11a	Bandwidth up to 54 Mbps and signals in a regulated frequency spectrum around 5 GHz
802.11b	Bandwidth up to 11 Mbps, and uses the unregulated radio signaling frequency (2.4 GHz)
802.11g	Bandwidth up to 54 Mbps, and it uses the 2.4 GHz frequency for greater range
802.11i	A standard for Wireless Local Area Networks (WLANs) that provides improved encryption for networks that use 802.11a, 802.11b and 802.11g standards
802.11n	Uses multiple input, multiple output (MIMO) technology to give Wi-Fi more speed (over 100Mbps) and range
802.16	A group of broadband wireless communications standards for Metropolitan Area Networks (MANs)
Bluetooth	Supports a very short range (~10 meters) and relatively low bandwidth (1-3 Mbps) designed for low-power network devices like handhelds



Find Wi-Fi Networks to Attack



How to Break WEP Encryption:

- Start the wireless interface in monitor mode on the specific access point channel
- Test the injection capability of the wireless device to the access point.
- Do a fake authentication with the access points, you can use tool such as aireplay-ng.
- Start a Wi-Fi sniffing with a bssid filter to collect unique IVs, airodump-ng or cain & Abel can be used to perform sniffing.
- Start a Wi-Fi packet encryption tool such as aireplay-ng in ARP request replay mode to inject packets.
- Extract encryption key from the IVs using cracking tools such as airecrack-ng.

How to Break WPA/WPA2 Encryption:

1. WPA PSK – WPA PSK uses a user defined password to initialize the TKIP, which is not crackable as it is a per-packet key but the keys can be brute-forced using dictionary attacks.
2. Offline Attack – you only have to near the AP for a matter of seconds in order to capture the WPA/WPA2 authentication handshake, by capturing the right type of packets, you can crack WPA keys offline.
3. Brute-Force WPA Keys – you can use such as aircrack, aireplay, KisMac to brute-force WPA Keys
4. De-authentication Attack – Force the connected client to disconnect, then capture the re-connect and authentication packet using tool such as aireplay, you should be able to re-authenticate in a few seconds then attempt to Dictionary Brute Force the PMK

Countermeasure:

While WPA2 is much more secure than WPA and therefore much more secure than WEP, the security of your router heavily depends on the password you set. WPA and WPA2 let you use passwords of up to 63 characters.

Use as many various characters in your Wi-Fi network password as possible. Hackers are interested in easier targets, if they can't break your password in several minutes, they will most likely move on to look for more vulnerable networks. Summary:

1. WPA2 is the enhanced version of WPA;
2. WPA only supports TKIP encryption while WPA2 supports AES;
3. Theoretically, WPA2 is not hackable while WPA is;
4. WPA2 needs more processing power than WPA;
5. **Use NetSpot to check your encryption!**

CONCLUSION :

If you leave your router with no security then anyone can steal the bandwidth, perform illegal actions out of your connection and name, monitor your web activity, and easily install malicious apps in your network. Both WPA and WPA2 are supposed to secure wireless internet networks from unauthorized access.

REFERENCES :

1. Kali Linux Wireless Penetration Testing by Vivek Ramachandran.
2. Penetration Tester's Open Source Toolkit by Johnny Long and others, Syngress Publishing, Inc., 2006.
3. Wireless Security Handbook by Aaron E. Earle, Auerbach Publications, 2006.
4. Hacking Wireless Networks for Dummies by Kevin Beaver and Peter T. Davis, Wiley Publishing, Inc., 2005.

CRYPTOGRAPHY: "ALPHA-NUMERIC CIPHER PROPOSED ALGORITHM"

MS. SHITAL ALHAT
RJSPM's Arts, Commerce & Science College, Bhosari, Pune- 39
Department of Computer Science
Email:-shital _1982@rediffmail.com

Abstract: As we know now a day's everyone is using Internet for so many purposes so that it is essential to get the security for their data. The data must be in coded format.

Cryptography is a method where we are transforming the data into the coded format so no one can easily read it and use it. Now everywhere we are using digitalization for transaction so security is important concern. There are so many techniques used for security purpose such as microdots, merging words with images, converting plain text into cipher text i.e.in coded format. In cryptography we are using encryption and decryption process. There are some algorithms that are used for converting plain text into cipher text such as Ceaser cipher. To give more security for the data I have introduced new proposed algorithm "ALPHA-NUMERIC CIPHER PROPOSED ALGORITHM" which will convert plain text in to cipher text.

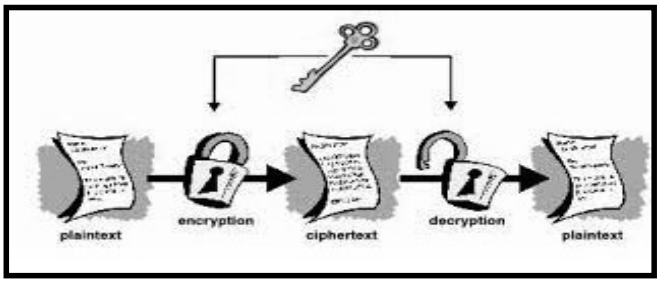
Keywords:Alpha-Numeric,Cryptology, Encryption, Decryption, Plaintext , Cipher text

INTRODUCTION:

As we know now a day's everyone is using Internet for so many purposes so that it is important to get the security for their data. The data must be in coded format .An accent people are using the same technology e.g. Different types of voice, coded words etc.

Cryptography is a method where we are transforming the data into the coded format so no one can easily read it and use it. In this technique we convert plain text into cipher text by using some algorithms.

The main aim of cryptography is to provide the Confidentiality, Integrity, Non-repudiation, and Authentication.



This word is consequent from the Greek word *kryptos*, meaning hidden. The origin of cryptography from Egyptian practice of hieroglyphics. The first known use of a modern cipher was by Julius Caesar, who did not trust his messengers when communicating with his governors and officers. For this reason, he created a system in which each character in was replaced by a character three positions ahead.

The earliest forms of secret writing are "Write in a way that no one can easily read it". The main technique of cipher types are transposition ciphers, which rearrange the order of letters (e.g., "hello world "becomes "olleh dlrow"), and substitution ciphers, letter is replaced by other letter or groups of letters (e.g., "my choice" becomes "nz cfpjdf"). The main reason of above techniques is give security to the data.

For the purpose to provide the additional security, we are modifying Caesar cipher and using substitution cipher technology develop a new technique that is proposed algorithm **"ALPHA-NUMERIC CIPHER PROPOSEDALGORITHM"**

Step I-In this technique firstly we will use **ALPHABET to NUMBER TABLE** i.e. we are substitute alphabets by numbers means alphabet 'A' will be substitute by the number 'A' by '0', 'B' by '1' and ……. 'Z' by '25'.After this we will get cipher text

ALPHABET to NUMBER TABLE

ALPHABET	NUMBER	ALPHABET	NUMBER
A	0	N	13
B	1	O	14
C	2	P	15
D	3	Q	16
E	4	R	17
F	5	S	18
G	6	T	19
H	7	U	20
I	8	V	21
J	9	W	22
K	10	X	23
L	11	Y	24
M	12	Z	25

Step II-After this whatever cipher text we get, again transform into cipher text by using**NUMBER to ALPHABETTABLE** As we know there are 26 alphabets in English language, first we will take mid that is number 13.We are starting assigning the alphabet from this mid position. So number 13 = alphabet A,14=B, 15=C…… Z=25 whatever value we will get that is our final cipher text.

NUMBER to ALPHABET TABLE

NUMBER	ALPHABET	NUMBER	ALPHABET
0	N	13	A
1	O	14	B
2	P	15	C
3	Q	16	D
4	R	17	E
5	S	18	F
6	T	19	G
7	U	20	H
8	V	21	I
9	W	22	J
10	X	23	K
11	Y	24	L
12	Z	25	M

Illustration of "ALPHA-NUMERIC CIPHER PROPOSEDALGORITHM"

Step I : Convert each alphabet into number using ALPHABET to NUMBER TABLE
Step II: Now cipher text that we get from first step again Convert into alphabet using NUMBER to ALPHABET TABLE
e.g. Convert the message "Hello this is my first program"
Step I : Convert each alphabet into number using normal ALPHABET to NUMBER table.
Plain text:-**H E L L O T H I S I S M Y F I R S T P R O G R A M**
Cipher Text: -**7 4 11 11 14 19 7 8 18 8 18 12 24 5 8 17 18 19 15 17 14 6 17 0 12**

Step II: Now cipher text that we get from first step again Convert into alphabet usingNUMBER to ALPHABET TABLE.

7 4 11 11 14 19 7 8 18 8 18 12 24 5 8 17 18 19 15 17 14 6 17 0 12
Final Cipher Text:-U R Y Y B G U V F V F Z L S V E F G C E B T E N Z
Comparative Analysis
ALPHA-NUMERIC CIPHER PROPOSED ALGORITHM is an extension to Caesar cipher. Caesar cipher can be analysed by any attacker easily, so new concept was implemented to complicate the Caesar Cipher & increase the complexity of the attacker to decode it.
The proposed system is an improvement over traditional plain encryption methods by using alphabet to number conversion and calculating mid. The text encrypted using proposed method, can't be decrypted using traditional crypto-analysis tools.
Advantages:

- It is hard to crack, less chances of getting hacked.
- One of the easiest methods to use in cryptography and can provide good security to the information.

Future Enhancement:

- We propose that this encryption method can be applied for data encryption in banks, defense, mobile networks, ATM networks, government sectors, etc. for sending confidential data**.**
- We propose to compose decryption algorithm for the same.

CONCLUSION :

- For proposing any new algorithm ,designer should being aware of security concern as well as must also consider probable future developments while working on their designs.
- This algorithm helps us to give more security for the data which will increase computer processing power.

REFERENCES :

1. https://www.learncryptography.com
2. https://www.ciphersbyritter.com
3. https://www.crypto101.io
4. "Modified Caesar Cipher for Better Security Enhancement" - International Journal of Computer Applications (0975 – 8887) Volume 73– No.3, July 2013
5. "Developing a Modified Hybrid Caesar Cipher and Vigenere Cipher for Secure Data Communication" - Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol.5, No.5, 2014

CLOUD COMPUTING DATA COMPRESSION TECHNIQUES: A REVIEW

HARSHADATALNIKAR,
Department Of Computer Science, S. N. Arts, DJM Commerce, BNS Science College, Sangamner
harshada.talnikar@gmail.com

Abstract: Cloud computing satisfies the need to connect resources and users without having physical connection. As the need for data processing at higher rate is increasing it leads a high computational requirements. Such resources for high computation are not available at each user end, in such situations cloud commuting proves very useful for users. Vast data manipulation at cloud requires some compression techniques to reduce size of data. Compression algorithms represents data in such a way that to increase data density per memory unit. It results in reduced data storage requirement and increased data transmission capacity.

Keywords: Cloud computing, resources, data computation, compression algorithms, data storage, data transmission.

INTRODUCTION:

Cloud computing [1, 2] is a used for the delivery of hosted services over the internet. Cloud computing allows companies to consume a computer resource, such as a virtual machine, storage or an application, It allows to utilize just like an electricity rather than having to build and maintain computing infrastructures in house (Figure 1).

Opposite to traditional computing where data is stored on your PC's local hard drive, the data in the cloud is stored on many physical and virtual servers which can be hosted by a third-party service provider. Example of a cloud computing file storage provider is Dropbox. Dropbox files can be accessed from any device via the Internet.

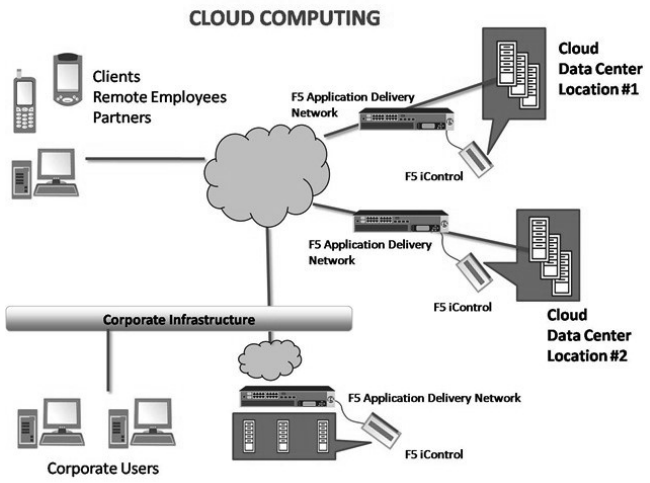


Figure 1: Cloud Computing

Cloud computing benefits

Cloud computing offers several attractive benefits for businesses and end users. Such as:

·Self-service: End users can use resources for almost any type of workload on demand. This eliminates the traditional need for IT administrators to manage compute resources.

·Elasticity: Companies can as per demand scale up as computing needs increase and scale down again as demands decrease. This eliminates the need for massive investments in local infrastructure.

· Pay per use: Compute resources are measured so that to enable users to pay only for the resources and workloads they use.

·Migration flexibility: Organizations can move certain workloads from the cloud or to different cloud platforms as desired or automatically for better cost savings or to use new services as they emerge.

Cloud computing deployment models

Cloud computing services are categorized as private, public or hybrid.

Private cloud services are used from a business's data center to its internal users. This model is useful for the versatility and convenience of the cloud, while preserving the management, control and security to local data centers. Internal users may or may not be billed for services. Some private cloud technologies and vendors include VMware and OpenStack.

In the public cloud services, a third-party cloud service provider delivers the cloud service over the internet. Public cloud services are sold as on demand, usually by the minute or hour, though long-term commitments for many services. Customers has to pay for the CPU cycles, storage or bandwidth they consume. Leading public cloud service providers include Amazon Web Services (AWS), Microsoft Azure, IBM and Google Cloud Platform.

A hybrid cloud service is a combination of public cloud services and an on-campus private cloud, with orchestration and automation between the two. Companies can use the private cloud to run mission-critical workloads or sensitive applications and use the public cloud to handle workload bursts or spikes in demand. The goal of a hybrid cloud is to provide a unified, automated, scalable environment that takes advantage of all that a public cloud infrastructure while still maintaining control over mission-critical data.

In addition, organizations are increasingly using a multi cloud model, or the use of multiple infrastructure-as-a-service providers. This helps applications to migrate between different cloud providers or to even operate concurrently across two or more cloud providers. Multicloud implementation can be a challenge because of the differences between cloud providers' services and application program interfaces (APIs).

Data compression

Data compression [3] is one of the useful technologies for multimedia applications. Unless data compression algorithms are applied, it would not be practical to put images, audio and video on websites. Mobile phones cannot be able to provide communication clearly without data compression. Using data compression techniques, one can reduce the consumption of resources, as hard disk space or transmission bandwidth. Data Compression is the process of representing data so that it takes less storage space or less transmission time. Compression is possible because the real world data is very redundant.

Classification of compression methods

We have two types of compression methods:

Lossless compression: - It is used to reduce the amount of source information to be transmitted in such a manner when compressed information is decompressed, there is not any loss of information.

Lossy compression: - Lossy compression is normally aimed at not to reproduce an exact copy of the information after decompression. Here some information is lost after decompression

Let's discuss some lossless compression techniques

Null Compression

This technique replaces a series of blank spaces with a compression code, which is followed by a value that represents the number of spaces

Run Length Encoding

This algorithm is especially used if data contains a large number of repeating characters, but does not work effectively if data file has less repeating of characters. Run length encoding is useful for images stored as solid black pixels.

Run length encoding expands the null compression technique by compressing any series of four or more repeating characters. The characters are replaced with a compression code, one of the characters, and a value that represents the number of characters to repeat.

In this technique first it reads file then it scans the file and find the repeating string of characters [4].When repetition of characters found it will store those characters with the help of escape character followed by that character and count the binary number of items it is repeated.

Huffman Coding

The Huffman coding algorithm is named after its inventor, David Huffman, [5]. Huffman Coding Algorithm uses a bottom-up approach

Steps in creation of Huffman tree.

1. Initialization: sort the existing nodes according to their frequency counts and store in a lists.

2. Repeat the following steps until the sorted list has only one node left

(a) From the list select two nodes with the lowest frequency counts.

Form a Huffman sub tree that has these two nodes as child nodes and

create a parent node.

(b) Assign the sum of the children's frequency to their parent node and

insert it into the list. Maintain the order of list.

(c) Delete these children from the sorted list.

3. Assign a 0 and 1codeword to the two branches of the tree on the path from the root.

After creating the Huffman tree, the method creates a prefix code for each node from the alphabet by traversing the tree from the root to the node. It creates 0 for left node and 1 for a right node.

LZW (Lempel-Ziv Welch) compression method

LZW is the most popular data compression method.

The main steps for this technique, firstly it will read the file until the characters in a file are null. It assigns a code to each character. If the same characters are found in a file then it will not assign the new code and then use the existing code from a dictionary. The process is continuous.

For, data transmissions, the dictionary is passed to a receiving system so it can decode the characters. For file storage system, the dictionary is stored along with the compressed file.

Spatiotemporal Compression

This technique explores spatial correlation of data, it partitions a data set into clusters so that, in one cluster all edges from the graph have similar time series of data. In each cluster, the workload can be shared by the inference based on time series similarity. Based on it, a data driven scheduling will be developed to allocate the computation and storage on cloud for better big data processing services. [6]

CONCLUSION :

This paper reviews the storage problems with cloud and some techniques used to overcome the problems in cloud computing. Though the research has been taken place by long time, the problems are not solved completely. As the increase in the amount of data computation enormously, it is not acceptable for efficient storage in cloud computing. So, there is need to develop some future technologies that will solve the storage problem.

REFERENCES :

1) M.Ali, S. U. Khan and A. V. Vasilakos, *Security in cloud computing: Opportunities and challenges, Information Sciences* 305 (2015).
2) L. Wei , H. Zhu , Z. Cao , X. Dong , W. Jia , Y. Chen and A. V. Vasilakos, *Security and privacy for storage and computation in cloud computing, Information Sciences,* 6 May 2013.
3) D.Salomon, 2004. *Data compression: The complete reference*. Springer Science & Business Media.
4) D.E. Asuquo, E.O. Nwachukwu and N.J. Ekong, *Application of Run-length Encoding Technique in Data Compression*.
5) S. Shanmugasundaram and R. Lourdusamy, *A Comparative Study of Text Compression Algorithms* International Journal of Wisdom Based Computing, Vol. 1 (3), December 2011.
6) C. Yang et al., *A spatiotemporal compression based approach for efficient big data processing on cloud, J. Computer. System Sci*. (2014).

NANOTECHNOLOGY: FUTURE OF COMPUTER SCIENCE

MS. GEETANJALIKOPLE
Assistant Professor
geetanjalikople@gmail.com
Haribhai V. Desai College Pune

Abstract : Actually nanotechnology is a very broad area of study and research at present. It has beendeveloped by many researchers and used in many fields of studies including physics, chemistry, biology, material science, engineering, and computer science. In this paper, we identify the nanotechnology development community and needs of nanotechnology in the field of computer science. In this paper, we are mainly concerned with top down approach and bottom upfabrication approach of nanotechnology that directly affects modern computer design and architecture.

Keywords: nanotechnology, nanofabrication, quantum dots, carbon nanotubes, Nano design, molecular nanotechnology.

INTRODUCTION:

In 1959, Richard Feynman, a future Nobel Laureate, gave a visionary talk entitled
"There's Plenty of Room at the Bottom on miniaturization to nanometer-scales. Later, the work of Drexler [1, 2] also gave futuristic visions of nanotechnology. Feynman and Drexler's visions inspired many researchers in physics, material science, chemistry, biology and engineering to become nanotechnologists. In a more general context nanotechnology can be seen as just the current stage ofa long-term ability to understand and manipulate matter at ever smaller scales as time goes by. Over the last century, physicists and biologists have developed a much more detailed understanding of matter at finer and finer levels. At the same time, engineers have gradually acquired the ability to reliably manipulate material to increasingly finer degrees of precision. Although we have long known much of what happens at the nanolevel, the levels of knowledge implied by; 1) knowing about the existence of atoms, 2) actually seeing them, 3) manipulating them, and 4) truly understanding how they work, are dramatically different.
The extensive use of computer and its wide application in the modern world have forced theresearchers to improve and manufacture a smaller, faster and a more reliable computer. This objective can be fulfilled by nanotechnology. Using nanotechnology we can design and manufacture electronic components and devices that can be used directly to make smaller, faster and reliable computer.

What is Nanotechnology?
Nanotechnology: It is defined as the engineering of functional systems at the molecular scale.
OR
Nanotechnology refers to the manipulation of matter on an atomic and molecular scale.
OR
The term nanotechnology is defined as "the design, characterization, production and application of structures, devices and systems by controlled manipulation of size and shape at the nanometre scale
(atomic, molecular and macromolecular scale) that produces structures, devices and systems with at least one novel/superior characteristic or property".

Types of Nano-Materials
There are many types of intentionally produced nanomaterial, and a variety of others are expected to appear in the future.Most current nanomaterial could be organized into four types:

(1) Carbon-based materials. These nanomaterials are composed mostly of carbon, most commonly taking the form of a hollow spheres, ellipsoids, or tubes. Spherical and ellipsoidal carbon nanomaterials are referred to as fullerenes, while cylindrical ones are called nanotubes. These particles have many potential applications, including improved films and coatings, stronger and lighter materials, and applications in electronics. Figures 1, 2, and 3 show examples of carbon-based nanomaterial.

Figure 1. Computer image of aC-60 Fullerene. U.S. EPA.

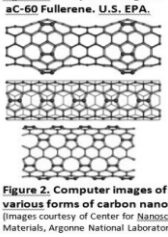


Figure 2. Computer images of various forms of carbon nanotubes. (Images courtesy of Center for Nanoscale Materials, Argonne National Laboratory)

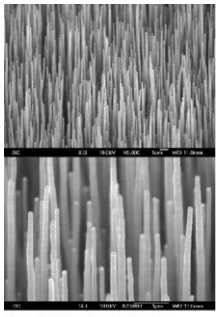


Figure 3. "Forest" of aligned carbon nanotubes.(Image courtesy David Carnahan of NanoLab, Inc.)

(1) Metal-based materials. These nanomaterials include quantum dots, nanogold, nanosilver and metal oxides, such as titanium dioxide. A quantum dot is a closely packed semiconductor crystal comprised of hundreds or thousands of atoms, and whose size is on the order of a few nanometers to a few hundred nanometers. Changing the size of quantum dots changes their optical properties. Figures 4 and 5 show examples of metal-based nanomaterials.

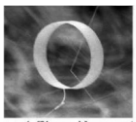


Figure 4. Zinc oxide nanostructure synthesized by a vapor-solid process. (Image courtesy of Prof. Zhong Lin Wang, Georgia Tech)

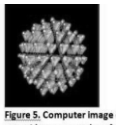


Figure 5. Computer image of a Gallium arsenide quantum dot of 465 atoms. (Image courtesy of Lin-Wang Wang, Lawrence Berkeley National Laboratory)

(1) Dendrimers. These nanomaterials are nanosized polymers built from branched units. The surface of a dendrimer has numerous chain ends, which can be tailored to perform specific chemical functions. This property could also be useful for catalysis. Also, because three-dimensional dendrimers contain interior cavities into which other molecules could be placed, they may be useful for drug delivery. Figure 6 shows an example a dendrimer.

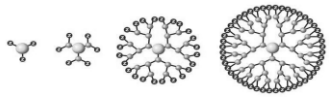


Figure 6.Computer image of generations of a dendrimer. Dendrimers are nanoscale branched polymers that are grown in a stepwise fashion, which allows for precise control of their size. (Image courtesy of Dendritic NanoTechnologies, Inc.)

(1) Composites combine nanoparticles with other nanoparticles or with larger, bulk-type materials. Nanoparticles, such as nanosized clays, are already being added to products ranging from auto parts to packaging materials, to enhance mechanical, thermal, barrier, and flame-retardant properties. Figure 7 shows an example of a composite.

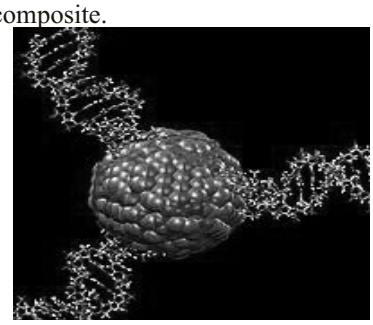


Figure 7.Computer image of a nano-biocomposite. Image of a titanium molecule (center) with DNA strands attached, a bio-inorganic composite. This kind of material has potential for new technologies to treat disease. (Image courtesy of Center for Nano scale Materials, Argonne National Lab)

Table 1. Examples of Products that Use Nanotechnology and Nanomaterials

Health and Fitness	Electronics and Computers	Home and Garden	Food and Beverage	Other
Wound dressing	Computer displays	Paint	Non-stick coatings for pans	Coatings
Pregnancy test	Games	Antimicrobial pillows	Antimicrobial refrigerator	Lubricants
Toothpaste	Computer hardware	Stain resistant cushions	Canola oil	
Golf club				
Tennis Racket				
Skis				
Antibacterial socks				
Waste and stain resistant pants				
Cosmetics				
Air filter				
Sunscreen				

Source: Woodrow Wilson Center Consumer Products Inventory. (http://www.nanotechproject.org/44/consumer-nanotechnology)

Generations of Nanotechnology

Nanotechnology has witnessed four generations till date.

a) **First generation** of nanotechnology,it is called passive nanostructures. Some applications are: Dispersed and contact nanostructures. Example Aerosols, colloids, coatings, nanoparticle reinforced composites, nano structured metals, polymers and ceramics.

b) **Secondgeneration** of nanotechnology, it refers to active nanostructures. Some applications are: Bio- active, health effects, physicochemical active (e.g--- 3D transistors, amplifiers, actuators, adaptive structures).

c) **Thirdgeneration** of nanotechnology,it is called systems of nanosystems. Some applications are: robotics, guided assembling: 3D networking and new hierarchical architectures.

d) **Fourth generation** of nanotechnology, it is called molecular nanosystems. Some applications are: molecular devices by design, atomic design and emerging functions. Fourth generation of nanotechnology basically deals with the manufacturing and development of nano-Computer.

NanotechnologyTechniques/Tools /Materials That Directly Affect Modern Computer

Nanofabrication
Quantum dots
Carbon Nanotubes
DNA computing
NVRAM (non volatile RAM)
NanoDesign (software system)

a) Nanofabrication:

It is a collection of technologies which are utilized in making micro devices. Micro fabrication is the term that describes processes of fabrication of miniature structures, of micrometer sizes and smaller. For instance, fabrication of IC (Integrated circuit).Nanofabrication or micro fabrication technologies originate from the microelectronics industry and the devices is usually made on silicon wafers. Nanofabrication methods can be divided into two categories: a) top down methods and b) bottom up methods

➢ Top down method: It involve carving out or adding a small number of molecules to a surface. This method is generally used by electronics industry in a process called photolithography. Photolithography is the process that transfers the geometric shape on a mask to the surface of a silicon wafer by exposure to UV (ultra violet) light through lenses.

➢ Bottom up method: This method is used to assemble atoms or molecules into nanostructures. In near future, the computer industry will use the above technology extensively to fabricate microprocessor chips. The microprocessor chips would be smaller, faster, reliable, efficient and lighter computers.

b) Quantum dots:

Quantum dots are crystals that emit only one wavelength of light when the electrons are excited. It is a new material made by bottom up method of nanofabrication. In future quantum dots could be used as quantum bits and to form the basis of quantum computers.

i. Working of quantum computers

In quantum computers, the binary rate in conventional computers is repeated by quantum bits or qubits, which can be in a state of 0, 1 and superposition (simultaneously both 0 and 1). As the quantum computer can hold multiple states simultaneously, it is assumed that it has the potential to perform a million computations at the same time. This would make the computer much more faster than before. The development of quantum computer is still under research.

ii. Limitations of quantum computer

Since quantum computers are based on quantum mechanical phenomenon, which are vulnerable to the effects of noise, coherence disappearance and loss of quantum bits. These problems are discussed below.

➢ Problem of coherence disappearance: A quantumcomputer can only function if the information existsfor long enough to be processed. The researchershave discovered that the coherence spontaneouslydisappears over the course of time. This could leadto a considerable problem for the development of a quantum computer.

➢ Simultaneousexistence of two states: In a quantumcomputer a superconducting quantum bits cansimultaneously exist in two states. Normally one ofthe two states disappears as soon as the systemcomes into contact with the outside world. Thecoherence then disappears as a result of thedecoherence process and the information in aquantum bit is lost.

iii. Solution to the above problem

More research needed. There is a need toclarify the issue that molecular dynamics simulationscarried out at finite temperatures of machinesof somedegree of complexity, in which both the mechanismitself and its mounting are subject to thermal noise.

c) Carbon nanotubes:

It is a tube shaped carbon material that ismeasured in nanometre scales. With the advancementof nanofabrication technique, researchers used thismaterial to create electronic components like transistors,diodes, relays and logic gates. These electroniccomponents can be directly applied in makingadvanced computer.

d) DNA computing:

It is an approach to nanocomputers. DNAcomputing uses bottom up approach or method tomake DNA molecules and DNA logic gates.

➢ Major Events:

In 1994, L. Adleman has tried to solve a complex travelling salesman problem by using DNA computing technique.

In 1997, researchers at the University of Rochester built DNA logic gates. This development is considered as a step towards a DNA computer.

Researchers have found that a DNA molecule can store more information than any conventional memory chip and DNA can be used to perform parallel computations.

The abovedevelopments make the idea of DNA computing very appealing to the current researchers and scientists of the world.

e) NVRAM (non volatile RAM):

Argonne research has developed a NVRAM (non volatile RAM) made up of tiny nano engineered ferro electric crystals. Since the tiny nano engineered ferroelectric crystals do not revert spontaneously, RAM made with them would not be erased should there be a power failure. Using NVRAM laptop computers would no longer need back up batteries, permitting them to bemade still smaller and lighter. This achievement of nano technology is considered as a long –standing dream of the computer industry.

f) Nanodesign (software system)

A research group at NASA has been developing a software system called Nano Design, for investigating fullerene nano technology and designing molecularmachines. The software architecture of Nanodesign is designed to support and enable their group to develop complex simulated molecular machines. The main purpose behind developing this software system is design and simulation of materials based onnano technology.

Advantages : Nanotechnology is helping to considerably improve, even revolutionize, manytechnology and industry sectors: information technology, energy, environmental science, medicine, homeland security, food safety, and transportation, among manyothers. Nanoscale transistors that are faster, more powerful, and increasingly energy efficient; soon your computer's entire memory may be stored on a single tiny chip.

Magnetic random access memory (MRAM) enabled by nanometer-scale magnetic tunnel junctions that can quickly and effectively save even encrypted data during asystem shutdown or crash, enable resume-play features, and gather vehicle accident data.

Displays for many new TVs, laptop computers, cell phones, digital cameras, andother devices incorporate nano structured polymer films known as organic light emitting diodes, or OLEDs. OLED screens offer brighter images in a flat format, as well as wider viewing angles, lighter weight, better picture density, lower power consumption, and longer lifetimes.

Other computing and electronic products include Flash memory chips for iPodnanos; ultra responsive hearing aids; antimicrobial/antibacterial coatings on mouse/keyboard/cell phone casings; conductive inks for printed electronics for RFID/smart cards/smart packaging; more life-like video games; and flexible displaysfor e-book readers.

Nanotechnology is improving the efficiency of fuel production from normal andlow-grade raw petroleum materials through better catalysis, as well as fuel consumption efficiency in vehicles and power plants.

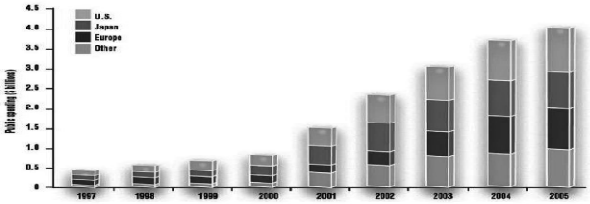
Researchers are developing wires containing carbon nano tubes to have much lower resistance than the high-tension wires currently used in the electric grid and thus reduce transmission power loss.

▪ Nanotechnology manufacturing has a promise of producing new materials ahundred times stronger than steel, and more efficient and cheaper to produceas compared to the existing

production techniques.

- Molecular manufacturing would greatly reduce water requirements, and also cheaply run greenhouses would be a means of saving water, land, and food.
- The efficient and inexpensive generation of electricity, using solar and thermal power, will make electric power available to basically everyone in the world.
- Faster, cheaper, and more powerful computers will be available that could help improve information and communication systems even in the remote stareas.
- Manufacturing of new technologies will be self-contained and clean, and will have less of an environmental impact.[5]
- Cheap and advanced equipment for medical research and health care will make improved medicine widely available. It will be feasible to restore humanorgan engineered tissue while simple products will greatly reduce infectious diseases prevailing in many parts of the world.
- Nanotechnology will enhance capabilities in space ventures and operations.[10]

Nanotechnology in Computers



2. Worldwide spending on nanotechnology tallied more than $4 billion last year. (courtesy of the National Science Foundation and Strategies Unlimited)

Literature survey

One of the most fundamental components in the manufacture of electronic devices,such as a CPU or memory, is a switch. Computers are constructed from thousands tomillions of switches connected together. In modern computers, components called transistors act as electronic switches. Transistors act as electronic switches, i. e. theyallow information to pass or not to pass under certain conditions. The development ofintegrated circuits (ICs) allowed the construction of a number of transistors on asingle piece of silicon (the material out of which IC's are made). IC's are also calledsilicon chips or simply chips. According to Johannes Swenson silicon chips are being used in computers. [6]

A silicon chip is an almost pure piece of silicon, usually less than one centimetersquare and about half a millimeter thick. It contains hundreds of thousands of microminiature electronic circuit components, mainly transistors, packed andinterconnected in layers beneath the surface. These components can perform control, logic, and/or memory functions.

The reason why silicon is used in computer chips is because it is easier, andconsequently less costly, to make complex circuits out of silicon than from any othermatter. With silicon, it is easy to make a high-quality insulator by adding someoxygen to create silicon oxide. Computer chips require precise regulation of voltageto manipulate data. According to the Cornell Center for Materials Research, silicon isideal for this because it can be made into either an effective insulator or asemiconductor, both essential for controlling electrical current. It is also one of the cheapest materials with this ability. [6]

But we think that instead of silicon chips we can use carbon nano tubes.

Silicon is subject to certain limitations, and industry is looking for a replacement.we don't think that it will be cheaper to build transistors from another material thansilicon, but carbon nano tubes can be used to produce smaller and faster components.This will also result in computers that consume less energy.

The most common semiconductor material in transistors is silicon, since it ischeap and easy to process. But silicon has its limitations. As the size of the transistorsis reduced in order to increase their speed, problems arise that lead to, among otherthings, increased energy consumption and large variation in the transistor properties.

By exchanging the silicon in the channel for a carbon nanotube, the transistors canbe made both smaller and faster than today's transistors. A carbon nanotube is amolecule in form of a hollow cylinder with a diameter of around a nanometer (roughly 1/50,000 of the width of a human hair) which consists of pure carbon. Some carbon nanotubes are semiconducting, and this means that they can be used intransistors, although there are several problems that must be solved before they can beconnected together to form large circuits. [7]

Components made with carbon nanotubes could endure greater heat thenconventional metal components, allowing computers to run hotter and reducing the pressure on the cooling systems. Graphite can be rolled into a cylinder with a diameterof about 1 nm. These strong but light 'carbon nanotubes' are being developed for araft of uses, such as sensors, fuel cells, computers and televisions.

The applications of nanotubes are set to expand even further now that scientistshave found that other materials besides carbon can form nanotubes.

Carbon nanotubes will be proving useful as it:

- Improves conductive, mechanical, and flame barrier properties of plastics andcomposites
- Optimizes processing fabrication, and reduces shipping costs.
- Enables eco-friendly anti-fouling paints, and other new applications.
- Enables clean, bulk micromachining and assembly of electronic components.
- Improves the true total cost of formulation, processing, and manufacturing.

CONCLUSION :

As the development of nanotechnology progresses in several fields including physics,chemistry, biology and material science, computer scientists, medical, military mustbe aware of their roles and brace themselves for the greater advancement of nanotechnology in the future.This paper is intended todescribe the role of nanotechnology in the developmentof a sophisticated small computer. Also, the paper isintended to describe the dependency of particularsection or field of nanotechnology which are directlyrelated to the development of an advance computer infuture.

REFERENCES :

[1] K. E. Drexler, 1986. Engines of Creation: the coming era of nanotechnology. Anchor Press.[2] K. E. Drexler, C. Peterson and G. Pergamit, 1991.Unbounding the Future: the Nanotechnology
Revolution.
[3]G. M. Whitesides and J. C. Love, 2001."The Art of Building Small". Scientific American, September.
[4] K.E. Drexler, 2001, "Machine phase nanotechnology", Scientific American, September
[5]http://www.nanowerk.com/news2/newsid=32466. Php
[6]http://www.sciencemediacentre.org/nanotechnology. html.
[7]http://www.intel.com/research/silicon/mooreslaw. Html
[8] Introduction to Microfabrication (2004) by S.Fanssila. ISBN 0-470-85106-6
[9] Fundamentals of Microfabrication (2nd ed,2002) by M.Madou ISBN 0-8493-0826-7
[10] http://leboaplac2008. wikispaces. com/Nano+%26+Virus+Technology
[11] K.E. Drexler, 1986. Engines of creation: the coming era of nanotechnology. Anchor Press

SIMULATION OF SAW DELAY LINE SENSOR USING MATLAB

P. S. VARADE, S. A. GANGAL, A. D. SHALIGRAM
Department of Electronic Science, Savitribai Phule Pune University,Pune (Maharashtra).
Email: psvarade1@gmail.com, adshaligram@gmail.com

Abstract: Surface acoustic wave (SAW) devices are used as sensors for different applications such as gas sensor, pressure sensor, vibration sensor etc .The first order model of SAW delay line sensor has been created using matlab. The model implements impulse response method to calculate radiation conductance, acoustic susceptance, frequency response, insertion loss. This paper presents the results of the model and results from the model for SAW delay line design.

Keywords: Surface acoustic wave, delay line, impulse response.

INTRODUCTION:

SAW devices are capable of detecting various physical quantities such as pressure, strain, stress, chemical species, mass loading, acceleration etc. The attractive features of SAW technology are: low power operation, light weight, ruggedness, low cost, and wireless operation is possible. SAW delay line consist of two sets of inter digitated transducer (IDT) made of suitable metal deposited on piezoelectric substrate as shown in figure 1(a). This configuration can be used as sensor by depositing sensing material between the two IDTs. During detection molecules of the input quantity get adsorbed on the sensing material, which results in mass loading. This causes change in phase velocity of surface acoustic waves and ultimately change in frequency. Simulation of these devices is necessary for predicting their behavior under different input conditions. Simulation of these sensors can be done using suitable software and models such as COMSOL multiphysics , Matlab, Xilinx. We faced the problem 'out of memory' using finite element modeling for above design using COMSOL multiphysics. So, we used here matlab for simulation of the same design using impulse response model. The same is also verified using hand calculations in frequency domain.

2. Model Implementation

The Impulse Response method developed by Hartmann, Bell and Rosenfeld [2] was used to model the SAW delay line. This is a first order model that does not take into account any second order effects such as reflections, however, it does model the piezoelectric, mechanical and electrical behaviors of the SAW device. This model is only valid for transducers where the IDTs are un-weighted (unapodized). The model assumes that the finger overlap or aperture is constant and that the metallization ratio between the fingers and spaces is 0.5. This model calculates the frequency response, conductance, susceptance, impedance, and other electrical parameters. The Impulse Response method uses the Mason equivalent circuit shown in Figure 1(b). In the circuit Ga(f) is the radiation conductance, Ba(f) is the acoustic susceptance, and C_T is the total capacitance.

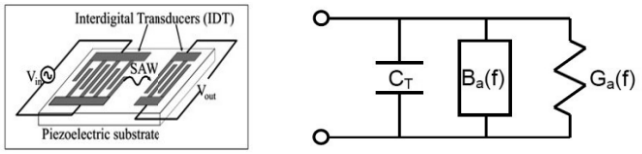


Figure1(a) Basic SAW delay line device (b) IDT equivalent circuit.

From the Impulse Response model one can calculate wavelength (λ), and the number of finger pairs (Np) using the following equations:

$$\lambda = v/fo \quad (1)$$

where ν is the acoustic velocity in the media, fo is the center or synchronous frequency, and NBW is the null bandwidth or fractional frequency.

2.1 Frequency Response

The frequency response of a SAW system can be implemented using the impulse response method. For a single IDT, the frequency response is shaped by the sinc function, and is calculated by the following equation [2]

$$|H(f)| = 2k\sqrt{(Cs.fo)}\ N_p\ (\sin(X)/X) \quad (2)$$

Where f is the frequency, k is the piezoelectric coupling coefficient; Cs is the capacitance for a finger pair per unit length. The variable x is used to simplify the equation and is defined as [2]:

$$X = N_p \pi (f-fo)/fo \quad (3)$$

When two IDTs are used as in the case of a SAW delay line, the frequency response for the system is found by combining the frequency response for each IDT [2]:

$$H_T(f) \cong H_1(f).\ H_2(f) \quad (4)$$

Generally the frequency response is normalized using the log equation:

$$H_n(f) = 20\log(|H_1(f).\ H_2(f) / H_1(fo).\ H_2(fo)|) \quad (5)$$

2.2 Radiation Conductance

The real part of the input admittance is called the radiation conductance. The radiation conductance is also shaped by the sinc function [2] and is found by

$$G_a(f) = 8k^2 C_s W_a fo N_p^2 |\sin(X)/X|^2 \quad (6)$$

Where Wa is the aperture or overlap height of the fingers. The radiation conductance is normalized by using the following equation:

$$G_n(f) = G_a(f) / G_a(fo) \quad (7)$$

2.3 Acoustic Susceptance

The third element of the model is the imaginary part of the input admittance which is also called the acoustic susceptance. The acoustic susceptance is the acoustic wave phenomena modeled as an electrical parameter. The acoustic susceptance is found by taking the Hilbert transform of the radiation conductance [2] and is given by:

$$Ba(f) = G_a(fo)\sin(2X) - 2X/2X^2 \quad (8)$$

The acoustic susceptance is normalized by using the following equation:

$$B_n(f) = Ba(f) / G_a(fo) \quad (9)$$

Notice that the acoustic susceptance is normalized using the radiation conductance since the acoustic susceptance at the synchronous frequency is zero.

2.4 Admittance and Impedance

The total admittance [3] is found by combining the radiation conductance, the acoustic susceptance and the total capacitance. The total admittance is given by

$$Y = G_a + j(2\pi f C_T + B_a(f)) \quad (10)$$

The total static capacitance (C_T) for the IDT is found by multiplying the capacitance per unit length for a pair of fingers (Cs) times the finger overlap or aperture (Wa) times the number of finger pairs (Np).

$$C_T = N_p C_s W_a \quad (11)$$

If equation (10) is inverted it will yield the impedance of the system [3]

$$Z(f) = 1/(G_a + j(2\pi f C_T + B_a(f))) \quad (12)$$

2.5 Insertion Loss

It is very useful to calculate the insertion loss of electrical systems. For SAW devices the insertion loss is a function of frequency, and is calculated using conductance, susceptance, and the load resistance Rg [4].

$$IL(f) = -10 \log [\, 2 G_a(f) R_g / (1 + G_a(f) R_g)^2 + [R_g 2\pi f C_T + B_a(f))]^2 \,] \quad (13)$$

The minimum insertion loss occurs when f = f0 the synchronous frequency.

2.6 Aperture Optimization

An optimal design must match the IDT resistance (real impedance) to the input resistance. The device aperture (Wa) is often adjusted so that the IDT design achieves the correct IDT resistance. The following equation was used to optimize the aperture in terms of the input resistance:

$$W_a = 1/R_{in}(\ 1/2 fo N_p C_s)(4k^2 N_p)/(4k^2 N_p)^2 + \pi^2 \quad (14)$$

where R_{in} is the input resistance.

2.7 Matching Network

If impedance matching is required, then a series inductor can be used to cancel out the static capacitance at the synchronous frequency. To calculate the series inductor value use the following [5].

$$\text{Inductor} = 1/(2\pi fo)^2 C_T \quad (15)$$

3. SAW Delay Line Design

SAW delay line that consist of two identical IDTs will be used to demonstrate the model. The synchronous frequency is 9.98 MHz . The substrate used is 128 degree YX-cut .LiNbO3.The phase velocity is 3992 m/s, Cs=5.6 pF/cm, K^2=4.8, The input .load resistance assumed is 50 Ohm, wavelength =400μm, finger width= finger gap=100 μm, Aperture length =100 *wavelength, Number of finger pairs = 10.

4. Results and Analysis

From the plots presented in figures 2 gives the values of acoustic susceptance (b(f)), normalized frequency response(h1(f)), radiation conductance(gn(f)),insertion loss, (il(f)).The radiation conductance and insertion loss are to be maximum & minimum respectively at synchronous frequency, the graph shows the same result. The bandwidth of designed device is 2MHz.

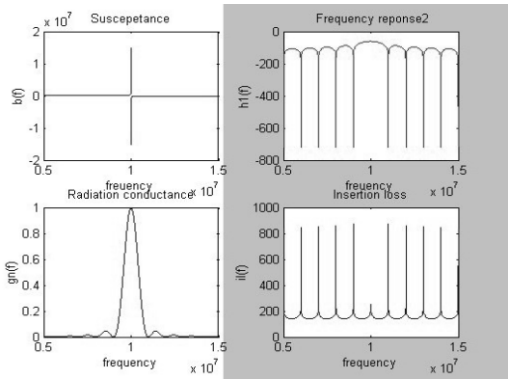


Figure 2: Results of simulation showing acoustic susceptance, frequency response , radiation conductance, insertion loss

CONCLUSION :

SAW delay line design can be implemented using Matlab. Other software's such as COMSOL multiphysics , Coventorware, VHDL are also used to implement the design ,but time required for analysis and memory required is more.

REFERENCES :

1. W. C. Wilson,G.M. Atkinson, 1st Order Modeling of a SAW Delay Line using MathCAD.

2. Clinton S.Hartmann, Jr., Bell D.T., Rosenfeld R.C., Impulse Model Design of Acoustic Surface-Wave Filters, *IEEE transactions on Microwave Theory and Techniques,* IEEE Transactions on,Volume 21, Issue 4, Apr 1973, pp. 162 – 175.

3. K. M. Lakin. T. Joseph, D. Penunuri, Planar SurfaceAcoustic Wave Resonators. *Ultrasonics Symposium,Proceedings*, IEEE, Volume 1, Milwaukee, Wisconsin, Nov.11-14, 1974, pp. 263 – 267.

4. Supriyo Datta, *Surface-Acoustic Wave Devices* (PrenticeHall, Englewood Cliffs, NJ, 1986).

5. V.M. Ristic, *Principles of Acoustic Devices*(New York,Wiley, 1983).

MY HEALTH PORTAL AND IMPLEMENTATION OF HEALTH CARD

ASHWINI KULKARNI
Assistant Professor,
Birla College of Science, Arts & Commerce
ashwini.k@gmail.com

ABSTRACT : Current concerns in medical and health care may be steered along several axes of classifications. One could perhaps consider the running around in case of medical emergencies, registrations with hospitals in case of surgeries, treatments, admissions, shifting hospitals, documentation, maintaining medical files and data of patients is absolutely cumbersome, not to mention the financing of medical care, examining questions of tax support, insurances, and many a times philanthropy and private purchase of services.

A single point of registration for patients and/or individuals; their entire medical history and data, be stored on central data server. Secondly it will be Geo distributed as per residence and location preferred for convenience. Further, pharmacies, hospitals, banks, Insurance can pitch in together on the portal and ensure best & fastest of medical help and enhancements.To quantify portal usage and explore potential differences in adoption and use according to patients' socioeconomic and clinical characteristics in a network of federally qualified health centres serving different countries.

Electronic patient portals offer patients access to information from their electronic health records and the ability to message their corresponding providers. These tools are becoming more widely used and are expected to promote patient engagement with health care.Its a one stop portal signifies the importance to patients ,doctors,Insurance companies. Retrospective analysis of data from portal and electronic health records

KEYWORDS : Health card | Digital Records | medical insurance | Medicine Stores | Labs | Reports

INTRODUCTION:

In accordance with the concept of "Digital India" initiated by honourable prime minister Mr. Narendra Modi; I take this opportunity to focus on the most important and inevitable part of human life, Health Care.

In today's era of Digitisation, automation, centralisation and cashless economy, it is imperative that citizens MUST get the facility of his/her Health Care digitally, keeping latest technologies, cost effectiveness and most importantly REACH. Widening our horizons, I visualised a concept of creating a one stop web portal, which essentially addresses below concerns from a health care initiative .

Our country is one of the most populous countries in the world and HEALTH has always been a most priority moment for any Government initiatives and funding per se, hence keeping abreast with the trends and patterns followed and implemented world wide, let us join and come together and bring EVERY citizen's health data embedded in ONE health portal, linked with the health card**.**

Problem Definition

Health Management system is essentially one of the primary areas of concern in our country. Awareness in overall Health care in this country is not developed and communication technology has not been introduced significantly to improve its quality. Ever increasing usage of information and communication technology (ICT) facilitates many countries to develop their ICT based e-health card system centrally.

The aim of MyHealthPortal is to improve efficiency, access and accountability of health-care services. The portal answers various avenues like research, design and implementation of e-Health card and web portal integration based solution that can be used to plug and bridge various web platform and infrastructure.

Using the e-Health card, patient's entire medical data, treatments, doctor's prescription, allergy, preferences details, family histories, past medical cases, patients present and previous health history is made accessible from ANY geographic location, without any sophisticated hardware limitations and higher bandwidth demands, this makes it a perfect timely MOVEMENT rather than just a project .

CASE STUDIES AND PERSONAL EXPERIENCES

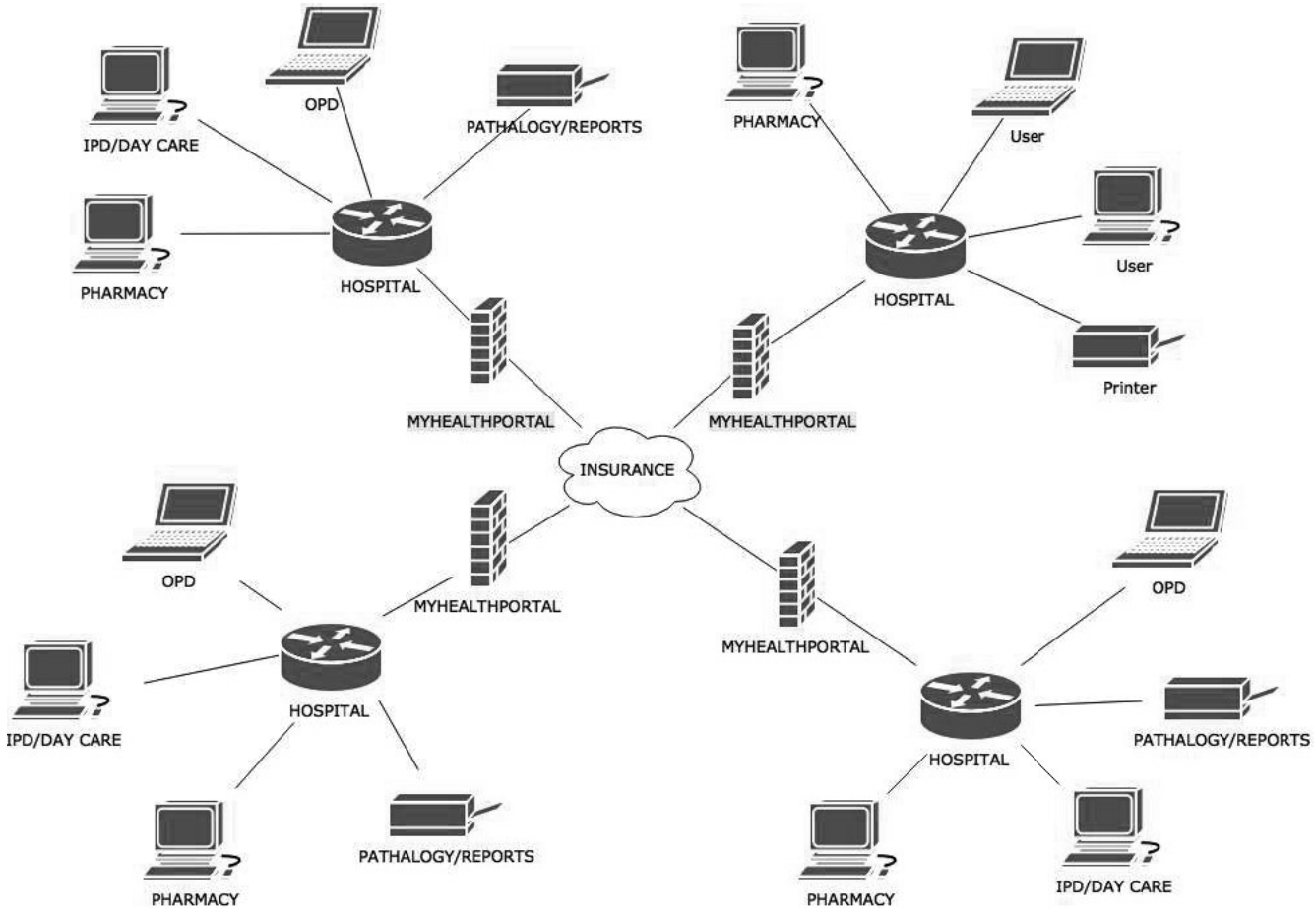
While I visited couple of hospitals, clinics and doctors in general, I found out that there is a huge gap in understanding patient to doctor psychology. Every now and then the patient is required to maintain and bring his/her medical file supplied by corresponding doctors/hospitals/labs/pharmacists et all. And eventually the patient ends up flooding their book racks with tens of files having absolutely no clues which file relates what? Besides the doctor too in times of medical situation need to refer to the tons of papers the patient dumps on doctors desk

Alternately I also took a peep on some of the health management systems, white papers of leading chain of hospitals on how they typically handle patient data. The main observation is a 180 degree difference in branded hospital chains and 70-80 bed stand alone hospitals

Most of the leading hospitals now have STOPPED using patient papers for IPD patients, the nurses now walk across with their sophisticated iPads/android tabs and by some clicks they access the entire patient data as per the role rights given by the head nurse and so on and so forth.

System design

Block Diagram of System Architecture



PROPOSED METHODOLOGY

- Selection of data from different areas like Doctors,Patients to collect their history.
- Retrospective analysis of data from portal and electronic health records Estimation
- Different recent methodology of Data Mining.
- electronic applications, typically web-based provided and maintained by health care institutions
- targeted towards providing functionality to all or a subgroup of patients
- basic functionalities to access (a subset of) a patient's clinical data
- optional, additional functionalities such as communication modules, prescription refills, appointment scheduling, or educational guidelines

CONCLUSION :

We may good early rates of adoption and use of an electronic patient portal during the this era of digitisation and its deployment among a predominantly low-income population, especially among patients with chronic diseases. Disparities in access to and usage of the portal will be evident. Continued efforts will be needed to ensure that portals are usable for and used by disadvantaged groups so that all patients benefit equally from these technologies.This will be the One stop portal in Medical World.

FUTURE WORK

The overall success of the concept is to generate massive awareness to corporatise medical sector, below is the envisaged future work where if likeminded professionals/Government/Private/Public sector and most importantly EVERY citizen of our great country pitches in, THIS would certainly become a movement rather than just a technical/innovative project for **DIGITAL INDIA**

- Distribution & Propagation of HEALTH CARD across length and breadth of country
- Usage of MyHealthPortal extensively by professionals and medical houses at large
- Data mining at mass level
- Free flow of funds to ensure that project does not fade
- Making it work in both ONLINE and OFFLINE mode seamlessly

REFERENCES :

1. Current problems in the Medical Care ROEMER MI.Public Health Rep. 1964 Aug; 79:720-2. PMID14201863
2. James A. Swartz, Ian Jantz. (2014) Association Between Nonspecific Severe Psychological Distress as an Indicator of Serious Mental Illness and Increasing Levels of Medical Multi morbidity. American Journal of Public Health 104:12, 2350-2358. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc. 2008;15(6):729–36.
3. Nazi KM, Hogan TP, Wagner TH, McInnes DK, Smith BM, Haggstrom D, et al.Embracing a health services research perspective on personal health records: lessons learned from the VA My HealtheVet System. J Gen Intern Med. 2009;25 (Suppl 1):62–7.
4. The Impact of Electronic Patient Portals on Patient Care: A Systematic Review of Controlled Trials Monitoring Editor: **Gunther Eysenbach**

Reviewed by Anne Moen, Petra Knaup-Gregori, Ursula Hübner, Bradley Crotty, and Daniel Firth, Elske Ammenwerth, PhD,[1] Petra Schnell-Inderst, PhD,[2,3] and Alexander Hoerbst, PhD[4]

CASE STUDIES

• Dr. Avadhut Shetye
Practising Physiotherapy/Acupuncture and Acupressure, Kalyan - East

• MAHER - Dr. Mhaskar Hospital
Dr. Vikas Mhaskar
Leading Gynecology and Maternity hospital, Kalyan(W)

•Mr. Suyog Kulkarni
MD - White Oak Interactive Pvt Ltd on Health Management System & Innovation

Mr. Mohammad Basher | Mr. Palash Roy
Introducing eHealthcard for developing countries, Bangladesh

STATISTICAL ANALYSIS OF CHANGE RATE OF SATARA DISTRICT'S POPULATION GROWTH

DR. MANISHA PATIL
Department of R & D Engineering,
NICE Systems Pvt. Ltd., Pune, India
manishap6@gmail.com

Abstract : Population is one of the biggest concerns of each county. India is also well known for high populated country. And India is nothing but integration of different states and in turn different cities. Metropolis cities like Mumbai, Delhi, Hyderabad, Chennai, Pune etc. these are well known cities where population growth rate seems high and the reason behind this high population growth rate is large scale of migration of people from villages to towns cities in search of employment opportunities , better education and good quality of life. But due to this migration of people from small town to metropolis cities, there is vast growth in the population of such metropolis. Along with big metropolis, small districts like Satara also facing the same issue. The main purpose of this research paper is throw lights on how the population of Satara city is increased over a decade and what the basic reasons behind that are.

Keywords: Growth, Population, Change, Period, Percentage

INTRODUCTION:

India is one of the largest country in the world. Huge volume of population made India as second largest country in the world. India's population made as an integration of populations of different states present in the India.

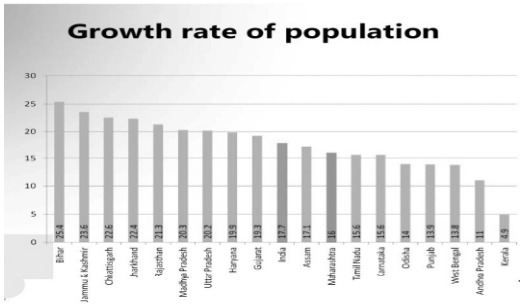


Figure 1: Population growth rate of different states of India[1]

From the above graph its seen that Maharashtra is one of the state which has population rate is 16. Following graph shows that population growth of Maharashtra state from 1995 to 2011. Till 2011, total population of Maharashtra was 11,23,74,333.

1. Analysis of Population of Maharashtra State

Maharashtra State has recorded 11,23,74,333 population against 9,68,78,627 in 2001, i.e., an addition of 1,54,95,706 during the decade 2001-11.Refer graph 2. This makes 15.99 percent growth rate during 2001-11. The same was 22.73 during 1991-01. This shows a reduction in growth rate by 6.74 percentage points. The total density is 365 persons per sq. Km in Maharashtra as against 315 in 2001.

From the senses of India 2011, following are the important factors published about Maharashtra State:

- There are 35 Districts in Maharashtra.
- There are 355 Tehsil in Maharashtra.
- There are 534 Towns (including 278 census towns) in Maharashtra.
- There are 35 Districts in Maharashtra.
- There are 43,665 total villages in Maharashtra out of which 40,960 villages are in-habited.

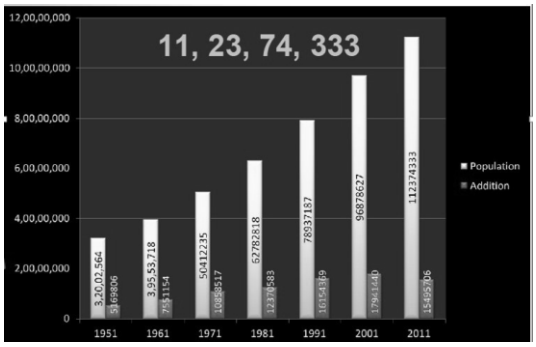


Figure 2: Population growth rate of Maharashtra State[1]

1.Analysis of Population of Different District of Maharashtra

As we seen above that there are 35 districts in Maharashtra, Satara is one of the famous district. Satara district belongs to Desh or Paschim Maharashtra Region. Satara District Administrative head quarter is Satara. It is Located 223 KM North towards State capital Mumbai. Satara District population is 3003922. It is 12th Largest District in the State by population. Refer Graph 3.

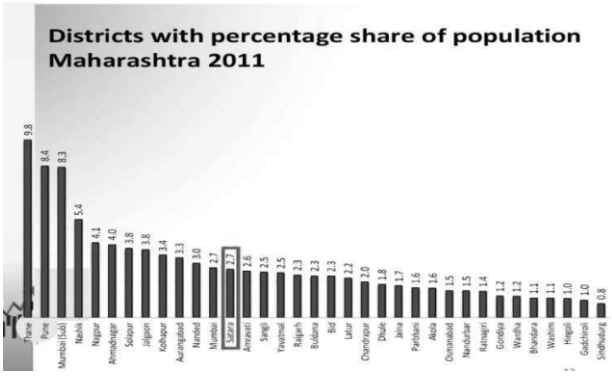


Figure 3: Population growth rate of Different Districts of Maharashtra State[1]

1. Geographical Location of Satara District

As we saw that Satara is one of the famous district in Maharashtra, India country. The geographical location of Satara is : It is Located at Latitude-17.6, Longitude-73.9. Satara District is sharing border with Pune District to the North , Raigad District to the west , Ratnagiri District to the South , Sangli District to the South . Satara District occupies an area of approximately 10484 square kilometres. . It is in the 99 meters to 550 meters elevation range.This District belongs to Western India.

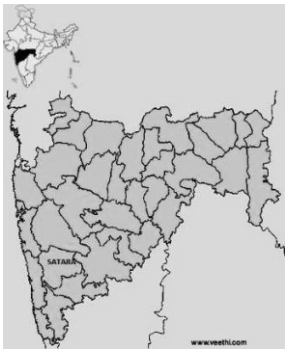


Figure 4 : Geography location of Satara in Maharashtra and in India [2]

1. Census 2011 of Satara District

According to census 2011 total population of Satara district is 3003922. Males are 1512549 and Females are 1491373 .Literate people are 1987395 among total. Its total area is 10484 kmÂ². It is the 12th largest district in the state by Population. But 11th Largest District in the state By Area.

121st Largest District in the Country By Population. 11th highest District in the State By literacy rate and 94th highest District in the Country By literacy rate. Its literacy Rate is 84.2

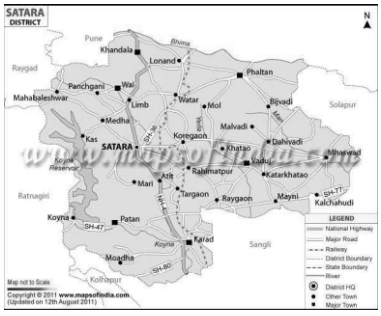


Figure 5: Talukas of Satara District (Source [2])

1. Analysis of Taluka Wise Population of Satara District (2001)

There are around total 11 talukas there in Satara districts and as per the Censes of 2001. The name of Talukas are: Mahableshwer, Wai, Khandala, Phaltan, Man, Khatav, Koregaon, Satara, Jaoli, Patan, Karad etc.

Figure 6: Talukas of Satara District .[2]

Sr. No.	Name of Taluka	Total Population	Male	Female
1	Mahableshwar	54546	7043	5694
2	Wai	189336	15894	15216
3	Khandala	119819	61047	57646
4	Phaltan	313627	25935	24865
5	Man	199598	51719	43338
6	Khatav	260951	76899	78200
7	Koregaon	253128	107596	107413
8	Satara	451870	55938	52110
9	Jaoli	124600	50473	52737
10	Patan	298095	5977	5621
11	Karad	543424	28198	27243

Table 1. Population data with respective Male, Female and No. of Taluka[3,4]

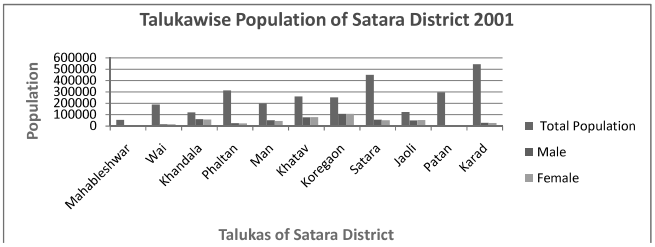


Figure 7: Chart for Talukas wise population of Satara District for the year 2001

1. Analysis of Taluka Wise Population of Satara District (2011)

As we saw there are around total 11 talukas there in Satara districts and as per the Censes of 2001 and 2011. This section provides detail analysis for the talukawise population of Satara district as per the census of 2011. Table 1.1 contains the total population for each taluka and population for Male, Female and Urban area. From the population data, it seems that Karad is the biggest Taluka in Satara District. Hence in Karad, total Male population is 230035 and Female population is 224012. Satara is also second most largest Taluka of Satara district in terms of Population. Hence in Satara, total Male population is 145145 and Female population is 144536.

Sr. No.	Name of Taluka	Total Population	Male	Female
1	Mahableshwar	72840	22354	22187
2	Wai	200741	81631	82266
3	Khandal	137450	61047	57646
4	Phaltan	342696	145881	136699
5	Man	225193	101719	99338
6	Khatav	275099	136899	138200
7	Koregaon	257327	107596	107413
8	Satara	501670	145145	144536
9	Jaoli	107890	50473	52737
10	Patan	299634	138102	147759
11	Karad	583360	230035	224012

Table 2. Population data with respective Male, Female and No. of Taluka

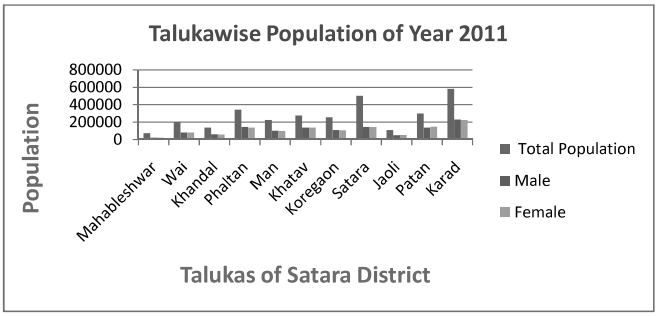


Figure 8: Chart for Talukas wise population of Satara District for the year 2011

From the table 1.1., it seems that Phaltan Taluka is on the 3rd rank in the terms of population of Satara district. So in Phaltan, total Male population is 145881 and Female population is 136699. After Phaltan, Patan comes in line for population growth. Hence in Patan, total Male population is 138102 and Female population is 147759. Khatav is also on the 6th rank for population growth. In Khatav, total Male population is 136899 and Female population is 138200. Like wise, Koregaon, Wai, Man etc. are the number of talukas in Satara district which are having between 2 to 2.5L populations. Mahableshwer is the smallest taluka in Satara district in term of population growth. Hence in Mahableshwer , Total Male population is 81631 and Female population is 82266.

1. Change in population growth from 2001 to 2011

As we saw that population statistics of Satara District for the year 2001 and 2011, there is need to understand what is the changing rate of population growth of Satara District between these years.

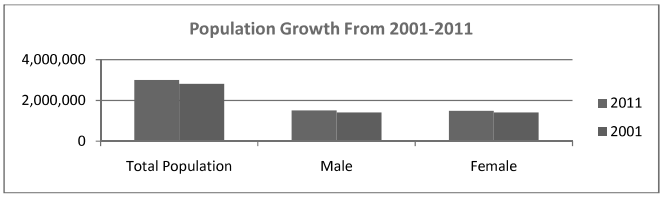


Figure 9: Chart for Total, Male and Female Population of Satara District for the years 2001- 2011

From the Figure 9, it has been seen that total population of Satara District has increased from 2.6L to 3.1 between this decades. Even Male population is also been increased and reached up to 1.5L and Female population is increased up to around 1.5L

1. Analysis of Taluka Wise Population Data between year 2001-2011

Change in population volume of Satara district based on their Talukas, from the figure 10, seems that there is also increasing rate in population growth. Mahabaleshwar has been increase around 18 K population through out this decade. Wai has been also increased around 20K but this growth is quite minimum as compare to the growth rate of Population. As we know that Mahabaleshwar is the famous place for tourism and hence this can be the reason migration of people has been done from small town, villages to Mahabaleshwar just to get better jobs and bread butter. Satara also biggest taluka of Satara district and it has been also increased its population drastically and again the reason behind this is migration of people to get better amenities for their life. Like wise Phaltan and Karad are the two biggest and commercial Talukas of Satara district and population growth rate of these two talukas are also reached quite high. Its around 40-45K.After these talukas, number comes of Patan and Khatav and then koregaon, Man and Khandala. From the Figure 10 it seems that overall population of Satara district gradually increased in a constant way.

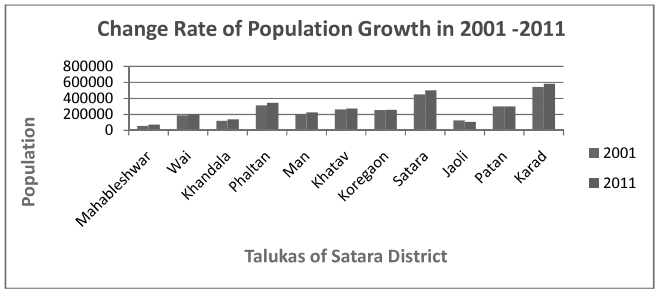


Figure 10: Chart for Total, Male and Female Population of Satara District for the years 2001- 2011

1. Analysis of Male Population Data between years 2001-2011

Change in population volume of Satara district based on their Male criteria, from the figure 11, it seems that there is drastic high increase in Male population in Karad taluka. Even in Patan also there is drastic change recognized in the Male population between year 2001 and 2011. Along with these Talukas, same observations are there for the Talukas like Satara, Wai, Phaltan and Khatav. Khandala, Jogali has the same rate of population growth in term of Male category.

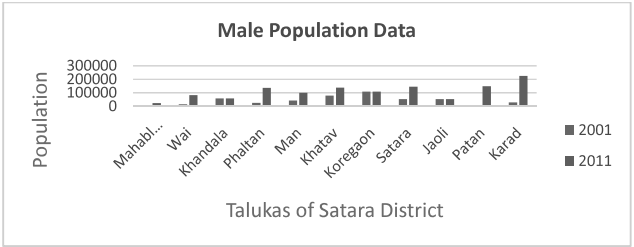


Figure 11: Chart for Total, Male and Female Population of Satara District for the years 2001- 2011

1. Analysis of Female Population Data between years 2001-2011

Change in population volume of Satara district based on their Female criteria, from the figure 12, it seems that there is drastic high increase in Female population in Karad taluka. Even in Patan also there is drastic change recognized in the Male population between year 2001 and 2011. Along with these Talukas, same observations are there for the Talukas like Satara, Wai, Phaltan and Khatav. Khandala, Jogali has the same rate of population growth in term of Male category.

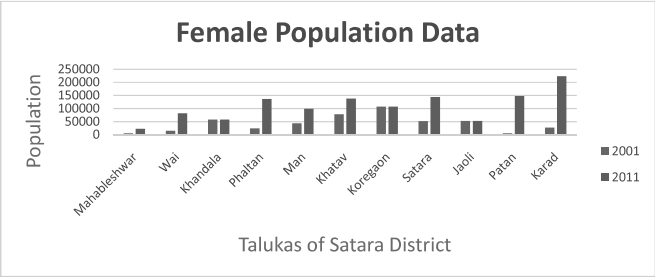


Figure 12: Chart for Total, Male and Female Population of Satara District for the years 2001- 2011

CONCLUSION :

India is well known country for volume of population, and it contains many state. Maharashtra is one of the biggest state of India and in turn Satara is well known state of Maharashtra. This papers shows that how the population growth of Satara seems in 2001 and 2011 and based on analyzed there is gradual change in population growth of Satara district between the year 2001 and 2011. The reason behind this population growth would be migration of people from small towns, village to get better jobs, education, aminities and better quality of life.

REFERENCES :

1. District Disaster Management Plan – Satara DDMP – Satara, http://satara.nic.i n/images/DMP%2 0satara.pdf

2.. http://www.mapsofindia. com/maps/maharashtra /tehsil/satara.html

3. BARAKADE A.J., TODKARI G.U., "A STUDY OF URBANIZATION IN SATARA DISTRICT OF MAHARASHTRA", Geoscience Research ISSN: 0976–9846 & E-ISSN: 0976–9854, Vol. 2, Issue 1, 2011, pp-38-44

4. http://www.slideshare.net/priteeg/taluka-population-2868728

5.BARAKADE A.J.1 AND TODKARI G.U."A study of urbanization in Satara district of Maharashtra" Geoscience Research, ISSN: 0976–9846 & E-ISSN: 0976–9854, Vol. 2, Issue 1, 2011, pp-38-44

A STUDY ON E- LEARNING ENVIRONMENT AND DIVERSION IN REPUTED ICT CITIES WITH SPECIAL REFERENCE TO PUNE CITY

PROF. MAHESH DESHPANDE
Vice Principal,
CCS-Wakad, Pune -57
maheshd.54@gmail.com

ABSTRACT : *Today's world is the system of Information, Communication and Technology (ICT). Everything is revolving around ICT. Education system is not an exception. Indian education system involves modern technologies for students to improve their knowledge. MOOC courses like Swayam is one of the Indian asynchronous courses of learning modes for the learners. Most of the students in computer technology and professional courses do not take care of such courses. These MOOC courses facilitates the students to get the knowledge in the form of Audio and Video methods, Study materials in the form of text and PDF which are downloaded by the registered candidates. The advantages of this course are that the student can complete his education at any place or any time he wish too. At the same time there are some disadvantages of these courses. In this paper the researcher focus on the pit falls of this technology in learning environment with Students having technical and non technical background. The diversification in learning and methods to increase their concentration on available online courses.*

KEYWORDS: *syllabi, variations, teacher, quality, directions, students, ICT*

INTRODUCTION:

From long ago Education is the hearts part of each and every kingdom till now a day's education has some importance or more than that. From punya bhumi Pune knowledge developer and expander Shri Sant Dnyaneshwar, Sant Shreshtha Tukaram Maharajaji and Ramdas Swami Maharajaji has given Direct education to the Social elements. They keep their Granths for betterment of coming social elements and now a day's big number of followers follows the road map. Lot of the Universities developed and they gave education to the students in various fields. Indian universities and non-Indian universities have major difference in complexity of syllabi. In India syllabus variation is the major problem in front of student. Lot of the colleges not generate proper bridging for the students due to this actual knowledge required for industries and knowledge developed by syllabi committee have lot of the difference. International universities have same type of problem but they solved in various ways by making bound with companies and directly given Education plus Job with good stipend. Indian students divert to the International Universities not only for different education but also have facilities and diversification in Education. In the India all things are available and possible but stakeholders of Educational sector not take part seriously. This is problem in front of professional education sector that how cope up speed with ICT and Industrial sector with which student cope up their knowledge regularly. As per consideration with second level metropolitan Pune have major Universities as Savitribai Phule University, Bharati Vidyapeeth and lot of the deemed universities working here but how programs to be design, develop, that Questions not specifically handle by these universities outcomes not specifically keep in front of students. This paper specifically concentrated on the colleges their teaching learning system elements and improvements required for adoption of E-learning synchronous/ asynchronous learning system.

OBJECTIVES:

1) To study various types of teaching learning system
2) To study supportiveness of this mechanism
3) Study on effective diversion towards asynchronous courses for better knowledge outcome
4) Study on primary survey related to syllabi & E- learning courses from student and its descriptive outcomes.

Scope:

- Syllabi of universities and E learning (Synchronous/Asynchronous) courses
- Students of various colleges (40) and working professionals (20)

Area: Pune city.
Primary data collection of students and working professionals by arranging Questionnaire via telephonic conversation

PRIMARY STUDY:

- Study concentrated on limited Questions given to students(PG 20+UG 20) MBA,MSC(CS)/ MCA,MCOM,BBA,BCA,BSC(computer),BCOM,BA MA each stream students taken and working professionals(20) .
- Only 60 members are taken for study from various colleges around and conclusions recorded.

SECONDARY STUDY:
1) E- learning methods
2) Synchronous and Asynchronous methods
3) View on Swayam course

HYPOTHESIS:
H1: Students require University Syllabi as well as E learning

asynchronous study for their development.
HO: Students doesn't require asynchronous study other than University syllabi.

DESCRIPTION:

In regular teaching learning and evolution there are 7 methods first four are teacher focussed and remaining are student focussed (Direct instruction, Drill and practice, lecture, Question and Answer, Discussion, Mental and modelling, Discovery, Inquiry). Whereas this view not only related to the regular methods but view same for ICT learning. Syllabi set by the university given on the web sites and on that syllabus each and every subject its course out come and program out comes given but teachers and students are only concentrated teaching-learning plan and its execution but not on its basic Ideology. E-learning deals with various methods of teaching and learning. In this concentrated on synchronous and asynchronous learning system.

Educationalist says that life is not complete when you drop out education. Now days various ways are open for the education there is no age limit for the E-learning also does not require pre education limit and no pre syllabi restriction for various courses. Its usefulness is varying as per students but gives us powerful knowledge of required subject teachers of various colleges are attracted towards asynchronous courses from Swayam. Here concentrated on the synchronous and Asynchronous learning methods for gaining knowledge but various methods also refer for effecting on students.

As per E-learning ICT uses following methods for effective learning:

1. **Flipped classroom**:

This is an instructional strategy in which computer-assisted teaching is integrated with classroom instruction. Students are given basic essential instruction, such as lectures, before class instead of during class. This frees up classroom time for teachers to more actively engage with learners. (This method deals with synchronous learning method PPT such work is presented related to syllabi.)

2. **Linear learning**:

Computer-based training (CBT) refers to self-paced learning activities delivered on a computer or handheld device such as a tablet or smartphone. CBT Computer-based training is conceptually similar to web-based training (WBT) which is delivered via Internet using a web browser. Assessing learning in a CBT is often by assessments that can be easily scored by a computer such as multiple choice questions, drag-and-drop, radio button, simulation or other interactive means. Assessments are easily scored and recorded via online software, providing immediate end-user feedback and completion status. Users are often able to print completion records in the form of certificates. (With server system it is asynchronous system of learning but when training is going on it is synchronous in nature.)

3. **Computer-supported collaborative learning**:

Computer-supported collaborative learning (CSCL) uses instructional methods designed to encourage or require students to work together on learning tasks, allowing social learning. CSCL is similar in concept to the terminology, "e-learning" and "networked collaborative learning" (NCL). Social networks have been used to foster online learning communities around subjects as diverse as test preparation and language education. Mobile-assisted language learning (MALL) is the use of handheld computers or cell phones to assist in language learning.(With network support various server support engines this method deals with asynchronous knowledge gainer with which student capture knowledge from NCL resources).

4. **Audio and video**:

Radio offers a synchronous educational vehicle, while streaming audio over the internet with webcasts and podcasts can be asynchronous. Classroom microphones often wireless can enable learners and educators to interact more clearly. Video technology has included CDs and DVDs, as well as on-demand and synchronous methods with digital video via server or web-based options such as streamed video from YouTube, Teacher Tube, Skype, Adobe Connect, and webcams. Telecommuting can connect with speakers and other experts. Interactive digital video games are being used at K-12 and higher colleges.(This is one way communication as like lectures recording lectures any time you can saw on the LCD,LED and Projector repeat any time as per requirements. Synchronous and Asynchronous both methods are used but maximally deal with Asynchronous.)

5. **Virtual Learning Environment**:

A virtual learning environment (VLE), also known as a learning platform, simulates a virtual classroom or meetings by simultaneously mixing several communication technologies. For example, web conferencing software such as go to Training, WebEx Training or Adobe Connect enables students and instructors to communicate with each other via webcam, microphone, and real-time chatting in a group setting. Participants can raise hands, answer polls or take tests. Students are able to whiteboard and screencast when given rights by the instructor, who sets permission levels for text notes, microphone rights and mouse control. (This method generally deals with synchronous learning environment RTCP such protocol working when stream line transmission is there.)

6. **Social networks** :

Asynchronous method of Group learning
Question answers are done on line as well as offline and can attain on time or off time. Group webpages, blogs, wikis, and Twitter allow learners and educators to post thoughts, ideas, and comments on a website in an interactive learning environment. Social networking sites are virtual communities for people interested in a particular subject to communicate by voice, chat, instant message, video conference, or blogs. The National School Boards Association found that 96% of students with online access have used social networking technologies, and more than 50% talk online about schoolwork. Social networking encourages collaboration and engagement and can be a motivational tool for self-efficacy amongst students. This communication when working in online mode via audio video then it deals like synchronous and when it is offline then it deal like Asynchronous.

7. **Synchronous Versus Asynchronous teaching**

methodology:

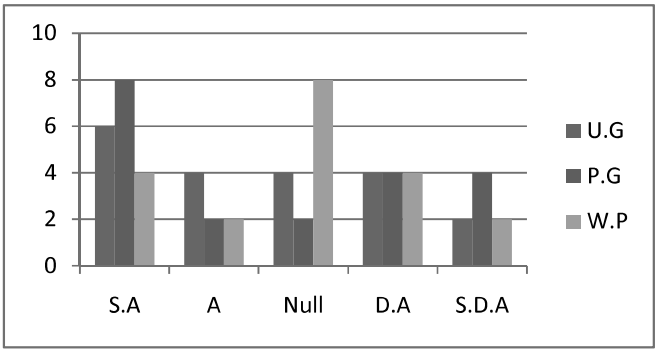
Synchronous is term utilized in E-learning system where deliverer and receiver access their data with salutations and can simultaneously communicated with each other chance of overpowering is negligible. Whereas Asynchronous teaching learning is communication system where deliverer and receiver not given any backup signal of acknowledgement data stored in given data base of E-mail or Blogs and have time get it for knowledge. The extent to which e-learning assists or replaces other learning and teaching approaches is variable, ranging on a continuum from none to fully online distance learning.

E-learning may either be synchronous or asynchronous. Synchronous learning occurs in real-time, with all participants interacting at the same time, while asynchronous learning is self-paced and allows participants to engage in the exchange of ideas or information without the dependency of other participants′ involvement at the same time.

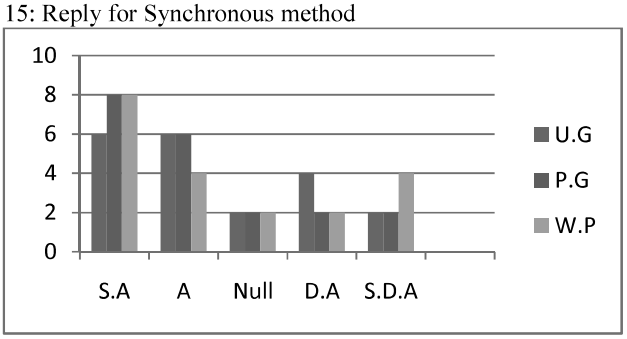
Synchronous learning refers to the exchange of ideas and information with one or more participants during the same period. Examples are face-to-face discussion, online real-time live teacher instruction and feedback, Skype conversations, and chat rooms or virtual classrooms where everyone is online and working collaboratively at the same time. Since students are working collaboratively, synchronized learning helps students create an open mind because they have to listen and learn from their peers. Synchronized learning fosters online awareness and improves many students' writing skills.

Asynchronous learning may use technologies such as email, blogs, wikis, and discussion boards, as well as web-supported textbooks, hypertext documents, audio video courses, and social networking using web support. At the professional educational level, training may include virtual operating rooms. Asynchronous learning is beneficial for students who have health problems or who have child care responsibilities. They have the opportunity to complete their work in a low stress environment and within a more flexible time frame. In asynchronous online courses, students proceed at their own-pace. If they need to listen to a lecture a second time, or think about a question for a while, they may do so without fearing that they will hold back the rest of the class. Through online courses, students can earn their diplomas more quickly, or repeat failed courses without the embarrassment of being in a class with younger students. Students have access to an incredible variety of enrichment courses in online learning, and can participate in college courses, internships, sports, or work and still graduate with their class. After studying various E-learning methodologies of synchronous and asynchronous Table no: 1 focuses on the syllabi and E-learning courses online and offline and their recommendations related to questionnaire.

14: Reply for Asynchronous course as like Swayam



15: Reply for Synchronous method



16: Reply for Asynchronous method

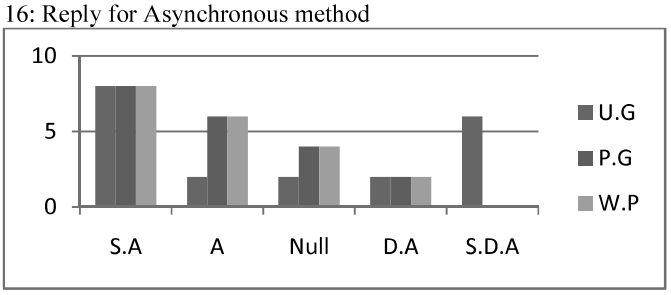


Table no 1: Survey of educational Stakeholders
(E-learning environment)

S.N.	Questions for stakeholders	Total 20/each	SA	A	Null	DA	SDA
1	The stakeholders know latest syllabus from their university?	U.G	10	05	02	02	01
		P.G	15	02	01	02	0
		W.P	12	4	02	02	0
2	Are you happy with specific program given by colleges & universities?	U.G	15	03	00	02	0
		P.G	07	02	0	07	04
		W.P	07	01	01	08	03
3	In colleges teaching learning board & other methods better than online education.	U.G	10	05	02	02	01
		P.G	07	07	02	04	0
		W.P	08	04	04	02	2
4	Student Got desired outcome effect from university syllabi or college education.	U.G	05	07	06	02	0
		P.G	07	03	05	05	0
		W.P	07	03	05	05	0
5	There is a Lacuna of industrial knowledge as required in each University & College.	U.G	15	03	02	0	0
		P.G	13	03	02	02	0
		W.P	10	07	02	01	0
6	Colleges and universities promote bridge course to the students, can they achieve course and job.	U.G	10	02	02	05	01
		P.G	08	06	02	04	02
		W.P	10	06	02	02	0
7	Skilfulness increased by colleges & universities for students via various programs	U.G	10	04	02	02	02
		P.G	10	04	02	01	03
		W.P	08	04	02	02	04
8	Skilfulness increased by industrial training & industrial online training efforts	U.G	12	04	02	01	01
		P.G	12	04	02	01	01
		W.P	10	04	04	01	01
9	New ICT & IT facility in colleges are sufficient for online courses	U.G	05	02	02	01	11
		P.G	05	02	01	08	04
		W.P	02	02	06	05	05
10	Colleges & universities promotes to various online syllabus by Swayam or any other Govt. Educational system.	U.G	05	02	02	10	01
		P.G	08	04	02	04	02
		W.P	05	02	02	07	04
11	Are you familiar with synchronous & asynchronous courses? (E-learning)	U.G	03	02	04	08	03
		P.G	02	02	02	10	04
		W.P	02	02	02	12	02
12	Seminars &workshops required for E-Learning promotion.	U.G	08	07	02	02	01
		P.G	10	02	2	02	04
		W.P	10	02	02	02	04
13	E-Learning(synchronous/ asynchronous) is effective in increasing knowledge	U.G	04	04	02	06	04
		P.G	08	06	02	04	00
		W.P	04	02	08	04	02
14	Courses designed by Swayam is supportive for progress and job	U.G	06	04	04	04	02
		P.G	08	02	02	04	04
		W.P	04	02	08	04	02
15	Synchronous method is better for effective learning? (by providing information)	U.G	08	06	02	02	02
		P.G	08	06	02	02	02
		W.P	08	04	02	02	04
16	Asynchronous method is better for effective learning?(By providing information)	U.G	08	02	02	02	06
		P.G	08	06	04	02	00
		W.P	08	06	04	02	00

Above Table no: 1 shows students their knowledge gaining methods and diversifications related to various questions

CONCLUSION :

i. From question no 1 of table 90% of the PG, UG, and working has idea of syllabi before their admission.
ii. Conclude that program and specific programs given by the colleges sufficient for their learning happiness 90% undergraduates are happy but PG and WP are 45%.
iii. 60%-75% are happy with regular teaching methods but 40% of post graduate and workings have required improving learning and teaching methods.
iv. 50%-60% are firm about the outcomes but 40% of student not knows what to do and the specific outcomes of this program.
v. Under graduate students, post graduates and working professionals are also in favour of industrial knowledge. 80% -90% are given favours towards involvement of industrial knowledge but universities and colleges have some limitations.
vi. More than 60% of respondents are like to add some bridge program for students and trainings from industrial experts with syllabi settled by colleges and universities . More than 80% PG students as well as Working professionals are in favour with industrial training already Universities are involved in it but it is for limited time.
vii. Other questions deals with the Internet learning method 60% are in favour of that with both asynchronous and synchronous learning methodology.
viii. Specific knowledge of Swayam Asynchronous syllabi their knowledge given to the respondents to view and reply. By viewing and studying students and working professionals are ready to register and saw same international asynchronous advanced syllabi designed by various universities and reply changes at next answers.
ix. H1: Satisfied: Hypothesis satisfies and lot of respondents divert towards E-learning environments and gives favours to the Synchronous/asynchronous learning methodology reaching information properly.
x. H0: Null hypothesis dissatisfied because lot of respondents not only favours to syllabi but also favours to the E-learning methodology.

As per the analysis hypothesis one is satisfy i.e. It is better way to choose E-learning (Asynchronous./synchronous) learning. Student requires E-learning which dissatisfies the second hypothesis. For effective better knowledge it is effective to divert towards the asynchronous courses. As per the step by step analysis of table no.1 maximum peoples are tends towards the e-learning (synchronous or asynchronous).

E-learning courses are supportive for all type of improvements in students. So it is required for students and working professionals to add advanced knowledge with E- learning for better and effective learning and understanding.

REFERENCES :

i. How to create Better E-learning Environment with E-Learning-
ii. http://elearningbrothers.com/how-to-create-better-elearning-environments.

iii. What is an innovative learning space - Government of India Ministry of education
iv. Report.

v. E-Learning Systems, Environments and Approaches. Theory and Implementation (year 2015). Editors: **Isaias**, P., **Spector**, J.M., **Ifenthaler**, D., **Sampson**, D.G. (Eds.)
vi. Learning Theories and the Design of E-Learning Environments
vii. First Publication: ૨૦૦૩ **Literature :** BIJAN B. GILLANI.

A STUDY: USE OF APP WITH REFERENCE TO ONLINE MONETERY TRANSACTION, ITS IMPACT ON BANKING CARDS AND ITS SERVICES

PRADIP PATIL AND ASHISH DHOKE
Indira Institute of Management, MCA, Pune, Indira College of Commerce & Science, Pune
pradip.patil@indiraiimpmca.edu.in, dhokeashish@iccs.ac.in

Abstract : In current era of the technology everything become a faster and fastest and in that competition banking sector is also in a race to popup with the upcoming trends and technology and to mean to achieve it every time banking sector comes with the innovative idea and one of the biggest and successful idea which will use the cards included credit card, debit card instead to go bank branch. But now a days, that service is replaced by the new generation of the development of the app and with help of those app bank can do the transaction, but for the replacement of the credit card now UPIN is one of the option available in a market.

In this paper we are going focus on the credit card, its uses, security issues with the credit card transactions, impact of the UPIN money transferring app and its advantages over the cards and its services, multiple platform/app/ UPIN available for the banking transaction and majorly in this paper we will focus on the future of the credit card i.e. its existence in future.

Keywords:
Plastic Money, UPIN, POP, UPI, NPCL, Mobile Banking, PSP, APBS.

INTRODUCTION:

1.1 Plastic Money in India:

The Plastic Money in the form of cards has been actively introduced by banks in India in 1990's, which was not very well-liked in Indian customers at the time of its introduction. As time goes, with importance of competency and technology, Consumers also changed their preferences and modified their viewpoint and decision regarding the acceptance, non - acceptance of specific product and services in the banking market. The Plastic Cards are getting popular with banking sectors as well as customers, also accepted in the market place. Yet at a large speed the Plastic Cards market is growing and it's a long journey for it to cover the complete market compare to other countries.

Hence, it is very important that the monitory transaction in India should be restructured and improved compared with other countries. In current scenario the bankers group of India are now realized their necessary requirement according to information technology for their future growth and survival in the market.

1.2 Problem and Need for the Study:

Day by day there is tremendous change in products and service in banking sectors using information technology. Over a period of time, in India, the usage of Plastic Money such as Credit Cards and Debit Cards are increasing with some insecurity in consumers mind. Most of the banking organizations issues Debit Cards with each of the account openings. Many of the ATM Cards are now converted into Debit cards or ATM cum Debit Cards.

The reason to issue the debit cards that it helps them to cut costs considerably. But this created many confusions regarding its usages and overcome the delayed process of payment of cheques. Still, there are many technical problems involved in transactions. Regardless of this the individual's benefits, the business gains advantages by accepting these medium of payment through Plastic Money. To purchase goods and services in easiest way, many of the banking organizations are trying the innovative ways which changes the buying habits of consumers.

In past few years the crime rates due to cybercrime such as banking cards frauds and ATM frauds affects the society in India. Traditionally, Indian business transactions are mostly carried out by using liquid cash. But, with the emergence of information technology, growth of internet, improved payment structure, E - Commerce (Electronic Commerce) and E - Banking development, entry of foreign banks and new generation banks, high level of competition, easy availability of credit, increase in affluence levels and income of middle group, greater amount of consumerism, fast changes with regard to life style etc., the Indian economy witnessed the growth of Plastic Money - credit cards and debit cards -in terms of usage as a mobile application. It also makes an attempt to make a comparative study between the plastic card and UPIN banking services.[9]

1.3 Mobile Money: A transformation

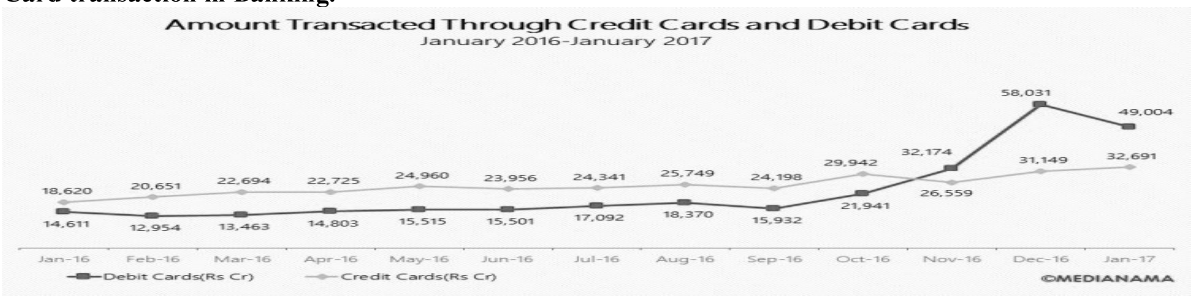
The mobile money is an idea which support for cashless transaction which emphasize the reduction of cash in circulation. Delivery of financial services is a convenience of the mobile phones which brings new values and opportunities in today's technical trends. The opportunities include reaching vast numbers of new customers and providing better service to existing customers. The telecommunications, software, and even retail industries propose the chance to build whole new business concepts. The aim of a business is to provide advantages of mobile money include affordability, security, and convenience to their customers. [10]

2) A future of banking cards

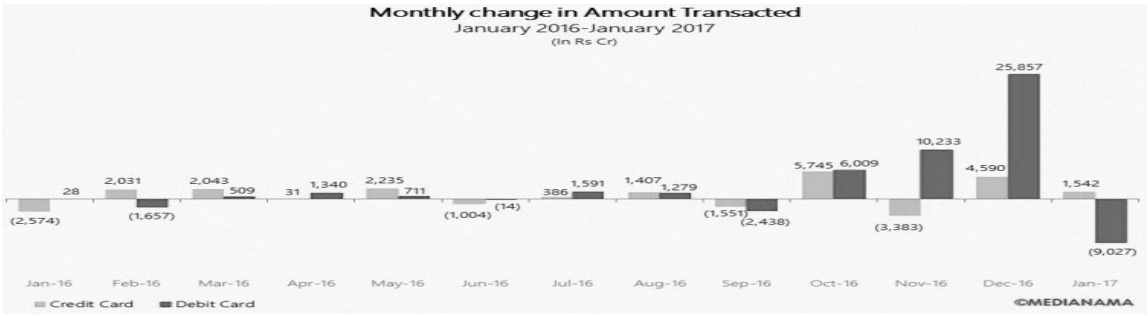
Biometrics (e.g. fingerprints, voice recognition) will become commonplace in transaction authorization, but will remain tied to a replaceable physical device (e.g. smartphone).

Biometrics are unique and unchanging, yet can be captured and replicated, so two-factor authentication (e.g. my fingerprint and my phone) will always be required.

Card transaction in Banking.



1. Total amount transacted through credit cards — which was Rs 32,691 crore in January 2017 — increased by 76% in the one year period ending January 2017. Amount transacted through debit cards increased by 235% during the same time. In January 2017, Rs 49,004 crore, were transacted through debit cards.

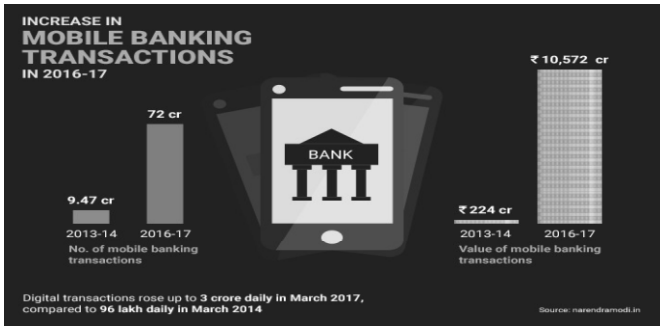


2. Total amount transacted through credit cards went up by Rs 1,542 crore in January 2017 while amount transacted through debit cards declined by Rs 9,027 crore.

Recent trends in mobile banking

India Sees 55% Increase In Digital Transactions In A Year; Mobile Banking Jumps 122%

Current UPA Government giving stress on adopting cashless methods of transactions as is will ensures that all the transactions leaves a trace.



When demonetisation had taken place and India was urged to go cashless, a concern was the availability of internet connectivity, a necessary requirement for cashless transactions. To address this issue, the government worked towards the laying of optical fiber cables for broadband networks and asked the public telecom operators BSNL, as well as private companies such as Reliance and Bharti Airtel to ensure better internet connectivity, especially in rural and semi-urban areas. The total broadband network in 2017 amounts to 2,05,404 kilometers, compared to the merely 358 kilometers in 2013-14 This will ensure that internet connectivity in many parts of the country is much better and helps people in conducting cashless transactions without any hiccups.

Future prospects of Mobile Banking in India

1. **Mobile Network Operators and Large Corporate Houses Co-Venture.**

With the popularity of collaborations between mobile network operators (MNO's) and banks live up to the promise of financial inclusion, the RBI and TRAI (Telecom Regulatory Authority of India) have announced that they will harmonize and coordinate with each other to avoid any form of regularity conflict. Many large Indian banks have partnered/co-ventured with large mobile network operations (MNO's) and handset vendors to facilitate their connection through mobile channel by providing access to financial services.

2. **Under-banked and Un-Banked Population**

Almost half of the country's population is unbanked. The large section of the Indian population not just in rural areas but also in many segments of urban markets, offers a large untapped market with a tremendous business potential.

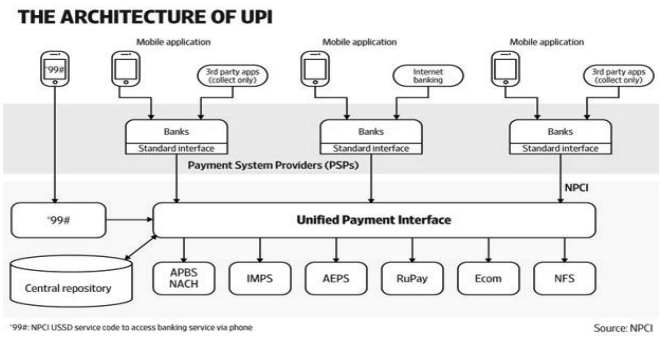
3. **Demographic Factors**

In India the population of youth (between the ages of 14-29) is the largest youth population globally, which is around 27% of the total 1.2 billion. Furthermore, adding the age group of 30-44, the proportion is 47%. %. Apart from the huge size of

this segment, they are among those who are the early adopters of latest technology and new services, which presents a huge opportunity for e/m-banking service providers. It has been observed that for the majority, access to financial services is a household need, and not only an individual need, and if the account holder is illiterate, other members of the family are competent enough to execute transactions and use electronic or mobile banking services.[12]

3. How UPI works

Unified Payment Interface is expected to a play a major role in achieving goals of universal electronic payments, a less-cash society, and financial inclusion



AEPS

An NPCI product, the Aadhaar Enabled Payment System (AEPS) offers instant, 24X7, interbank electronic fund transfer service through mobile phone. It allows Aadhaar biometric authentication-based transactions from a bank account that is linked with the Aadhaar number.

APBS

The Aadhaar Payments Bridge System (APBS) is a system allowing remittances to be made to an Aadhaar number without providing any other bank or account details. It uses the NPCI central mapper as a part of National Automated Clearing House (NACH) to enable government user departments to electronically transfer subsidies and direct benefit transfers to individuals on the basis of their Aadhaar number.

PSPs

Payment Support Providers (PSPs), as defined by RBI, collectively cover all RBI-regulated entities under the Payments and Settlement Act of 2007. These include banks, payments banks, PPIs, and other regulated entities. In addition to the Aadhaar and the mobile number as global identifiers (mapped by NPCI), PSPs can offer any number of virtual addresses to customers so that they can use the virtual address for making and receiving payments.

IMPS

Immediate Payment Service (IMPS), launched on 22 November 2010, is now available to the Indian public from over 65 banks.

USSD

Unstructured Supplementary Services Data

NPCI

National Payment Corporation of India

UIDAI

The Unique Identification Authority of India (UIDAI) which issues digital identity (called Aadhaar number) to residents of India and offers online authentication service.

Value of UPI

UPI is a cheap, secure, reliable, mobile-first, interoperable, open-source, instantaneous settlement and both pull and push platform.

Moreover, while pre-paid wallets can't do more than Rs.10,000 worth of transactions without KYC (know your customer) norms in a given month, a UPI-enabled platform bank account can transfer up to Rs.1 lakh instantaneously.[12]

4. Mobile Wallets & UPI Payment Apps in India

India is facing a cash currency crunch, and until the gap is met by new currency, situation will remain for most of the population. Nonetheless, thanks to the evolution of digital payments there are several ways to skip the paper currency with the help of our smartphones. Here is a list of mobile apps that can be used to making payments online, at certain offline merchants, P2P transfers, bank transfers etc. You can take your pick, and also promote the usage of such apps to help others who are dependent on making or receiving cash payments.

App	Wallet	UPI	Bank Transfer	Send On Mobile #	Android App	Rating	iOS App	Rating
Airtel Money	Yes	No	Yes	Yes	Link	4.2	Link	3
Axis Bank Lime	Yes	No	No	No	Link	3.5	N/A	N/A
BHIM App	No	Yes	Yes	Yes	Link	4	N/A	N/A
Chillr	No	Yes	Yes	Yes	Link	4.4	Link	4
Citrus Pay	Yes	No	No	Yes	Link	3.9	Link	4
Freecharge	Yes	No	No	No	Link	4.3	Link	3.5
FTcash	No	Yes	No	Yes	Link	3.9	N/A	N/A

HDFC PayZapp	No	No	No	Yes	Link	4	Link	2
ICICI Pockets	Yes	Yes	Yes	Yes	Link	4.1	Link	3.5
Itzcash	Yes	No	Yes	Yes	Link	4.4	Link	N/A
Jio Money	Yes	No	No	No	Link	4.3	Link	3.5
Mobikwik	Yes	Yes	Yes	Yes	Link	4.2	Link	4
mRupee	Yes	No	No	Yes	Link	3.7	N/A	N/A
Oxigen Wallet	Yes	No	Yes	Yes	Link	3.7	Link	4
Paytm	Yes	No	Yes	Yes	Link	4.4	Link	4.5
PhonePe	No	Yes	Yes	Yes	Link	4.1	N/A	N/A
SBI Buddy	Yes	No	No	Yes	Link	3.9	Link	2
Trupay	No	Yes	Yes	Yes	Link	4	Link	4.5
Vodafone M-Pesa	Yes	Yes	Yes	Yes	Link	4.2	Link	4.5

Fig : List of Mobile Wallets & UPI Payment Apps in India

If you are new to payments via mobile apps, then you can mostly look at going either with mobile wallets or apps supporting UPI payments, depending upon your requirement. Mobile wallets are digital instruments where you can store money for instant payments. You load money by transferring from your bank account via credit/debit cards or net banking. Most of the wallets listed here are semi-closed wallets, i.e. you can transfer money to people who have the same wallet, or make payments at merchants who are authorized to accept from that particular instrument. RBI has increased the monthly balance limits for mobile wallets to Rs 20000.UPI or Unified Payment Interface, is an electronic funds transfer instrument that enables all bank account holders to send and receive money from their smartphones without the need to enter bank account information or net banking userid/ password. This requires only the recipient's mobile number or Virtual Payment Address (VPA). Apart from the above mentioned apps, banks like SBI & HDFC have launched their own UPI apps. 31 banks are on UPI platform. You can read all about UPI here, and see if your bank has launched one too. Most wallets also allow making payments by entering credit/debit card information, i.e. without first adding money. This would be time consuming if you have to make payments often. UPI is faster, if you are not comfortable storing your money in a 3rd party app. Apps like Paytm, SBI Buddy, The Mobile Wallet are available in multiple languages. Mobikwik allows users to add cash to their wallets by visiting any of the nearby outlets.

4.1 Tips to Use Mobile Banking in the Right Way

From our work with leading players worldwide, from our research into the macro-trends impacting banking and from our survey of global banking executives, we have identified the following six priorities for retail banks to win in 2020:

Security measures:

a. Don't store important personal information on your smartphone

If you visit your bank's website using your phone, make sure to clear out the cookies and cache regularly. Never store any usernames or passwords in your phone, regardless of how difficult they are to remember. If you were to lose the phone where this information is stored, someone could easily access your bank account. It is also important to change your password regularly. If you use your phone for business purposes, be doubly careful on protecting confidential information.

b. Sign up for SMS alerts

Most banks now offer text message alerts via SMS to alert of a low balance or a paid bill. Larger banks like Bank of America use SMS texts to aid in fraud prevention. Nine times out of ten, consumers can easily sign up for these phone alerts with online banking. Receiving alerts on your phone is a great way to stay on top of your account activity around the clock. It's also an added convenience for times when a computer isn't close at hand.

c. Before downloading an app, check its authenticity

Most banks offer tailored mobile banking applications to their customers which can be downloaded to a Smartphone and used to manage bank accounts. Unfortunately, this has also invited potential fraud in the form of carefully duplicated applications created by scam artists. Before downloading any app to your phone, make sure that it's an authentic application released by your bank. Avoid third party software that asks for any personal information if you can't verify the source. While most apps are legit, it never hurts to be too careful. [12]

5. Probable Future enhancement

As per the details through which we have seen above, we can able to suggest some methodology through which credit card can be replaced by the mobile and UPI, and in a future we can extend it with some of the security majors like finger print scanner, retina scanning, face recognition, and with the help of those extended technology we can able to make it thread proof and able to overcome all the security related issues.

CONCLUSION :

Although the cards are getting increase day by day for the banking transaction but then also within short duration of the time it will be get evaporated like anything and that place will be taken by the mobile UPI transition service technology, demographics, changing customer expectations, greater competition and issues with banks' own legacy business and operating models. The challenges are clear, even if the ultimate endgame is not. Banks need to get ahead of these challenges and retool to win in 2020. They need to make hard choices about which customers to serve, how to win and where not to play. They need to rebuild their organizations around the customer, simplify and structurally reduce cost. They need to learn to be agile, innovative and adaptable in order to execute effectively – and deal with uncertainty as the future unfolds. They need to do things differently. Each bank's unique response will depend upon the bank's current position, aspirations for the future, desired customer focus, organizational capabilities, brand promise, regulatory situation and capital constraints. Banks should consider the posture they wish to adopt. Do they want to shape this future, rapidly follow, or manage defensively, putting off change? Staying the same is not an option. Every bank needs to develop a strategy to tackle these challenges.

Though every coin has two sides in similar manner we can able to make the system 99 % error proof, but it will be in the hand of the user how to make use of it in positive or in negative manner.

REFERENCES :

[1]J. Book, Modelling design and control of flexible manipulator arms. A tutorial review, *Proc. 29th IEEE Conf. onDecision and Control*, San Francisco, CA, pp500-506, 1990.

[2] Gajendra Singh , Jarikre Amos Oghengweke, Kankam Boa , The Impact and Challenges of Mobile Money TransferServices in Ghanaian Telecom Industries, Journal of IMS Group | Volume 10 No. 2 | July - December 2013 pp. 10-18

[3]Dhanush J. Nair , Sunny Nahar , ATM Transaction : A New Time Based Approach International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 6, June 2015 2254pp 2254-2256

[4] RenjuChandran ,pros and cons of mobile banking International Journal of Scientific and Research Publications, Volume 4, Issue 10, October 2014 1 ISSN 2250-3153 pp 1-5

[5]M. Ozaki, Y. Adachi, Y. Iwahori, & N. Ishii, Application of fuzzy theory to writer recognition of Chinese characters, *International Journal of Modelling and Simulation, 18*(2), pp 112-116, 1998.

[6] R.E. Moore, Interval analysis (Englewood Cliffs, NJ: Prentice-Hall, 1966).

[7]http://www.freepressjournal.in/over-rs-490-cr-involved-in-cyber-fraud-cases-since-2011-govt/

[8]https://www.amfiindia.com/

[9]http://www.fraud-magazine.com/default.aspx

[10]https://www.rbi.org.in/scripts/BS_CircularIndexDisplay.aspx?Id=9322

[11]http://ncrb.gov.in/

[12]http://bit.ly/1SYOShJ.

[13]http://special.ndtv.com/cashless-bano-india-14/news-detail/india-sees-55-increase-in-digital-transactions-in-a-year-mobile-banking-jumps-122-1724624/7

FUZZY LOGIC IN RELATION TO USE POTASSIUM IN GRAPE VINEYARDS

VILAS B. MORE[1], D. N. KASHID[2] AND B. T. JADHAV[3]
Department of Computer Science, Shivaji University, Kolhapur - 416004 (MS)
Department of Statistics, Shivaji University, Kolhapur - 416004 (MS)
Rayat Institute of Research and Development, Satara- 415001(MS)

Abstract: The grape is one of the major important fruit crops of the Maharashtra. At present, the use of fertilizer e.g. Potassium, requirement is increasing, so cost also increases. Doses of Potassium are decided by the experts after Petiole testing results from laboratory. In the present study, the petiole samples are collected from different vineyards of the Maharashtra and Karnataka State. In this paper to decide the actual requirement of Potassium, the Fuzzy Information System (FIS) is developed. The data obtained from the analysis of petiole testing report is used as an input to FIS and output results are compared with the results suggested by the expert. It is found that the Potassium requirement suggested by FIS of any grape vineyard is less than the laboratory expert suggestions. Results help to reduce the cost of production without affecting the yield level of the grape vineyard and to maintain the petiole quality. Fuzzy Inference System is developed by using Mat Lab software.

Key Words: Potassium, Grapes, Petiole, Fuzzy Logic**,** Mat Lab.

INTRODUCTION:

Grape is one of the major important fruit crops of the country grown on an area of about 1.19 lakh hectares with an annual production of 25.85 lakh metric tons (National Horticulture Database, India, 2015). In Maharashtra, it is grown on an area of 90,000 hectares with production of 21.60 lakh metric tons (National Horticulture Database, India, 2015). Under the tropical condition, the vine is pruned twice in a year i.e., once after the harvest of fruits during April (back pruning) and again for fruits during October (forward pruning) [4,7].Potassium is one of the major nutrients supplied to the vineyard for cane maturity and fixation of bunch developed in the bud during fruit bud differentiation stage [15].

Potassium (K) helps in increasing the sugar content in berries. The increase in sugar content ultimately leads to increase in yield per vine [4]. The sweet berries containing high amount of sugar in the form of total soluble solids are generally preferred by the consumers. This is achieved by application of potassium. The main growth stages at which Potassium is needed at optimum for March/April pruning are bud fixing stage, cane maturity stage and period before fruit pruning. Adequate status of Potassium has been emphasised for formation of fruitful buds at bud initiation and differentiation stages [1] and at bud fixation after differentiation (50 to 55 days after pruning) and at cane maturity [19].After October pruning, adequate Potassium is needed for translocation of sugars to the berries.

The nutrient requirement of grapevine is assessed by the researcher using the petiole testing report. It is observed that the suggestions for requirement of the nutrients from the same vineyard may vary from laboratory to laboratory. The nutrient requirement of each vineyard is different and is based on the nutrient status of each garden. This shows that vagueness is present in nutrient suggestion. There is also vagueness in the interpretation of the test results. To avoid this we have developed FIS for suggestion of Potassium requirement. The principal contribution of fuzzy logic- is its high power of precision [11].Most of the practical applications of fuzzy logic are associated with its relational facet. Considering this, Fuzzy Inference System (FIS) suggests the Potassium requirement for grape vineyard based on the petiole testing report. The objective of this paper is to avoid the vagueness in Potassium treatment to petiole and reduce the cost of production.

2. Laboratory Data Base:

The Petiole is collected from the fields of Solapur, Sangli, Nasik, Pune districts of Maharashtra and Belgaon, Bijapur, Gulbarga districts of Karnataka. The petiole testing laboratory of Maharashtra Rajya Draksh Bagaitdar Sangh, Manjri Farm, Pune (Maharashtra) has been used for the investigation by the researcher. During the period of study (2014 to 2016), **685 samples** of different grape vineyards and different locations are collected and analyzed in the laboratory using laboratory standard methods. The observations are grouped into 10 classes.

Researcher developed database from laboratory results as given in table1. After analysis K observed is called as input and suggestions about nutrients to farmer is called as output.

Table 1: Database developed for input and output parameters of Potassium

Class No. (Sample Farmer)	INPUT Potassium(%) Observed in analysis	OUTPUT (Suggestions to farmers)			
		ORGANIC	INORGANIC Nutrient		
		Before pruning (Kg/acre)	Before Pruning (Kg/acre)	30-60 days after pruning (Kg/acre)	105-135 days after pruning (Kg/acre)
	K	NL	AK 1	AK 2	AK3
1	0 – 0.50	84	5	63	63
2	0.50 – 0.80	70	5	53	53
3	0.80 – 1.10	62	5	46	46
4	1.10 – 1.65	56	5	42	42
5	1.65 – 2.20	56	5	42	42
6	2.20 – 2.40	51	5	38	38
7	2.40 – 2.60	45	5	34	34
8	2.60 – 2.80	39	5	29	29
9	2.80 – 3.00	34	5	25	25
10	3.00 – 5.00	28	5	21	21

(K: Potassium, NL: Natural Potassium, AK1: Artificial Potassium 1(Inorganic1), AK2: Artificial Potassium 2, AK3: Artificial Potassium 3)

The data analyzed for different growth stages of grape for organic and inorganic compound. The variation for organic compound is in the range of 28 to 84 kg/acre. However, for inorganic compounds (30 - 60 days and 105 -135 days) the values are in the range of 21 to 63 kg/acre. There is no variation in the inorganic nutrients suggestion before pruning. To avoid such variations in nutrient suggestions fuzzy inference system has been developed.

3: Development of Fuzzy Inference System for Potassium Nutrient Suggestions

Fuzzy Inference System is developed by using triangular membership functions for input and output variables as shown in fig.1.For defuzzification,centroid method is used.The system is developed by using Mat Lab Software.

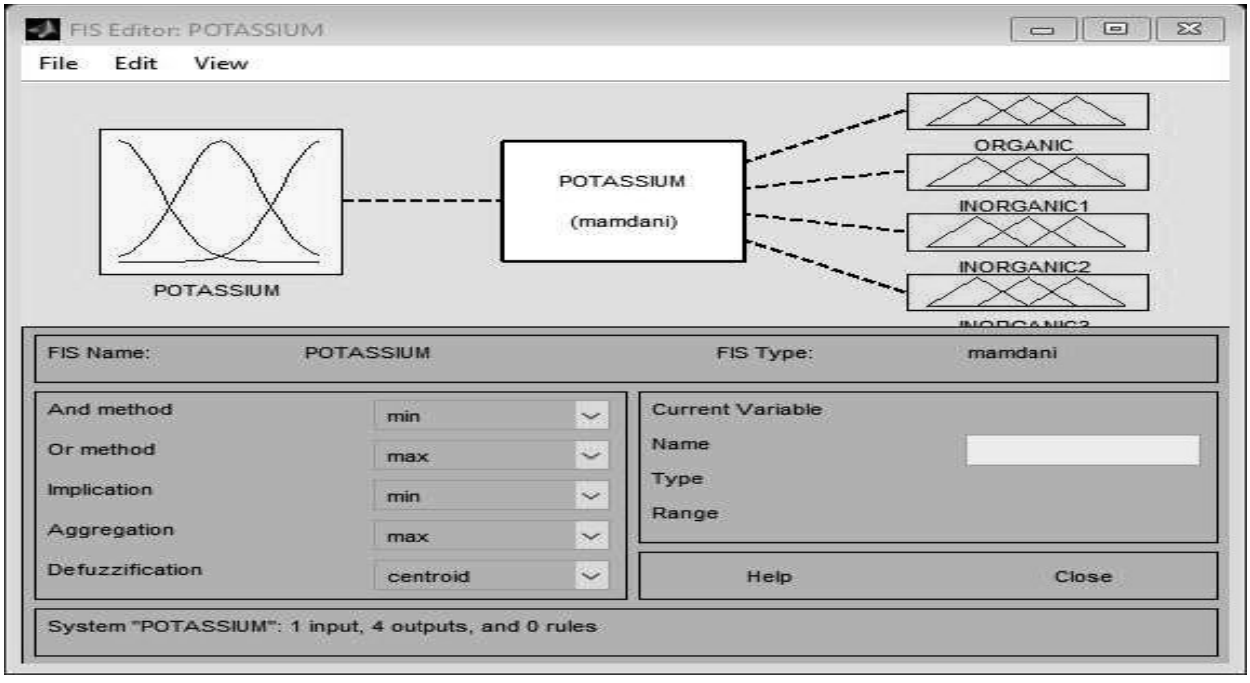


Fig.1: Fuzzy Inference System

I) Fuzzification:

Potassium is used as input for FIS which may carry out the Fuzzification for suggestion of Potassium for the petiole. The requirement of Potassium varies in the range of 0 to 5 %. Fuzzy Set for input variable Potassium measured from petiole is given in table 2.

Table 2: Fuzzy Set for input variable Potassium measured from petiole

Sr.No.	Potassium Levels	Fuzzy sets (%)
1	μ_{MOL}	L (0.00, 0.00, 0.50)
2	μ_{VL}	Λ (0.00, 0.50, 0.80)
3	μ_{JL}	Λ (0.50, 0.80, 1.10)
4	μ_{ML}	Λ (0.80,1.10, 1.65)
5	μ_{L}	Λ (1.10, 1.65, 2.20)
6	μ_{N}	Λ (1.65, 2.20, 2.40)
7	μ_{H}	Λ (2.20, 2.40, 2.60)
8	μ_{MH}	Λ (2.40, 2.60, 2.80)
9	μ_{JH}	Λ (2.60, 2.80, 3.00)
10	μ_{VH}	Λ (2.80, 3.00, 5.00)
11	μ_{MOH}	⅃ (3.00, 5.00, 5.00)

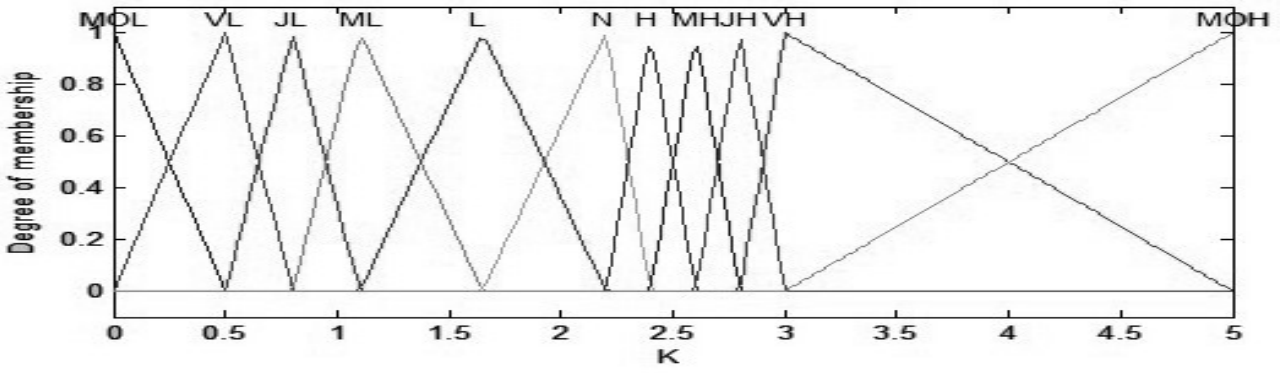
The values of NL are within the range of 23 to 84 kg/acre. Here AK1 is constant as 5 kg/acre, while AK2 and AK3 are within the range is 17 to 63 kg/acre. Fuzzy set for output variables of Potassium is given in table 3.

Table 3: Fuzzy set for output variables of Potassium

Sr.No.	Organic and Inorganic fertilizer (Membership Functions)	Fuzzy set		
		ORGANIC	INORGANIC	
		Before pruning (NL) (kg /acre)	30-60 days after pruning(AK2) (kg /acre)	105-135 days After pruning (AK3) (kg /acre)
1	μ_{MOH}	L (84, 84, 70)	L (63, 63, 53)	L (63, 63, 53)
2	μ_{VH}	Λ (84, 70, 62)	Λ (63, 53, 46)	Λ (63, 53, 46)
3	μ_{JH}	Λ (70, 62, 56)	Λ (53, 46, 42)	Λ (53, 46, 42)
4	μ_{MH}	Λ (62, 56, 56)	Λ (46, 42, 42)	Λ (46, 42, 42)
5	μ_{H}	Λ (56, 56, 51)	Λ (42, 42, 38)	Λ (42, 42, 38)
6	μ_{N}	Λ (56, 51, 45)	Λ (42, 38, 34)	Λ (42, 38, 34)
7	μ_{L}	Λ (51, 45, 39)	Λ (38, 34, 29)	Λ (38, 34, 29)
8	μ_{ML}	Λ (45, 39, 34)	Λ (34, 29, 25)	Λ (34, 29, 25)
9	μ_{JL}	Λ (39, 34, 28)	Λ (29, 25, 21)	Λ (29, 25, 21)
10	μ_{VL}	Λ (34, 28, 23)	Λ (25, 21, 17)	Λ (25, 21, 17)
11	μ_{MOL}	⅃ (28, 23, 23)	⅃ (21, 17, 17)	⅃ (21, 17, 17)

Membership functions for input variable Potassium measured in petiole is shown in Fig.2.

Fig 2: Fuzzy membership functions for Potassium measured



Membership functions for output variables i.e. Organic Potassium, Inorganic Potassium1, Inorganic Potassium 2 and Inorganic Potassium 3 are shown in fig.3a, 3b and 3c respectively.

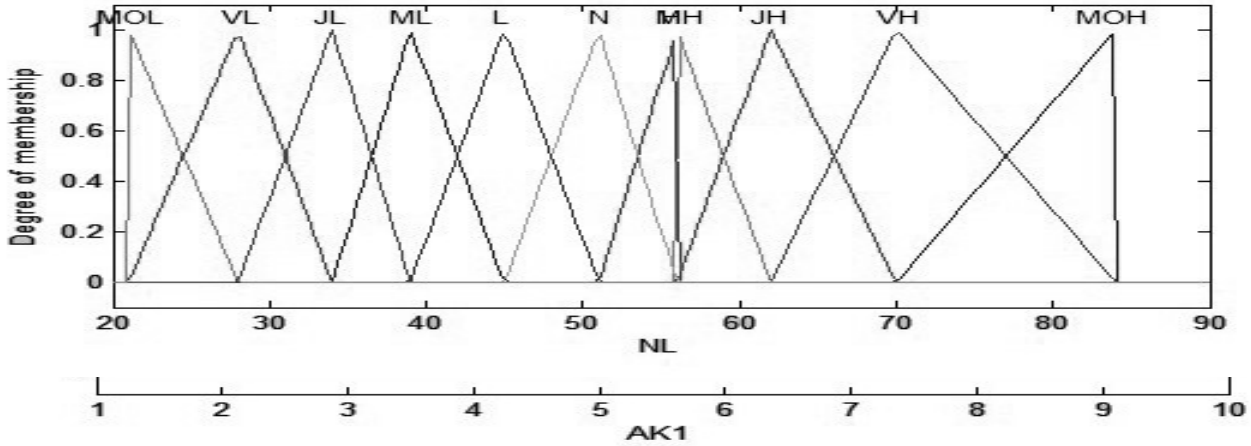


Fig.3a: Fuzzy membership functions for Organic Potassium

Fig.3b: Fuzzy membership functions for Inorganic Potassium1 (AK1)

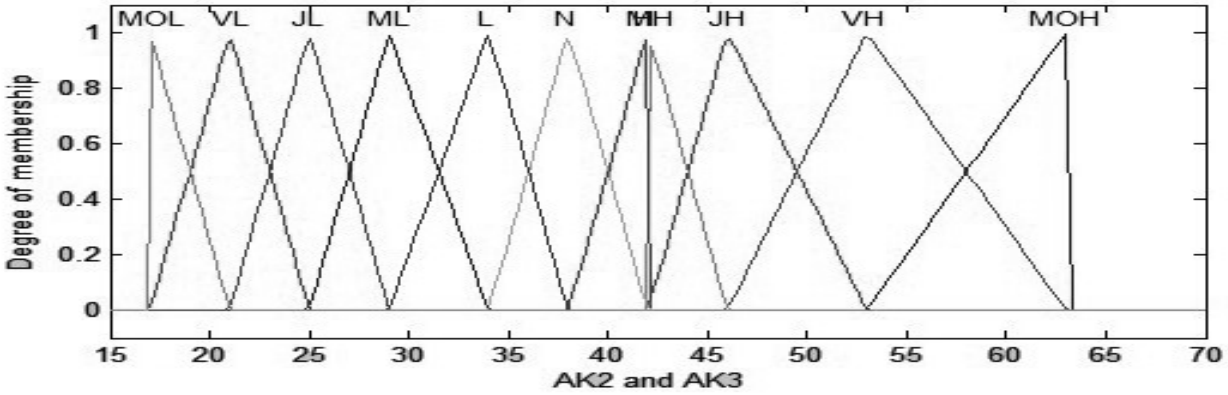


Fig.3c: Fuzzy membership function for Inorganic Potassium (AK2 and AK3)

Rule base developed by the researcher for potassium is given in table 4.

Table 4: Fuzzy rule base for Potassium

If K is	MOL	Then NL is	MOH	And AK1 is	C	And AK2 is	MOH	And AK3 is	MOH
	VL		VH		C		VH		VH
	JL		JH		C		JH		JH
	ML		MH		C		MH		MH
	L		H		C		H		H
	N		N		C		N		N
	H		L		C		L		L
	MH		ML		C		ML		ML
	JH		JL		C		JL		JL
	VH		VL		C		VL		VL
	MOH		MOL		C		MOL		MOL

The fuzzy rule base is read as:

If Potassium is Most Low Then Natural Potassium is Most High and Artificial Potassium1 is constant and Artificial Potassium 2 and Artificial Potassium 3 is Most High.

Fig.4 shows the rule base for fuzzy inference system of Potassium.

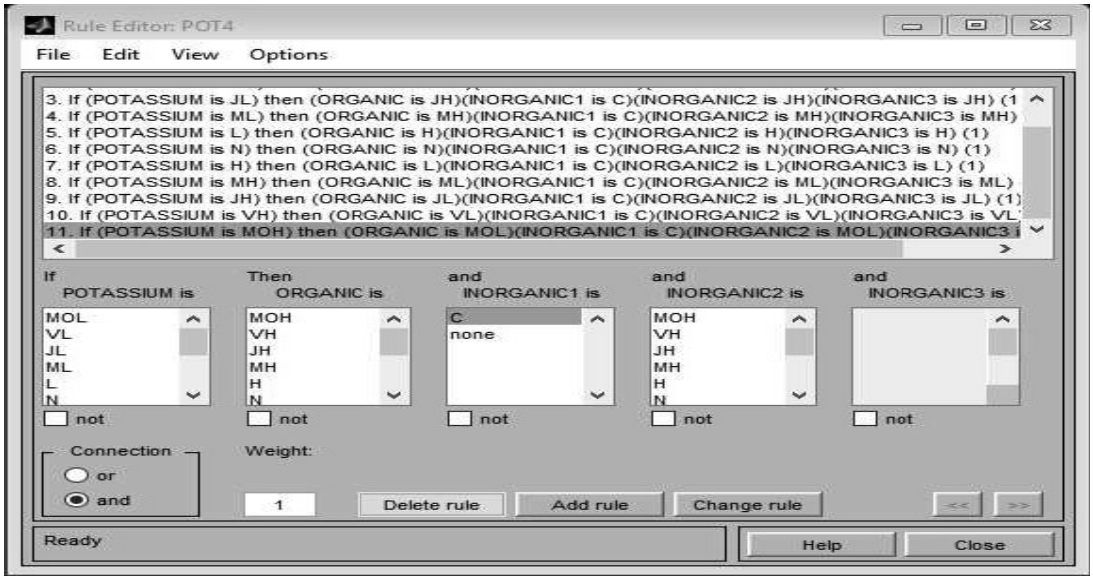


Fig. 4: Rule base for fuzzy inference system

II) Defuzzication:

Conversion of fuzzy quantities measured K, NL, AK1, AN2 and AK2 are converted into precise quantity. Here centroid defuzzification method is used and shown in fig.5.

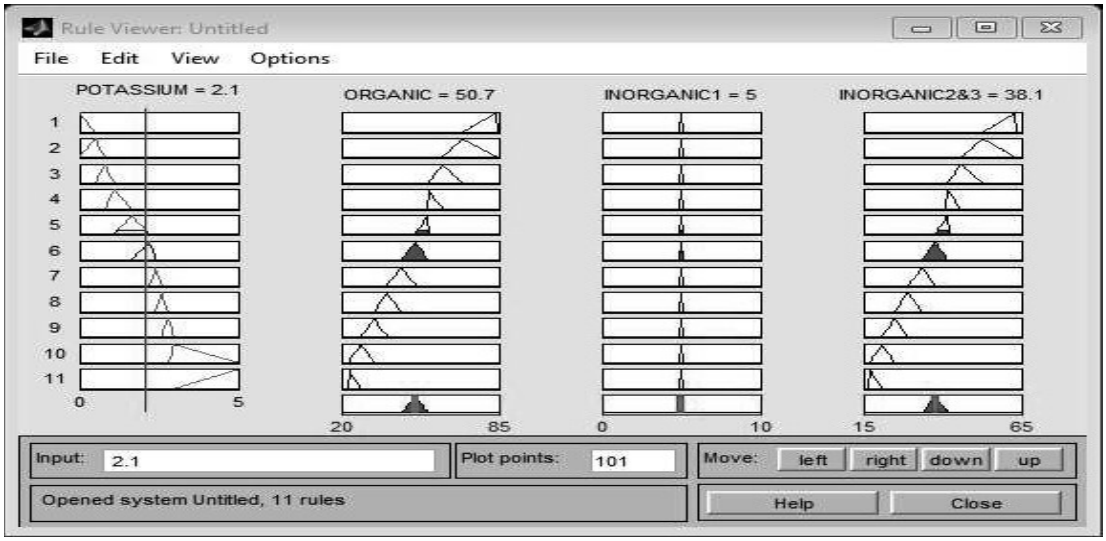


Fig.5: Defuzzication of the System

The FIS developed is based on the expert database. The results obtained through FIS are given in the table 5.

Table 5: Potassium suggested by FIS

Sr. No.	Farmer No.	INPUT (%)	OUTPUT			
			ORGANIC	INORGANIC		
			Before pruning (kg/acre)	Before Pruning (kg/acre)	30-60 days after pruning (kg/acre)	105-135 days after pruning (kg/acre)
		K	NL	AK 1	AK 2	AK3
1	21	0.50	72.0	5	54.0	54.0
2	495	0.75	66.0	5	49.4	49.4
3	552	1.05	60.4	5	45.2	45.2
4	365	1.30	56.6	5	42.2	42.2
5	360	1.60	55.1	5	41.2	41.2
6	379	2.10	50.7	5	38.1	38.1
7	326	2.25	48.9	5	36.6	36.6
8	335	2.65	37.7	5	28.2	28.2
9	207	2.87	31.7	5	23.5	23.5
10	165	4.25	27.4	5	20.1	20.1

4. Statistical Analysis:

Potassium suggested by laboratory for different vineyards are given in the ranges, while statistical analysis gives exact quantity for Potassium. To evaluate the performance of proposed FIS through statistical analysis and simulation is carried out by using Mat Lab software programming.

The analyzed database is classified into 10 different classes. Minimum and maximum outputs for corresponding input classes are given in table 6.

Table 6: Analyzed database for Potassium

Class No.	Potassium measured INPUT (%)	OUTPUT			
		ORGANIC	INORGANIC (K)		
		Before pruning (kg/acre)	Before Pruning (kg/acre)	**30-60** days after pruning (kg/acre)	**105-135** days after pruning (kg/acre)
	K	NL	AK 1	AK 2	AK3
1	0.00 – 0.50	72.00 – 72.00	3.43 – 3.43	53.99 – 53.99	53.99 – 53.99
2	0.50 – 0.80	64.30 – 70.04	3.23 – 3.43	48.16 – 52.57	48.16 – 52.57
3	0.80 – 1.10	58.03 – 62.66	3.25 – 3.43	43.49 – 47.00	43.49 – 47.00
4	1.10 – 1.65	55.09 – 57.10	3.18 – 3.43	41.21 – 42.58	41.21 – 42.58
5	1.65 – 2.20	50.66 – 55.09	3.20 – 3.43	37.99 – 41.21	37.99 – 41.21
6	2.20 – 2.40	45.00 – 49.49	3.18 – 3.43	33.66 – 37.01	33.66 – 37.01
7	2.40 – 2.60	39.34 – 45.00	3.18 – 3.43	29.33 – 33.66	29.33 – 33.66
8	2.60 – 2.80	33.66 – 39.34	3.18 – 3.43	25.00 – 29.33	25.00 – 29.33
9	2.80 – 3.00	27.65 – 33.66	3.18 – 3.43	20.99 – 25.00	20.99 – 25.00
10	3.00 – 5.00	23.66 – 27.65	3.18 – 3.43	18.52 – 20.99	18.52 – 20.99

The values for organic compound before pruning are observed and it is found that the class 10 recorded minimum values with a range of 23.66 to 27.65 kg/acre, while class1 recorded higher value of 72.00 kg/acre. The same trend was also observed for inorganic compound at both the stages (30 to 60 days and 105-135 days after pruning).

The paired t-test is applied for different parameters of Potassium and the results are given in table 7.

Table 7: Paired t-test for different parameter

Parameter	t-statistic	p-value	Decision
Natural Potassium	3.30	0.005	Reject H_0
Artificial Potassium 1	134.76	0.000	Reject H_0
Artificial Potassium 2& 3	3.08	0.007	Reject H_0

The results given in the table 7 supports our claim. Hence, we conclude that the output suggested by statistical analysis is less as compare to the output suggested by the expert for all different parameters of Potassium.

5: Statistical Fuzzy Model

Fuzzy Inference System for analyzed database has been developed.

I) Fuzzification for analyzed database: Fuzzification is carried out for analyzed potassium by using triangular membership functions. The analyzed database for input and output parameters of potassium are given in table 8.

Table 8: Analyzed database for input and output parameters of Potassium

Class No.	Potassium measured INPUT (%)	OUTPUT suggested			
		ORGANIC	INORGANIC (K)		
		Before pruning (kg/acre)	Before Pruning (kg/acre)	**30-60** days after pruning (kg/acre)	**105-135** days after pruning (kg/acre)
	K	NL	AK 1	AK 2	AK3
1	0 – 0.50	72.00	3.43	53.99	53.99
2	0.50 – 0.80	70.04	3.43	52.57	52.57
3	0.80 – 1.10	62.66	3.43	47.00	47.00
4	1.10 – 1.65	57.10	3.43	42.58	42.58
5	1.65 – 2.20	55.09	3.43	41.21	41.21
6	2.20 – 2.40	49.49	3.43	37.01	37.01
7	2.40 – 2.60	45.00	3.43	33.66	33.66
8	2.60 – 2.80	39.34	3.43	29.33	29.33
9	2.80 – 3.00	33.66	3.43	25.00	25.00
10	3.00 – 5.00	27.65	3.43	20.99	20.99

Potassium measured from petiole is used as input domain for FIS. Fuzzy system carried out fuzzification for requirement of potassium for petiole. The input potassium varies in the range of 0 to 5 %.The values of NL are in the range of 27.65 to 72.00 kg/acre. Here AK1 is constant as 3.43 kg/acre while AK2 and AK3 varies from 20.99 to 53.99 kg/acre. It is given in table 9.

Fig.6: Defuzzication of analyzed database

Results obtained through FIS for analyzed database is shown in table 10.

Table 10: Results obtained through FIS for analyzed database

Class No.	Farmer No.	INPUT (%)	OUTPUT			
			ORGANIC	INORGANIC (K)		
			Before pruning (kg/acre)	Before Pruning (kg/acre)	30-60 days after pruning (kg/acre)	105-135 days after pruning (kg/acre)
		K	NL	AK 1	AK 2	AK3
1	21	0.5	68.2	3.25	51.2	51.2
2	495	0.75	63.7	3.25	47.7	47.7
3	552	1.05	60.2	3.25	45	45
4	365	1.3	56.5	3.25	42.3	42.3
5	360	1.6	54.6	3.25	40.8	40.8
6	379	2.1	50.3	3.25	37.6	37.6
7	326	2.25	48.3	3.25	36.1	36.1
8	335	2.65	37.6	3.25	28.2	28.2
9	207	2.87	31.1	3.25	23.2	23.2
10	165	4.25	26.4	3.25	19.7	19.7

From table 6 and 10, we observe that the results obtained by using FIS are less than the results obtained by the expert. This method reduces the quantity of the fertilizer.

6: Sampling and Simulation for large database for more accuracy

In this part of performance evaluation, we used simulated data in large scale. The 1000 observations are generated from U (0, 5) because sample input range is 0 to 5 %. Output is obtained using FIS for each class. The same experiment conducted 1000 times and minimum and maximum values of output of the FIS corresponding to input classes are given in table 11.

Table 11: Analyzed database using simulation for Potassium

Class No.	Measured INPUT (%)	OUTPUT suggested by statistical simulation			
		ORGANIC	INORGANIC (K)		
		Before pruning (kg/acre)	Before Pruning (kg/acre)	**30-60** days after pruning (kg/acre)	**105-135** days after pruning (kg/acre)
	K	**NL**	**AK 1**	**AK 2**	**AK3**
1	0.00 – 0.50	72.00 –79.21	3.18 – 3.43	54.00 – 59.60	54.00 – 59.60
2	0.50 – 0.80	63.09 –71.89	3.18 – 3.43	47.30 – 53.92	47.30 – 53.92
3	0.80 – 1.10	58.36 – 62.66	3.18 – 3.43	43.83 – 47.01	43.83 – 47.01
4	1.10 – 1.65	54.44 – 57.96	3.18 – 3.38	40.85 – 43.42	40.85 – 43.42
5	1.65 – 2.20	50.66 – 54.21	3.18 – 3.43	37.99 – 40.73	37.99 – 40.73
6	2.20 – 2.40	45.17 – 50.45	3.19 – 3.43	33.79 – 37.80	33.79 – 37.80
7	2.40 – 2.60	39.56 – 44.85	3.19 – 3.43	29.46 – 33.53	29.46 – 33.53
8	2.60 – 2.80	33.86 – 39.11	3.19 – 3.43	25.18 – 29.19	25.18 – 29.19
9	2.80 – 3.00	27.83 – 33.39	3.18 – 3.43	21.14 – 24.86	21.14 – 24.86
10	3.00 – 5.00	23.40 – 27.65	3.18 – 3.43	18.31 – 20.99	18.31 – 20.99

From table 7 and 11, we observe that the result obtained from analyzed data and simulation for large database is less than the result obtained by the expert given in table 1. The result indicated that the analysis done through simulation technique gives the more accuracy. The stat-fuzzy model used for calculating the required output is simple and easy. The output ranges for different classes obtained from table 7 and 11 are not exactly same but closer to each other, because 1000 samples are simulated 1000 times.

7: Random sampling for result analysis and comparison

Table 12 shows that suggested value for natural potassium (NL), AK1, AK2 and AK3 based on the status of potassium available in the petiole sample analyzed in laboratory. Similarly, the data using FIS gives requirement of exact quantity of particular fertilizer during the season. FIS developed for potassium suggestion by using statistically analyzed database gives better result, which helps for saving the cost required for the use of fertilizer which ultimately reduces cost of production.

Table 12: Suggestion of Potassium (NL, AK1, AK2 and AK3)

Sr. No.	Farmer No.	Input Data (%)	Potassium suggested by Expert			Potassium suggested by FIS			Potassium suggested by FIS(Stat. Fuzzy Model)		
			NL (kg/ acre)	AK1 (kg/ acre)	AK2 and AK3 (kg/ acre)	NL (kg/ acre)	AK1 (kg/ acre)	AK2 and AK3 (kg/ acre)	NL (kg/ acre)	AK1 (kg/ acre)	AK2 and AK3 (kg/ acre)
1	21	0.5	84	5	63	72	5	54	68.2	3.25	51.2
2	495	0.75	70	5	53	66	5	49.4	63.7	3.25	47.7
3	552	1.05	62	5	46	60.4	5	45.2	60.2	3.25	45
4	365	1.3	56	5	42	56.6	5	42.2	56.5	3.25	42.3
5	360	1.6	56	5	42	55.1	5	41.2	54.6	3.25	40.8
6	379	2.1	56	5	42	50.7	5	38.1	50.3	3.25	37.6
7	326	2.25	51	5	38	48.9	5	36.6	48.3	3.25	36.1
8	335	2.65	39	5	29	37.7	5	28.2	37.6	3.25	28.2
9	207	2.87	34	5	25	31.7	5	23.5	31.1	3.25	23.2
10	165	4.25	28	5	21	27.4	5	20.1	26.4	3.25	19.7

Potassium suggested by statistical analysis FIS method is less as shown in graph 7a, 7b and 7c.

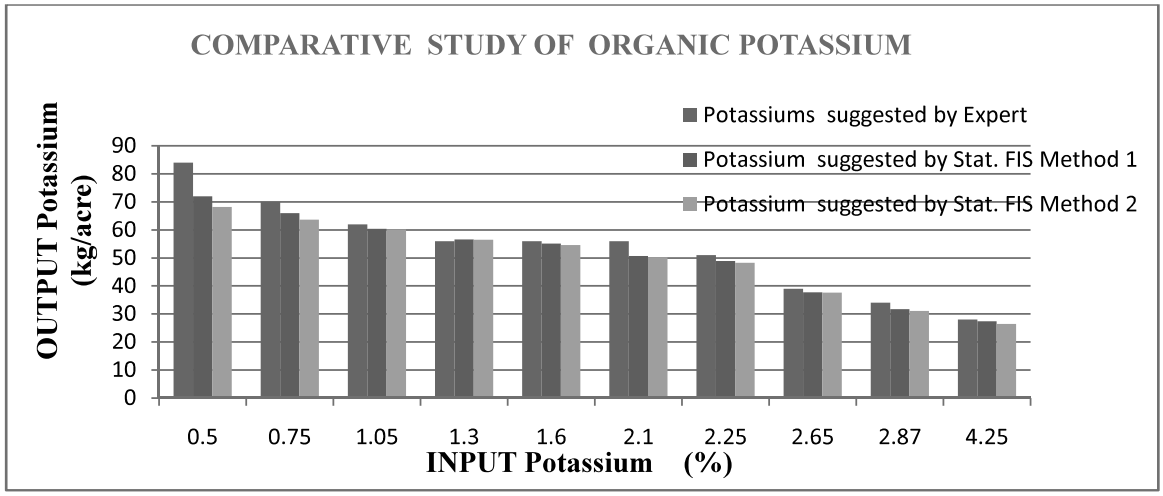


Fig.7a: Comparative study of Organic Potassium

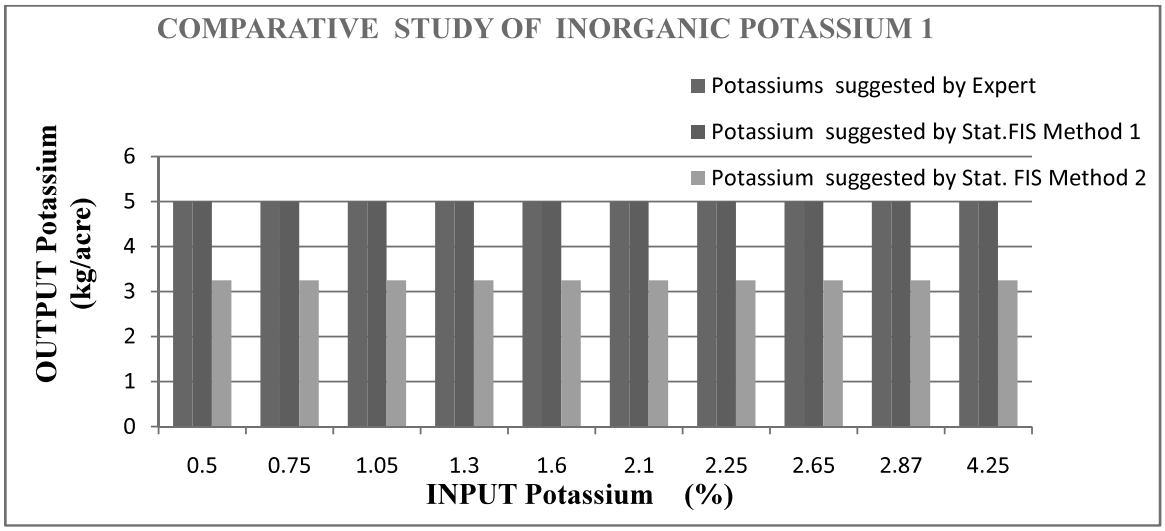


Fig.7b: Comparative study of Inorganic Potassium 1

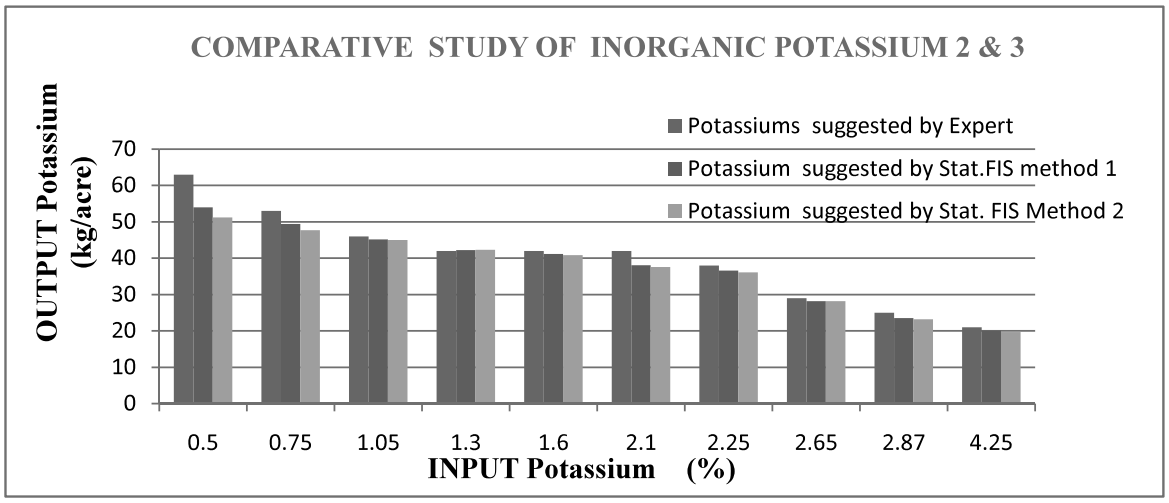


Fig.7c: Comparative study of Inorganic Potassium 2 & 3

CONCLUSION :

Laboratory expert does not inform exact quantity of fertilizer (Potassium) required for the petiole. The FIS system suggests accurate quantity of potassium for the petiole. This reduces the cost of fertilizers and it increases the yield.

REFERENCES :

Bhargava, B.S. and Sumner, M.E., "Proposal for sampling grape (Vitis vinifera L.)
Petiole for nutritional diagnosis", Communications in Soil Science and Plant Analysis", 18**,** pp**.**581-591, 1987.

2. Christensen L. P., A. N. Kasimatis and F. L. Jensen. "Grapevine nutrition and fertilization in the San Joaquin Valley", Calif. Div. Sci. Publication 4087, 40 pp. 1978.
3. Dr. T. V. Sathe, Dept. of Zoology, Shivaji University, Kolhapur. "Grapes and Their Management", Shruti Publication, Jaipur, 2010.
4. "Functions of Potassium in Plants", Better Crops, Vol. 82 (No. 3, 1998).
5. G.J. Klir and B. Yuan, "Logic Fuzzy Sets and Fuzzy," Prentice-Hall of India, New Delhi, ch.1-12, pp.11-338, 1997.
6. G. S. Nijjar, Panjab Agriculture University, Ludhinana, "India Nutrition of Fruit trees", Kalyani Publisher, New Delhi, 1990.
7. "GRAPE NEWS", Publication of National Research Centre for Grapes, Pune Vol. 9 (2), April – June, 2005.
8. Harshavardhan G. Naganur, Sanjeev S. Sannakki, Vijay S Rajpurohit, Arunkumar R "Fruits Sorting and Grading using Fuzzy Logic", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), ISSN: 2278 – 1323, Volume 1, Issue 6, August 2012.
9. Harimranjit Singh and Narinder Sharma, **"**A Review of Fuzzy Based Expert System in Agriculture" INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, ISSN: 2277-9655, Singh, 3(7): July, 2014.
10. Ian Cavers "An Introductory Guide to MATLAB", University of British Columbia, 1998.
11. Ibrahim, "Fuzzy Electronics", PHI Edition, 2000.
12. Kwang H. Lee, "First Course on Fuzzy Theory and Applications". ISBN: 3-540 - 22988- 4, 2005.
13. Monika Sharma, Vibhuti P N Jaiswal, Amit Goyal, "Fruit Detection with Multiple Features using Fuzzy Logic", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 3, Issue 11, November 2013.
14. N. Alavi, "Date grading using rule-based fuzzy inference system", Journal of Agricultural Technology, Vol. 8(4): 1243-1254, 2012.
15. Peacock, W. L., L. P. Christensen, and F. E. Broadbent, "Uptake, storage, and utilization of soil applied nitrogen by Thompson Seedless as affected by time of application", Am J. Enol.Vitic. 40:16-20 (1989).
16. "Production Guide for Organic Grapes", NYS IPM Publication, No. 224, Integrated Pest Management, Cornell University, New York, 2010.
17. Rajeshwar G Joshi, Parag Bhalchandra, Dr.S.D.Khmaitkar, "Predicting Suitability of Crop by Developing Fuzzy Decision Support System", International Journal of Emerging Technology and Advanced Engineering, Volume 3, Special Issue 2, January 2013.
18. R.G. Maiti, "Fruit Crops of India", Kalyani Publisher, New Delhi, 2007.
19. Satyendra Nath Mandal, J. Pal Choudhury, Dilip De, and S. R.Bhadra Chaudhuri "Roll of Membership functions in Fuzzy Logic for Prediction of Shoot Length of a Mustard Plant Based on Residual Analysis", World Academy of Science, Engineering and Technology, Vol.14, 2008.
20. Winkler A. J., J. A. Cook, W. M. Kliewer, and L. A. Lider. "General viticulture". University of Calif. Press, Berkeley and Los Angeles, 710 pp. 1974.

A PEEK INTO COMPUTER GENERATED IMAGERY PIPELINE

PROF. ANKIT JAIN
Assistant Professor, Suryadatta Group of Institutes
G-1, Vimal Classic Building, Lane No. 13Q, Tingre Nagar, Pune
a12345.ankit@gmail.com / ankit333@yahoo.com
MRS. NISHITA JAIN
Teaching Assistant, Symbiosis Center of Media & Communication
G-1, Vimal Classic Building, Lane No. 13Q, Tingre Nagar, Pune
nishitawassan@gmail.com

Abstract : Computer Graphics or Computer Generated Imagery or CGI is a subfield of Computer Science. CGI produces fascinating and jaw dropping results. At times CGI diminishes the thin line between reality and virtuality , as it gets difficult to identify the original from CGI. The jaw dropping effects however require a lot of efforts and follow a strict process and procedure. This paper tries to broadly explain the process and procedure commonly known as pipeline of CGI.

Keywords- Computer Generated Imagery (CGI), 3D, Digital Image, Computer Graphics (CG)

INTRODUCTION:

A sub-field of computer science responsible for displaying images, graphics and films created using computers is known as Computer graphics. Computer graphics require specialized hardware and software to create data. In 1960, Verne Hudson and William Fetter of Boeing coined the term Computer Graphics (CG), also referred as Computer-Generated Imagery (CGI). Computer Graphics consists of UI Design, Vector Graphics, 2D & 3D Models. Science of optics, geometry and physics is the overall methodology for computer graphics.

Computer Generated Imagery

With the evolution of mankind, time has changed and new technologies have evolved. Changes in technology over time has led to digitalization and digitalization led to development of cameras, computers, smart phones and virtual reality. As we move into digital world we see a lot of pictures, images, films, graphics, etc. These images, pictures, etc., that we see are either generated through digital medium or converted to digital medium through digital image processing. We see these digital images every now and then and with the advancement of technology it is becoming difficult to identify and differentiate between the original and computer generated images just by looking at them. Computer Graphics or Computer Generated Imagery as we know it consists of 3D & 2D Graphics, Vectors, UI Designs, etc., and is widely used in animation, movies, advertising, video games and graphic design. Computer graphics can be subdivided into Geometry, Animation, Rendering, Imaging and Topology depending on the steps involved in the generating computer graphics. CGI requires a prescribed workflow.

Workflow

CGI follows a particular workflow. After completing the shoot or obtaining the original shoot the actual workflow starts. CGI workflow can be broadly classified into three parts namely

1. Pre Production
2. Production
3. Post Production

1. Pre Production

Pre Production also known as the Planning Stage. During pre production all the elements of CGI are defined and decided after obtaining the original footage on which CGI is required. As planning for CGI is done before shooting begins, the plans are handed over to the pre production department and then they further decide their requirements as per the desired shots. References are collected and desired characters/objects/effects are planned. From the references collected, characters/objects are designed. Character sheet and turnaround is created to proceed to the second stage. During pre production, match moving the footage and generating cloud data is also done.

2. Production

Production also known as the Creation Stage. During production all the elements of CGI, defined during planning stage are created. Production can further be divided into

a. Modeling
b. Texturing
c. Rigging
d. Animation
e. VXF
f. Lighting
g. Rendering

Modeling

Modeling is the study of three-dimensional objects digitally. Appearance of an object depends on its exterior, silhouettes and boundaries are used to represent the objects. Representing object with uses basic geometrical shapes is done in this stage. The representations are Lagrangian & Eulerian.

Geometry can further be classified into

a. Implicit surface modeling
b. Digital geometry processing
c. Discrete differential geometry
d. Point-based graphics
e. Subdivision surfaces
f. Out-of-core mesh processing

Texturing

Texturing is the study of look and feel development. During this stage the models or geometries created are developed and given a look and feel depending on the requirement of the scene as per desired CGI. Textures can be realistic or semi realistic depending on the requirement of the shot.

Rigging, Animation & VFX

Rigging is the art of making models ready for movements. Animation studies the ways to represent and manipulate motion. Movement or deformation on surface of object over time is described in animation. Physical simulations have become popular over parametric and data driven models with the advancement of computers and powerful computing.

Animation can be subdivided into

a. Performance capture
b. Character animation
c. Physical simulation (e.g. cloth, fluid, dynamics, etc.)

Lighting & Rendering

Generating images from a model is known as rendering. In layman's term, rendering is converting models generated from software to an image which can be viewed easily. Rendering requires light to create realistic image. Non-photorealistic rendering can also be done for artistic style. Realistic rendering requires transport and scattering.

Transport

Transfer of illumination from one place to another is known as transport. As illumination transfers it creates visibility. Transport is travelling of light from one place to another by means of direct or indirect illumination. It is necessary to light up the scene and create visibility in order to render a realistic scene.

Scattering

Scattering can be defined as the relationship between the incoming and outgoing light at a given point on the object. Scattering is how the light reacts with the object's surface at a given point. Scattering depend on the material and shading of the surface. Shading defines how material properties vary across the surface of the object. Shading and scattering describe the appearance of the object. Scattering is usually described in terms of a bidirectional scattering distribution function or BSDF. BSDF addresses scattering function application on the surface with the help of a shader.

Types of Rendering

a. Physically based rendering
b. Real time rendering
c. Non-photorealistic rendering
d. Relighting

Imaging

Study of image acquisition and image editing is known as imaging. Rendering builds an image on the display screen. However the image built is still in the software and hence acquiring the image in correct format and extension is very important. All the required details and metadata needs to be acquired from the software. Post acquiring image, necessary corrections and editing is required to match the image to the required scene. This process of acquiring and editing the image is known as Imaging.

3. Post Production

Post Production also known as the Finishing Stage. During post production all the elements of CGI rendered during the production stage are further obtained and composited using layers rendered from CGI. After compositing layers, the final image is further edited to match the look and feel of original image. After matching the two footages scene is further matched and lined up with the proceeding scene to check for errors.

Software

As discussed earlier CGI requires a few software and hardware. Software requirements of CGI can broadly be classified into three parts

Pre Production Software

Pre production is more of a manual work where collection of references is done, characters and props are designed and production for CGI is planned. It mostly requires artistic approach and manual work. However, software usage is minimal. Software required for pre production, but not limited to, are

- Storyboard Pro
- Harmony
- Flash
- Photoshop
- Illustrator

Production Software

There are various 3D Production software available for usage. Some are free software and some are professional paid software. To name a few

• Daz Studio	Wings 3D	SketchUp
• Blender	Sculptris	TodMod
• Maya	Modo	Fuse
• LightWave 3D	Poser	Cinema 4D
• 3ds Max Design	LuxRender	Maxwell
• Mitsuba	Octane	V-Ray
• Mental Ray	Iray	Lumion
• KeyShot	Corona	Silo
• ZBrush	Mudbox	MatchMover
• PFTrack	Boujou	

Post Production Software

Post production being one of the crucial part in CGI, software and hardware requirement is also very crucial. Software required for post production are

• After Effects	Fusion	Nuke
• Flame	Smoke	Shake
• Natron	Motion	Combustion

CONCLUSION :

CGI or CG is a very vast and intense field of computer science. It requires a lot of specialized study and knowledge of various software. CGI requires and follows a particular workflow and any deviation from workflow may cause undesired results and untimely delay in work. Pro Production defines the flow and work required to create desired CGI. Production process actually creates complete CGI and effects and Post Production gives the finishing touch by integrating it with the original shoot.

REFERENCES :

Web links

1. https://en.wikipedia.org
2. www.blenderguru.com

DIAGNOSIS OF HERPES(NAGIN) ON EYE THROUGH IMAGE PROCESSING AND MACHINE LEARNING

KAKASAHEB RANGNATH NIKAM
Dept. of Computer Science, HPT Arts and RYK Science College, Nashik
krnikam05@gmail.com

Abstract: *Automatic eye detection is very important to diagnose various eye disease in early stages. Like Diabetic Ratinal(DR) effect and Glaucoma, other dangerous disease is Herpes(Nagin). Image mining is the process of searching and discovering valuable knowledge from data. Image processing is having significance for disease detection on infected eyes. With help of image processing and machine learning techniques, it is possible to automate and/or assist ophthalmologist in diagnosis. This paper describes various image processing and machine learning techniques for detection of eye diseases. It gives information about Herpes(Nagin), how it affects.*

Keyword: Herpes(Nagin), keratitis, ophthalmologist, Machine learning, Image Processing.

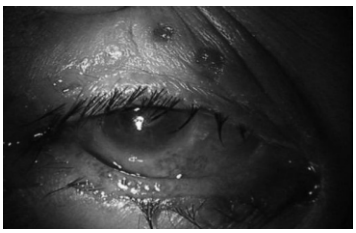
INTRODUCTION:

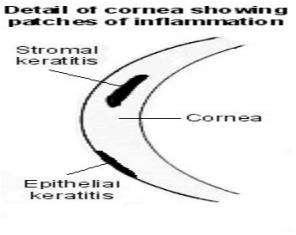
Clinical studies generally undertaken through examination of eyes for various clinical measurements. Unlike kidney or lungs, both eyes are easily examine for diagnosing diseases.

Research shows "the findings in the left eye are generally likely to be more similar to those in the right eye of the same individual" [1].

Herpes(Nagin) "is a virus that causes cold sores, that also causes eye infections. This virus lives inside the nerves in your face and can travel down the nerves to your eye." It may damaging your eye and causing permanent eyesight problems [2].

The purpose of paper is to find and study the ways available to detect Herpes(Nagin)
in early stages by applying image processing techniques.





Body of Paper:

If ophthalmic data is properly collected and analysis is done by applying Deep Learning it provide a way to detect Herpes(Nagin) in early stages.

How my eye affect by Herpes(Nagin) virus?

The cornea(front part of the eye) is infected first is called as Keratitis. i.e. infection of Cornea. When infection is in superficial layer(top) of cornea, is called epithelial keratitis.

But if dipper layer of cornea is involved, is called stromal keratitis.

Impact after Nagin(Herpes): Vision is good enough to drive in about 9 in 10 eyes. However severe and recurrent infections may lead to impairment vision and even severe sight impairment in some cases. Then only option is corneal transplant to restore vision.

Different Image Processing Methods to detect HERPES(NAGIN)

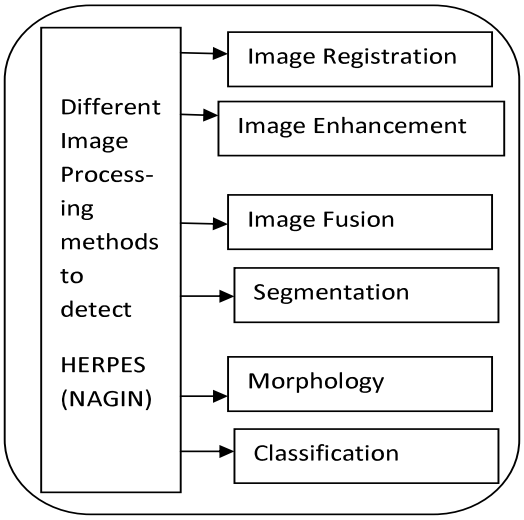
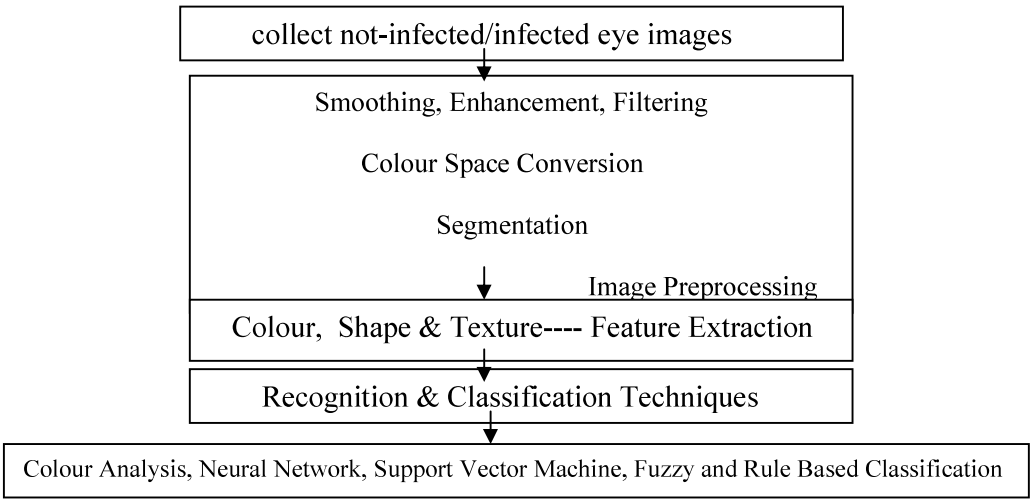


Image Enhancement- Image enhancement includes varying brightness and contrast of image. It also includes filtering and histogram equalization. It comes under pre- processing step to enhance various features of image.

Image Registration- Image Registration is an important technique for change detection in retinal image diagnosis. In this process, two images are aligned onto a common coordinate system. Images may be taken at different times and with imaging devices In medical diagnosis, it is essential to combine data from different images and for better analysis and measurements images are aligned geometrically.

Image Fusion- Image fusion is a process of combining information acquired from number of imaging devices. Its goal is to integrate contemporary, multi sensor, multi-temporal or multi-view information into a single image, containing all the information so as to reduce the amount of information.

Following block diag. shows image processing system with steps.



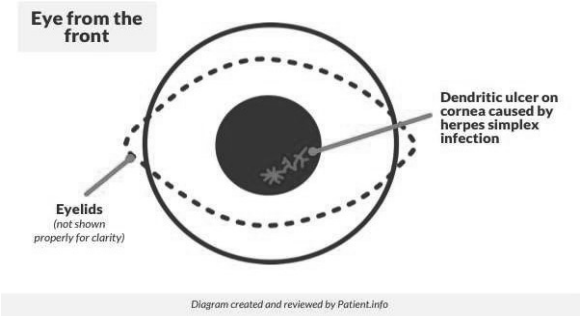
Which people are infected by Herpes(Nagin) on eye?

* Generally it affects about 2 to 3 person in 1000.
* People those suffered from cold sores.
* Adults of age between 30 to 40 years.

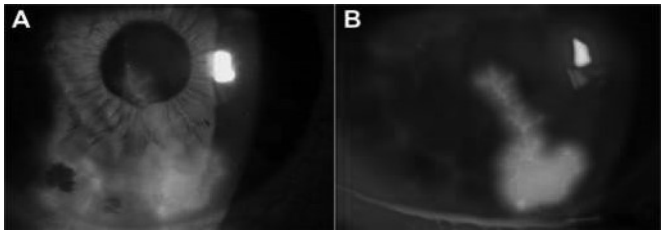
How the Herpes(Nagin) infection is diagnosed?

Generally our family doctor examine our eyes with a magnifier. She/he may put stain on the front of our eye. It shows any irregular areas on the transparent front part of the eye.

Due to Herpes(Nagin) Simplex infection, shows small scratch on the cornea. There is ulcer develop which called as dendritic ulcer. Dendritic means Branching.



As shown in figure, ulcer is not round with a smooth edge, but like a tree with many finger-like branches.



We must urgently referred to an ophthalmologist(eye specialist). They immediately start appropriate treatment.

MACHINE LEARNING[8] techniques KNN, SVM, HMM are applicable here.

a) KNN: **K - Nearest Neighbor**[7] is a kind of instance-based learning, where the function is only locally approximated and all computation is referred until classification[6]. This technique is called lazy learning because, it does not need any training or minimal training phase. All the training data is needed only during the testing phase and this technique uses all the training data so that if we have a large data set then we need special method to work on part of data which is the algorithmic approach[6]. Although classification is the primary application of KNN, we can also use it for density estimation also. The k-nearest neighbor algorithm is one of the simplest algorithm of all machine learning algorithms.[7] KNN classification[7] was formulated from the requirement to perform several analysis when reliable parametric estimates of probability densities are not known or difficult to determine.

b) SVM (Support Vector Machine): In ML support vector machines (SVMs also referred as Support Vector Networks) are supervised learning models with correlated learning algorithms that learns data and determines patterns, used for regression and classification analysis[7]. Given a set of training examples, each marked as referring to one category for one of two categories, an SVM training algorithm creates a model that divides new examples into one category or the other devising it a non-probabilistic binary linear classifier[6]. An SVM model is a representation of the example as points in space assigned so that examples of the different categories are divided [6]. In addition to performing linear classification, SVMs can expeditiously perform a nonlinear classification using the trick called the kernel trick, implicitly mapping their into high-dimensional feature spaces.

c) HMM (Hidden Markov Model): An embedded HMM [6] - based approach for face recognition and detection uses an effective set of observation vectors gained from the 2D-DCT coefficients. The embedded HMM can sculpture the two dimensional data finer than the one-dimensional HMM and is computationally less difficult than the two-dimensional HMM [6]. This model is well suited for face images since it exploits important facial characteristics, structure of "states" inside each of these "super states".

CONCLUSION :

Herpes on eye make blind single infected eye. Different ML techniques like KNN, SVM, HMM applicable here with given image processing steps including Colour Analysis, Neural Network, and more. Which potentially reduce workload on doctors and increase the efficiency of limited healthcare resources. Various methods used by different scientists and constant research happening in this field. Here is an attempt to learn and understand some of the techniques used till now for diagnosis of eyes using image processing and ML. It is possible to apply Feed Forward Back Propagation Neural Network for detecting herpes in coming future.

REFERENCES :

[1] People and Eyes: Statistical approaches in ophthalmology I E Murdoch, S S Marris, S N Cousens.

[2] Eye Infection Herpes Simplex, Author by Dr. Oliver Starr 23 Nov 2017.

[3] Researchers published by Byers Eye Institute at Stanford University.

[4] a study published online in *Ophthalmology*, the journal of the American Academy of Ophthalmology.

[5] Rishab Gargeya, Theodore Leng. Automated Identification of Diabetic Retinopathy Using Deep Learning. *Ophthalmology*, 2017.

[6] A Study on Retinal Disease Classification and Filtration Approaches. Parul, Neetu Sharma.

[7] An Improved k Nearest Neighbor Classifier Using Interestingness Measures for Medical Image Mining. J. Alamelu Mangai, Satej Wagle, and V. Santhosh Kumar.

[8] Review of Image Processing and Machine Learning Techniques for Eye Disease Detection and Classification Mr. Langade Umesh, Ms. Malkar Mrunalini, Dr. Swati Shinde.

THE COMPARATIVE STUDY OF DIFFERENT TECHNIQUE TO FIND SOFTWARE COST ESTIMATION

SHIVENDU BHUSHAN
Indira College of Commerce and Science
Shivendu@iccs.ac.in
SUMIT S. SASANE
Indira college of commerce and science.
sumit.sasane@iccs.ac.in
NADAF B.G
Indira College of commere and Science College.
nadaf@iccs.ac.in

Abstract : The software cost estimation is most important factor in project management. Accurate cost estimation helps us complete the project within time and budget. There are different techniques to find the software cost estimation depending on different factors that affect the cost of software. Cost estimation process is explained in this paper. This paper showing the comparison of various software cost estimation methods and some cost estimation models that can mostly use for the software cost estimation projects. The benefits and drawbacks of the existing cost estimating techniques have been highlighted in this paper. There is lot of research to be done on experience based metrics where domain knowledge and skills are required. Todays need to develop new cost estimation models.

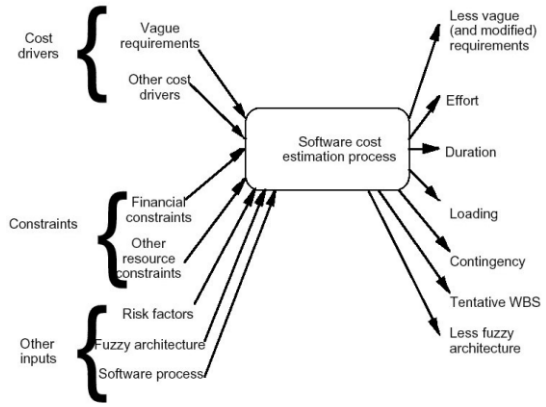
Keywords: Non-algorithmic techniques, algorithmic techniques, LOC, Artificial neural network ,COCOMO,FPA

INTRODUCTION:

In modern days according to requirement the size as well as complexity of the project software is increasing. For that purpose planning, requirements time management and other factors are to consider is very important. It is difficult to say that which method is useful for finding the cost of software because according to different parameter different method of cost of estimation. Still their goal is to develop software which is cheap and at the same time deliver good quality. At first we estimate software size, then the needed effort after this we derive the schedule and at last calculate cost of the software. Various techniques are used in software cost estimation and we can broadly classify these techniques into two categories namely Algorithmic and Non-Algorithmic techniques.

2. Cost Estimation Process:

Generally we find the software cost estimation but It is very important that to understand what is software cost estimation process. Definition of software cost estimation process is a set of techniques and procedures that is used to derive the software cost estimate.



3. Software Cost estimation:

Generally, there are many methods for software cost estimation, which are divided into two groups: Algorithmic and Non-algorithmic. Using of the both groups is required for performing the accurate estimation. If the requirements are known better, their performance will be better.

3.1 Non Algorithmic methods

These models first compare the project under consideration with the previously done projects by the organization and analyses the information from the most similar projects to make the cost estimates. Basically, these methods make use of the past experiences.

3.1.1 Analogy-based estimation

The basic idea behind estimation by analogy is that whenever we get a new software project for cost estimation, it is just compared to historical similar projects to arrive at the nearest similar software project through which we can estimate our current project cost. The values and data from previously complete projects are deduced to calculate cost of our current project. We can use this technique both at system or component level. The details of the exact steps followed for estimation based on Analogy are given below:

i. Determining the attributes of our current project.
ii. Finding a historical similar project as compared to our current project whose attributes are already stored in the database.
iii. Calculating the cost of the current project from the historical similar project.

3.1.2 Expert judgement method

Estimation based on Expert judgement captures the knowledge of experts and the estimation of cost is dependent on those projects which involved the inclusion of the expert.

Usually there are some scenarios when we have limitations to gather and find data. Expert Judgement method is good to be used in these situations. It is the widely used estimation strategy for software projects. Wideband Delphi technique is one of the cost estimation technique based on expert judgement. Here the participants are involved in two assessment rounds. Work breakdown structure is another example of the expert judgement method.

3.1.3 Top-down estimation

The top down estimation method also known as macro model, considers effort as a function of size of the project. That is, EFFORT = a*SIZEb , where a and b are constants. At first, an overall cost is estimated, the project is then partitioned into various levels and the cost estimation of every level of the project is derived from the global properties of the software project. The overall cost estimation of the project makes it very easy to estimate costs at the start, however, one needs to revise the initial estimates as the project progresses, which leads to delays if the revisions lead to varying results from the earlier estimates. Due to the fact that very little detailed information is available at the start, this method is highly regarded in early cost estimation

3.1.4 Bottom up:

This is the exact opposite of the top down approach. In this method, we first estimate the cost for each and every small components of the project, which is then combined to the cost of **overall project. It aims to consolidate the small information available and how they interact to arrive** at the overall cost. COCOMO method uses this approach for cost estimation. Although a much consolidated technique, bottom up cannot be applied to projects where much detail is not known during the start of the project. Trying to apply bottom up in these situations can lead to bad estimations.

3.1.5 PRICE TO WIN:

Here we are focused more on the budget of customer rather than the functionality of the software. Overall software cost is agreed on the basis of an outline proposal and the development of software is restricted by that cost.

3.1.6 Artificial neural network based estimation

Artificial Neural Network based estimation models are trained by the use of historical data. They produce good results and the algorithmic parameter values are adjusted in such a way that the differences between the actual and predicted estimates are reduced.

3.1.7 Fuzzy logic based estimation

Fuzzy Logic based estimation is also called soft computing technique. Soft computing techniques are emerging software estimation techniques. Fuzzy logic has evolved as an important tool to solve such problems, for which mathematical models cannot be created or we can say that it is difficult to create. Development of software is many a times characterized by parameters that exhibit fuzziness. Application of fuzzy logic in cost estimation helps in overcoming many problems which exists in the already available cost estimation techniques.

3.2 Non Algorithmic methods

In these methods, software cost estimation is provided using mathematical equations. A lot of research is done on historical data and inputs such as skill levels, risk assessments, the number of functions to perform source lines of code etc. in order to arrive at these mathematical equations. These methods have developed a lot of commendable models such as Putnam model, COCOMO model, and function points based models.

3.2.1 Function point Analysis

Measuring software size in terms of line of code in analogous to measuring a car stereo by the number of registers, capacitors and integrated circuits involved in its production. At first, Alan Albrecht (1983) presented function point metric to measure the functionality of project, which appeared to be a solution to the size measurement problem to measure the functionality of project. In this method estimation is done by determination of below indicators

User Inputs: - Information entering the system.
User Outputs: - Information leaving the system. Logic
Files: - Information held within the system.
Enquiries: - Request for instant access to information.
Interfaces: - Information held by other system that are used by the system being analyzed.

Table 1: Functional Units and Weighting Factors

Functional units	Weighting Factors		
	Simple	Medium	Complex
User Inputs	3	4	6
User Outputs	4	5	7
Files	3	4	6
Enquiries	5	6	8
Interfaces	7	6	10

At first, the number of each mentioned indicators should be tallied and then complexity degree and weight are multiplied by each other. Generally, the unadjusted function point count is defined as below:

$UFC=\sum\sum NijWij \qquad i=1 \text{ to } 5, j=1 \text{ to } 3$

Where Nij= is the number of indicator i with complexity j

Wij= is the weight indicator i with complexity

According to the previous experiences, function point could be useful for software estimations because it could be computed based on requirement specification in early stages of the project to compute the FP, UFC should be multiplied by the technical complexity factor which is obtained from the component in the table 2

Table 2: Technical Complexity Factor Component

F1	Reliable back-up and recovery	F5	Complex interface	F9	Performance
F2	Reliable back-up and recovery	F6	Reusability	F10	Online Data Entry
F3	Heavily used configuration	F7	Multiple rites	F11	Complex Processing
F4	Operational ease	F8	Data Communication	F12	Online Update
F12	Installation ease	F14	Facilitate Change		

Each component can change from 0 to 5 and 0 indicate that the component has no effect on the project and the component is compulsory and very important respectively. Finally, the TCP is calculated as:

$$TCF=0.65+0.01(\sum(Fi))$$

The range of TCF is between 0.65 and 1.35. Ultimately, function point computed as:

$$FP=UFC*TCF$$

3.1.1 Constructive Cost Estimation Model:

COCOMO (Constructive Cost Estimation Model) was proposed by Boehm [1981]. According to Boehm, software cost estimation should be done through three stages: Basic COCOMO, Intermediate COCOMO, and Complete COCOMO.

Basic COCOMO Model

The basic COCOMO model gives an approximate estimate of the project parameters. The basic COCOMO estimation model is given by the following expressions:

$$Effort = a1 \times (KLOC)^{a2} \quad PM$$

$$Tdev = b1 \times (Effort)^{b2} \quad Months$$

Where

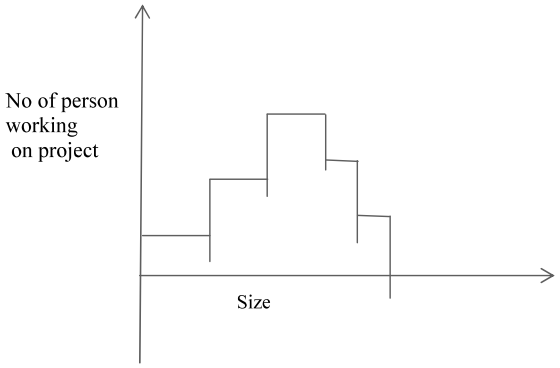
KLOC is the estimated size of the software product expressed in Kilo Lines of Code,

a1, a2, b1, b2 are constants for each category of software products.

Tdev is the estimated time to develop the software, expressed in months,

Effort is the total effort required to develop the software product, expressed in Person- months

The effort estimation is expressed in units of person-months (PM). It is the area under the person-month plot (as shown in fig. 11.3). It should be carefully noted that an effort of 100 PM does not imply that 100 persons should work for 1 month nor does it imply that 1 person should be employed for 100 months, but it denotes the area under the person-month curve (as shown in fig. 1.1).



According to Boehm, every line of source text should be calculated as one LOC irrespective of the actual number of instructions on that line. Thus, if a single instruction spans several lines (say n lines), it is considered to be nLOC. The values of a1, a2, b1, b2 for different categories of products (i.e. organic, semidetached, and embedded) as given by Boehm [1981] are summarized below. He derived the above expressions by examining historical data collected from a large number of actual projects.

Estimation of development effort

For the three classes of software products, the formulas for estimating the effort based on the code size are shown below:

Organic : $Effort = 2.4(KLOC)^{1.05}$ PM

Semi-detached : $Effort = 3.0(KLOC)^{1.12}$ PM

Embedded : $Effort = 3.6(KLOC)^{1.20}$ PM

Estimation of development time

For the three classes of software products, the formulas for estimating the development time based on the effort are given below:

Organic : $Tdev = 2.5(Effort)^{0.38}$ Months

Semi-detached : $Tdev = 2.5(Effort)^{0.35}$ Months

Embedded : $Tdev = 2.5(Effort)^{0.32}$ Months

some insight into the basic COCOMO model can be obtained by plotting the estimated characteristics for different software sizes. Fig. 1.2 shows a plot of estimated effort versus product size. From fig. 11.4, we can observe that the effort is somewhat superlinear in the size of the software product. Thus, the effort required to develop a product increases very rapidly with project size.

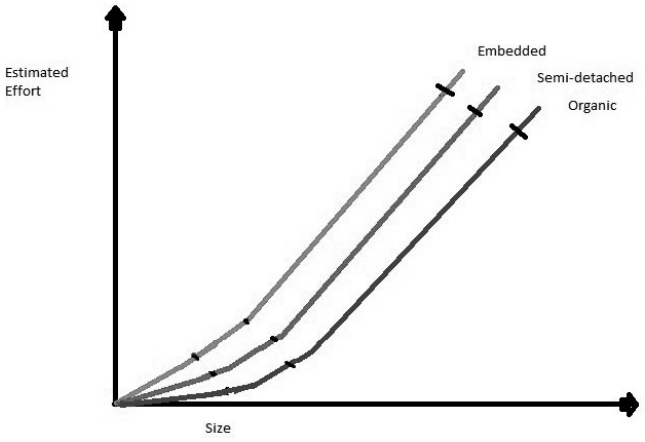


Fig. 1.2: Effort versus product size

Putnam model

3.1.1 :

Another popular software cost model is the Putnam model. The form of this model is:

Technical constant $C = size * B^{1/3} * T^{4/3}$

Total Person Months $B = 1/T^4 * (size/C)^3$

T= Required Development Time in years

Size is estimated in LOC

Where, C is a parameter dependent on the development environment and it is determined on the basis of historical data of the past projects.

Rating: C=2,000 (poor), C=8000 (good) C=12,000 (excellent)

The Putnam model is very sensitive to the development time, decreasing the development time can greatly increase the person-months needed for development.

An estimated software size at project completion and organizational process productivity is used. Plotting effort as a function of time yields the TimeEffort Curve. The points along the curve represent the estimated total effort to complete the project at some time. One of the distinguishing features of the Putnam model is that total effort decreases as the time to complete the project is extended. This is normally represented in other parametric models with a schedule relaxation parameter.

This estimating method is fairly sensitive to uncertainty in both size and process productivity.

4. COMPARISON BETWEEN VARIOUS MODELS

Sr. No.	Method	Type	Advantages	Disadvantages
1	Analogy-Based Estimation	Non-Algorithmic	Work based on the past experiences and knowledge base. No new resources or experts required.	A large amount of information about the project is required and sometimes similar project data might not be available.
2	Expert based Judgement	Non-Algorithmic	Experts bring experience to the project and can make the process faster	Cannot be quantified. Difficult to document the factors used by experts. Chances of biased decisions.
3	Price to Win	Non-Algorithmic	Often wins the contract.	May cause delay of delivery or force the development team to work overtime.
4	Bottom-Up	Non-Algorithmic	Stable as the estimation errors in the various components might balance out	Time consuming, might be inaccurate due to shortage of information, not feasible with limited resources.
5	Top-Down	Non-Algorithmic	Doesn't miss the cost of system level functions, faster and easier	Overlook low-level costs, no justification for decisions, less stable
6	Putnam Model	Algorithmic	Efficient for very large projects	Uncertainty in the software size may result in inaccurate cost estimations.
7	COCOMO	Algorithmic	Repeatable results can be generated, easily modifying input data, easy refining and customization of formulae, clear results.	Inaccurate cost estimate as the size is uncertain at beginning. A lot of data required, not suitable in practice.
8	Function Point	Algorithmic	Estimation based on requirements or design specifications, tool independent	Hard to implement, not considered good enough.
9	Neural Net	Machine Learning	Superior cost estimate, consistency.	Training data required, no standard guidelines.
10	Fuzzy logic	Machine Learning	Flexible, no training needed.	Difficult to use and understand the concept.

CONCLUSION :

We saw various techniques for cost estimation which have their own advantages and disadvantages. It is very hard to point out a failsafe method, as most of the cost estimation is scenario dependent. Some methods are highly accurate but cumbersome to implement, while some only work best on certain known parameters. In order to pick a method for any given project, a proper analysis of the project is much needed. A wrong estimation technique can significantly delay a project, and a right technique can make a project breeze through its deadlines. A detailed analysis of global factors must be made, or else, it becomes very difficult to give correct deadlines. Also, an emphasis on smaller details is necessary, as they can cause delays when they add up together.

REFERENCES :

1. http://www.computing.dcu.ie/~renaat/ca421/LWu1.html
2. http://www.tandfonline.com/doi/full/10.1080/20550340.2016.1154642
3. http://smallbusiness.chron.com/cost-estimation-methods-accounting-52025.html
4. https://www.project-management-skills.com/project-cost-estimating.html
5. http://moodle.autolab.uni-pannon.hu/Mecha_tananyag/szoftverfejlesztesi_folyamatok_angol/ch13.html

REAL TIME FACE DETECTION AND TRACKING SYSTEM IN PUBLIC OR COMMERCIAL PLACES

DEEPAK S DHARRAO
Research Scholar, JJTU, Rajasthan
Email:- deepakdharrao@gmail.com
NILESH J UKE
Research Guide, JJTU, Rajasthan.
Email:- nilesh.uke@gmail.com

Abstract : Real time face detection and tracking with PTZ camer has been widely applicable in public or commercial places for surveillance scenarios in this days. After detecting a face in real time tracking is an important task for which we are going to use PTZ camera Features in this proposed system. The PTZ camera in the System is supposed to trace and recognise an human being which may considered as a threat or we can identify a person for commercial aspect also for e.g to identify a old customer. The main objective is that the camera must be capable to detect and keep track of a target objects. Even though the place is crowded the system should track the target object efficiently.in this this we are going to use an ASIFT algorithm which is complements SIFT by simulating the two parameters. Which is used to model camera optical axis direction so that it covers all six parameters while SIFT covers only four parameters.

Keywords – ASIFT, SIFT, PTZ control, Face Detection, Object Tracking.

INTRODUCTION:

Object detection and tracking is one of the most significant task of automated video surveillance system. The importance of automatic object tracking comes from the fact that it has a wide range of real time applications, including surveillance and monitoring of human behaviour in public or commercial places, residential areas, smart rooms, hand gesture recognition, etc. The main problem in detection process is the influence of uncontrolled and dynamic environment. In the outdoor scenario it completely differ for the day and night vision of the camera. The viola-jones algorithm [1] used for face detection explicitly from last decade, but it cannot be used for object detection apart from the human face. The algorithms to detect and track the object varies from the color, shape, appearance and structure of the object[15]. Object Tracking is one of the most important task of automated video surveillance system. For efficient tracking system it is need to use efficient camera which should have Pan, Tilt and Zoom feature. Efficient tracking can only possible if the detection and recognition algorithm should be that much effective. In proposed system we are going to use effective background subtraction and noise removal technique as the part of pre-processing method. Also we are going to use A fully affine invariant image matching algorithm (ASIFT) needs to cover the all six affine parameters. The SIFT method covers only four parameters by normalizing rotations and translations, and simulating all zooms out of the query and of the search images.

2. Background

In video surveillance system object detection takes place at lower level of processing while tracking takes place at higher level of application [3]. P. Kamavisdar, S. Saluja, and S. Agrawal elaborated a survey on different method of image classification. Their paper aims to classify single object image using an efficient technique[17]. Sometimes object tracking involves tracking of a single interested object and that is done using normalized correlation coefficient and updating the template. Object tracking is an important job within the field of computer vision. Linear Discriminant Analysis (LDA), Symbolic Approach, Symbolic Representation are some of the techniques for supervised classification(18). The aim of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video. But tracking has two definition one is in literally it is locating a moving object or multiple object over a period of time using a camera [1]. Tracking is the problem of estimating the trajectory or path of an object in the image plane as it moves around a scene. A comprehensive survey on intra-camera tracking algorithms is found in [3] and it can be classified into two categories in terms of tracking strategy: deterministic and probabilistic tracking.

The tasks of detecting the object and establishing a correspondence between the object instances across frames can either be performed separately or jointly. Detection and recognition of continuous activities from video is a core problem to address for enabling intelligent systems that can extract and manage content fully automatically. Recent years have seen a concentration of works revolving around the problem of recognizing single-person actions, as well as group activities [7]. Firsttly object is detected detected and recognised for finalizing the target object.In the next stepe, the object region and correspondence are jointly estimated by iteratively updating object location and region information obtained from previous frames. There are different methods of Tracking like kernel tracking, Silhouette Tracking, Counter based object Tracking, Feature based object Tracking [8].

It gives us an accurate shape description of the target objects. The goal of silhouette tracker is to find the object region in each frame by means of an object model generated using the previous frames. There are two types of Silhouette Tracking, first is the Shape matching and second one is the Contour tracking[16].

It plays a vital role to select a proper feature in tracking. Therefore, the feature selection is closely related to the object representation. For example, color is used as a feature for histogram appearance representations, the contour-based representation, the object edges are usually used as features [4]. Basically, many tracking algorithms use a combination of

these features. There are various details of common visual features and they are as follows:

Color-Color of an object is influenced by two factors. These are mainly based on Spectral power distribution of the illuminant and Surface reflectance properties of the object. Different color models are RGB, L*u*v and L*a*b used to represent color.

Edges-Edge detection is used to identify strong changes in image intensities generated by object boundary. Edges which are less sensitive to illumination changes are compared with color features. Most popular edge detection approach is Canny Edge detector.

Optical Flow-It is defined as a dense field of displacement vector which defines the translation of each pixel in a region. It is calculated using the brightness constraint, through them they assume brightness constancy of corresponding pixels in consecutive frames. Optical Flow is basically used as a feature in which motion is based on segmentation and tracking application.

Texture-Texture is used in the measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity. It requires a processing step to generate the descriptors. There are different kind of texture descriptors: They are as follows : The Gray-Level Co-occurrence Matrices, The loss texture measures, The wavelets and The steerable pyramids. E.

3. Design of Proposed System

This paper aimed to design a Real-time object detection and automated tracking system. The objects in this proposed system we are considering a human face. Our primary objective is a tracking of a target object with ease with efficient recognition and detection system. Once the object is detected and successfully classified then the tracking of the target is done. Here we are going to provide a phases of object detection and tracking system diagram of proposed system.

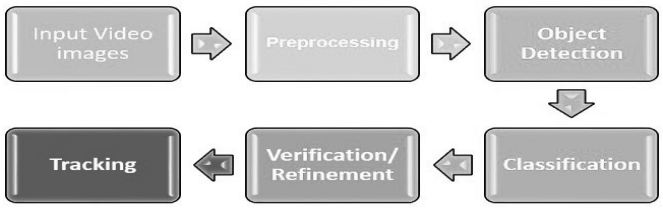


Figure 1. *Phases of object detection and tracking system*

A. Input Image

The first stage of any computer vision system is the image acquisition stage. In real time surveillance system the image is taken from video stream. After the image has been obtained, various methods of processing can be applied to the image to perform the many different vision tasks required today. However, if the image has not been acquired acceptably then the proposed tasks may not be achievable, even with the aid of some form of image enhancement. PTZ cameras are widely prominent to track and gain high resolution object and face image at a distance. The image resolutions differ according to camera used. But typically, 720×486 is considered to be sufficient for PTZ views. Image acquisition and pre-processing can be done using library functions based on OpenCV and MALAB[2].

B. Object Detection and Classification

It is the process firstly we have to find out the moving object in each. For that we are using simple background subtraction method that in video processing the background does not change with the number of frames. So that it is assumed that background is static. To get required foreground object the difference between the object Ok and the background Bk is calculated using the formula,

$$Dk(x, y) = Ok(x, y) - Bk(x, y)$$

Then threshold the difference is calculated[14]. While Noise removal focuses on improving the frame quality. Since pixels with better strength are easier to be processed upon. Target Detection based on GMM model is the technique used in this paper. Considering the moderate cost and requirement and relative good foreground detection result in both grayscale and color video sequence, so Gaussian Mixed Model (GMM) is implemented to target detection in this system.

$$p(x|ë) = Ói=1^{M} wi\, g(x|\mu i, Ói),$$

where x is a D-dimensional continuous-valued data vector (i.e. measurement or features), wi, where i = 1, . . ., M, are the mixture weights, and g (x|μi , Ói), i =1. . . . M, are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form. The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

$$ë = \{wi, \mu i, Ói\} \text{ where } i = 1, \ldots, M.$$

C. Object Recognition:

After noise removal and background subtraction object recognition important task before tracking of the targeted object. It is a technology in the field of computer vision for finding and identifying objects in an image or video sequence. To recognise a human we need to extract a features facial part like eyes, nose, lips etc. the important part of face recognition system is to recognise a face in real time for effective tracking system. In that we have to consider a viewpoints, in many different sizes and scales or even when they are translated or rotated. Objects can even be recognized when they are partially obstructed from view. This task is still a challenge for computer vision systems. Many approaches to the task have been implemented over multiple decades.

- Statistical approaches, which consider patterns as points in d-dimensional space.
- Structural approaches, which use arcs and segments to represent shapes.
- Neural network approaches.
- Hybrid approaches, which combine statistical and structural.

D. Tracking

In the proposed system we are going to use kalman filter technique. The Kalman filter is a recursive predictive filter

that is based on the use of state space techniques and recursive algorithms. Kalman filtering which is composed of two steps [4].

- The prediction step.
- The correction steps.

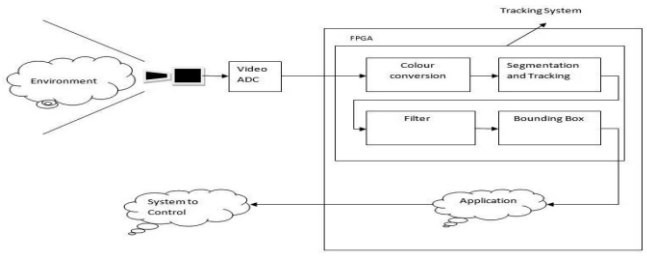


Figure 2 . *Block Diagram Tracking System*

The first step state that it is predicted with the dynamic model. The prediction step is the step which uses the state model to predict the new state of the variables. In this sense it is an optimal estimator. Kalman filter is used in the vision community for tracking.

Also we are going to use A fully affine invariant image matching algorithm (ASIFT) needs to cover the all six affine parameters. The SIFT method covers only four parameters by normalizing rotations and translations, and simulating all zooms out of the query and of the search images.

4. Methodology

A. Efficient Filter for Impulse Noise Removal:

A novel morphological mean (MM) filter is proposed for high-density impulse noise removal. The flow of MM filter consists of two important modules: first one is the noise-free pixel counter (NPC) module and second one morphological pixel dilation (MPD) module. In the NPC module, all pixels of the input image are examined in order to collect both the position and number of the noise-free pixels. Next, the dilatation operation of these pixels is performed to fill into the neighbour noise pixels in the MPD module for best recovery of the image corrupted by a high density of noise.

1. Noise Free pixel counter module

The proposed NPC module aims to detect noise-free pixels and as a result it restores the neighbour noise pixels for images with high densities of noise. The position of the noise-free pixels in the input corrupted image are detected by using Eq. (a).

Steps:

a) Input: Noise Image I.
Output: Number of noise free pixels Ni,
Binary noise mask Bi

b) Initialize Bi ←0

i. Ni ←Number of pixels P

ii. i ←1

c) **for** each pixel x **do**

d) **if** I(x) is noise pixels **then**

e) Bi ←1

f) **else**

g) Bi ←0

h) end if

i) Ni ←Ni – Bi(x)

j) end for loop

2. Morphological Pixel dilation module

After the noise-free pixels of the corrupted image are detected in the NPC module, the detected pixels are then used to replace the eight-connected noise pixels within the scanned window. By doing so, the image corrupted by high densities of noise can be effectively restored[10].

B. Background Subtraction Algorithm

1. The Universal Multimode Background Subtraction:

In this proposed system we are going to use an efficient background subtraction method which uses pixel classification, model update and multiple color spaces technique. The system firstly creates multiple background models of the scene followed by an initial foreground/background probability estimation for each pixel. Next, the image pixels are merged together to form mega-pixels, which are used to spatially de-noise the initial probability estimates to generate binary masks for both RGB and YCbCr color spaces[12]. The masks generated after processing these input images are then combined to separate foreground pixels from the background. It includes the following Steps:

BG Model Selection

Binary Mask (BM) Generation

Binary Masks Aggregation/Fusion

Binary Masks Purging

Foreground Mask

C. Object Recognition:

In this paper object recognition is mainly based upon matching the attributes stored by using the SIFT algorithm. SIFT stands for Scale Invariant Feature Transform. SIFT can be incorporated with OpenCV. SIFT can efficiently identify the object in between the image, using feature matching, rotation, translation, Zoom. Moving objects are characterized by their color-histograms.

D. Object Tracking:

For object Tracking we are going to use kalman filter technique which works in two parts. The first step state that it is predicted with the dynamic model. The prediction step is the step which uses the state model to predict the new state of the variables. In this sense it is an optimal estimator. Kalman filter is used in the vision community for tracking. In this system we are using a fully affine invariant image matching algorithm (ASIFT) which needs to cover the 6 affine parameters. ASIFT complements SIFT by simulating the two parameters that model the camera optical axis direction, and

then applies the SIFT method to compare the simulated images, so that all the 6 parameters are covered. In other words, ASIFT consist of three parameters: the scale, the camera longitude angle and the latitude angle (which is equivalent to the tilt) and normalizes the other three (translation and rotation) [5]. ASIFT can thus be mathematically shown to be fully affine invariant. Against any prognosis, simulating the whole affine space is not prohibitive at all with the proposed affine space sampling[9].

5. Dataset

The datasets employed in this system contain synchronized and static views, captured at a frame rate of 30 fps from un-calibrated camera installed at a corner/wall in an auditorium. Each image frame has a size of 320×230 pixels. A total of 5000 frames (≈ 2-hour in length) from camera view can be used for activity based scene decomposition. Detailed description on dataset is given as follows: A camera placed at the most effective position can cover most of the floor area of the room. The PTZ features extends the vision of the camera and allows the higher success rate of capturing the unidentified events occurred within the range[13]. The current proposed system mainly focuses on an Auditorium where large number of crowd may gather within a very short span of time. The camera must keep a watch at any instance of time. The goal of an objective is to locate out the threatful objects at a run time (The type of object is covered under object-set). As soon as objective 1 gets complete. 2nd Objective (Object Tracking) should be done within the relay.

CONCLUSION :

In this system we are going to use efficient noise removal and background subtraction technique. Which provides a noise free output for the recognition system. The novel recognition system will recognised a face with great accuracy in real time which is very useful for fact predicting the path of targeted object in the frame for tracking purpose. Here we are going to use a PTZ camera which supports for accurate recognising and fast tracking of intended object. This system assures responsive and easy to use surveillance model with a great accuracy and more scope to integrate new and innovative techniques in future.

REFERENCES :

[1] PTZ Design and Implementation of a Detection and Tracking of Moving Object Based on PTZ Camera,2015

[2] Multiple Object Detection using OpenCV on an Embedded Platform.2012

[3] Kumar P, Bindu NS. A Comparative study on object detection and Tracking in video. 2013;2(12):1784–9..

[4] ASIFT: An Algorithm for Fully Affine Invariant Comparison.

[5] New Object Detection, Tracking and Recognition Approaches for Video Surveillance over Camera Network,2015

[6] Online Human Interaction Detection and Recognition with Multiple Cameras,2016

[7] Visual Attention Based Background Modelling for Detecting Infrequently Moving Objects,2016

[8] Salient Point Tracking for Key Frames Selection in Outdoor Image Sequences,2016

[9] Occlusion aware Real time Object Tracking by Integrated Circulant Structure Kernels Classifier,2016

[10] Multiple Sensor Fusion and Classification for Moving Object Detection and Tracking,2015

[11]Face Tracking and Recognition Data Distance: A Coaxial and Concentric PTZ Camera System,2013

[12] Neural Background Subtraction for Pan-Tilt Zoom Cameras,2014

[13] Hierarchical Ensemble of Background Models for PTZ-Based Video Surveillance,2014

[14] Uke NJ, Thool RC. Objects tracking in video: A object-oriented approach using Unified Modeling Language. Int J Comput Vis Robot [Internet]. 2015;5(2):202–16.

[15] Piccardi M. Background subtraction techniques : a review *. 2004;3099–104.

[16] del-Blanco C, Jaureguizar F, Garcia N. An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications. IEEE Trans Consum Electron [Internet]. 2012;58(3):857–62.

[17] Zelnik-Manor L, Irani M. Event-based analysis of video. Proc 2001 IEEE Comput Soc Conf Comput Vis Pattern Recognition CVPR 2001. 2001;2.

[18] Hurdale P, Kamble S. Comparative Study of Face Recognition Techniques. 2014;3(3):926–8.

VIDEO AFFECTIVE CONTENT ANALYSIS : A CONCISE INTRODUCTION

DHAMMJYOTI V. DHAWASE
dhammjyoti.dhawase@gmail.com
Indira College of Engineering & Technology, Pune

Abstract : The ultimate goal of Affective Computing [AC] is to develop systems, which interact with the users in a natural manner [14]. It can be defined as, to produce "emotion aware machines". The study of human feelings represents an interesting on-going topic which involves multidisciplinary expertise including psychology, neurophysiology, and cognitive neuroscience. It has benefits in numerous fields including video retrieval. Some previous algorithm uses video extraction on basis of different audio visual features.
Our work go ahead to extract videos on the basis of emotions or affective state. In accordance to that we go through the survey of different methods, algorithms and techniques of many authors'. In this paper, for physiological and sociological research foundation affective computing area researchers, we give a brief idea of literature review .On the basis of previous and ongoing video content analysis techniques. This should be helpful to the advance research in same area.

Keywords: Affective computing, Video retrieval, content analysis, Emotions

INTRODUCTION:

With the increasing amount of multimedia data, indexing and retrieval is becoming a crowded research area. Lots of research has been done on video structuring, event detection, and semantics modeling. These works try to provide an effective and efficient way to manage and access multimedia databases. Since last some years, researchers have revealed the significance of affective analysis from a personalized media point of view. For example, many users favor a flexible tool to quickly browse the funniest or the most sentimental segments of a movie, as well as the most exciting parts of a sports game video. Compared with traditional video indexing, affective content analysis puts much more emphasis on the audience's reactions and emotions. The affective content of a given video clip can be defined as the intensity and type of feeling or emotion (both are referred to as *affect*) that are expected to arise in the user while watching that clip. Clearly, each way of perceiving video content requires a particular type of information in order to index, classify, filter or organize the video collection correspondingly. In order to get better results in video searching, tagging and retrieval on the basis of state of emotions important phases like input video clips, content analysis, feature extraction, and classification are carried out.

Review of Feature extraction and its mapping to appropriate emotion and classification with the use of efficient classifier/ classifiers should be focused here. Here our aim is that, user can easily, efficiently and within less time access the more appropriate video contents or videos information as per their requirement. As Expressions, Emotions and related feelings are our base of video access, we find out different access models related to it. From a technical point-of-view, Affective Computing outcome is to effectively map features extracted from human signs like physiological signals, behavioral correlates, facial expressions, movements, etc into a well-defined "emotional space"[13].

Literature Review

Video Affective Content Analysis for retrieval or tagging is the very challenging and interested area for researchers. User requirement is an improvement in retrieval speed; avoid duplication of frames and result of only relevant data. To fulfill these requirements many algorithms and technologies are invented by different researchers. These works in different phases like feature extraction, analysis, classification and retrieval of result. A brief Survey of video content analysis is given below as well as we go through the survey of main contents which related to our system.

Video Affective Content Analysis

Esra Acar, et.al. [57], makes the consideration of the ever-growing available multimedia data, annotating multimedia content automatically with feeling(s) expected to arise in users is a challenging problem. In order to solve this problem, author proposed deep learning methods. It consists of particular convolutional neural networks (CNNs) in order to learn mid-level representations from automatically extracted low-level features. They exploit the audio and visual modality of videos by employing Mel-Frequency Cepstral Coefficients (MFCC) and color values in the RGB space in order to build higher level audio and visual representations. Author chooses multi-class support vector machines (SVMs) for classifying video clips into four affective categories representing the four quadrants of the Valence-Arousal (VA) space. Results on a subset of the DEAP dataset (on 76 music video clips) show that a significant improvement is obtained when higher level representations are used instead of low-level features, for video affective content analysis [57].

Min Xu, et, al. [28], In this work, Author focus on comedy and horror films to extract the affective content by detecting a set of so-called audio emotional events (AEE) such as laughing, horror sounds, etc. Those AEE can be modeled by various audio processing techniques, and they can directly reflect an audience's emotion. They use the AEE as a clue to locate corresponding video segments. Their experimental dataset consists of 40-minutes comedy video and 40-minutes horror film. An average recall and precision of above 90% is achieved.

Affective Feature Extraction

The video content can be captured by various visual and audio features. Specifically, the affective content of a video consists of two main categories of data: visual data and auditory data. The visual data can be further divided into visual image, print, and other graphics, while the auditory signal can be divided into speech, music, and environmental sound.

Audio Features

Audio features are essential in characterizing a video's affective content. Author Wang and Cheong's [15] study shows that audio features are often more informative than visual ones with respect to affective content characterization. Author thinks the first step in acoustic feature extraction is audio type segmentation (also called audio source separation), since the audio part of a video often contains a mixture of sounds from different sources. Audio type segmentation divides the audio part of a video into speech, music, and environmental sound. They used two features (chroma difference and low short time energy ratio) to distinguish music sound from environmental sound with a simple SVM for every two second segment of audio signal.

Lu et al. [51] introduced a technique to segment and classify audio signals into speech, music, and environmental sounds. Their method first segments a signal into speech and non-speech using such features as high zero crossing rate (ZCR) ratio, low short time energy ratio, linear spectral pairs, and spectrum flux. The non-speech signal is then further divided into music and environmental sound using band periodicity, noise frame ratio, and spectrum flux.

Bachu et al. [52], proposed to use zero-crossing rate and energy features to separate speech and non-speech signals. More recently, Radmard et al. [53] proposed a clustering method to separate speech and non-speech signals based on the analysis of campestral peak, zero-crossing rate, and autocorrelation function peak of short time segments of the speech signal. Zhang and Kuo [54], proposed to use zero crossing rate to separate audio signals into music, speech, and environmental sounds. Wang and Cheong [15] extracted 12 audio features to capture a film's affective content, including energy statistics, Log Frequency Power Coefficients (LFPC), MFCC, and zero-crossing rate statistics.
These features are used for both emotional dimension prediction and emotional category classification.

Visual Features

The visual elements which generally used by filmmakers, manipulate to inject emotion include tempo, lighting, motion and color. Tempo is an important feature of films and has significant power to attract viewers' attention and to affect viewers' emotion intensity [17]. It captures the amount of camera and subject movement in each shot and between shots. Various features have been proposed to capture a video's tempo. Film tempo can also be changed by varying the camera position and movement speed in order to inject different types of emotion into the movie.

Motion is another important film element that can control a video's tempo. Motion-related features include motion intensity, motion dynamics, and visual excitement. Motion intensity [12], [22], [38], [39], [42], [48], [52], [53], reflects the smoothness of transitions between frames. Motion is highly expressive able to evoke strong emotional responses in viewers [2], [18]. Detenber et al. [18] and Simmons et al. [19] concluded that an increase of motion intensity on the screen causes an increase in the audience's arousal.

Lighting, the spectral composition of the light, is another powerful cinematography tool to manipulate visual elements. Lighting measures the contrast between dark and light, and influences the appearance of every element in the scene. For example dim lights, shadow play, and predominantly dark backgrounds to create sad, surprising, frightening, or suspenseful scenes [15], [40].

Color is also an important element that can be changed to affect the viewers' emotion. Specifically, color brightness is often used to affect valence while color saturation is used to influence arousal. E.g. Sad or frightening videos commonly consist of gray frames.
To better represent the movie's color, color features are typically computed in the Hue, Saturation, and Value (HSV) space, since psychological studies have shown that humans can better perceive emotions in HSV space than others [22], [8]. Zhang et al. [52] characterized valence using color saturation and color energy. Teixeira et al. [16] adopted color heat, color activity, and color weight to characterize a video's color.

Video features may be mapped to emotional descriptors using a classifier for categorical descriptors or a regressor for dimensional descriptors. Here we mentioned some classifier's use in brief.

Classifiers

Many machine learning methods have been investigated to model the mapping between video
features and discrete emotional descriptors, including support vector machines [35], multi-layer feed-forward neural networks (NNs) [22], Adaboost [8], Gaussian Mixture Models (GMMs) [49], KNearest Neighbor (KNN) [50], hidden Markov models (HMMs) [36], [61], Dynamic Bayesian Networks (DBNs) **,** and conditional random fields (CRFs) [48]. A classifier is divided into static or dynamic based on temporal information. Multi-layer feed-forward neural networks [22], SVMs [12], [15], [23] and GMMs [49] are used for static modeling. NNs are known to be effective for nonlinear mappings, and achieve good performance given effective features. For example, Watanapa et al. [22] proposed to classify movie clips into excitement, joy, or sadness using a two stage sieving artificial neural network, in which the first stage specialized in filtering the excitement class and the second stage classified joy and sadness.

Wang and Cheong [15], adopted a specially adapted variant of SVM to classify films into anger, sadness, fear, joy, surprise,

and neutral. Both SVM and NN are deterministic approaches. Yazdani et al. [49], proposed to use GMM for affective content analysis of music video clips. SVM, NN, and GMM use the input features to perform classification only. They do not perform any feature selection. In contrast, Ada boost performs feature selection, constructs a weak classifier with each selected feature, and combines the weak classifiers to perform the final classification. Teixeira et al. [11], proposed to detect pleasure, arousal, and dominance coefficients as well as six emotion categories using two Bayesian network topologies, a hidden Markov model and an autoregressive hidden Markov model. Their system first extracts a set of low-level audiovisual features from video shots, and then feeds them into two Bayesian networks to estimate the values of pleasure, arousal, and dominance for each video segment.

CONCLUSION :

Affective Computing is a significant development in computing; it studies and enhances human- machine-interaction in natural manner. Due to the difficulty in defining objective methods to automatically access the emotions of a video, the research topic of video affective content analysis has not been thoroughly explored until. Previous video affective retrievals have been done according to the audiovisual features extracted. In our work, we have done literature survey regarding to the video affective content analysis. Because of this, anyone can understand all about video affective content analysis in one go. In future we will explore more research findings which are beneficial to retrieve videos on the basis of affective states.

REFERENCES :

118

[1] Yimin Hou, Ting Xiao, Shu Zhang, Xi Jiang, Xiang Li, et al, "Predicting Movie Trailer Viewer's "Like/Dislike" via Learned Shot Editing Patterns" , IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 7, NO. 1, JANUARY-MARCH 2016

[2] B.-L. Yeo and B. Liu, "Rapid scene analysis on compressed videos," IEEE Trans. Circuits Syst. Video Technol., vol. 5, no. 6, pp. 533–544, Dec. 1995.

[3] T. G. Dietterich, R. H. Lathrop, and T. Lozano-P"erez, "Solving the multiple instance problem with axis-parallel rectangles," Artif. Intell., vol. 89, no. 1, pp. 31–71, 1997.

[4] S. Wang and X. Wang, "Emotional semantic detection from multimedia: A brief overview," in Kansei Engineering and Soft Computing: Theory and Practice. Hershey, PA, USA: IGI Global, 2010.

[4] S. Wang and X. Wang, "Emotional semantic detection from multimedia: A brief overview," in Kansei Engineering and Soft Computing: Theory and Practice. Hershey, PA, USA: IGI Global, 2010.

[5] Aayush Bhardwaj aayush,bhardwaj Ankit Gupta,"Classification of human emotions from EEG signals using SVM and LDA Classifiers", 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)

[6] P. Ekman, W.V. Friesen and J.C. Hager, "FACS Manual," salt Lake City, UT: A Human Face, May 2002

[7] Housb Lee, et al, "Towards Unobtrusive Emotion Recognition for Affective Social Communication" , The 9th Annual IEEE Consumer Communications and Networking Conference, 2012

[8] P. Valdez and A. Mehrabian, "Effects of color on emotions," J. Experimental Psychol.: sGeneral, vol. 123, no. 4, pp. 394–409, 1994.

[9] M. Soleymani and M. Pantic, "Human-centered implicit tagging:Overview and perspectives," in Proc. IEEE Int. Conf. Syst., Man Cybern., 2012, pp. 3304–3309.

[10] L. Canini, S. Benini, P. Migliorati, and R. Leonardi, "Emotional identity of movies," in Proc. 16th IEEE Int. Conf. Image Process., 2009, pp. 1821–1824.

[11] G. Irie, T. Satou, A. Kojima, T. Yamasaki, and K. Aizawaki, "Affective audio-visual words and latent topic driving model for realizing movie affective scene classification," IEEE Trans. Multimedia, vol. 12, no. 6, pp. 523–535, Oct. 2010.

[12] M. Xu, X. He, J. S. Jin, Y. Peng, C. Xu, and W. Guo, "Using scripts for affective content retrieval," in Proc. Adv. Multimedia Inf. Process. 11th Pacific Rim Conf. Multimedia, Shanghai, China, 2010, pp. 43–51.

[13] Gaetano Valenza , Riccardo Barbieri , et-al, "A Nonlinear Heartbeat Dynamics Model Approach for Personalized Emotion Recognition", 35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July, 2013

[14] G. Irie, T. Satou, A. Kojima, T. Yamasaki, and K. Aizawaki, "Affective audio-visual words and latent topic driving model for realizing movie affective scene classification," IEEE Trans. Multimedia, vol. 12, no. 6, pp. 523–535, Oct. 2010.

[15] H. L. Wang and L.-F. Cheong, "Affective understanding in film," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 6, pp. 689– 704, Jun. 2006.

[16] R. M. A. Teixeira, T. Yamasaki, and K. Aizawa, "Determination of emotional content of video clips by low level audiovisual features," Multimedia Tools Appl., vol. 61, no. 1, pp. 21–49, 2012.

[17] C. R. Plantinga and G. M. Smith, Passionate Views: Film, Cognition, and Emotion. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1999.

[18] B. H. Detenber, R. F. Simons, and G. G. Bennett Jr, "Roll em!: The effects of picture motion on emotional responses," J. Broadcasting Electron. Media, vol. 42, no. 1, pp. 113–127, 1998.

[19] R. F. Simons, B. H. Detenber, T. M. Roedema, and J. E. Reiss, "Emotion processing in three systems: The medium and the

[20] J. Wang, B. Li, W. Hu, and O. Wu, "Horror video scene recognition via multiple-instance learning," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., 2011, pp. 1325–1328.

[21] D. Swanberg, C.-F. Shu, and R. Jain, "Knowledge- Guided Parsing in Video Databases," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases*, SPIE Press, Bellingham, Wash., 1993.

[22] S. C. Watanapa, B. Thipakorn, and N. Charoenkitkarn, "A sieving ANN for emotion-based movie clip classification," IEICE Trans. Inf. Syst., vol. 91, no. 5, pp. 1562–1572, May 2008.

[23] M. Xu, J. Wang, X. He, J. S. Jin, S. Luo, and H. Lu, "A three-level framework for affective content analysis and its case studies," Multimedia Tools Appl., vol. 70, no. 2, pp. 757–779, May 2014.

[24]B. Adams, C. Dorai, and S. Venkatesh, "Novel approach to determining tempo and dramatic story sections in motion pictures," in Proc. Int. Conf. Image Process., 2000, vol. 2, pp. 283–286.

[25] Z. Wang, Y. Li, S. Wang, and Q. Ji, "Capturing global

semantic relationships for facial action unit recognition," in Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 3304–3311.
[26] S. Nie and Q. Ji, "Capturing global and local dynamics for human action recognition," in Proc. 22nd Int. Conf. Pattern Recog., 2014, pp. 1946–1951.
[27] L. Canini, S. Benini, and R. Leonardi, "Affective recommendation of movies based on selected connotative features," IEEE Trans. Circuits and Systems for Video Technol., vol. 23, no. 4, pp. 636–647, Apr. 2013.
[28] Min Xu, Liang-Tien Chia and Jesse Jin, "Affective Content Analysis in Comedy and Horror Videos by Audio Emotional Event Detection ", *©2005 IEEE*
[29] Ruiz-del-Solar, J., Navarrete, P.: FACERET: An interactive face retrieval system based on self-organizing maps. In: International Conference on Image and Video Retrieval, Lecture Notes in Computer Science, vol. 2383, Springer (2002) 149–156
[30] Y. Cui, S. Luo, Q. Tian, S. Zhang, Y. Peng, L. Jiang, and J. Jin, "Mutual information-based emotion recognition," in The Era of Interactive Media. New York, NY, USA: Springer, 2013,pp. 471–479.
[31] Oh, K., Zaher, A., Kim, P., "Fast k-nn image search with self-organizing maps" , International Conference on Image and Video Retrieval, Lecture Notes in Computer Science, vol. 2383, Springer (2002) 288–297
[32] Hussain, M., Eakins, J., Sexton, G.: "Visual clustering of trademarks using the self-organizing map. In: International Conference on Image and Video Retrieval", Lecture Notes in Computer Science, vol. 2383, Springer (2002) 139–148
[33] Huang, J., Umamaheswaran, D., Palakal, M., "Video indexing and retrieval for archeological digital library", CLIOH. In: International Conference on Image and Video Retrieval, Lecture Notes in Computer Science, vol. 2383, Springer (2002) 278–287
[34] The why, the where and the when to move the camera [Online]. Available: http://filmmakeriq.com/2013/04/the-why-the-whereand- the-when-to-move-the-camera/, 2013.
[35] C. Y. Wei, N. Dimitrova, and S.-F. Chang, "Color-mood analysis of films based on syntactic and psychological models," in Proc. IEEE Int. Conf. Multimedia Expo, 2004, pp. 831–834.
[36] Kai Sun and Junqing Yu, "Video Affective Content Representation and Recognition Using Video Affective Tree and Hidden Markov Models", A. Paiva, R. Prada, and R.W. Picard (Eds.): ACII 2007, LNCS 4738, pp. 594–605, 2007. © Springer-Verlag Berlin Heidelberg 2007
[37] Hanjalic, A." Extracting Moods from Pictures and Sounds: Towards truly personalized TV", IEEE Signal Processing Magazine 3, 90–100 (2006)
[38] *Zhang1,2, Qi Tian1, Qingming Huang3, Wen Gao2, Shipeng,* " UTILIZING AFFECTIVE ANALYSIS FOR EFFICIENT MOVIE BROWSING", , 978-1-4244-5654-3/09/$26.0 ©2009 IEEE
[38] Y. Cui, J. S. Jin, S. Zhang, S. Luo, and Q. Tian, "Music video affective understanding using feature importance analysis," in Proc. ACM Int. Conf. Image Video Retrieval, New York, NY, USA, 2010, pp. 213–219.
[39] Akira Kojima, Toshihiko Yamasaki,et-al., "Affective Audio-Visual Words and Latent Topic Driving Model for Realizing Movie Affective Scene Classification", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 12, NO. 6, OCTOBER 2010 523
[40] Z. Herbert, Sight, Sound, Motion: Applied Media Aesthetics. Belmont, CA, USA: Wadsworth 1999.
[41] J. A. Russell, "A circumplex model of affect," *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.
[42] A. Hanjalic and L.-Q. Xu, "Affective video content representation and modeling," IEEE Trans. Multimedia, vol. 7, no. 1, pp. 143–154, Feb. 2005.
[43] L. Canini, S. Benini, P. Migliorati, and R. Leonardi, "Emotional identity of movies," in Proc. 16th IEEE Int. Conf. Image Process., 2009, pp. 1821–1824.
[44] Nevenka Dimitrova, Lalitha Agnihotri and Gang Wei, "Video classification based on HMM using text and faces". European Conference on Signal Processing, Finland, September 2000
[45] S. Wang, Y. Zhu, G. Wu, and Q. Ji, "Hybrid video emotional tagging using users EEG and video content," Multimedia Tools Appl., vol. 72, no. 2, pp. 1257–1283, Sept. 2014.
[46] S. Wang and X. Wang, "Emotional semantic detection from multimedia: A brief overview," in Kansei Engineering and Soft Computing: Theory and Practice. Hershey, PA, USA: IGI Global, 2010.
[47] A. Yazdani, K. Kappeler, and T. Ebrahimi, "Affective content analysis of music video clips," in Proc. 1st Int. ACM Workshop Music Inf. Retrieval User-Centered Multimodal Strategies, New York, NY, USA, 2011, pp. 7–12.
[48] M. Xu, C. Xu, X. He, J. S. Jin, S. Luo, and Y. Rui, "Hierarchical affective content analysis in arousal and valence dimensions," Signal Process., vol. 93, no. 8, pp. 2140–2150, 2013.
[49] L. Canini, S. Benini, and R. Leonardi, "Affective recommendation of movies based on selected connotative features," IEEE Trans. Circuits and Systems for Video Technol., vol. 23, no. 4, pp. 636–647, Apr. 2013.
[50] A. Yazdani, K. Kappeler, and T. Ebrahimi, "Affective content analysis of music video clips," in Proc. 1st Int. ACM Workshop Music Inf. Retrieval User-Centered Multimodal Strategies, New York, NY, USA, 2011, pp. 7–12.
[51] F. Nack, C. Dorai, and S. Venkatesh, "Computational media aesthetics: Finding meaning beautiful," IEEE MultiMedia, vol. 8, no. 4, pp. 10–12, Oct.–Dec. 2001.
[52] S. Zhang, Q. Huang, S. Jiang, W. Gao, and Q. Tian, "Affective visualization and retrieval for music video," IEEE Trans. Multimedia, vol. 12, no. 6, pp. 510–522, Oct. 2010.
[53] L. Canini, S. Gilroy, M. Cavazza, R. Leonardi, and S. Benini, "Users' response to affective film content: A narrative perspective," in Proc. Int. Workshop Content-Based Multimedia Indexing, Grenoble, France, Jun. 2010, pp. 1–6.
[54] H. Katti, K. Yadati, M. Kankanhalli, and C. Tat-Seng, "Affective video summarization and story board generation using papillary dilation and eye gaze," in Proc. IEEE Int. Symp. Multimedia, 2011, pp. 319–326.
[55] J. J. Kierkels, M. Soleymani, and T. Pun, "Queries and tags in ffect-based multimedia retrieval," in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2009, pp. 1436–1439.
[56] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; using physiological signals," IEEE Trans. Affective Comput., vol. 3, no. 1, pp. 18–31, Jan.–Mar. 2012.
[57] Esra Acar, Frank Hopfgartner, and Sahin Albayrak, "Understanding Affective Content of Music Videos Through Learned Representations", Technische Universit¨at Berlin,

ANALYSIS OF BEHAVIOUR OF STUDENTS AWARENESS REGARDING CYBERCRIME AND ONLINE SECURITY

PROF. DEEPALI N. BHOSKAR ,
Asst. Professor in BCA ,Sinhgad College of Science ,Ambegaon(BK) ,Pune
deepali.scos@sinhgad.edu

Abstract : Students are every day using Information technology for different reasons like chatting, social networking, communication, purchasing online products etc. During use of this are they aware of cybercrime and security measures to be taken. With this paper we want to analyze the behaviour of students regarding cybercrime and online security.
To analyze the behaviour we have conducted online survey by sending questionnaire to students. It has been analyzed that students are not aware about what is cybercrime and online security.

Keywords: Friend List, Auto tagging, , Phishing

INTRODUCTION:

[1]Cybercrimes can be defined as: "Offences that are committed against individuals or groups of individuals with a criminal motive to intentionally harm the reputation of the victim or cause physical or mental harm, or loss, to the victim directly or indirectly, using modern telecommunication networks such as Internet (networks including but not limited to Chat rooms, emails, notice boards and groups) and mobile phones (Bluetooth/SMS/MMS)"
Students are hanged out online continuously. They can't leave without their smart phones through which they can access the digitized world. It's a very important part of their life. They can live without other daily needs but not without internet. The objective of this research aims at finding out are they aware of cybercrime and security measures and precautions to be taken while being online.
[2]Online Security (Internet security) is a branch of computer security specifically related to the Internet, often involving browser security but also network security on a more general level, as it applies to other applications or operating systems as a whole. Its objective is to establish rules and measures to use against attacks over the Internet.[1] The Internet represents an insecure channel for exchanging information leading to a high risk of intrusion or fraud, such as phishing[2], online viruses, trojans, worms and more.

Due to the technology advancement the destructive mind people are trying to disturb someone's life. People want to make easy money so they are misusing the advanced technology. This study is focusing on student's online habits and their awareness level regarding online security and cybercrimes. Awareness plays a major role in preventing from such people and we may enjoy this advanced technology .

Research Methodology

An online survey is conducted to know how many of them are aware of cybercrime and online security, focusing students by sending mails to visit the link and give their responses to collect data.

Objectives of our study are:

1. To understand Internet usage habits
2. To know students awareness about cybercrime
3. To know students awareness about online security

Important Terminologies:
1. Friend List: Friends' list is a list of users who are a part of your social network.
2. Location: Location is the accurate/inaccurate position of the user which is derived from the cellular network used by the user, IP address or directly using Global Positioning System (GPS)
3. Auto tagging: Auto-tagging is a feature available in many camera applications. It allows the user to auto tag each image with the location (derived using GPS)
4. Phishing: Phishing is a process of harvesting usernames and passwords using web pages which appear to be legitimate, but infact are fake and created by the attacker

RESULTS AND DISCUSSION :

1. Purpose of accessing the internet

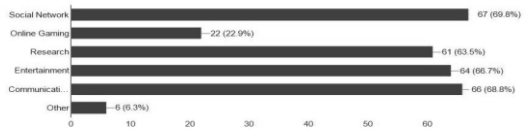
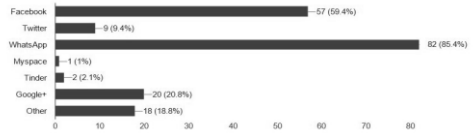


Fig.1
According to the above graph maximum number of users access internet for visiting social networking websites.

2.Popular Social Networking sites/apps



WhatsApp is the most popular app amongst people. Followed by Facebook and Google+

1. Friend Lists

In Fig 3 ,90.6% people use Facebook and out of these 20.8 % people have more than 800 users in their friend's list. Majority of the people have fewer than 500 friends

Fig 3

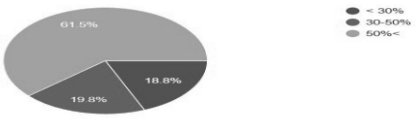


Fig 4

In Fig 4 Out of 90.6% facebook users 61.5 % users claim that they know more than 50% people from their friends' list personally. However, 38.5% people know less than 50% people personally. This informs us that people are willingly interacting with total strangers. There's a chance that these strangers can take undue advantage of innocance of such users. This is a popular social engineering technique for commiting fraus. Such attacks can be prevented only if the user is vigilant and aware of such techniques and attackers.[9,10]

1. Usage of Mobile Apps (Services and Permissions)

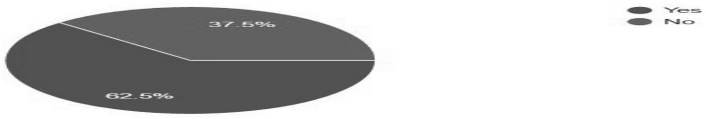


Fig 5

- Fig 5 says that 62.5% people use Facebook Messenger (Facebook chat application) on their Android/iOS device.

Fig 6

Fig 7

- Out of the total people surveyed 11.5% users have their location services on 24/7 while 24% are not sure whether their location services are on or off.
- Switching on location services allows 3rd party apps to gather a user's location data.
- Facebook messenger used to append location data to the messages which were sent using messenger app. This data can be used to pin point the location of a person with whom you are chatting.

Fig 8

- Fig 8 says that 10.4% people auto tag their images with their location obtained from GPS, while 10.4% are not aware of such a thing.
- Location data tagged to such images can be easily extracted to determine the location of the image and the person inside. If the person is uploading images live, then one can track the person in real-time. This is an open invitation to stalkers
- Out of all the facebook users 12.5% use the facebook's check in feature. Facebook's check in feature allows the user to announce his/her location to his/her friends. People usually use this feature whenever they enter a restaurant, hotel or any other travel related place. It's a good way to stay connected with your friends but, it also enables the stalkers to pin point your location. It has been observed that 62.5% facebook users use the check in feature sometime or the other, especially while they are travelling. Eg. "ABC Checked into Pune International Airport"
- This information can be used against you to commit various frauds.

1. Phishing and Fake Page Detection

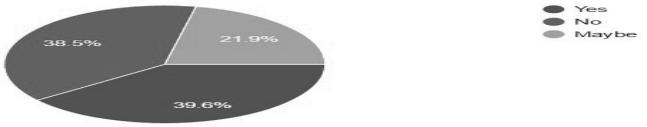


Fig 9

- In Fig 9 s urprisingly 38.5% users do not check webpages for anomolies while entering their login credentials. 21.9% users are not sure. This means, 38.5% people can be victims of phishing attacks.
- On the other hand success rate of spear phishing attacks is close to 42.5%

Fig 10

- All the above mentioned elements have to be checked for detecting anomolies on any webpage
- Checking just one of the element will lead to inaccurate results, thus, a fake page might pass as a legitimate one. Here, again success rate of spear phishing will be high since, 50% people look for one of the element instead of all

Fig 11

- As per Fig 11 Only 30.2% are aware of the true meaning of the green padlock sign
- Websites with a green padlock sign need not be a safe website, people can use fake certificates
- 19.8% users are not aware of any green padlock sign!
- Details of the certificate must be verified before entering login credentials.

1. Indian IT Act

Fig 12

- Ultimately the users were asked if they were aware of Indian IT Act
- Surprisingly 53.1% answered "NO"
- IT Act is the primary law in India that deals with cybercrime and electronic commerce

CONCLUSION :

We conclude from online survey and data analysis that students are less aware about cybercrime and online security. Even we analyzed their Internet usage habits, online behaviour and which Social networking sites are regularly used by them.

REFERENCES :

[1]https://en.wikipedia.org/wiki/Cybercrime
[2]https://en.wikipedia.org/wiki/Internet_security
[3]https://www.facebook.com/help/204604196335128?helpref=faq_content
[4]http://searchsecurity.techtarget.com/definition/cyberstalking
[5]http://cybercellmumbai.gov.in/html/write-ups/index.html
[6]http://www.cyberlawsindia.net/internet-crime.html
[7]http://cyber-law-enforcement.blogspot.in/
[8]http://meity.gov.in/sites/upload_files/dit/files/downloads/itact2000/it_amendment_act2008.pdf
[9]http://www.makeuseof.com/tag/facebook-places-location-tagging-safely-weekly-facebook-tips/
[10]https://www.theguardian.com/technology/2015/may/28/marauders-map-chrome-app-tracks-facebook-messenger
[11]https://exposingtheinvisible.org/resources/obtaining-evidence/image-digging
[12]http://smallbusiness.chron.com/make-check-in-feature-work-facebook-41278.html

WHAT IS WATER QUALITY INDEX AND DEVELOPMENT OF WQI CALCULATOR?

DHERE AMAR M.,
Assistant Professor, Dept. of Science,
Sir Vithaldas Thackersey College of Home Science (Autonomous),
S.N.D.T. Women's University, Santacruz (W), Mumbai, MS
Email- prof.amardhere@gmail.com
PONDHE GORAKSH M.
Head, P.G.Dept. of Environmental Science,
PVP College of Arts, Commerce and Science, Pravaranagar- Loni

ABSTRACT : *Present article aims to analyse the application of Water Quality Index for survey of river pollution and development of calculator. Secondary resource are widely referred to draw the results. With .xlc software simple arithmetic calculator developed. This calculator is tested though the secondary data and errors are corrected accordingly. This calculator gives the accurate result of water quality without any manual arithmetic and calculation and save the time.*

Key words : water quality index, calculator, river water.

INTRODUCTION:

Water Quality Index (WQI)- Globally there is no accepted index of water quality. Many nations have regulated their select some parameters to fix the water quality this creates a confusion to aggregate the water quality data all over the world (UNEP, 2015). Water quality parameters generally normalizing according to the expected concentration which resulted good versus bad (Pesce and Wunderlin, 2000).

Definitions of water quality index: Water quality index have numerous definitions following are the few of those.

- Water quality index defined as numerical representation of water for overall water quality (Cude C. 2001).
- Water quality is a phrase to describe the chemical, physical and biological characteristics of water. Defining water 'good' or 'bad' water quality is not as simple as it seems because its depends on the context in which it used (Global Environment Centre, 2017).
- Water quality index provides a single number that expresses overall water quality at a certain locations and time based on several water quality parameters (Ram K.S., Anadh H. 1996).
- Water quality index reflects the composite influence of different water quality parameters on the overall quality of water (Latha S.P., Rao N. K., 2010).
- Water quality index is helpful to summaries the large amount of water quality data into simple term (Ujjania N.C. and Dubey M., 2015).
- Water quality index is defined as technique of rating that provides the composite influence of individual water quality parameters on overall quality of water for human consumption (Mitra B.K., 1998, Reza R., Singh., G, 2010)
- Water quality index is defined as a rating reflecting the composite influence of different water quality parameters (Ramkrishnaiah C.R., Sadashivaaiah C. and Ranganna G., 2009).
- Water quality index is a single numerical expression which reflects the composite influence of nine significant physical, chemical and microbiological parameters of water quality (EPA, 1974). This definition is proposed by the National Sanitation Foundation Water Quality Index (NSFWQI).

Definitions of WQI have wide verities of have some common things i.e it was a numerical expression calculated with scientific equations include several parameters and gives expression for usages of water based on the results.

Water Quality index: In 1959 State Sanitation Engineer in USA proposed uniform method for water quality during conference. Meanwhile in 1965 Environmental Pollution Panel initiate the process to develop index of chemical pollution (EPA, 1974). Whereas 1965 Horton develop water quality index by selecting commonly considered 10 parameters in USA (Otto, 1978). However Brown and others build up Water Quality Index in the early 1970, were hundreds of expert members come together with Brown. Horton's water quality index is modified by the group of people waking with Brown (Brown et al, 1970). They choose nine criteria for calculating the degree of the water quality. Aim of water quality index to establish the correct mean of class for usages of water for various purposes (Stambuk-Giljanvoc, 1999). More complex date of water analysis is simple converted to the number with the help of numerical methods is major significance of the water quality index (Bordalo A.A., 2001). The calculation of water quality though the standard parameters consider for the drinking purpose. Water quality index is most comprehensive method to communicate information about overall quality of water for public domain and reliable information for the policy regulations (UNEP, 2010).

Liou et al (2004) develops water quality index in Taiwan based on the standard score of temperature. pH, toxic substances, organic and particulate matter. Tsegaye et al (2005) find the water quality index of 18 streams by using seven water quality parameters. Kin and Cardone (2005) studies water quality around the mines in USA. First Indian Water Quality Index use to find the water purity of Ganaga for drinking water supply (Bhargava, 1985). In 1990 with the respect to guideline of Central Pollution Control Board

(CPCB) for water quality NSFWQI is used by Vedparkash to detect the water quality of Ganga river (Abbasi, 2002).

There are several water quality indexes applicable to measure the quality of water. Main water quality index now a day used all over the world namely National sanitation foundation water quality index (NSFWQI), Canadian Council Ministers of the Environment Water Quality Index (CCMEWQI) and Oregaon Water Quality Index (OWQI). However not a single index universally accepted but NSFWQI is widely accepted in many research studies (EPA, 1974, Abbasi, 2002, Ramkrishnaiah C.R., Sadashivaaiah C. and Ranganna G., 2009, Sanrgaonkar A., Deshpande V., 2003, WU, 2008, Bhargava, 1985). This index has merit to ralted the index value to potential water use and easy to communicate with the layman (Tyagi S., Sharma B., Singh P., Dohal R., 2013). For present study National sanitation foundation water quality index is used to calculate the water quality index of Bhima river from origin o the Ujjani reservoir on the different sampling station.

National Sanitation Foundation Water Quality Index: This is most commonly used water widely index developed by the national sanitation foundation in 1970 were 142 water quality scientist surveyed the 35 water quality tests out of these 9 parameters are selected for the to include in the index. These are dissolved oxygen (DO), fecal coliform, pH, biochemical oxygen demand (BOD), temperature, total phosphate, nitrate, turbidity and total dissolved solids. This method is rigors because of its parameters, scale and assigning weights (Tyagi S., Sharma B., Singh P., Dohal R., 2013). This method is developed by the national sanitation foundation in USA have strength to implement for calculating the WQI of critically polluted water bodies (EPA, 1974).

NSFWQI Equation:

$$WQI = \sum_{i=1}^{n} w_i q_i$$

Where

qi = sub index for ith water quality parameter;

wi= weight associated with ith water quality parameter;

n= number of water quality parameter

Table No. 1: Significance ratings and weights for nine paramereres include in WQI

Parameter	Weight
pH	0.11
Change in temp	0.10
DO	0.17
BOD	0.11
Turbidity	0.08
Phosphate	0.10
Nitrate Nitrogen	0.10
E. coli*	0.16
Total Dissolved Solids	0.07

Source: Brown et al, 1970

In the NSFWQI expression n=9 and includes

- ✓ Dissolved oxygen expressed as percent saturation
- ✓ Fecal coliform density (FC), n/100ml
- ✓ pH
- ✓ nitrate (NO_3) mg/L NO_3-N
- ✓ phosphate (PO_4) mg/L PO_3-P
- ✓ temperature (T) °C
- ✓ total Dissolved Solids TDS mg/L
- ✓ turbidity (NTU)

Calculation of oxygen saturated percent:

$$\text{DO \%} = \frac{DO\ Vlaue\ from\ Chart\ at\ temp.and\ salinity}{Dissolved\ oxygen\ (mg/lit)} \text{ x } 100$$

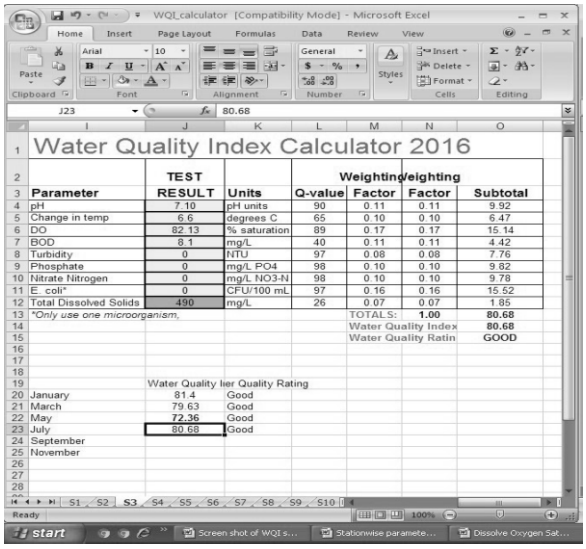
For Chart at temperature and salinity see Annexure- A

Table No. 2: Water quality ratings of NSFWQI.

WQI Value	Rating of Water Quality
91-100	Excellent
71-90	Good
51-70	Medium
26-50	Bad
0-25	Very Bad

Source: Brown et al, 1970

WQI Calculation: For present study NSFWQI is use through the calculator with the help of simple programming on Microsoft Excel 7. Although Wilkis University in USA bring online calculator for NSFWOI. The screen image of the NSFWQI is shown in the picture number 1. This is most effective method to resolve calculation hazels as well it use offline. This software gives instant results even find the dominant parameter responsible for degradation of water quality.



Picture No. 1: NSFWQI Software based Calculator.

CONCLUSION :

Water Quality Index was not just comprehensive representation of overall water quality but it helpful for determining the quality of water. WQI calculator is tested and gives the reliable results without any hassle from manual calculations.

REFERENCES :

1) Abbasi, S. A. (2002). Water quality indices, state of the art report. *Indian National Committee on Hydrology. Secretariat, National Institute of Hydrology. INCOH Secretariat, India*.

2) Bhargava, D. S. (1985). Water quality variations and control technology of Yamuna River. *Environmental Pollution Series A, Ecological and Biological*, *37*(4), 355-376.

3) Bordalo, A. A., Nilsumranchit, W., & Chalermwat, K. (2001). Water quality and uses of the Bangpakong River (Eastern Thailand). *Water Research*, *35*(15), 3635-3642.

4)Brown, L. C., & Barnwell, T. O. (1987). *The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual* (p. 189). US Environmental Protection Agency. Office of Research and Development. Environmental Research Laboratory.

5)Cude, C. G. (2001). Oregon water quality index a tool for evaluating water quality management effectiveness. *JAWRA Journal of the American Water Resources Association*, *37*(1), 125-137.

6)National Research Council (US). Committee on the Biological Effects of Ionizing Radiations, & United States. Environmental Protection Agency. Office of Radiation Programs. (1980). *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation, 1974* (Vol. 3095). National Academy

7)Kim, A. G., & Cardone, C. R. (2005). Scatterscore: a reconnaissance method to evaluate changes in water quality. *Environmental monitoring and assessment*, *111*(1-3), 277-295.

8)Rao, N. K., Latha, S. P., Kumar, A. P., & Krishna, H. M. (2010). Morphometric analysis of Gostani River basin in Andhra Pradesh state, India using Spatial information technology. *International journal of geomatics and geosciences*, *1*(2), 179.

9) Liu, J. (2004). An overview on chronic arsenism via drinking water in PR China. *Toxicology*, *198*(1), 25-29

10) Mitra, B. K. (1998). Spatial and temporal variation of ground water quality in sand dune area of aomori prefecture in Japan, Paper number 062023. In *2006 ASAE annual meeting*.

11)Reza, R., & Singh, G. (2010). Heavy metal contamination and its indexing approach for river water. *International Journal of Environmental Science & Technology*, *7*(4), 785-792.

12)Otto, A. (1978). Raman spectra of (CN)-adsorbed at a silver surface. *Surface Science*, *75*(2), L392-L396.

13)Pesce, S. F., & Wunderlin, D. A. (2000). Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River. *Water Research*, *34*(11), 2915-2926.

14)Ramakrishnaiah, C. R., Sadashivaiah, C., & Ranganna, G. (2009). Assessment of water quality index for the groundwater in Tumkur Taluk, Karnataka State, India. *Journal of Chemistry*, *6*(2), 523-530.

15) Ramakrishnaiah, C. R., Sadashivaiah, C., & Ranganna, G. (2009). Assessment of water quality index for the groundwater in Tumkur Taluk, Karnataka State, India. *Journal of Chemistry*, *6*(2), 523-530.

16) Sargaonkar, A., & Deshpande, V. (2003). Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental monitoring and assessment*, *89*(1), 43-67.

17)Štambuk-Giljanović, N. (1999). Water quality evaluation by index in Dalmatia. *Water Research*, *33*(16), 3423-3440.

18)Tsegaye, T., Sheppard, D., Islam, K. R., Tadesse, W., Atalay, A., & Marzen, L. (2006). Development of chemical index as a measure of in-stream water quality in response to land-use and land cover changes. *Water, Air, & Soil Pollution*, *174*(1), 161-179.

19) Tyagi, S., Sharma, B., Singh, P., & Dobhal, R. (2013). Water quality assessment in terms of water quality index. *American Journal of Water Resources*, *1*(3), 34-38.

20) Ujjania, N. C., Dubey, M., Sharma, L. L., Balai, V. K., & Srivastva, R. M. (2015). Bio-invasion of exotic fish tilapia (Oreochromis mossambicus P. 1852) in Lake Jaisamand, India.

21) UNEP. 2010.". *Cambodia REDD+ Roadmap." Area, no. November*, 1-54.

IMPORTANCE OF DATA WAREHOUSING, DATA MINING, OLAP AND OLTP TECHNOLOGIES IN INDUSTRIES.

DANAVE BHARATI MURLIDHAR
Assistant Professor
BHINGARE MADHURI BALASAHEB
Assistant Professor
Dept. of Computer Science
New Art's,Commerce and Science College
Ahmednagar, M S., India

Abstract: This paper provides an overview of Data warehousing, Data Mining technologies, exploring the features, applications and the architecture of Data Warehousing. The data warehouse supports on-line analytical processing (OLAP), the functional and performance requirements of which are quite different from those of the on-line transaction processing (OLTP) applications traditionally supported by the operational databases. Data warehouses provide on-line analytical processing (OLAP) tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data mining. Data warehousing and on-line analytical processing (OLAP) are essential elements of decision support, which has increasingly become a focus of the database industry. OLTP is customer-oriented and is used for transaction and query processing by clerks, clients and information technology professionals. An OLAP system is market-oriented and is used for data analysis by knowledge workers, including managers, executives and analysts. Data warehousing and OLAP have emerged as leading technologies that facilitate data storage, organization and then, significant retrieval. Decision support places some rather different requirements on database technology compared to traditional on-line transaction processing applications.

Keywords: Data Warehousing, OLAP, OLTP, Data Mining, Decision Making and Decision Support

INTRODUCTION:

A data warehouse is a "subject-oriented, integrated, time varying, non-volatile collection of data that is used primarily in organizational decision making. Typically, the data warehouse is maintained separately from the organization's operational databases. There are many reasons for doing this. The data warehouse supports on-line analytical processing (OLAP), the functional and performance requirements of which are quite different from those of the on-line transaction processing (OLTP) applications traditionally supported by the operational databases. Data warehousing is a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, analyst) to make better and faster decisions. It serves as a physical implementation of a decision support data model and stores the information on which an enterprise needs to make strategic decisions. The data can be stored in many different types of databases. One data base architecture that has recently emerged is the "data warehouse", a repository of multiple heterogeneous data sources, organized under a unified schema at a single site in order to facilitate management decision-making. Data warehouse technology includes data cleansing, data integration and online Analytical processing. OLAP stands for analysis techniques with functionalities such as summarization, consolidation and aggregation, as well as the ability to view information from different angles. Ten years ago, Data Warehousing was largely unknown. Today, many companies are receiving considerable business value from their warehousing efforts. First American Corporation (FAC), a regional bank located in the Southeast, lost $60 million in 1990 and was operating under letters of agreement with regulators. A new senior management team developed a customer intimacy strategy with a data warehouse at the heart of the strategy. Using warehouse data, FAC was able to determine the profitability of all of their clients and products; develop programs to attract, maintain, and enhance their customer base; create profitable new product and service offerings; and redesign their distribution channels to increase profitability and better meet customers' needs. Data warehousing helped FAC to become a profitable, innovative leader in the financial services industry. Data warehouse systems are now an established component of information systems landscape in most companies. Due to high failure rates of data warehouse projects, several procedure models for building data warehouse systems were published considering their special requirements. In addition, most development methodologies are lacking concepts to ensure long-term evolution and establishment of data warehouse systems which are organizational challenges.

2. Data Warehousing

2.1 Definition of data warehousing

According to W.H.Inmon, a leading architect in the construction of data warehouse systems, A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process . So, data warehouse can be said to be a semantically consistent data store that serves as a physical implementation of a decision support data model and stores the information on which an enterprise needs to make strategic decisions. So, its architecture is said to be constructed by integrating data from multiple heterogeneous sources to support and /or adhoc queries, analytical reporting and decision-making. Data warehouses provide on-line analytical processing (OLAP)

tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data mining. The functional and performance requirements of OLAP are quite different from those of the on-line transaction processing applications traditionally supported by the operational databases. Data can now be stored in many different types of databases. One type of database architecture that has recently emerged is data warehouse, which is a repository of multiple heterogeneous data sources, organized under a unified schema at a single site in order to facilitate management decision-making. Data warehouse technology includes data cleaning, data integrating, and on-line analytical processing (OLAP) that is, analysis techniques with functionalities such as summarization, consolidation and aggregation, as well as the ability to view information from different angles. A data warehouse is defined as a "subject-oriented, integrated, time variant, non-volatile collection of data that serves as a physical implementation of a decision support data model and stores the information on which an enterprise needs to make strategic decisions. In data warehouses historical, summarized and consolidated data is more important than detailed, individual records. Since data warehouses contain consolidated data, perhaps from several operational databases, over potentially long periods of time, they tend to be much larger than operational databases. Most queries on data warehouses are ad hoc and are complex queries that can access millions of records and perform a lot of scans, joins, and aggregates. Due to the complexity query throughput and response times are more important than transaction throughput. Data warehousing is a collection of **decision support** technologies, aimed at enabling the **knowledge worker** (executive, manager, analyst) to make better and faster decisions. Data warehousing technologies have been successfully deployed in many industries: manufacturing, retail, financial services, transportation, telecommunications, utilities and healthcare.This paper presents a roadmap of data warehousing technologies, focusing on the special requirements that data warehouses place on database management systems (DBMSs).

2.2 DATA WAREHOUSING FUNDAMENTALS

A data warehouse (or smaller-scale data mart) is a specially prepared repository of data designed to support decision making. The data comes from operational systems and external sources. To create the data warehouse, data are extracted from source systems, cleaned (e.g., to detect and correct errors), transformed (e.g., put into subject groups or summarized), and loaded into a data store (i.e., placed into a data warehouse).

The data in a data warehouse have the following characteristics:

Subject oriented — The data are logically organized around major subjects of the organization, e.g., around customers, sales, or items produced.

Integrated — All of the data about the subject are combined and can be analysed together.

Time variant — Historical data are maintained in detail form.

Nonvolatile — The data are read only, not updated or changed by users.

A data warehouse draws data from operational systems, but is physically separate and serves a different purpose. Operational systems have their own databases and are used for transaction processing; a data warehouse has its own database and is used to support decision making. Once the warehouse is created, users (e.g., analysts, managers) access the data in the warehouse using tools that generate SQL (i.e., structured query language) queries or through applications such as a decision support system or an executive information system. As the organizational domain of data warehouse systems still lacks attention of data warehouse researchers compared to technical aspects. Therefore this paper aims at providing deeper insights in the current organizational situation of data warehouse departments in practice. The organizational domain of companies can be divided in a structural, human resource, political, and symbolic dimension and each dimension requires different design instruments. The structural dimension focuses on goals, formal roles and relationships. Structures are created to achieve the company's goals considering technological and environmental factors. Rules, policies, processes, and hierarchies are the design elements of the structural dimension. Drawing from psychology, the human resource dimension takes care about the needs, feelings, prejudices, and limitations of all individuals. The political dimension sees organizations as arenas. Different interest groups cause conflicts while competing for power and resources and the organizational life is characterized by bargaining, negations and compromises. The symbolic dimension abandons the assumptions of rational behaviour and views organizations as some kind of theatres.

2.3 Architecture and End-to-End Process

Figure 1 shows a typical data warehousing architecture.

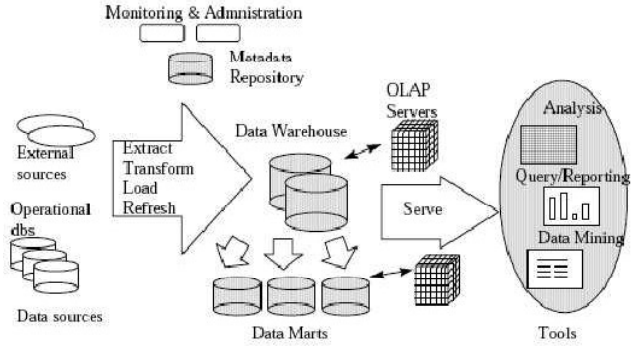


Figure 1: Data Warehousing Architecture

It includes tools for extracting data from multiple operational databases and external sources; for cleaning, transforming and integrating this data; for loading data into the data warehouse; and for periodically refreshing the warehouse to reflect updates at the sources and to purge data from the warehouse, perhaps onto slower archival storage. In addition to the main warehouse, there may be several departmental data marts. Data in the warehouse and data marts is stored and managed by one or more warehouse servers, which present multidimensional views of data to a variety of front end tools: query tools, report writers, analysis tools, and data mining tools. Finally, there is a repository for storing and managing

metadata, and tools for monitoring and administering the warehousing system. Designing and rolling out a data warehouse is a complex process, consisting of the following activities:-

Define the architecture, do capacity planning, and select the storage servers, database and OLAP servers, and tools.

Integrate the servers, storage, and client tools.

Design the warehouse schema and views.

Define the physical warehouse organization, data placement, partitioning, and access methods.

Connect the sources using gateways, ODBC drivers, or other wrappers.

Design and implement scripts for data extraction, cleaning, transformation, load, and refresh.

Populate the repository with the schema and view definitions, scripts, and other metadata.

Design and implement end-user applications.

Roll out the warehouse and applications.

3. OLTP and OLAP:

The job of earlier on-line operational systems was to perform transaction and query processing. So, they are also termed as on-line transaction processing systems (OLTP). Data warehouse systems serve users or knowledge workers in the role of data analysis and decision-making. Such systems can organize and present data in various formats in order to accommodate the diverse needs of the different users. These systems are called on-line analytical processing (OLAP) systems.

3.1 Major distinguishing features between OLTP and OLAP

i) Users and system orientation: OLTP is customer-oriented and is used for transaction and query processing by clerks, clients and information technology professionals. An OLAP system is market-oriented and is used for data analysis by knowledge workers, including managers, executives and analysts.

ii) Data contents: OLTP system manages current data in too detailed format. While an OLAP system manages large amounts of historical data, provides facilities for summarization and aggregation. Moreover, information is stored and managed at different levels of granularity, it makes the data easier to use in informed decision-making.

iii) Database design: An OLTP system generally adopts an entity –relationship data model and an application-oriented database design. An OLAP system adopts either a star or snowflake model and a subject oriented database design.

iv) View: OLTP system focuses mainly on the current data without referring to historical data or data in different organizations. In contrast, OLAP system spans multiple versions of a database schema, due to the evolutionary process of an organization. Because of their huge volume, OLAP data are shared on multiple storage media.

v) Access patterns: Access patterns of an OLTP system consist mainly of short, atomic transactions. Such a system requires concurrency, control and recovery mechanisms. But, accesses to OLAP systems are mostly read-only operations, although many could be complex queries.

3.2 Need of data warehousing and OLAP

Data warehousing developed, despite the presence of operational databases due to following reasons:

An operational database is designed and tuned from known tasks and workloads, such as indexing using primary keys, searching for particular records and optimizing 'canned queries'. As data warehouse queries are often complex, they involve the computation of large groups of data at summarized levels and may require the use of special data organization, access and implementation methods based on multidimensional views. Processing OLAP queries in operational databases would substantially degrade the performance of operational tasks.

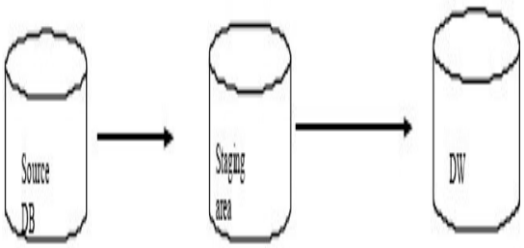
An operational database supports the concurrent processing of multiple transactions. Concurrency control and recovery mechanisms, such as locking and logging are required to ensure the consistency and robustness of transactions.

Decision support requires historical data, whereas operational databases do not typically maintain historical data. So, the data in operational databases, though abundant, is always far from complete for decision-making.

Decision support needs consolidation (such as aggregation and summarization) of data from heterogeneous sources; and operational databases contain only detailed raw data.

4. Data Flow

The steps for building a data warehouse or repository are well understood. The data flows from one or more source databases into an intermediate staging area, and finally into the data warehouse or repository (see Figure 2).



At each stage there are data quality tools available to massage and transform the data, thus enhancing the usability of the data once it resides in the data warehouse.

5. Data Mining

Data Mining is the extraction or "Mining" of knowledge from a large amount of data or data warehouse. To do this extraction data mining combines artificial intelligence, statistical analysis and database management systems to attempt to pull knowledge form stored data. Data mining is the process of applying intelligent methods to extract data patterns. This is done using the front-end tools. The spreadsheet is still the most compiling front-end application for Online Analytical Processing (OLAP). The challenges in supporting a query environment for OLAP can be crudely summarized as that of supporting spreadsheet operation effectively over large multi-gigabytes databases.

6. Decision making using a Data Warehouse

A Decision Support System (DSS) is any tool used to improve the process of decision making in complex systems. A DSS can range from a system that answer simple queries and

allows a subsequent decision to be made, to a system that employ artificial intelligence and provides detailed querying across a spectrum of related datasets. Amongst the most important application areas of DSS are those complicated systems that directly "answer" questions, in particular high level "what-if" scenario modeling. The data warehouse environment is more controlled and therefore more reliable for decision support than the previous methods. The data warehouse environment supports the entire decision support requirements by providing high-quality information, made available by accurate and effective cleaning routines and using consistent and valid data transformation rules and documented presummarization of data values. It contains one single source of accurate, reliable information that can be used for analysis. Data Warehouses (DW) integrate data from multiple heterogeneous information sources and transform them into a multidimensional representation for decision support applications.

CONCLUSION :

Data warehouse can be said to be a semantically consistent data store that serves as a physical implementation of a decision support data model and stores the information on which an enterprise needs to make strategic decisions. So, its architecture is said to be constructed by integrating data from multiple heterogeneous sources to support and /or adhoc queries, analytical reporting and decision-making. Data warehouses provide on-line analytical processing (OLAP) tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data mining. Data warehousing and online analytical processing (OLAP) are essential elements of decision support, which has increasingly become a focus of the database industry. OLTP is customer-oriented and is used for transaction and query processing by clerks, clients and information technology professionals. The job of earlier on-line operational systems was to perform transaction and query processing. Data warehouse systems serve users or knowledge workers in the role of data analysis and decision making. Such systems can organize and present data in various formats in order to accommodate the diverse needs of the different users. OLAP applications are found in the area of financial modeling (budgeting, planning), sales forecasting, customer and product profitability, exception reporting, resource allocation, variance analysis, promotion planning, market share analysis. Moreover, OLAP enables managers to model problems that would be impossible using less flexible systems with lengthy and inconsistent response times. More control and timely access to strategic information facilitates effective decision-making. This provides leverage to library managers by providing the ability to model real life projections and a more efficient use of resources. OLAP enables the organization as a whole to respond more quickly to market demands. Market responsiveness, in turn, often yields improved revenue and profitability. And there is no need to emphasize that present libraries have to provide market-oriented services.

REFERENCES :

[1] Devlin, B. & Murphy, P. (1988) An Architecture for a Business and
Information System, IBM Systems Journal, 27 (1), 60-80.

[2] Inmon, W.H. (1996) Building the Data Warehouse, Second Edition,
New York: John Wiley & Sons.

[3] Kimball, R. (1996) The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses, New York:John Wiley & Sons.

[4] Gardner, S.R. (1998) Building the Data Warehouse, Communications
of the ACM , 41 (9), 52-60.

[5] Simon, A. (1998) 90 Days to the Data Mart, New York: John Wiley & Sons.

[6] Frolick, M.N. & Lindsey, K. (2003) Critical Factors for Data Warehouse Failure, Journal of Data Warehousing,8 (1).

[7] Hwang, H.-G., Kua, C.-Y., Yenb, D. C., & Chenga, C.-C. (2002)
Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan, Decision
Support Systems, In Press, Corrected Proof.

[8] Finnegan, P. & Sammon, D. (1999) Foundations of an Organisational
Prerequisites Model for Data Warehousing, Proceedings of the 7th
European Conference on Information Systems (ECIS 1999),Copenhagen, June.

SECURE SHARING OF SENSITIVE DATA FOR EDUCATIONAL ORGANIZATION USING LIGHTWEIGHT ENCRYPTION OVER CLOUD COMPUTING

PRIYA DUDHALE – PISE
Research Scholar , JJT University , Rajasthan
priya.pise@gmail.com

Abstract- Cloud has been around for two decades and it consists of the vast amount of data from all over the world. Most of the people at a personal level and organization level have moved their data to the cloud and share data across all around the world. The main challenge faced by everyone is to share the data all over the world or at organizational level securely without giving away the important data to any exploiters. To overcome the challenge to share the data securely over the cloud, an efficient data encryption algorithm for encrypting data before sending it to the cloud. In this proposed we are using a combination of Attribute-Based Encryption and Byte Rotation Encryption Algorithm for encrypting the data before sending it to the cloud. This will help the user to securely store and share the data in encrypted form.

Keywords- Cloud Computing, Data Privacy, Encryption, Data Security, Data Sharing, Access Control.

INTRODUCTION:

Cloud computing means storing data and accessing that data from the Internet instead of Using Traditional hardware for most of the operations. More than 50% of IT companies have moved their Business to the cloud. Sharing of data over the cloud is the new trend that is being set on. The amount of data generated on a day to day life is increasing and to store that all of the data in traditional hardware is not possible because of limited storage capacity. Therefore transferring the data to the cloud is a necessity where the user can get unlimited storage. Security of that data over is the next big concern for most of us. After uploading the data to the cloud use loses its control over that data. [1]Since personal data files are sensitive, data owners are allowed to choose whether to make their data files public or can only be shared with specific data users. Therefore privacy of the personal sensitive data is a big concern for many data owners. When any of the people upload the data onto the cloud they are leaving their data in a place where monitoring over that data is out of their control, The cloud service provider can also spy on the personal data of the users. When someone has to share data over the data they have to share the password to each and every user for accessing the encrypted data which is cumbersome. Therefore to solve this problem data should be encrypted before uploading it onto the cloud which can be safe from everyone. Now the data encryption part bring some new problems such as we have to provide an efficient encryption algorithm such that if the data Is in encrypted format it cannot be easily to get break or get accessed by any exploiters. The next big concern is time consumption for encryption. Traditional Hardware with big configuration can encrypt data in short amount of time but limited resource devices suffers from this problem. They require more amount of time of encryption and decryption. So, an efficient crypto system is to be proposed which can worked equally or heterogeneously on all of the devices.

1. RELATED WORKS

Cryptography is the study of transmitting secret messages securely from one party to another. To accomplish this task, the original text, called plaintext, is translated into an encrypted version called cipher text, which is sent to the intended recipient. The recipient decrypts the text to obtain the original message. Cryptography is considered not only a part of the branch of mathematics, but also a branch of computer science.
Attribute-based encryption (ABE) is proposed by Sahai and Waters. Attribute-based encryption (ABE) is a moderately late approach that re-evaluates the idea of public key cryptography. Attribute-based encryption is also referred to as ABE is a sort of public-key encryption wherein the secret key of a person and the cipher-text is established upon attributes. In an ABE, a person's keys and cipher-texts are labeled with units of descriptive attributes and a symmetric key can decrypt a selected cipher-text only if there's a match between the attributes of the cipher-text and the person's key. It reduces the quantity of key used and hence makes encryption and decryption technique faster.

2. PROPOSED SYSTEM

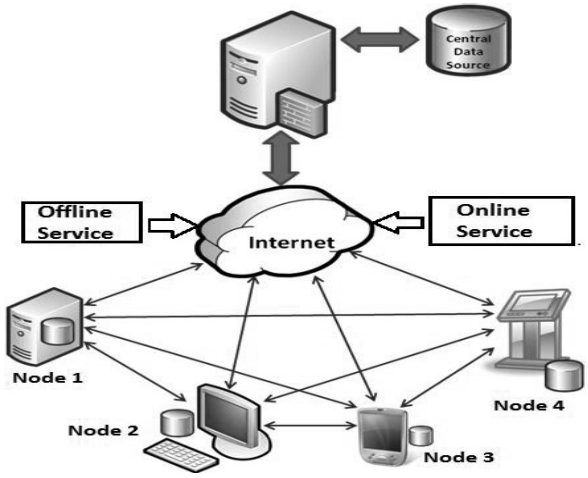
To address privacy issue in existing system we propose a crypto-system for secure sharing of data over the cloud, which uses combination Attribute Based Encryption and Byte Rotation Encryption Algorithm for secure encryption of the data over cloud. The main three works are as follows:

1. Identify the issues in cloud system for data storage on cloud. Since data is not secure on cloud user can upload the data in encrypted format.
2. Propose a crypto-system which can run on all limited resources devices. It can take data from the user and provide off-line-online service.
3. Apply Attribute Based Encryption Algorithm and Byte Rotation Algorithm for encryption of data to securely transfer the data between the users.

2.1 ADVANTAGES

- Here data can be transferred from one user to another securely over the cloud.
- The system cost will be decreased.
- It will work on all limited resource devices.

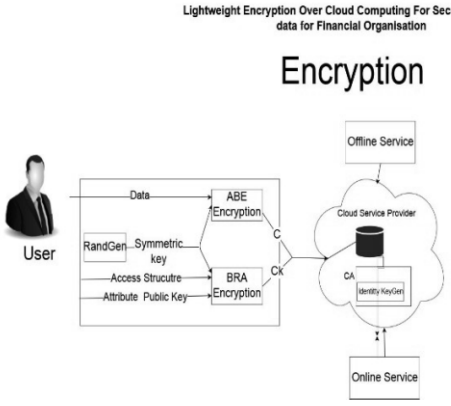
3.2 ARCHITECTURE AND MODULES DETAILS



The architecture of the proposed system is shown in the figure which shows the users and the operations involved. The detailed description of the architecture is explained as follows:

- **Nodes**: The User is responsible for uploading and sharing its personal data on the cloud.
- **On-line and Off-line Services**: In On-line Service data will encrypted and directly transfer to the respective user. In Off-line Service if there is no Internet Connection the data will get encrypted first and then it will get stored in Main Server. Until the system does not comes on-line the data will not be shared over the cloud
- **Cloud Service Provider**: Cloud service provider is responsible for providing all the required services to its users according to their demands.
- **Encryption and Decryption**: Here we are using the combination of ABE and BRE algorithm to encrypt and decrypt the files.
- **File Upload and Download**: The file which are uploaded on cloud are encrypted form. Users can download the file which are decrypted if he is authorized.

3.3 PROPOSED SYSTEM ENCRYPTION ALGORITHM



In our proposed system data is encrypted before uploading to the cloud.Combination of Attribute Based Encryption and Byte Rotation Algorithm are used for the encryption of the data.ABE will help to identify the attributes of the data and BREA will perform matrix operations on the block of the data to be encrypted.

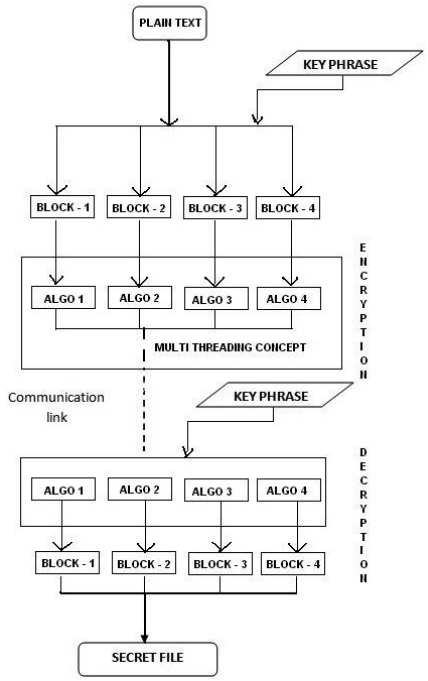
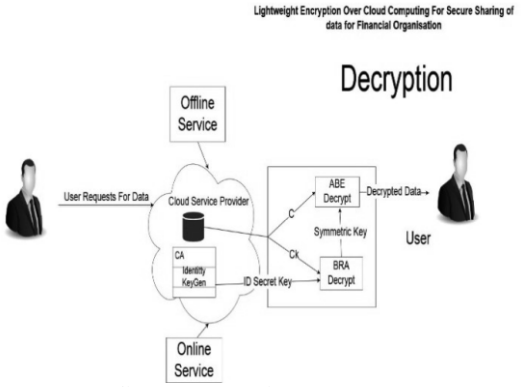


Fig: Flow Diagram

After performing encryption operation a random key is generated alongside the encrypted data.Data will be send in encrypted format to respective user.To decrypt this data reciever has to enter the One Time Password(OTP) which will be matched with key generated using ABE algorithm.



Proposed System Algorithm

Step-1: Start

Step-2: Accept the data from the user.

Step-3: The Attributes of the data from the users formats are obtained by the Attribute-Based Encryption.

Step-4: With the help of these Attributes ,Random Key is generated and type of data is obtained for encryption by BRE algorithm.

Step-5: The data is converted into equal number of blocks and N x N matrix will be generated on the basis of these blocks.

Step-6:Based on no. of blocks, pool of threads will be created.

Step-7:Run the threads in multi core system to create encrypted data in short amount of time.

Step-8:A secret key is generated in order to open the encrypted file which is stored in the cloud.

Step-9:The secret key is shared to the user via email or mobile number of the authorised user. This key will be used to decrypt the encrypted file.

Step-10:The file selected will be decrypted in the original form using the key. Step-11:Stop.

1. IMPLEMENTATION

This period of the venture is critical in light of the fact that at this stage the hypothetical plan is changed over into functional one. This stage is a basic stage since this stage require exceptionally exact arranging and need the learning of existing framework and its detriments.

The execution stage ought to be created by considering every one of the prerequisites, imperatives. The new framework ought to be successful and work appropriately.

CONCLUSION :

In this paper, the issue of sharing the data in cloud computing securely is resolved. Data privacy can be maintained by combination of ABE and BRE algorithm.. Authentication is used to guarantee data privacy and data integrity. This indicates that the proposed system can be used to enhance privacy preservation in cloud services.

REFERENCES :

1. A Lightweight Secure Data Sharing Scheme for Mobile Cloud Computing Ruixuan Li, Member,IEEE, Chenglin Shen, Heng He, Zhiyong Xu, and Cheng-Zhong Xu, Member, IEEE
2. Towards Se-cure Data Sharing in Cloud Computing Using Attribute Based Proxy Re-Encryption with Keyword Search Hanshu Hong; Zhixin Sun
3. X. Liang, Z. Cao, H. Lin, and 1. Shao, "Attribute based proxy re-encryption with delegating capa-bilities," in Proc. 4th ACM Int. Symp.
4. Priya Dudhale Pise,Dr. Nilesh J Uke,"Efficient Security Protocol for Sensitive Data Sharing on Cloud Platforms" in 2017 IEEE.
5. K. Liang et aI., "A OFA -based functional proxy re-encryption scheme for secure public cloud datasharing," IEEE Trans. Inf. Forensics Security,vol. 9, no. 10, pp. 1667-1680, Oct. 2014.
6. H. Hong, Z. Sun. "An efficient and traceable KP-ABS scheme with untrusted attribute authority in cloud computing", JoCCASA, 5(2).pp.I-8,201 6.
7. J. Liu, X. Huang, and 1. K. Liu, "Secure sharing of personal health records in cloud computing: Ciphertext-policy attribute-based signcryption," Future Generat. Com put. Syst., vol. 52, pp. 67-76,Nov. 2015.

E-COMMERCE IN INDIA - AN OVERVIEW

MRS. RITIKA R. KULKARNI
Ritika176@gmail.com

Abstract- E-commerce or electronic commerce is a process of buying and selling goods and services using computer network. This paper attempts to provide a review of various research studies carried out on E-commerce. This paper gives the overview of the future of E-commerce in India as well as the opportunities associated with E-commerce. This paper elaborates the difference between E-commerce and E-business and different challenges Indian E-business industry may face in future.

Keywords: E-commerce, E-business, online marketing, Internet, challenges

INTRODUCTION:

With an increased use of internet, Indian E-commerce industry has been growing rapidly. E-commerce market plays an important role in Indian economy. According to Morgan Stanley's report, Indian E-commerce market will grow at 30% annually and predicted gross merchandise value is expected to cross $120 billion by 2020.

The major role in ecommerce is played by online shopping trend in which people are started buying goods online using internet. Online shopping is now become extremely popular because of simplicity of buying and convenience.

E-business: Similar to E-commerce industry, other emerging market in India is E-business. Apart from buying and selling goods, E-business or electronic business includes activities such as providing services to the clients, communicating with employees and business partners with the help of internet or other communication network.

2. Difference between e-commerce and e-business

E-commerce involves selling and purchase of goods, whereas E-business involves all the internal processes apart from selling and purchasing, such as finance management, inventory management, risk factors and additional offline sales. In short, E-commerce is emphases more on trading over the internet, and E-business is more running entire business activities using computer network or internet. The basic requirement for E-commerce is a website; apart from that, marketing, advertisement, selling of product, any monetary transaction has to be done using Internet. There are four types of E-commerce

- B2B – stands for Business to Business. In this buying and selling of goods and services takes place between businesses to business.

Example: Oracle, Alibaba.

- B2C – Business to Customer. In this, goods are sold by the business to customer.

Example: Intel, Dell etc.

- C2C – The transaction between customers to customer.

Example: OLX, Quicker etc.

- C2B-includes transaction between customers to the business.

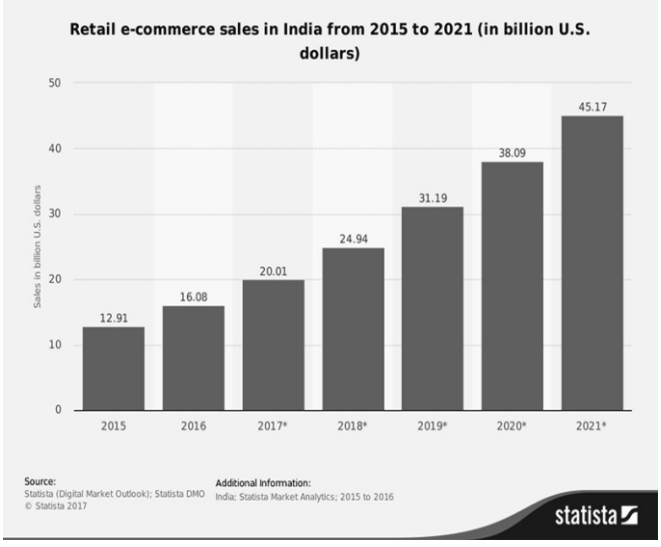
E-business includes E-commerce which covers all the internal process of such as production, inventory management, product development, risk management, finance, and human resources. E-business strategy is more complex, more focused on internal processes, and aimed at cost savings and improvements in efficiency, productivity and cost savings. Amazon and eBay are the examples of world's biggest E-businesses.

E-commerce in India

India is a developing country. With the increased use of internet, E-commerce is also growing rapidly. Key factors to be consider for this growth are increased used of personal computers, increasing number of Smartphone users, rising standards of living, competitive Internet Service providers and cheap data usage rates are main factors working as fuel to E-commerce industry. As a result of this, E-commerce is showing tremendous growth in India.

E-commerce provides multiple benefits to the customer. Goods are available online, with wider options to choose from and doorstep delivery is also available which saves overall purchase time. Customers can surf in the internet and search for products, customer can compare different models. They can also go to other online product review sites and review the product. Once the customer likes a product after all research, he can order for it online. Most of the E-commerce business companies provide flexible payment modes and return policy and warranty. E-commerce has also made it easy to make payments using various options such as credit cards, debit cards, direct online money transfers, all this attracts majority of people to opt for online shopping.

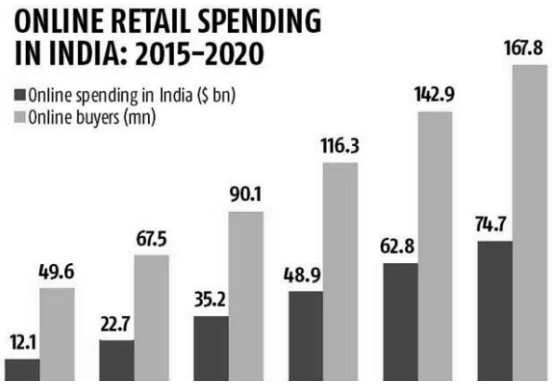
According to Google India, online shoppers may cross 100 million by the end of the year 2017.



"According to World pay projections, India will be the world's second largest E-commerce market by 2034.

Top contributors of India's E-commerce ecosystem are:

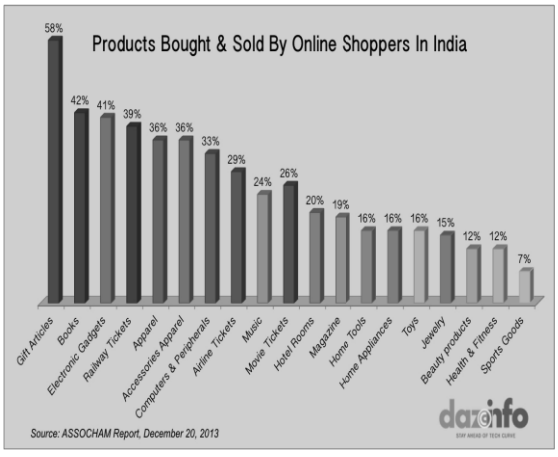
Online Retail: study shows than more than 60% of citizens in India visited a retail site in Nov 2017. Almost everything such as groceries, electronic goods, computers, laptops, mobile sets, apparels, jewelries can be purchased online the most popular online shopping sites are Amazon in, flipkart.com, paytmMall.com, snapdeal.com, eBay. In, jabong.com and many more.



Online Travel: This is second popular sector in E-commerce. It includes online ticket booking, hotel reservations

Online Classified: This is another category making its mark in E-commerce. Online job portals, matrimonial sites, yellow pages, online couponing are the largest contributor in this sector.

Online Banking and Other Services: besides online shopping, people started doing all the banking transactions online. People also avail services such as online bill payment, online insurance payment, and online recharge with the help of credit or debit card.



Challenges for the e-commerce sector in India

The phenomenal growth of the E-commerce sector is accompanied by certain challenges:

Poor Infrastructure Facility: Infrastructure is the biggest problem in India. There are many states and districts where infrastructure facilities such as transport, electricity supply, roads, etc. are the major bottleneck

Lack of knowledge of new Technology: Poor literacy rate leads to less use of advanced technology such as Internet especially many people in rural areas are illiterate. They don't have knowledge about E-commerce. Due to illiteracy they are not able to do online transactions.

Absence of E-commerce laws: one biggest problem in E-commerce is absence of cyber law to regulate the online transactions.

Logistics or Shipping Challenges: Issues related with high delivery charges for products, delay in delivery and lack of proper courier services in some areas. Logistics is a concern for both buyers and sellers. Sellers have to deliver the product safe and secure to the right person and in right time frame. Regular postal services do not offer an acceptable level of service. Insurance for high value articles leads to higher cost.

Risk in online shopping and quality assurance: There is inherent risk in online shopping as regards the authenticity of the web sites. Fraudsters may hack web sites and mislead customers.

Privacy and Security: Many Business owners fail to protect their online business through installation of authentic protection services like antivirus and firewall protection, it may lead to inherent risk in online shopping as the web sites may get hacked by hackers.

Collection of payment and transfer of payment: It is still difficult to make payments through internet in India due to unstable internet connection. Also most of the online buyers prefer cash on delivery payment mode; this manual cash collection is tedious, requires more manpower.

Customer satisfaction: In India, most of the customers prefer to buy product physically. Customers prefer to touch and feel the product before buying. This may lead to the problem for companies dealing with product like online jewelry, antique and handicraft items.

Product return policy and refund: This is another factor which may lead to overall loss of revenue. If customer is not satisfied with the product, it must be replaced or returned back. It may lead to loss of money, loss of shipping cost and more is loss of company image or reputation.

Constant reinvention: As the technology is changing rapidly, E-commerce companies need to constantly update their websites in order to provide better service to the customer. It may lead to high maintain and advertisement cost.

CONCLUSION :

E-commerce is continuously growing. Due to the increasing use of Internet, online shopping trend is becoming popular in people. The future of E-commerce in India would be bright in the upcoming years if all essential factors would be implemented. There are various segments that would grow

in the future like: Travel and Tourism, electronic appliances, hardware products and apparel. There are some challenges in front of the E-commerce companies including customer uncertainties, which can be resolved using good decision making and business strategies. People could found various opportunities of employment. Overall the future of E-commerce in India would be bright in the upcoming years if all essential factors would be implemented.

REFERENCES :

[1] www.researchmanuscripts.com
[2]E-commerceGuide.Com
[3]www.slideshare.net/.../research-paper-on-ecommerce-challenges-and-opportunities
[4]http://www.indianmba.com/Faculty_Column/FC822/fc822.html [5]https://en.wikipedia.org/wiki/E-commerce_in_India
[6]http://www.investorwords.com/1637/e_commerce.html
[7]http://www.ibef.org/download/The-Riseand- Rise-of-E-commerce-in-India.pdf

IMPERATIVE TECHNIQUE TO IMPROVED PERFORMANCE OF ETL (EXTRACTION, TRANSFORMATION AND LOADING) LOADING MULTIPLE FILES IN DATA WAREHOUSE ENVIRONMENT

MR. KALYAN C. JAGDALE
Assistant Professor,
Prof. Ramkrishna More College, Akurdi, Pune-44.,
E-mail: kalyan.jagdale7@gmail.com

MR. SANTOSH T. JAGTAP
Assistant Professor,
Prof. Ramkrishna More College, Akurdi, Pune-44.,
E-mail: st.jagtap@gmail.com

Abstract : Data warehouse (DW) are complex computer systems whose main goal is to facilitate the decision making process of knowledge workers. Extraction-Transformation-Loading (ETL) tools have pieces of software. It is responsible for extraction of data from heterogeneous operational several sources, their transformation (conversion, cleaning, normalization, etc.) and loading data into Data Warehouse (DW). There are several quality issues when data loading into target environment. So, it is essential to improve data quality in the initial design stages of the Data Warehouse environment. In this paper, we have suggested technique to loading multiple file into SQL server faster, more smoothly and more easily.

Keywords : DW, ETL, DSA, Data Warehouse

INTRODUCTION:

The data warehouse has historical data to provide analytical process, decision-making and data mining tools. The data stored by different organizations often differs from the original, content, and representation area and express concerns about different areas. The data partly derived from internal transactions of an administrative, logistical, and commercial nature and partly from external sources. It is necessary to process them through proper extraction tools and analytical methods, which are capable of converting them into information and knowledge, resulting in subsequently used by decision makers. The extracting data from source systems and bringing it into the data warehouse environment is called as ETL. It stands for Extraction, Transformation and Loading. ETL process is an important phase where data extracted from different data sources and propagated to the DSA (Data Staging Area) where it transformed and cleansed before being loaded into the Data Warehouse environment. Source, Staging area and target environments have many different data structure like flat files, XML data sets, relational tables, non-relational sources, web log sources, legacy systems, and spreadsheets.

Business Intelligence at the Data Warehouse system environment requires users to get strategic information is new patterns of data warehousing. This new system environment supports the daily process and enables it to make strategic decisions. Data warehouse is the only viable solution and it based on the data extracted from operational systems are all totally unsatisfactory. At a high level, the data warehouse includes an important measure of business processes stored on business dimensions. For example, a data warehouse might have units of sales, by product, day, customer group, sales district, sales region and promotion. The business dimensions are product, day, customer group, sales district, sales region, and promotion. Data is derived from the operational systems that support the basic business processes of the organization. During operational systems and data warehouses, there is data staging area. In staging area, the operational data is cleansed and transformed into a suitable form for placement in the data warehouse environment for easy retrieval.

Data Extraction

- Select data sources and determine the types of filters to be applied to individual sources
- Generate automatic extract files from operational systems using replication and other techniques
- Create intermediary files to store selected data to be merged later
- Transport extracted files from multiple platforms
- Provide automated job control services for creating extract files
- Reformat input from outside sources
- Reformat input from departmental data files, databases, and spreadsheets
- Generate common application code for data extraction
- Resolve inconsistencies for common data elements from multiple sources

Data Transformation

- Map input data to data for data warehouse repository
- Clean data, deduplicate and merge/purge
- Denormalize extracted data structures as required by the dimensional model of the data warehouse
- Convert data types
- Calculate and derive attribute values

- Check for referential integrity
- Aggregate data as needed
- Resolve missing values
- Consolidate and integrate data

Data Staging

- Provide backup and recovery for staging area repositories
- Sort and merge files
- Create files as input to make changes to dimension tables
- If data staging storage is a relational database, create and populate database
- Preserve audit trail to relate each data item in the data warehouse to input source
- Resolve and create primary and foreign keys for load tables
- Consolidate datasets and create flat files for loading through DBMS utilities
- If staging area storage is a relational database, extract load files

Data Storage

This area covers the process of loading the data from the staging area into the data warehouse environment. All functions for transforming and integrating the data are completed in the data staging area. The prepared data in the data warehouse is like the finished product that is ready to be stacked in an industrial warehouse.

Once succeeded in getting a set of data in SQL Server. Loading multiple flat files from one or more directories use SSIS and a Foreach Loop container as follows:

1. Open SSIS and create a new package.
2. Add a variable at package level.
3. Right-click in the Connection Managers tab and select New Flat File Connection.
4. Name the connection manager.
5. Click the Browse . . . button and navigate to the directory containing the files to load.
6. Add an OLEDB connection manager at project level that connects to the database.
7. Add a Foreach Loop container on to the Control Flow pane. Name it and double-click to edit it.
8. Select Collection on the left, and choose Foreach File Enumerator as the enumerator type.
9. Click the Browse . . . button and navigate to the directory containing the files to load.
10. In the Files field, enter the file and/or extension filter to limit the files that will be enumerated Click the Fully Qualified radio button to return the full path and file name.
11. Select Variable Mappings on the left. Select FileName as the variable to use.
12. Click OK to close the dialog box.
13. Add a Data Flow task inside the Foreach Loop container, and configure it to load from the Flat File source to a destination table.
14. Click the Flat File connection manager that created in step 4, and display the Properties window (by pressing F4), unless it is already visible.
15. In the Expressions property, click the Ellipses button to display the Expressions dialog box.
16. In the Property Column, select Connection String.
17. Click the Ellipses button for the Connection String property to display the Expression Builder dialog box.
18. Expand the variables list and drag the User::FileName variable down into the Expression Field.
19. Click OK to close the dialog box. It returns to the Property Expressions Editor dialog box.
20. Click OK to close the dialog box and use the file name found in the FileName variable instead of the hard-coded file name that originally used when setting up the connection manager.
21. Run the package.

CONCLUSION:

If suggested technique set up correctly then all files in source directory should into destination environment. This technique is essentially a loop through all or a selection of the files in a directory. It needs a string variable that is used to hold the name of each file that is loaded and it replace the actual file that defined when creating the SSIS task. The technique is almost identical with a single file load. It presumes that all the loading files are in same format and will be loaded into the same destination environment.

REFERENCES :

1. *PAULRAJ PONNIAH*, "Data Warehousing Fundamentals A Comprehensive Guide for IT Professionals", *John Wiley & Sons, Inc. ISBNs: 0-471-41254-6 (Hardback); 0-471-22162-7 (Electronic)*
2. W. H. Inmon, "Building the Data Warehouse, Third Edition", *Wiley, ISBN: 0-471-08130-2.*
3. Jiawei Han and Micheline Kamber , "Data Mining: Concepts and Techniques Second Edition", *Morgan Kaufmann Publishers, ISBN 10: 1-55860-901-6.*
4. Adam Aspin, "SQL Server 2012 Data Integration Recipes", *Apress, ISBN-13 (pbk): 978-1-4302-4791-3, ISBN-13 (electronic): 978-1-4302-4792-0.*
5. Ralph Kimball and Joe Caserta, "The Data Warehouse ETL Toolkit", *Wiley Publishing, Inc., eISBN: 0-764-57923-1.*
6. Claudia Imhoff, Nicholas Galemmo and Jonathan G. Geiger, "Mastering Data Warehouse Design Relational and Dimensional Techniques", *Wiley Publishing, Inc., ISBN: 0-471-32421-3.*
7. Ralph Kimball and Margy Ross, "The Data Warehouse Toolkit Second Edition The Complete Guide to Dimensional Modeling", *Wiley Computer Publishing, ISBN 0-471-20024-7.*

ASSESSMENT OF IMPACT OF WATERSHED ACTIVITIES ON GROUND WATER QUALITY IN ARID REGION (KR22, KR25, KR 34 AND BM 114) OF MAHARASHTRA, INDIA

DR. SATISH S. PATIL*, DR. PAWAR C.B. [+], THAKARE S.Y. [+], MORE S.B. [+]
*Professor and Head, Department of Environmental Science, Dr.BAMU, Aurangabad.
+ Assistant Professor, Department of Computer Science, Indira College of Commerce and Science, Tathwade, Pune.

Abstract : Several physical and chemical causes are active in changing the composition of the groundwater percolating through the soil and rocks. The watersheds development is proven technology for harvesting good quality water. Thus, in this investigation we analyzed the impact of watershed interventions, 48 water samples each has been studied for pre monsoon and post monsoon condition. The statistical summary of 19 hydro-geochemistry parameters has been presented in this research investigation. This research deals with the study of impact of pre-monsoon and post-monsoon season on the various parameters of groundwater. Also an attempt is made to study the type of changes in the groundwater parameters during pre and post monsoon season. The research shows that the parameter which has maximum variation for its content in groundwater is potassium (K). Using data mining and statistical techniques, some interesting results about the changes in the groundwater parameters are studied.

INTRODUCTION:

Lotic ecosystems helpful in maintaining the natural cycles in the environment. Lentic ecosystems are standing ecosystems works contrary to lentic ecosystems. Especially dams are better example of lentic ecosystem. Dams and human development are two sides of one coin but converting lotic ecosystems in lentic ecosystems largely threats the human life in entire world. Salinity, alkalinity, decreasing water quality, significant health issues and agriculture allied problems such as loss of soil fertility, significant decline in crop yields are the principal problems in many countries. Watershed development for harvesting good quality water is always found to be effective tool in some recent decades. Watershed development activities are small lotic ecosystems have very negligible impact over native environment.

In some recent decades watershed development and management have got immense importance in drought prone areas of India. Salinity and alkalinity hazards are associated in both excessive irrigation areas and drought prone areas of India. Hence harvesting good quality water at its own source is one of best intervention to overcome the challenge.

Objectives – To analyze the impact of watershed development activities on ground water quality.

Methodology –

Hydro-geochemistry and Ground Water Characteristics in the Study area –

Generally the rainwater contains the Carbon dioxide, which is treated as a powerful agent because of its properties to break up the all mineral when it reaches with ground water. The new compounds formed were CO_3, HCO_3, SO_4, Na, K, Ca and Mg, Soluble silicates and free Sio_2.

A number of studies have been proved that, several physical and chemical causes are active in changing the composition of the groundwater percolating through the soil and rocks. Evaporation (Generally it is tend to be 7-8 mm/day in the Deccan trap area), Base Exchange, adsorption, oxidation of sulphides, and reduction of sulphates cause changes in their chemical composition. Based on the geochemistry about 48 samples were collected and analyzed using APHA-AWWA-(1975) manual for the dug-wells of the study area for the pre-monsoon season (May 2010) and 48 samples for post monsoon (November 2010), the data was presented in Table No.1 and 2. Following parameters have been considered for analysis.

RESULTS AND DISCUSSION :

1.Variability in the ground water parameters during pre-monsoon season.

From the analysis of ground water parameters it is revealed that, lowest variation is reported for pH. The coefficient of variation (CV) is reported 3.49%. The maximum variation is observed for K for which CV is 320 %. The iron content is also not showing significant variation. Na and Na+k changing from sample to sample to much extent i.e 101% and 103% respectively. The analysis of another parameters for their variations among different samples is given in Table No.1. The less variation have been observed in most of parameters, which reveals the importance of watershed activities.

Table No. 1 Pre - monsoon: Variability in the ground water parameters during pre-monsoon season.

Parameters	Mean	Standard Deviation	Coefficient of Variation %
pH	8.31	0.29	3.48
EC	1030.33	585.04	56.78
TDS	660.86	374.4	56.65
T/HAR	351.36	188.34	53.60
T/ALK	260.82	123.66	47.41
Ca	83.62	62.27	74.46
Mg	34.58	19.43	56.18
Na	88.82	89.97	101.2
K	7.04	22.52	319.8
Na+K	95.84	99.09	103.3
Fe	0.21	0.06	28.57
Cl	151.78	132.4	87.23
SO4	96.67	126.88	131.2
CO3	2.36	3.27	138.5
HCO3	296.24	159.24	53.75
NO3/N	8.55	4.7	54.97
F	0.74	0.31	41.89
SAR	2.2	2.07	94.09
Na%	32.41	18.94	58.43

1. **Variability in the ground water parameters during post-monsoon season.**

From the analysis of ground water parameters it is revealed that, lowest variation is reported for pH. The coefficient of variation (CV) is reported 4.18 %. The maximum variation is observed for K for which CV is 380 %. The variation in water samples is observed. The lot of factors contribute to variations in values of K. There may be possible chance of contamination of sample water from water drained from agriculture field. The iron content is also not showing significant variation. Na and Na+k changing from sample to sample to much extent i.e 106% and 106% respectively. The analysis of another parameters for their variations among different samples is given in Table No.1. The less variation have been observed in most of parameters, which reveals the importance of watershed activities.

There is no impact of pre monsoon or post monsoon season on the variability in the various parameters of ground waters.

Table No. 2- Post – monsoon- Variability in the ground water parameters during post-monsoon season.

Parameters	Mean	Standard Deviation	Coefficient of variation
pH	8.13	0.34	4.18
EC	908.96	485.16	53.3
TDS	596.84	332.33	55.6
T/HAR	354.67	188.1	53.0
T/ALK	247.06	106.33	43.0
Ca	87.78	72.91	83.0
Mg	35.57	18.32	51.5
Na	65.71	69.8	106.3
K	3.51	13.37	380.9
Na+K	71.02	75.49	106.2
Fe	0.22	0.07	31.81
Cl	133.39	118.49	88.82
SO4	88.39	104.42	118.1
CO3	2.82	3.19	113.1
HCO3	274.44	136.67	49.79
NO3/N	8.57	4.92	57.40
F	0.67	0.29	43.28
SAR	1.68	1.71	101.7
Na%	27.67	16.7	60.35

1.Comparative analysis of ground water parameters between pre and post monsoon season.

The values of total hardness, calcium, magnesium, iron, carbonate and nitrate is found to be reduced during post monsoon season whereas the content of remaining ground water parameters increased.

Table No.3: Changes in groundwater parameters during pre-monsoon and post-monsoon season

Parameters	average of pre-monsoon(x)	Average of post-monsoon(y)	Difference(x-y)
pH	8.31	8.13	0.18
EC	1030.33	908.96	121.37
TDS	660.86	596.84	64.02
T/HAR	351.36	354.67	-3.31
T/ALK	260.82	247.06	13.76
Ca	83.62	87.78	-4.16
Mg	34.58	35.57	-0.99
Na	88.82	65.71	23.11
K	7.04	3.51	3.53
Na+K	95.84	71.02	24.82
Fe	0.21	0.22	-0.01
Cl	151.78	133.39	18.39
SO4	96.67	88.39	8.28
CO3	2.36	2.82	-0.46
HCO3	296.24	274.44	21.8
NO3/N	8.55	8.57	-0.02
F	0.74	0.67	0.07
SAR	2.2	1.68	0.52
Na%	32.41	27.67	4.74

Also it can be observed that, most of the GW parameters are decreased during post monsoon season as compared to pre monsoon season.

1.To study whether there is significant impact of pre-monsoon and post-monsoon season on change in groundwater parameters

To study this, we use paired –t test for testing the null hypothesis that
H_0 : There is no significant change in groundwater parameters after monsoon
Against alternative hypothesis that
H_1 : There is significant change in groundwater parameters after monsoon
The test statistic used here is given as

$$t = \frac{\bar{d}}{s/\sqrt{n}} \rightarrow t\ (n\text{-}1)$$

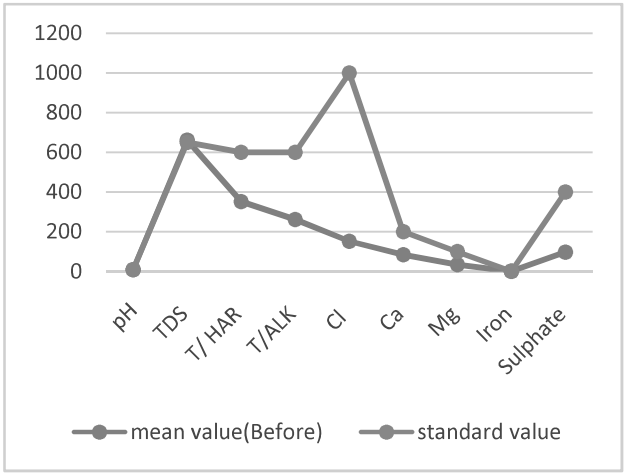
Thus, here we get, calculated value of t = 2.239137
And the critical value of t = 2.100922

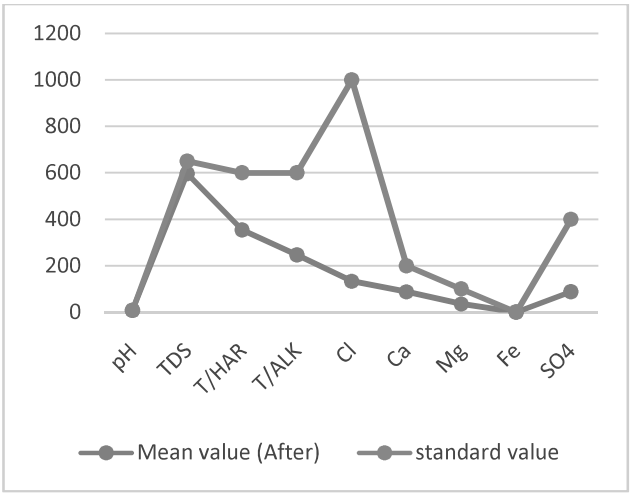
Thus, since $t_{cal} > t_{tab}$
Ho is rejected and hence it means there is significant change in groundwater parameters after monsoon season.

5. Comparative study of groundwater parameters with permissible limits

Parameter	mean value(Before)	standard value
pH	8.31	8.5
TDS	660.86	650
T/ HAR	351.36	600
T/ALK	260.82	600
Cl	151.78	1000
Ca	83.62	200
Mg	34.58	100
Iron	0.21	1
Sulphate	96.67	400

Parameters	Mean value (After)	standard value
pH	8.13	8.5
TDS	596.84	650
T/HAR	354.67	600
T/ALK	247.06	600
Cl	133.39	1000
Ca	87.78	200
Mg	35.57	100
Fe	0.22	1
SO4	88.39	400





It is evident that all the ground water parameters value fall under the permissible limit. In this investigation we compare these values with Ground Water Survey and Development Agency (GSDA), Pune. It is found that, the values of TDS, Total hardness, Calcium, Magnesium, Iron and Sulphate are found within limit for both pre and post conditions. These values reflects the importance of watershed development activities on ground water quality.

CONCLUSION :

In some recent decades watershed development and management have got immense importance in drought prone areas of India. Drought is striking large part of world and almost maximum part of India. About 338 districts of 14 states have been fall under drought prone area of country.

Agriculture is a backbone of our country and therefore land management is very crucial for sustaining the agriculture productivity. Most of the DPAP area is possess degraded lands, water scarcity, fragile ecosystem, depleted ground water levels and extreme poverty. In India agro

ecologies contribute 56.7 percent net sown area, 40 percent of food production, and 66 percent of the livestock. About 85 percent of coarse cereals, 83 percent pulses, 42 percent of rice, 70 percent of oilseeds and 65% cotton are cultivated is rain-fed. (CRIDA, 2011:31). In India most of irrigated regions are coped with salinity, alkalinity decline in fertility of soil. Contrary to this drought prone region have huge potential to harvest good quality rainwater which has no adverse impact on land deterioration. Hence watershed development and effective management is best tool for drought prone region development.

REFERENCES :

1. Eggleton R.E., C. Foundoulis and Varkevisser (1987): Weathering of basalt: changes in rock chemistry and mineralogy. Clays and Clay Minerals, v. 35, p. 161-169.
2. Duraiswami Raymonad., Babaji Maskare and Uday Patankar (2012): Geochemistry of the Groundwater in the Arid region of Deccan Trap of country, Maharashtra, India, Memoir, Indian Society of Applied Geomatics, June 2012, p.61-67.
3. Kulkarni, H., S.B. Deolankar, A. Lalwani, B. Joseph and S.Pawar (2000): Hydro-geological framework of Deccan basalt groundwater systems, India, *Hydrogeology Jour*., v. 8, p. 368-378.
4. APHA-AWWA-(1975): Standard methods for examination of water and waste water, 20th Ed.
5. BSI- Bureau of Indian Standard (BIS), Specification IS - 10500-91, Revised (2003): Central Board for the prevention and control of water pollution, New Delhi and GSDA, 2008.
6. ICMR- Indian Council of Medical Research,(1975): Ground Water Survey and Development Agency.
7. Indian Standards institution (1979): Methods of sampling and test (Physical and chemical) for water used in industry., IS: 3025-1964.
8. Indian Standards Institution, (1983): Indian standard specification for drinking water. ISI: 10500.
9. Uppal, H. L. (1962): Reclamation of saline and alkali soils. Seminar on Soil Salinity and Alkalinity Problems. Water Technology Center, Indian Agriculture Research Institute, New Delhi, p. 86-92.
10. Saleh A., F. AL-Ruwaih, and M. Shehata (1999): Hydrochemical processes operating within the main aquifers of Kuwait. *Jour Arid Environ*. v. 42,p. 195-209.
11. Richards, (1954): Diagnosis and improvement of saline and alkali soils. U.S. Department of Agriculture Handbook, p.60.

SURVEY: ANALYSIS ON VISUALLY IMPAIRED GUESTS TO FEEL FIREWORKS

PRAVIN AVHAD[1], AKSHAY PHALKE[2], RAVINDRA TAMBE[3],
Department Of Computer Engineering[1,2,3].
Shri Chhatrapati Shivaji Maharaj College Of Engineering
Savitaribai Phule Pune University. India.
pravin123.avhad@gmail.com[1], akshayphalke3@gmail.com[2],ravindra.tambe01@gmail.com[3]

Abstract : The aim of this paper is to represent a method where a blind person can get feeling about a unique haptics-supported fireworks display experience specially tailored to visually impaired users. Fireworks provide dazzling night displays that people all around the world continue to enjoy. This paper is specially designed for visually impaired visitors, the idea being to offer an aesthetic technology for the blind and visually impaired community. This paper works on haptic feedback, which essentially involves a series of tactile vibrations that translate the experience.

Keywords : Visual communication, Latex screen Display, Water jets spray, Visual effects, and data visualization.

INTRODUCTION:

Visual impairment is the major disability faced by millions of people around the world. Humans are very sensitive to feeling about fireworks display experience to visually impaired users. Humans easily identify a wide variety of fireworks such as New Year celebration, Party Celebration etc. Exploiting this human capability we can convey many kinds of information to the visually impaired. It is possible to create an impression of objects or things that a person visually feels about a unique haptics-supported fireworks display. Water jets strike on the rear surface of a flexible screen. After that it produces tactile effect on the front surface for all blind, visually impaired and sighted people. It is therefore possible to substitute the eyes to represent the physical world to some extent. This paper works on blind person feeling about visually impaired guests to feel fireworks. Feeling fireworks for blind person is one of analysis on visually impaired guests.

II. RELATED WORK

Many analysis have been done in the area of visually impaired guests. Mukhtar Masood Rana, Marcian Cirstea and Tim Reynolds [1] proposed mechanism in which impaired person using mobile phone with the help of prototype SWANS-M (Wireless). But this mechanism is not possible to available wireless technology everywhere. Kazushige Magatanil, Koji Sawa and Kenji Yanashima [2] proposed mechanism in which they suggested that every subject could follow the navigation voice but some subject couldn't go across the wide pathway straight, and then the system lost their position. M. P. Menikdiwela, K.M.I.S.Dharmasena and A.M. Harsha S. Abeykoon [3] proposed mechanism in which new product concept of an electronic travelling aid with haptic percept- ion for the visually impaired people which are having very costly. Kevin Kumar D, Senthil Kumar Thangavel [4] model is proposed for detecting text from natural scene video and informing the user through audio to guide visually impaired people in a library but it works only library concepts and related character. Frederico Ferreira, Sofia Cavaco [5] proposed mechanism in which the main goal of the game is to help and motivate blind teenage students to learn and like mathematics but the plan to make the question-answer files editable, which will drawback the game's usability. André Lima, Daniela Mendes & Sara Paiva [6] proposed a case study that targets the two mentioned problems, developed in the Historical Center of Viana do Castelo, a city on the north of Portugal, made in cooperation with a Visually Impaired Association but they having some limitation regarding to do with public urban transportation. Daniel Vera Yánez, Diego Marcillo[7] proposed a mechanism in which they put that use of a system that detects and recognizes nearby obstacles, giving an audible warning in order to avoid a collision. But they didn't find the long distance obstacle for blind person. k.gopala krishnan, c.m.porkodi, and k.kanimozhi[8] proposed an algorithm for image recognition by speech sound. Blind people face a number of challenges when interacting with their environments because so much information is encoded visually. But it works only related to sound.

This paper addresses that blind person who not only hearing sound of fireworks but also able to visually feel the fireworks on the latex screen display with the help of visual communication and visual effect which are provided to latex screen display with the help of some effects and technology as shown in fig. 1

Fig.1. **visually feel fireworks on the latex screen display**

III. TECHNOLOGY USED

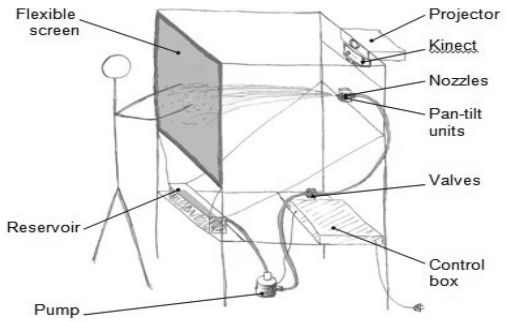


Fig.2 **Picture drawing of feeling firework to user & how it work**

As shown in above fig. 2 a latex screen is available in front Of blind person. The screen is latex and has dimensions of 1 x 1 m. Water is moving through the system in a closed by a controllable pump that provides variable water pressure. An Arduino controls the pump, the pan-tilt of the nozzles and the opening and closing of the nozzle valves. We use different nozzles for different firework effects. There are three jet nozzles for creating rockets and explosions, one shower nozzle for creating crackle effects, and one nozzle with a novel design for a blooming flower effect. Compound fireworks are created by sequencing effects. Rear projection provides visual content, and a rear-mounted Kinect detects a user's finger press on the flexible screen to provide interaction. A laptop con- trols the projector & Kinect.

Flower Nozzle

A central effect in firework shows is the *flower* effect an explosion in the sky that starts at a single point and forms an expanding circle of light points as shown in We have developed a novel custom nozzle with a spray patt-ern that mimics the flower firework by starting off as a sin-gle column of water and creating an expanding cone as the pressure is increased. The nozzle has a silicone membrane with small holes through which the water flows. The membrane is flat when the nozzle is off and it is tuned so that the nozzle pro- duces a single column of water at the lowest operating water pressure . Further increasing the pressure creates continuously-expanding cone of water.

Practicalities Of deployment

The device is free standing, on caster wheels for ease of Mov- ement, & runs off one standard power cord. It has been designed to be modular and transportable with easy setup. the pump is a near-silent medical product, while the water jets make a light drumming sound on the plastic screen which is audible but unobtrusive? The back-projected visual fireworks are clearly visible outdoors at night or in regular in door lighting condition. The device sits in a plastic tray to prevent accidental water leakage.

CONCLUSION :

This work has presented a novel design of tactile screen that is scalable and economic, and which has been used to make a tactile firework show. It provides an extra dimension for blind people at a traditional firework show, in the context of an inclusive experience shared by all. A user study showed that the tactile fireworks are meaningful analogs of visual fire- works.

REFERENCES :

[1] Mukhtar Masood Rana, Marcian Cirstea and Tim Reynolds, "Developing a Prototype using mobile devices to assist visually impaired users".
[2] Kazushige Magatanil, Koji Sawa and Kenji Yanashima, "DEVELOPMENT OF THE NAVIGATION SYSTEM FOR THE VISUALLY IMPAIRED BY USING OPTICAL BEACONS".
[3] M. P. Menikdiwela, K.M.I.S.Dharmasena and A.M. Harsha S. Abeykoon, "Haptic based walking stick for visually impaired people"
[4] Kavin Kumar D, Senthil Kumar Than gavel, "Improving Capabilities for Visually Challenged Person in Library Environment".
[5] Frederico Ferreira, Sofia Cavaco," Mathematics for all a Game-Based Learning Environment for Visually Impaired Students".
[6] André Lima, Daniela Mendes & Sara Paiva," Mobile solutions for visually impaired people."
[7] Daniel Vera Yánez, Diego Marcillo," *Portable Navigation Device for Blind People*".
[8] K.GOPALA KRISHNAN, C.M.PORKODI, and K.KANIMOZHI,"IMAGE RECOGNITION FOR VISUALLY IMPAIRED PEOPLE BY SOUND".
[9] Kingsley, P. eTable: A haptic elastic table for 3D multi-Touch Interactions. Master's thesis, University of Bristol, 2012.
[10] Nature News. Deaf people use 'mind's ear' to process Vibrations.

http://www.nature.com/news/2001/011127/full/news011129-10.html, 2001.

METHODS USED FOR DETECTION AND CORRECTION OF SKEW IN DOCUMENT IMAGES

SARIKA T. DEOKATE [#1], NILESH J. UKE [*2]
[#]Research Scholar JJTU Rajashtan, Research Guide JJTU Rajashtan
[1]sarikajaankar@gmail.com
[2]nilesh.uke@gmail.com

Abstract: In document image processing skew detection is one of the main factors to consider for further processing. There may be inclination or angle rotation of the document images after scanning it. It is required to remove this inclination before using this image for normalization and segmentation in pre-process. To segment the characters and to check the characters accurately, skew detection and correction at word level is also important task in any document analysis. In this system skew detection and correction of word is also considered to increase the efficiency and accuracy of recognition. The main motto of this work is to increase the efficiency of the proposed system. In proposed system Hough Transform and rotation transformation will be proposed based on the text line for Devnagari Marathi text.

Keywords: skew detection, skew correction, document analysis, Hough transform

INTRODUCTION:

Now a day almost all the private and government organization are moving ahead with digital era.The past year's paper documents having equal importance in current digital era. Digitization of these documents is becoming necessity and for that lot of efforts are taken by the organizations. Optical Character Recognition (OCR) is becoming first choice of all the organization to convert these paper documents to digital formats. But many OCR systems are not proving results up to the mark and hence need tremendous research. In the recent years recognition of handwritten documents is a challenging task due its variety of feature. In handwritten documents two skew are detected prominently. One skew is at page level, which is mostly occurred at the time of scanning or taking picture by camera. Second skew is at word level, which is occurred due to writing style of the person. Due to these skews segmentation and recognition rate is decreased. So it is required to detect and correct the skew for accuracy improvement [1, 11]. In following section some of the techniques are specified which are used by the researchers for skew detection and correction.

2. Literature Survey:

In past decades, lot of work has been done for skew analysis, detection and correction of document images. Skew detection of printed document is easier task as compared to the handwritten text documents. Though this skew is accidental, it reduces the recognition rate dramatically and so need to be removed or reduced. Different methods like Fourier Transform, Hough Transform, Cross Correlation, Nearest neighbour method, Gradient Analysis, histogram profile projection methods are used for the same.

When considering the skew detection and correction following information need to be considered:

2.1 Skew: Skew is described as the angle of falsehood in the recognized word or given image. Mostly document gets rotated or tilted at the time of scanning the document. Due to the writing style, there can be skew in words too [7].

2.2 Skew angle: A line of data/text is likely to be in horizontal; its variation from horizontal line is described as skew angle. If this angle is greater than 2, then recognition and accuracy rate is decreased [8, 5].

2.3 Skew Detection: In document analysis skew can be detected using its text lines. These lines are perfectly vertical or horizontal without any skew. Skew angle can be detected as slope/angle of rotation on its line w.r.t. it's X-axis. Alignment of the document and word is must to for next processing stages and good results [fig.1].

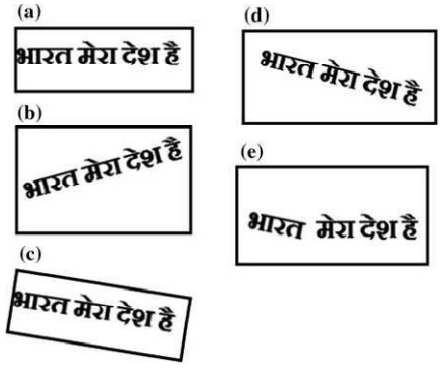


Fig.1 Document and Word skew [7]

Many researchers considered the mobile taken images, scanned printed or handwritten documents. These documents also contains the different nontextual objects and so efficient techniques need to be evaluated for good skew correction [2,6]. Harris Corner detection with Hough Transform has been used by the author to detect the skew angle of the images taken by blind people using mobile phone. Author used 3630 documents images of the ICDAR 2015 database and concentrated on printed documents with nontextual elements[6]. Mostly researchers used Hough Transform for the detection and correction of skew for different languages[3]. In [3,4], researcher used Hough Transform word level skew detection for Devnagari Script. This

deviation is calculated using the inclination at horizontal lines of words and then image is rotated accordingly either in clockwise direction or anticlockwise direction.

In [9], author proposed a system for skew correction of complex document images with a skew angle of -45° to +45°. These images contains nontexual elements also like graphs, tables, figures etc. In this system monochromatic and color images are scanned from top to bottom searching for the leftmost pixel of the image. This leftmost points' horizontal distance with deviation is used to find the skew angle which is then applied for skew correction.

A Roy et al. proposed the candidate path to perform both Skew detection as well as the segmentation. This candidate path is nothing but a graphical path which is generated using the segmenting points obtained from the handwritten patterns of bangla letters. Around 150 dataset is referred for skew detection and correction of handwritten bangla script [7, 10].

In Nearest neighbour method, histogram is calculated to find the skew of the document. The connected components (CC) from the document are searched and for all these nearest neighbour connected components (CC), direction vectors histogram is calculated. To calculate the skew angle dominant histogram peak value and centroid of the nearest neighbour CC is calculated [5].

2.4 Skew Correction:

The skewed page will be corrected once the angle of the document image is found. For this text recognition and skew angle estimation is required [4]. Headline of the word and documents are detected at the time skew angle estimation [3]. To get the good results, rotation algorithm must be choose with its performance parameters.

3. Proposed Methodology : Hough Transform is mainly used to identify the curves, lines, shapes which are termed by some parametric equations. To detect the image points on required curve using edge detection is inefficient technique [4, 11]. Hough Transform is used to overcome this problem. Hough Transform is simply used to identify the straight lines. In Devnagari Marathi text, the words contain the Shirorekha on the words. So to detect the tilt or skew of such document images or words can be efficient using Hough Transform.

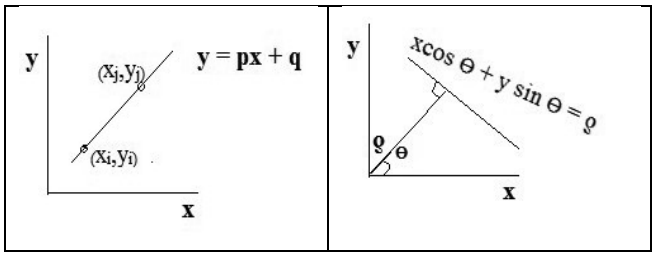


Fig 2.a Image Line Space Fig. 2.b Angular Parameter space

Straight line is defined as y= px + q, where (p, q) are angular parameters represented in parametric space. Linear Hough Transform finds the presence of line by using

$$x\cos\Theta + y\sin\Theta = \varrho$$

where ,

ϱ = the distance of the origin point to its closet point upon the straight line.

Θ = angle between the line connecting the origin with closet point and x axis.

The (ϱ, Θ) is nothing but a Hough Transform space for straight lines. Peak point from image space is identified to get the end points of line. Using the head line i.e. Shirorekha of words and egde detection method, Θ can be calculated. Once this skew angle is detected, page and word is rotated. To correct this skew rotation transformation will be used. using angle Θ, word will be rotated in horizontal direction.

CONCLUSION :

In optical character recognition it is very essential to perform the skew detection of the page and words to perform the accurate recognition. In this paper we studied the methods utilized by different researchers like Hough transform, projection profile, Fourier transform etc. In proposed system Devnagari Marathi Text is considered, so Hough transform and rotation transform method has been proposed. Further study is going on to get the accuracy in skew detection and correction.

REFERENCES :

[1] R. Jayadevan , S. R. Kolhe, P. M. Patil , Umapada Pal . Offline Recognition of Devanagari Script: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* Volume: 41, Issue: 6, Nov. 2011.

[2] A.K. Das, B. Chanda. A fast algorithm for skew detection of document images using morphology *IJDAR* 4: 109–114 2001.

[3] B.B.Chaudhuri and U.Pal. Skew Detection of Digitized Indian Script Documents. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 19, pp. 182-186 ,1997.

[4] *T. A.Jundale,R. S.Hegadi.* Skew Detection and Correction of Devanagari Script Using Hough Transform. *Procedia Computer Science*, Volume 45, Pages 305-311, 2015.

[5] Arwa AL-Khatatneh, Sakinah Ali Pitchay, Musab Al-qudah. A Review of Skew Detection Techniques for Document. *2015 17th UKSIM-AMSS International Conference on Modelling and Simulation*. 2015 .

[6] Ahmed GARI, Ghizlane KHAISSIDI, Mostafa MRABTI, Driss CHENOUNI, Mounim El YACOUBI. Skew detection and correction based on Hough Transform and Harris Corners. *IEEE* 2017.

[7] Deokate S., Uke N. Various Traditional and Nature Inspired Approaches Used in Image Preprocessing. In: Pawar P., Ronge B., Balasubramaniam R., Seshabhattar S. (eds) *Techno-Societal 2016.* ICATSA 2016. Springer, Cham.2016

[8] Aparna Kokku and Srinivasa Chakravarthy. V. Govindaraju, S. Setlur. A Complete OCR System for Tamil Magazine Documents. *Guide to OCR for Indic Scripts, Advances in Pattern Recognition*, C _ Springer-Verlag London Limited 2009.

[9] Rafael Dueire Lins and Bruno Tenório Ávila A. Campilho, M. Kamel. A New Algorithm for Skew Detection in Images of Documents. *ICIAR 2004*, LNCS 3212, pp. 234–240, 2004. © Springer-Verlag Berlin Heidelberg 2004

[10] Roy A, Bhowmik TK, Parui SK, Roy U (2005). A novel approach to skew detection and character segmentation for handwritten Bangla words. *In: Digital imaging computing: techniques and applications. DICTA IEEE* 2005.

[11] http://en.wikipedia.org/wiki/Hough_transform.

AUDIO STEGANOGRAPHY USING LSB FOR HIGH CAPACITY TEXT FILE

MR. HARSHAL V. PATIL
DOCS Bhusawal Arts, Science & P.O.Nahata Commerce College, Bhusawal, India
harshal4patil@gmail.com
DR. B. H. BARHATE
DOCS Bhusawal Arts, Science & P.O.Nahata Commerce College, Bhusawal, India
barhate_1@yahoo.com

Abstract : Audio steganography is one of the security technique in which any type of data can be hidden in audio. The technique is so powerful since anybody else cannot easily detect the error in audio. This paper covers all the details about audio steganography in which LSB is used. Here all rules are covered about where to hide data in audio.

Key words: Audio Stegnography, Stego-audio, LSB, PSNR, MSE.

INTRODUCTION:

Today's large demand of internet applications requires data to be transmitted in a secure manner. Data transmission in public communication system is not secure because of interception and improper manipulation by eavesdropper. Steganography is defined as the study of invisible communication. Steganography usually deals with the ways of hiding the existence of the communicated data in such a way that it remains confidential.

"Steganography" is a Greek origin word which means "hidden writing". Steganography word is classified into two parts: Steganos which means "secret or covered" (where you want to hide the secret messages) and the graphic which means "writing" (text). However, in the hiding information the meaning of Steganography is hiding text or secret messages into another media file such as image, text, audio, and video.

Audio steganography is the technique in which hiding information inside audio signals or hiding secret audio file in cover file. There are number of types of audio files available like WAV file (.wav) and MPEG layer 3 file (mp3) etc. In audio Steganography, secret message is embedded into inside audio signal which results from slender shifting of the binary sequence of the equivalent audio file. Audio Steganography methods can embed messages like texts, Images, small audio clips in WAV, AU, and even MP3 sound files. There are some methods like LSB coding, spread spectrum, Phase coding, Echo hiding which are being used for audio Steganography [3, 4, 5].

1. Basic terms for Steganography:

- **Cover-object** - The original object where the message has to be embedded.
- **Stego-message** - Secret message, cover text and cover image that has to be embedded in the cover object.
- **Stego object** - The cover object, once the message has been hidden or embedded.
- **Stego Key**- The secret code to be shared between Sender and receiver to embed and retrieve the message.
- **Embedding algorithm**: It is the way or the idea that is often used to embed the secret information in the cover message. [3]

2. Characteristics of Steganography system:

To measure the security strengths and weaknesses of Steganography system, we have to go through following features which are concretely defined by its application

- **Capacity**- The notion of capacity in data hiding indicates the total number of bits embedded and successfully recovered by the Stego system.
- **Robustness** - After hiding data on the cover, how much this data intact in the cover by applying various attacks like noise, transformations, filters and cropping.
- **Security**: It is said that the embedded algorithm is secure if the embedded information does not hack or remove by an attacker.
- **Imperceptibility** - The MSE (Mean Squared Error) and PSNR (Peak Signal to Noise Ratio) both factors are useful for checking the perfectness of cover and stego files. These ratios help to find how both files are more identical or different with each other.

3. Types of Stegnography:

1. **Image Steganography: -** For hiding the secret message into carrier image, which is then converted into stego image.
2. **Text Steganography: -** In this, the message that is to be sent is rooted firstly in a text file by formatting. The format it based on line-shift coding, word-shift coding, feature coding etc. Reformatting of the text destroys the rooted content hence the technique is not robust.
3. **Audio Steganography:-** The secret message is embedded into unused audio bits as every file contains some unused bits or unused area of bits where secret message can be hided.
4. **Video Steganography: -** Video steganography divides the video into audio and image frames where embedding is performed in the audio file.

4. Basic Model of Audio Steganography:

The model for steganography is shown in Fig 1. Message is the data that the sender wishes to remain it confidential. Message can be plain text, image, audio or any type of file. Password is known as a stego-key, using the stego key the receiver can extract the message from cover file if receiver knows stego key. The cover-file with the secret information is known as a stego-file. [1] [2]

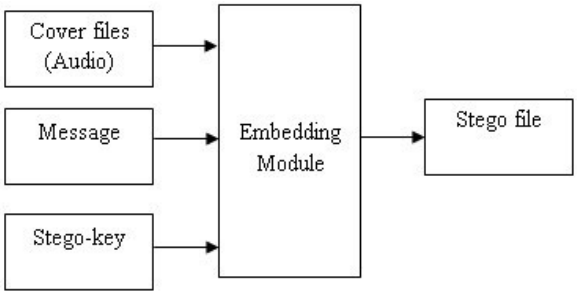


Fig. 1 Basic Audio Stegnography Model

The components of steganographic system are:

Emb Message: The message to be embedded.

Cover: The data in which emb will be embedded.

Stego file: A modified version of cover that contains the embedded message emb.

Stego-Key: Additional secret data that is needed for the embedding and extracting processes and must be known to both, the sender and the recipient.

1. Audio Steganography:

Audio Steganography is the most relevant for Information Security Media. Important authentication is embedded in digital sound. In Audio Steganography, the vulnerability of the Human Auditory System (HAS) is used to hide information in the audio. In earlier years, several Information hiding algorithms are proposed as well as implemented. All these algorithms exploit the characteristics of the human auditory system (HAS) for hiding information in a transparent manner [6].

Embedding Algorithm:

1. Select the audio file for embedding the secret message.
2. Select key file, as a text file
3.Select the text file or type the text in the text box containing the secret message.
4. Compare text file and audio file size.

If text file size > audio file contents Error message displayed indicating cannot embed secret message. Else Embed secret message in the audio file in the 3rd and 4th LSB bit of every sample.

5.After embedding secret message, it creates Stego-audio file

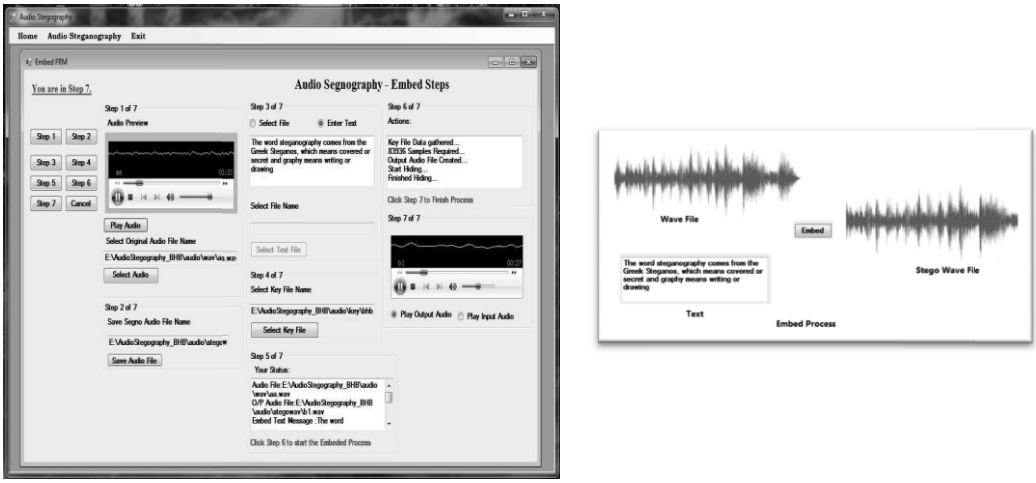


Fig. 2 Embed Steps of Audio Stegnography

Extracting Algorithm:

1. Select the Stego- audio file for extracting the secret message.
2. Select key text file, if matches then proceed else error message will occurs
3.Extract the secret message from the Stego-audio file from the 3rd and 4th LSB bit of every sample.
4.Extracted text will display in text box or separate text file

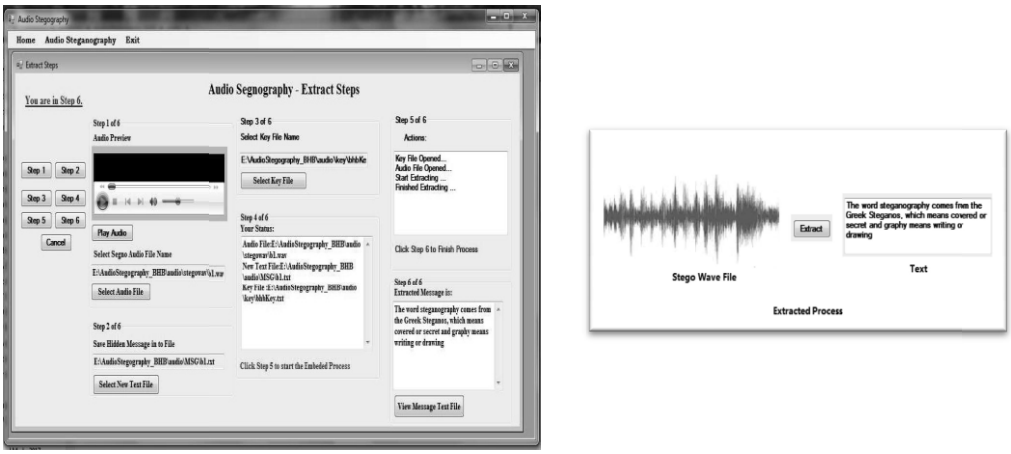


Fig. 3 Extract Steps of Audio Stegnography

Peak Signal-to-Noise Ratio (PSNR):

The term peak signal-to-noise ratio (PSNR) is an expression for the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. Because many signals have a very wide dynamic range, (ratio between the largest and smallest possible values of a changeable quantity) the PSNR is usually expressed in terms of the logarithmic decibel scale.

Mean Square Error (MSE) could be estimated in one of numerous approaches to quantify the contrast between values implied by an evaluation and correct quality being certified.

1. Results for Audio Steganography for Text files:

Cover	Secret	Audio In size in KB	Key Text File in KB	MSE	PSNR in DB
Audio	Text	150	10	$2.191E^{-08}$	129.7565
		200	20	$3.112E^{-08}$	122.4819
		153	30	$2.177E^{-07}$	117.4827
		250	40	$3.341E^{-08}$	126.2730
		300	50	$3.249E^{-08}$	117.2179
		310	60	$3.356E^{-08}$	120.8761
		400	70	$2.771E^{-08}$	120.5520
		450	80	$3.861E^{-08}$	119.8008
		479	90	$8.834E^{-08}$	118.7028

Table – 1: Quality measures using MSE & PSNR

CONCLUSION :

The message signal is transmitted with utmost security and can be retrieved without any loss in transmission in this method. This proposed system will not change the size of the file even after encoding and also suitable for any type of audio file format. This method also helps to hide big messages. It has considerably low robustness against the attacks. Therefore maintain the robustness during the substitutions of bits.

CONCLUSION :

[1] Jayaram , Ranganatha , Anupama , "Information Hiding using audio stegnography - A Survey" The International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.3, August 2011 ,Pg 86-96.

[2] Burate D. J., M.R.Dixit, "Performance Improving LSB Audio Steganography Technique",
International Journal of Advance Research in Computer Science and Management Studies, Volume 1, Issue 4, September 2013, Pg 67-75.

[3] Yusuf Perwej, FirojParwej, Asif Perwej, "An Adaptive Watermarking Technique for thecopyright of digital images and Digital ImageProtection", The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.2, April 2012, PP 21-38.

[4] Mohammad Abdullatif, Akram M. Zeki,JalelChebil,Teddy Surya Gunawan, "Properties of Digital Image Watermarking", 2013 IEEE 9th International Colloquium on Signal Processing and its Applications, 8 - 10 Mac. 2013, Kuala Lumpur, Malaysia, 978-1-4673-5609-1/13/$31.00 ©2013 IEEE,PP 235-240.

[5] Lalit Kumar Saini ,Vishal Shrivastava, "A Survey of Digital Watermarking Techniques and its Applications", International Journal of Computer Science Trends and Technology (IJCST) – Volume 2 Issue 3, May-Jun 2014, PP 70-73

[6] B. H. Barhate, Prof.Dr. R. J. Ramteke," Audio Steganography using Combination of LSB and Key for Images, Audio and Text Messages", International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 5, Issue 7, July 2015),PP 231-236.

OVERVIEW OF NATURAL LANGUAGES PROCESSING

PROF. PALLAVI A. PATIL
DOCS, Dhanaji Nana Mahavidyalaya, Faizpur, MS
patil.pallavi2710@gmail.com

Abstract : Natural Language Processing (NLP) is an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things. NLP researchers aim to gather knowledge on how human beings understand and use language so that appropriate tools and techniques can be developed to make computer systems understand and manipulate natural languages to perform the desired tasks.

The aim of this paper is to cover basic idea of NLP like types of languages processing, approach and techniques, applications etc.

Keywords: NLP, Phonology, Morphology, Lexicon

INTRODUCTION:

Natural language processing (NLP) can be defined as the automatic (or semi-automatic) processing of human language. Natural Language processing (NLP) is a field of computer science and linguistics concerned with the interactions between computers and human (natural) languages. In theory, natural-language processing is a very attractive method of human-computer interaction. Natural language processing is the task of analyzing and generating by computers, languages that humans speak, read and write. NLP is concerned with questions involving three dimensions: language, algorithm and problem.

Natural Language processing (NLP) is a field of computer science and linguistics concerned with the interactions between computers and human (natural) languages. Natural language generation systems convert information from computer databases into native language. Language is meant for Communicating about the world. By studying language, we can come to understand more about the world. If we can succeed at building computational mode of language, we will have a powerful tool for communicating about the world. We look at how we can exploit knowledge about the world, in combination with linguistic facts, to build computational natural language systems.

The goal of natural language analysis is to produce knowledge representation structures like predicate calculus expressions, semantic graphs or frames. This processing makes use of foundational tasks like morphology analysis, Part of Speech Tagging, Named Entity Recognition, both shallow and deep Parsing, Semantics Extraction, Pragmatics and Discourse Processing.

Aims &Objective of NLP:

The major aims and objectives of Natural Language Processing are as follows:

1. To study the nature of language. (Linguistics),
2. Window into cognition (Psychology),
3. Human interface technology,
4. Text translation,
5. Information management.

History of NLP:

Before the 1970s, most NLP researchers were concentrating on MT as an application (see above). NLP was a very early application of CS and started about the same time as Chomsky was publishing his first major works in formal linguistics (Chomskyan linguistics quickly became dominant, especially in the US). In the 1950s and early 1960s, ideas about formal grammar were being worked out in linguistics and algorithms for parsing natural language were being developed at the same time as algorithms for parsing programming languages. However, most linguists were uninterested in NLP and the approach that Chomsky developed turned out to be only somewhat indirectly useful for NLP.

NLP in the 1970s and first half of the 1980s were predominantly based on a paradigm where extensive linguistic and real-world knowledge was hand-coded. Statistical NLP became the most common paradigm in the 1990s, at least in the research community. Speech recognition had demonstrated that simple statistical techniques worked, given enough training data.

Classification of NLP: Natural Language Processing basically can be classified into two parts i.e. Natural Language Understanding and Natural Language Generation which is shown in following fig.1

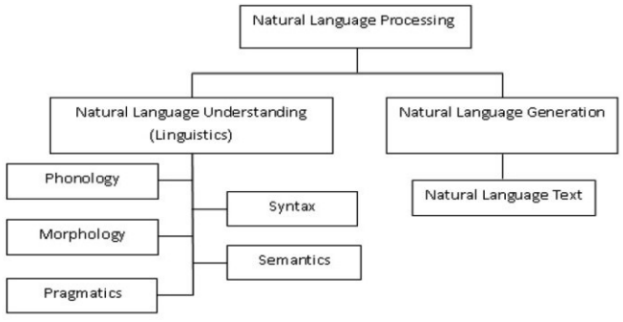


Fig. 1 Classification of NLP

Components of NLP:

There are following two components of NLP

- **Natural Language Understanding**

– Mapping the given input in the natural language into a useful representation.

– Different level of analysis required:

▪ morphological analysis,

- syntactic analysis,
- semantic analysis,
- discourse analysis

• **Natural Language Generation**

– Producing output in the natural language from some internal representation.

– Different level of synthesis required:

- deep planning (what to say),
- syntactic generation

Stages of NLP and associated ambiguities:

Traditionally, NLP - of both spoken and written language has been regarded as consisting of the following stages:

• Phonology and Phonetics (processing of sound)

At this stage utterances are processed. Apart from many challenges due to noise, two common problems are homophony and word boundary recognition.

Homophony arises when two words sound the same, though their meanings are widely different.

Example - I got up late vs. I got a plate, both of which sound very much the same.

• Morphology (processing of word forms)

Words form from root words or lexemes through processes of inflexion, derivation, back formation. The main ambiguity at the level of morphology arises from choices available in breaking the word into stem and suffix as well as from choices of features.

Example – Jaaii – Jaa and ii also Ja and aii.

• Lexicon (Storage of words and associated knowledge)

Words are stored in the lexicon with a variety of information that facilitates the further stages of NLP, like question answering, information extraction etc. For example, the word dog might be stored in the lexicon with information like:

POS (Noun)

Semantic Tag (Animate, 4-legged)

Morphology (takes 's' in plural)

Words typically have multiple meanings even in the same part of speech. Dog, for example, means an animal and a very detestable person.

•Parsing / Syntactic Analysis (Processing of structure)

Parsing or syntactic processing refers to uncovering the hierarchical structure behind a linear sequence of words.

Example - No smoking areas will allow hookahs inside.

Here no can qualify the rest of the sentence, meaning thereby there isn't a smoking area that will allow hook as inside.

The two meanings are sort of opposite of each other.

• **Semantics (Processing of meaning)**

It is concern with what words mean and how these meaning combine in sentence to form sentence meaning.

• **Pragmatics (Processing of user intention, modeling etc.)**

It is concern with how sentence are used in different situation and how use affects the interpretation of the sentence.

This is one of the hardest problems of NLP and has seen Very little progress. The problem involves processing user intention, sentiment, belief world, modals etc.- all of which are highly complex tasks.

• **Discourse (Processing of connected text)**

It is concern with how immediately preceding sentence affects the interpretation of the next sentence.

NLP Tools and Techniques

A number of researchers have attempted to come up with improved technology for performing various activities that form important parts of NLP works. These works may be categorized as follows:

• Lexical and morphological analysis, noun phrase generation, word segmentation, etc. (Bangalore & Joshi, 1999; Barker & Cornacchia,2000; Chen & Chang, 1998; Dogru & Slagle, 1999; Kam-Fai et al.. 1998; Kazakov et al.. , 1999; Lovis et al.. 1998; Tolle & Chen, 2000; Zweigenbaum & Grabar, 1999).

• Semantic and discourse analysis, word meaning and knowledge representation (Kehler, 1997; Mihalcea & Moldovan,1999; Meyer & Dale, 1999; Pedersen & Bruce, 1998; Poesio & Vieira,1998; Tsuda & Nakamura, 1999)

• Knowledge-based approaches and tools for NLP (Argamon et al.., 1998; Fernandez & Garcia-Serrano, 2000; Martinez et al.., 2000, 1998).

Noun phrasing is considered to be an important NLP technique used in information retrieval. One of the major goals of noun phrasing research is to investigate the possibility of combining traditional keyword and syntactic approaches with semantic approaches to text processing in order to improve the quality of information retrieval.

Many researchers worked on NLP, building tools and systems which makes NLP what it is today. Tools like Sentiment Analyser, Parts of Speech (POS) Taggers, Chunking, Named Entity Recognitions (NER), Emotion detection, Semantic Role Labelling made NLP a good topic for research.

Sentiment analyser (Jeonghee etal.,2003) [3] works by extracting sentiments about given topic. Sentiment analysis consists of a topic specific feature term extraction, sentiment extraction, and association by relationship analysis. Sentiment Analysis utilizes two linguistic resources for the analysis: the sentiment lexicon and the sentiment pattern database. It analyses the documents for positive and negative words and try to give ratings on scale -5 to +5.

Parts of speech taggers for the languages like European languages, research is being done on making parts of speech taggers for other languages like Arabic, Sanskrit (Namrata Tapswi , Suresh Jain ., 2012) [4], Hindi (Pradipta Ranjan Ray et al., 2003) [5] etc. It can efficiently tag and classify words as nouns, adjectives, verbs etc. The most procedures for part of speech can work efficiently on European languages, but it won't on Asian languages or middle eastern languages. Sanskrit part of speech tagger is specifically uses treebank technique. Arabic uses Support Vector Machine (SVM) (Mona Diab etal.,2004) [6] approach to automatically tokenize, parts of speech tag and annotate base phrases in Arabic text.

Application of NLP

The applications can be divided into two major classes: Text-based applications and Dialogue-based applications.

1. Text-based applications:

Text-based applications involve the processing of written text, such as books, newspapers, reports, manuals, e-mail messages, and so on. These are all reading-based tasks. Text-based natural language research is ongoing in applications such as

• finding appropriate documents on certain topics from a database of texts (for example, finding relevant books in a

library)

• extracting information from messages or articles on certain topics (for example, building a database of all stock transactions described in the news on a given day)

• translating documents from one language to another (for example, producing automobile repair manuals in many different languages)

• summarizing texts for certain purposes (for example, producing a 3-page summary of a 1000-page government report)

• One very attractive domain for text-based research is story understanding. In this task the system processes a story and then must answer questions about it. This is similar to the type of reading comprehension tests used in schools and provides a very rich method for evaluating the depth of understanding the system is able to achieve.

2. Dialogue-based applications:

It involves human-machine communication. Most naturally this involves spoken language, but it also includes interaction using keyboards.

Typical potential applications include

• question-answering systems, where natural language is used to query a database (for example, a query system to a personnel database)

• automated customer service over the telephone (for example, to perform banking transactions or order items from a catalogue)

• tutoring systems, where the machine interacts with a student (for example, an automated mathematics tutoring system)

• spoken language control of a machine (for example, voice control of a VCR or computer)

• general cooperative problem-solving systems (for example, a system that helps a person plan and schedule freight shipments)

Some NLP Task

There are following NLP Task:

• Word segmentation

• Topic segmentation and recognition

• Part-of-speech tagging

• Word sense disambiguation

• Named entity recognition (NER)

• Parsing

CONCLUSION :

NLP is the mode by which interaction between computer & human being is become possible. We conclude that it converts information from computer databases into human readable form. Describe the successful use of NLP in a limited domain. NLP researchers aim to gather knowledge on how human beings understand and use language so that appropriate tools and techniques can be developed to make computer systems understand and manipulate natural languages to perform the desired tasks.

REFERENCES :

1. Ahonen, H.; Heinonen, O.; Klemettinen, M. & Verkamo, A.I. (1998). Applying data mining techniques for descriptive phrase extraction in digital document collections. IEEE International Forum on Research and Technology. Advances in Digital Libraries - ADL'98, 22-24 April 1998, Santa Barbara, CA. Los Alamitos, CA: IEEE Computer Society, pp. 2-11
2. Amsler, R.A.(1984). Machine-readable dictionaries. In: M. E. Williams, (ed.) Annual Review of Information Science and Technology (ARIST: Volume 19, White Plains, NY: Knowledge Industry Publications Inc. for the American Society for Information Science. pp.161-209.
3. Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003, November). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In Data Mining, 2003. ICDM 2003. Third IEEE International Conference on (pp. 427-434). IEEE.
4. Tapaswi, N., & Jain, S. (2012, September). Treebank based deep grammar acquisition and Part-Of-Speech Tagging for Sanskrit sentences. In Software Engineering (CONSEG), 2012 CSI Sixth International Conference on (pp. 1-4). IEEE.
5. Ranjan, P., & Basu, H. V. S. S. A. (2003). Part of speech tagging and local word grouping techniques for natural language parsing in Hindi. In Proceedings of the 1st International Conference on Natural Language Processing (ICON 2003).
6. Diab, M., Hacioglu, K., & Jurafsky, D. (2004, May). Automatic tagging of Arabic text: From raw text to base phrase chunks. In Proceedings of HLT-NAACL 2004: Short papers (pp. 149-152). Association for Computational Linguistics.
7. Ogallo, W., & Kanter, A. S. (2017, February 10). Using Natural Language Processing and Network Analysis to Develop a Conceptual Framework for Medication Therapy Management Research. Retrieved April 10, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/28269895?dopt=Abstract
8. Ochoa, A. (2016, May 25). Meet the Pilot: Smart Earpiece Language Translator. Retrieved April 10, 2017, from https://www.indiegogo.com/projects/meet-the-pilot-smart-earpiece-language-translator-headphones-travel.
9. "Natural Language Processing." Natural Language Processing RSS. N.p., n.d. Web. 25 Mar. 2017

IMPLEMENTATION OF CLOUD BASED IN EDUCATIONAL DIGITAL LIBRARY

SACHIN.E.TAYDE
Department of Computer Science
S.S.S.K.Innani.Mahavidyalaya . Karanja(lad) Dist. Washim MS, India
setayde_24@rediffmail.com

Abstract : During the last few years, there is innovative development in the field of multimedia communication and wireless technology. Cloud computing is fundamentally altering the expectations for how and when computing, storage and networking resources should be allocated and managed. Cloud computing technology came up as a boon for libraries and is offering various opportunities for libraries to connect their services with cloud. The emergence and development of cloud computing have a great effect on the development and application of digital library. Libraries may soon be building and managing their own data centers. This model would let libraries maintain more control over the applications and data stores that contain sensitive, private information about patrons. In the recent years growth of computer technology contribute to the progress towards the application of the cloud computing. The paper describes architecture, use of cloud computing in libraries and how cloud computing works which helps understanding the need of implementing virtualization using cloud computing in digital library.

Keywords- Cloud Services, Cloud Platforms, Educational Cloud computing, digital library, IaaS,PaaS,SaaS.

INTRODUCTION:

Rapid development of information technology (IT) industry for the last several decades has introduced us with many new terms. Nowadays, we are doing the same tasks but in a flexible, much cheaper, and are in a portable manner, either by using desktop computer or mobile devices to several types of servers tied together to create a so called Cloud Computing System (CCS). Cloud Computing has emerged as a phenomenon that represents the way by which IT services and functionality are charged for and delivered. Cloud computing can be defined as the aggregation of computing as a utility and software as a service [1] where the applications are delivered as services over the Internet and the hardware and systems software in data centers provide those services [2].The new concept of cloud and libraries has generated a new model called cloud libraries. Though the usages of cloud computing may vary with the libraries nature, services and information needs but most common usages of cloud computing with in libraries can be development of digital libraries, corporate cataloging, acquisition, storages and sharing the resources on virtual environment. This Cloud computing technology provides almost everything as service using Internet, and every resource is highly scalable. Resources in cloud computing are provided as service based on data centers. Educational institution is moving towards adopting new developing technology for providing the student new and faster means of resources through which they can adopt the higher level knowledge. Cloud computing might be an area for the educational institution to provide faster and much chipper resources for student with globalization. This paper proposed an idea to develop various clouds for educational sectors which help different students and faculty to research on the various subjects globally.

1. OVERVIEW OF EDUCATION, CLOUD COMPUTING AND DIGITAL LIBRARY

Educational institution consist a large I.T. infrastructure and for managing it requires many services. Student required software for simulation, experimental performance and manipulation of statically data. These software's are high end software and required skilled people to work with, thus rather this software can be installed in the virtualized environment on the clouds data centre. Internet can be medium to provide access to this software with high reliability. Virtualization technology helps creating multiple client nodes for student, these nodes can be easy access through Internet [3].

The basic principle of Cloud Computing is making tasks distributed in large numbers of distributed computers but not in local computers or remote servers. In other words, by collecting large quantities of information and resources stored in personal computers, mobile phones and other equipment. There is a serial of problem in digital library, such as resource independent of each other. Low level of information technology, non-uniform resource form and hardware limitation. In order to solve these problems, it proposes a new digital library platform based on cloud computing, which can provide personal service to different terminal users, such as computer, PC etc. The educational sectors are looking for options which are chipper and more convenient in terms to improve the performance and ranking of the students [4].

2. CLOUD COMPUTING

Cloud computing is emerging as one of the most important branch for providing seamless application on mobile devices.Cloud computing is not a new technology that suddenly appeared on the web but it is a new form of computing. It is a web-based processing, whereby shared resources, software and information are provided on demand to computers and other similar devices. Cloud computing can be defined as ,"It refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services" [5]. Cloud computing data centers can be sliced into various servers and virtual machines can be a solution to their problems, and these are easy to access to the computing services on demand. Cloud computing can also be used for research work for various platform [6].Cloud computing has become a significant technology emerging trend, and many experts, researchers and academicians expect that cloud computing will reshape information technology (IT) sector and the IT marketplace in world. With the cloud computing

technology, users use a wide variety of devices, including PCs, Laptops, Smart Phones, and PDAs to access different kinds of utility programs, storage, and application development platforms over the Internet, via services offered by cloud computing providers. This section provides an overview of Cloud computing including definition and service oriented cloud architecture. In overall cloud computing revolves around two things one is Cloud platforms and other is Cloud services.

2.1 Cloud computing Platforms

Cloud platforms are basically the hosts that provide the required resources to the clients. It is an arrangement for executing software applications in a logically abstract environment comprising of various utility cloud services [7]. Cloud computing is being driven by cloud providers including Amazon, Google, Salesforce and Yahoo as well as traditional vendors including IBM Microsoft and are adopted by different users. Few well-known cloud platforms are,

- Amazon Elastic Cloud Computing (EC2) [8]
- Microsoft Azure[9]
- Hyrax[10]
- Google App Engine[11]
- Force.com[12]

2.2 Cloud Services

Cloud services are hosted services. Cloud service [13] is a software system which is responsible providing interoperable machine-to -machine interaction over a network or internet which is further accessed by other cloud computing components ,clients, software or end users directly like,

- Integration (Amazon simple Queue Service)
- Mapping (Google maps, Yahoo! Maps)
- Payments (, Google Checkout)
- Search (Google Custom search)

3. SERVICE ORIENTED CLOUD ARCHITECTURE

Here, we focus on a layered architecture of cloud computing. In cloud computing there are different categories of cloud services. The cloud services are generally classified based on layer concept (fig.1) .This architecture is commonly used to demonstrate the effectiveness of the cloud computing model in terms of meeting the users' requirements [14]. In the upper layer of this paradigm, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) are stacked. These services delivered to the users in real time via internet. The service model has been explained in below fig.1.

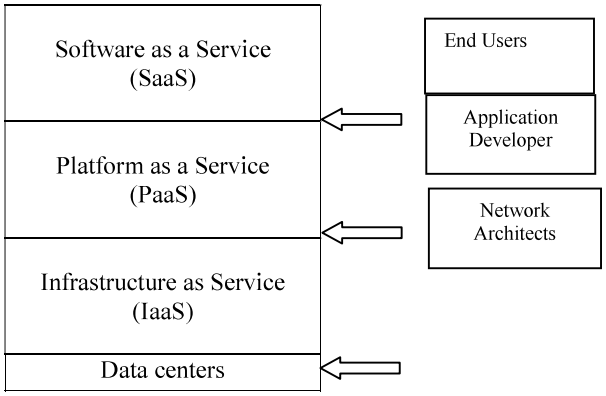


Fig.1. Service cloud computing architecture

- Software as a Service (SaaS)

It is a model of software deployment whereby the provider licenses an application to the consumers for use as a service on demand. The capability provided to the consumer is to use the provider applications running on a cloud infrastructure. In this service, users can avail the facilities to access and use any software available with cloud vendors.

- Platform as a Service (PaaS)

Platform as service helps in generating the computing platforms to run the software and other tools over the internet without managing the software and hardware at the end of user side.The capability provided to the consumer is to deploy onto the cloud infrastructure consumer created or acquired applications created using programming languages and tools supported by the provider. It is the delivery of computing platform and solutions stack as a service.

- Infrastructure as a Service (IaaS)

This service comprises a wide range of features, services and resources which support to build a virtual infrastructure for computing. Organizations can be developed entire infrastructure on demand. The capability provided to the consumer is to provision processing, storage, networks and other fundamentals computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications IaaS enables the provision of storage, hardware, servers and networking components. Amazon Web Services, HP,IBM, Google Base are the examples of IaaS.

- Data Centers layer

This layer provides the hardware facility and infrastructure for clouds. In data layer, a number of servers are linked with high speed networks to provide services for customers.

The Cloud computing architecture can be divided these layers as in fig.1.Given this architectural model, the users can use the services flexibility and efficiently.

5 Cloud computing realization

The Cloud in cloud computing are categorized in three types they can be used in various services as educational public cloud, educational private cloud and educational hybrid cloud.

- Public clouds: It provides services like application, storage and makes resources available to public through the Internet.
- Private clouds: Private cloud is special infrastructure dedicated to a single educational organization for services, resources and data storage.
- Hybrid clouds: It is the combination of one or more public and private educational cloud.

The educational sector can be implemented using public cloud, as they are chipper compare to private and hybrid, for some of the institution and university, which carry out the research work in depth, they can use the private and hybrid cloud after undergoing through the various terms. The components used in the architecture as in fig.2.

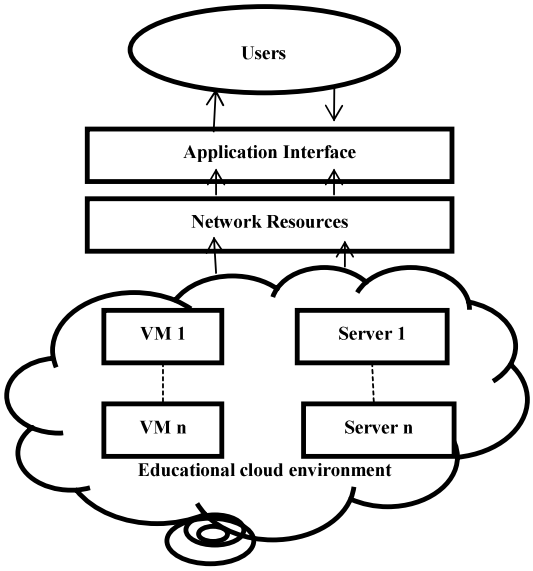


Fig. 2 Architecture for Cloud Computing Educational Environment

6. IMPLIMENTATION

The implementation work consists of a thin client which acts as node for the client to access and work with the services provided by the cloud computing model. A thin client is a low configuration computer system which consist hardware only to access information online. Secondly an application interface is required for working which act as GUI interface through which user can interact with cloud. The user interface will give a dashboard panel and a configuration window which help user to communicate and configure the services. Operating system and network management software are also required, this software are powerful enough to handle the connectivity and to provide a standard bandwidth through which the thin client can communicate with the cloud. The services are provided through Internet, this Internet connectivity and network device layer. Cloud environment is made up of shared resources, these shared resources are none other than the same computing resources which are used for computing, but with a slight change that the shared resources can be located at remote location and accessed using an Internet connection. Better resources management will lead to maximize the usability of cloud resources.

CONCLUSION :

Cloud computing represents an exciting opportunity to bring on-demand applications to digital library. This technology provides a better used in educational system much more reliable platform for handling computing resources, it appears to the users of high-quality service and high security.. This paper has discussed about introducing cloud computing and virtualization in educational sector in digital libraries.. This provides a platform independent of infrastructure with much more flexible model. Thus considering the educational cloud environment proposed model in terms educational sector in digital library will help in making a better model for student with powerful functional capabilities which can be further implemented in the real world cloud computing environment.

REFERENCES :

[1] Khan, M. Othman, S. Madani, and S. Khan, "A survey of mobile cloud computing application models," *IEEE Communications Surveys &Tutorials*, issue 99, 2013.

[2] A. N. Khan, M. M. Kiah, S. U. Khan, and S. A. Madani, "Towards secure mobile cloud computing: A survey," *Future GenerationComputer Systems*,vol.29,issue 5,pp. 1278–1299, July 2013.

[3] Virtualization and Cloud Computing group.virtualization and cloud computing .China Electronic Industry Press.October,2009.

[4] Pandya, M. (2012).Cloud computing for libraries: A SWOT analysis. *8th Convention PLANNER-2012 Sikkim University, Gangtok,* 387-394.

[5] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, Above the clouds: a Berkeley view of cloud computing, Technical Report UCB/EECS-2009-28, 2009.

[6] Fan Bingsi, Cloud Computing and Libraries: Defense for Research on the Cloud Computing, Library and Information Service, 2009.

[7] Adrian Otto's Blog.' What is a Cloud Platform?' http://adrianotto.com/2011/02/cloud-platform/

[8] "Amazon elastic compute cloud(EC2),'AWS.[online].Available :http://www.amazon.com/ec2/

[9] Microsoft azure.[online].Available : http://www.microsoft.com/azure/

[10] E. Marinelli, "Hyrax: cloud computing on mobile devices using MapReduce", Master thesis, Carnegie Mellon University, 2009.

[11] Google app engine. [Online]. Available: http://appengine.google.com/

[12] http://www.force.com/why-force.jsp

[13] Cloud Computing. October 11th 2008, http://www.reference.com/browse/cloud+computing?s=t

[14] W. Tsai, X. Sun, and J. Balasooriya, "Service-Oriented Cloud Computing Architecture," in Proceedings of the 7th International Conference on Information Technology: New Generations (ITNG), pp. 684-689, July 2010.

A STUDY OF VIRAL DISEASES AND THEIR IMPACT ON PUBLIC HEALTH USING KNOWLEDGE BASED SYSTEM

AUTHOR : ASHWINI RAJENDRA KULKARNI, DR. SHIVAJI D. MUNDHE
ashwinibrahme@sinhgad.edu, director_mca_simca@sinhgad.edu
1. Asst. Prof. SIMCA(MCA), Narhe, Pune, 2. Director , SIMCA, Narhe Pune

Abstract : The knowledge Based system Plays vital role in diverse domains such as healthcare, Pharmacy, Telecommunication, Agriculture , Banking and Finance , Information Security , Agriculture, etc. Healthcare is one of the important domain where public health plays vital role and there are many healthcare is challenging for daily heath care activities. The paper is focused towards the study of various viral diseases and their patterns. Also the research paper explains about the challenges, origin of research and significance of the proposed research. The paper also confide towards the knowledge management methods and techniques, applications of KM in healthcare, current and future challenges of healthcare. The research paper is also spotlighted on Knowledge discovery which is mainly based on data mining techniques. This paper also describes on the research done in text mining healthcare by various researcher, scholars and scientists.

Keywords: Data Mining, Healthcare, Knowledge Based system, Viral Diseases, Outbreaks

INTRODUCTION:

Healthcare is a general domain into which a great deal of effort in terms of knowledge management placed and knowledge discovery can be especially beneficial in the healthcare field where manual analysis and generating effective knowledge discovery from useful information is not possible because of huge availability of information on website. Over the past decades, several infectious diseases have increased in incidence and expanded into new geographic areas. There are multiple factors that contribute to the spread of disease, including increasing urban population density, more international travel, and widespread international import/export of goods. The various viral and bacterial diseases are spread through various types of viruses and bacteria's. The some of the common viral diseases are dengue, Chikungunya, malaria, swine flu, H1N1, Influenza Flu, Cholera, Rabies, etc. The viral and bacterial diseases had great impact on public health. Therefore the present research will focus on Viral and Bacterial diseases controls and knowledge discovery for healthcare organizations, medical experts which will be useful for the prediction of outbreaks of these diseases so that the remedial measures can be considered.

1. Literature Review

An author Christo El Morr, Julien Subercaze described in his paper entitled **"Knowledge Management in Health care",** about the adoption of knowledge management in healthcare is challenging for daily heath care activities. Author also explained about knowledge management methods and techniques, applications of KM in healthcare, current and future challenges of healthcare. The researcher spotlighted on Knowledge discovery which is mainly based on data mining techniques. [8]

A research paper "**Prediction and Decision Making in Health Care using Data Mining**" of Boris Milovic, Milan Milovic illustrates on KDD and data mining methods as classification, clustering, association rule, text mining, link analysis. Paper also focus on applications of Data mining in healthcare, healthcare is new complex era where new knowledge is being accumulated daily on growing rate. The main challenges of data mining in healthcare are large volume of heterogeneous data, converting information to knowledge etc. [10]

V.Jayarajand V.Mahalakshmi in the paper entitled "**Text Mining Template Based Algorithm for Text Categorization for Improving Business Intelligence"**, focus on extracting and processing information and deriving knowledge for decision making and improving the scope of business intelligence. The aim of this paper is IRFC (information retrieval based on configuration file) for extracting data from any source of data in the form of configuration file to support all kind of information. This **IRFC** technique is compared with **KNN Text Classification Algorithm** where time optimization is compared and results are interpreted for selecting the proper candidate for job from number of resumes.

K.L.Sumathy and M. Chidambaram in the paper entitled **"Text Mining: Concepts, Applications, Tools and Issues – An Overview** "talk on the information retrieval from unstructured text is very complex requires specific processing methods and algorithm to extract useful patterns. This paper describes about general framework of text mining which contains two main steps text refining and knowledge distillation, text mining process, areas of text mining such as information retrieval, information extraction, data mining and natural language processing, applications of text mining like telecommunication, bank, IT, media, insurance, political analysis, pharmaceutical, health care, bioinformatics, business intelligence, national security and many more.

A research paper entitled "**Dengue disease prediction using Weka data mining tool**" by an author KASHISH ARA SHAKIL, SHADMA ANIS AND MANSAF ALAM detailed about how the various Data Mining techniques and WEKA tool is used for prediction of Dengue disease. This paper focuses on comparison five algorithms i.e. Naïve Bayes, J48, SMO, REP Tree and Random tree and estimation of performance of these algorithms for prediction of Dengue diseases. Researcher

has considered Dengue dataset for performance evaluation and prediction.

2. Significance of the study

- The proposed study is important for different stakeholders of healthcare domain such as doctors, medical experts, patients, society, government, world health organization, hospitals, social media agencies like newspapers, various health related websites, etc.
- The present study is significant for healthcare organizations, hospitals, world health organization, doctors and medical experts to predict the outbreaks of viral infective diseases and demographics to understand and aware to focus the public health.
- The doctors and medical experts will get valuable guidelines for viral infective diseases control and better remedial actions.
- The present study is also helpful to government for effective and efficient decision making.
- The present study is more beneficial to society, social medias to make awareness regarding the outbreaks of viral infective diseases

3. Objectives:

1. To study the existing problems in healthcare.
2. To understand the viral and bacterial diseases.
3. To identify the various patterns of diseases.
4. To predict the outbreaks of diseases.
5. To design and implement the knowledge discovery system.

4. Methodology:

This research is carried out on Filed survey in which the questionnaire is used for collecting the information from doctors and hospitals from Western Maharashtra. The data will be collected, hypothesis will be formulated and the analysis will be done with the help of statistical tools. Based on the analysis the result and Interpretation will be done. The data collection method is based on nature, scope, availability of money and time, precision factor etc. Data is collected by using primary and secondary data collection method.

Primary Data: The primary data related to the proposed study will be collected using questionnaire and formal and informal discussion with the doctors, medical experts.

Secondary Data: The secondary data related to the proposed study will be prepared by from newspaper, Published Reports, Periodicals, Internet, various research papers, thesis etc.

CONCLUSION :

There are various challenges in the field of healthcare as Intermediate form of medical data, Multilingual text refining, Domain knowledge integration , Large dimension of healthcare data , Complexity of natural language processing , Ambiguity and context sensitivity in data , Discover the relationship between the medical terms and concepts etc. The Prediction or outbreak of viral infective diseases is one of the biggest challenge in government hospitals. Most of the hospitals are unwilling to share their data due to privacy concern, patients also don't want to disclose their data; to resolve the same as well as to get the better knowledge based system for the prediction of viral infective diseases is very important so that the medical experts , healthcare's, government can take better remedial action.

REFERENCES :

1.Peter B. Jensen, Lars J. Jensen & Soren Brunak, "Mining electronic health records: towards better research applications and clinical care", Nature Reviews Genetics13, 395-405 (June 2012) http://www.nature.com/nrg/journal/v13/n6/full/nrg3208.html

2. Shaker H. El-Sappagh, Samir El-Masri, A. M. Riad, Mohammed Elmogy, "Data Mining and Knowledge Discovery: Applications, Techniques, Challenges and Process Models in Healthcare", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 3, May-Jun 2013, pp.900-906

3.Bijan Raheemi, "Data Mining and Knowledge Discovery in Healthcare and Medicine", April 30, 2014http://www.ieeeottawa.ca/aicn/data-mining-and-knowledge-discovery-in-healthcare-and-medicine/

4."Data Mining in Biomedicine: Current Applications and Further Directions for Research", S. L. TING, C. C. SHUM, S. K. KWOK, A. H. C. TSANG, W. B. LEE, J. Software Engineering & Applications, 2009, 2: 150-159 doi:10.4236/jsea.2009.23022 Published Online October 2009 (http://www.SciRP.org/journal/jsea)

5."A Survey on Medical Text Mining", Revathi M Nair, Sindhu L., International Journal of Computer Applications (0975 – 8887) Volume 108 – No 15, December 2014

6.Dengue Disease Prediction Using Weka Data Mining Tool, KASHISH ARA SHAKIL, SHADMA ANIS AND MANSAF ALAM, Department of Computer Science, Jamia Millia Islamia New Delhi, India

7.Unlocking the Power of Unstructured Data, Susan Feldman, Judy Hanover, Cynthia Burghard and David Schubmehl, June 2012, IDC Health Insights #HI235064

8.Knowledge Management in Health care, Christo El Morr, Julien Subercaze

9.A survey on Data Mining approaches for Healthcare, Divya Tomar and Sonali Agarwal International Journal of Bio-Science and Bio-Technology Vol.5, No.5 (2013), pp. 241-266 http://dx.doi.org/10.14257/ijbsbt.2013.5.5.25

10.Prediction and Decision Making in Health Care using Data Mining, Boris Milovic, Milan Milovic, International Journal of Public Health Science (IJPHS), Vol. 1, No. 2, December 2012, pp. 69~78, ISSN: 2252-8806 _ 69 , Journal homepage: http://iaesjournal.com/online/index.php/IJPHS

11." Rule Based Expert System for Viral Infection Diagnosis", Maitri Patel, Atul Patel , Paresh Virparia , International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, May 2013, ISSN: 2277 128X, www.ijarcsse.com

12. Text and Data Mining: Technologies under Construction, January 22, 2016, © 2016 Outsell, Inc.

13.**TEXT MINING IN HEALTHCARE: APPLICATIONS AND OPPORTUNITIES,** Uzma Raja, Tara Mitchell, Timothy Day and J. Michael Hardin, ResearchGate, (https://www.researchgate.net/publication/24182770_Text_mining_in_healthcare_Applications_and_opportunities)

14. Online Tutorial: Business Intelligence, Predictive Analytics, and Data Mining Contenthttp://wps.pearsoned.co.uk/ema_ge_turban_elec_comm_2012/217/55592/14231612.cw/index.html)

15. E-book: Introduction to Data Mining and Knowledge Discovery Third Edition By Two Crows Corporation, ISBN: 1-892095-02-5

16. "A Conceptual Study Of Knowledge Discovery Using Text Mining And Its Applications", International Journal of Management, IT & Engineering (ISSN: 2249-0558) Online , Volume 5, Issue 9, September 2015, page no. 174-181

17."A Review of Knowledge Based System in Healthcare Using Text Mining", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 5, Issue 6, June 2015, Page No. = 1326-1328

PRACTICAL APPROACH OF ORANGE TOOL FOR DATA MINING

MRS.CHARUSHILA PATIL
RSSP's Maharashtra College Of Science and Commerce , Pune.
charub.patil@gmail.com

Abstract: Data Mining is a process of finding out hidden patterns from large collection of data. There are various tools available to analyze, visualize and extract data using data mining. However all the tools are not compatible to perform all analysis operations, In this paper we have attempted data mining tool Orange to analyze data .

Keywords : Orange,KNN,Precision,Recall,F-measure,Confuison Matrix.

INTRODUCTION:

Data Analysis is a process of performing three major operations cleaning of the data, integration of the data and modeling data. However there are various tools of data mining to perform data visualization, data analysis and data extraction. Comparison of some tools along with parameters and features and decided to use for analysis. Some of Data Mining Tools: Orange, Weka, R, Rapid Miner, Knime, Data Melt

Orange : A data mining tool which is useful for visual programming and explorative data analysis. It can be written in Python. Orange has multiple components are known as widgets. This data mining tool supports macOS, Windows and Linux.

1. Proposed Work

The orange data mining is useful for analysis of the data.It is open source tool. It supports programming languages like C, C++ and Python that also supports data validation, comparison and prediction. Orange is very user friendly. Orange uses for practical Implementation. As shown in figure (a) we have taken **breast_cancer.tab** data from Orange which is available in orange data set.In this data we are having **286** instances and **9 features** of all.

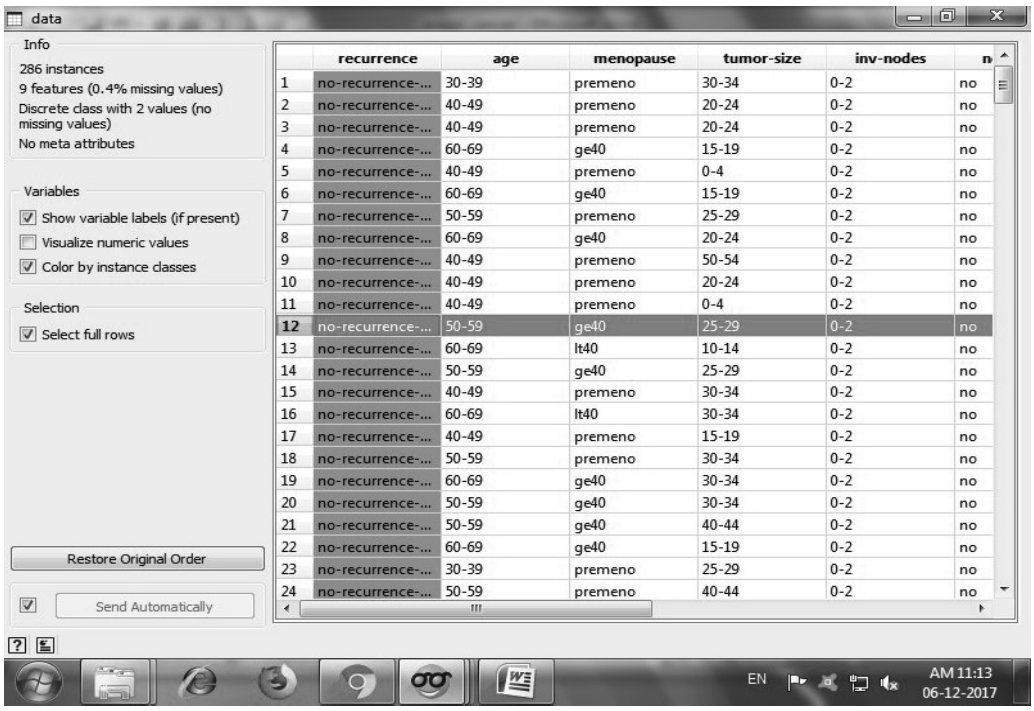
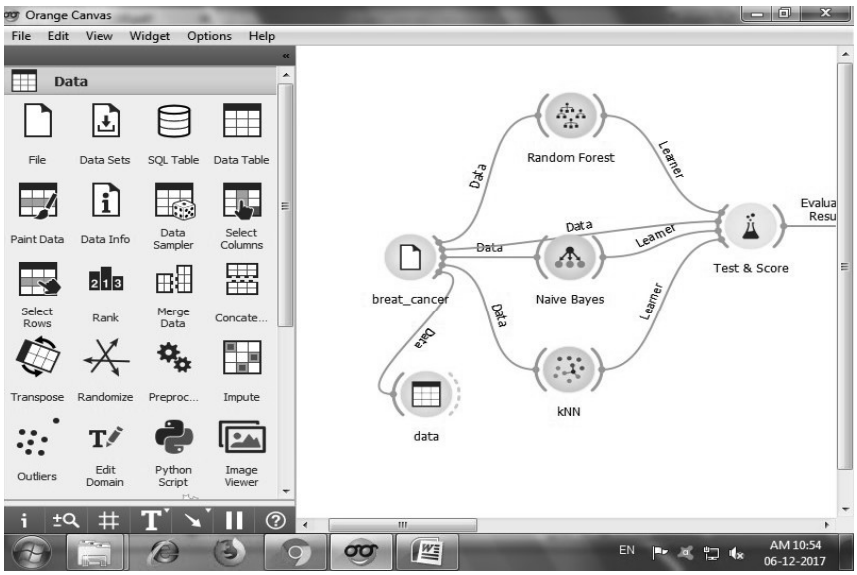


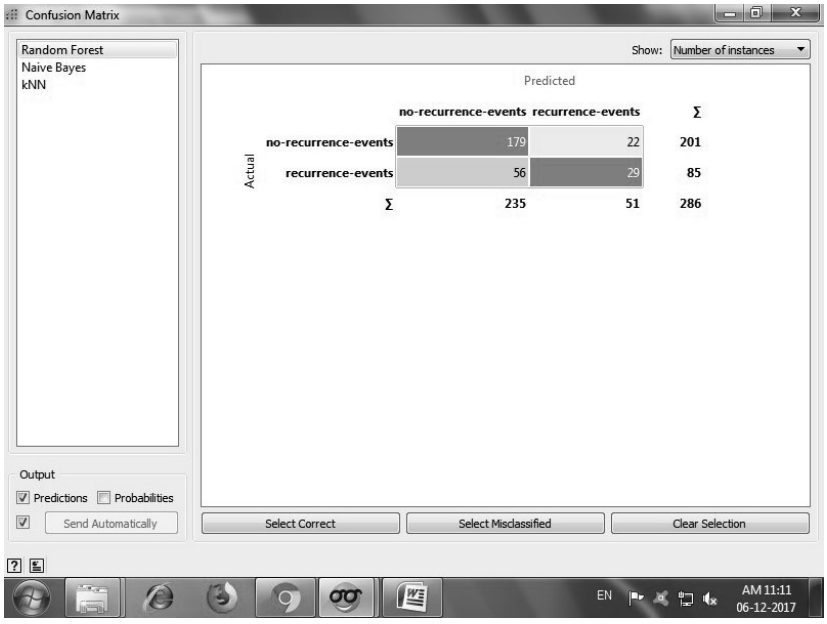
Figure (a)

In the following screen shot we can see the actucal implemantation of Orange tool.We can simple drag file icon to import the file from datasets. then we have applied three different algorithms Random Forest , Naive Bayes and KNN on the same data.We can visualize our reesult by Test and Score icon- figure(b).



Figure(b)

Confucion Matrix: After applying the algorithms on the dataset we can see the confusion matrix as shown in thefigure (c) A confusion matrix is a table that is often used to **describe the performance of a classification model** (or "classifier") on a set of test data for which the true values are known. Diagonal elements are showing the values which are correctly classified and of diagonal elements of matrix are wrongly classifies instances.



Figure(c)

CONCLUSION :

We can simply say that Orange tool can by used for data mining for data analysis. this tool will give you best visualization results which will help to improve data analysis .This tool is effective for data anylsis. We have applied algorithms like Random Forest , KNN,Naive Bayes algorithm. Orange tool is giving best results for these algorithms.

REFERENCES :

1. Zhang, H., Raitoharju, J., Kiranyaz, S., & Gabbouj, M. (2016). Limited random walk algorithm for big graph data clustering. Journal of Big Data, 3(1), 26.

2. Pempek, Tiffany A., Yevdokiya A. Yermolayeva, and Sandra L. Calvert. "College students' social networking experiences on Facebook." Journal of applied developmental psychology 30.

3 (2009): 227-238. 3. Griffiths, Mark D. "Facebook addiction: concerns, criticism, and recommendations— a response to Andreassen and colleagues." Psychological Reports 110.2 (2012): 518- 520.

STUDY AND ANALYSIS OF FAILURES OF SOFTWARE PROJECT AND ITS IMPACT ON BUSINESS OF IT COMPANIES

PROF. ASHVINI SHENDE
PhD Student , Management TMV Pune
ashwinishende@iccs.ac.in
PROF. SARITA BYAGAR
PhD Student , Computer Studies Symbiosis Technical University ,Pune
sarita.byagar@iccs.ac.in

Abstract : Most software projects can be considered at least partial failures because few projects meet all their cost, schedule, quality, or requirements objectives. Failures are rarely caused by mysterious causes, but these causes are usually discovered post-mortem, or only after it's too late to change direction. This research paper is based on survey carried out on different software consultants, developers and practitioners who were asked to provide reasons of failed projects with which they have been acquainted. Many factors has been considered in this research which generally leads to failure of software project and after analysis it has been proven that "Lack of user involvement' in the requirement phase leads more impact on software project failures. Hence, in this research, researcher has recommended that user should get involved completely in requirement gathering, designing phase. In this paper researcher has collected data from IT companies for checking impact of requirement changes on its Business.

Keywords: Software project, factors responsible for software project failures etc.

INTRODUCTION:

21st century known for computerization of all manual works, that human being was doing so far. Computerization made man life easy and this computerization become possible because of integration of hardware and software. Software plays most important role in the automation of most of electronic appliances. Hence, in current market, demand for all types of software is increasing day by day. This demand leads to development of thousands of software applications in turn increase in software industries.

Every year many software industries are spending billion on IT application development. Statistically, 31% of projects will be cancelled before they ever get completed. 53% of projects will cost twice as of their original estimates, overall, the success rate is less than 30% [1]. Why did the project fail? From symptom to root cause -what are the major factors that cause software projects to fail? What are the key ingredients that can reduce project failure?

Project failure can be defined as one or a combination of cost overruns, late deliveries, poor quality, and/or developing a product that does not get used. Regardless of their involvement during the planning stages, more often than not, software developers bear the brunt of the responsibility for such situations; after all, they're the ones who built the application. However, closer examinations of the projects do not always show evidence of incompetence [2].

1. Scope of the Study

The study is related to the study and analysis of failures in software project. Pune city has been considered for this research work. As this research mainly focused analysis of failures in software project, the scope of this research is decided to have software companies resides in following two areas of Pune City.

1. PMC area
2. PCMC area

Pune is the second largest city in Maharashtra and well known for educational facilities, research institutes and software industry. Due to the good educational facilities, Pune is called as "**The Oxford of the East**" and hence students from all over the world are getting attracted towards pune city. Due to big software industry, pune is transforming into vibrant modern city with bubbling activities in the IT and Hi-Tech sectors. Thousands of software companies can be found in pune city. And as there is software development industries, SDLC process surely gets followed by all software companies.

During the course of the present study the main focus has been given on the study and analysis of failures in software project. In this research, survey has been carried out for analysis of current scenarios which leads to software project failures in many software industries.

Objectives of the study

The main objective is to study impact of collecting needs from customer on business of IT companies.

3. Research Methodology

This research study is related to study and analysis of failures of software project. It utilizes both primary and secondary data. The secondary data utilizes already available information both published as well as unpublished. For primary data however such a facility is not available and it has to be collected by using the survey method. The scope of research is limited; the survey is undertaken by obtaining a purposive and quota sample. The description of the research methodology required for the process of obtaining a sample as well as the nature and size of sample should be adequately explained. Purposive, quota and convenience sampling techniques involves the selection of respondents based on the

important characteristics under study such as where they work, position in organization, specific knowledge related to the research problem etc.

4. Data Presentation, Analysis and Interpretation

Survey based research methodology has been used to carry out this research. This research is related to the study and analysis of failures of Software project and its impact on software cost if requirements are changing frequently special reference in Software companies of Pune city. The researcher has tested positively the hypotheses of this research study, with the help of primary and secondary data. For the purpose of the study, samples have covered all software companies present under PMC and PCMC area. Hence, the researcher has selected one sample viz. software companies present under PMC and PCMC area and collected data from the employees working in these software companies.

4.1 Common SOP Issues that may affect Software business.

For this, researcher has collected data from 10 companies, where every company's 5 clients details have been collected for measuring development time or duration span during year 2015-2016.

According to SDLC phase, first phase where client's SOP is finalized. 30 % to 35 % time of total development time is only required to collect SOP and freezed it. [10]

From each company 5 clients data is collected and analyzed their SOP collection duration. From table 1 researcher has observed that for each company out of 5 clients atleast 3 clients are taking more time for SOP as their SOP is not freezed and eventually it has impact on software business.

Sr.No	Company Name	Client Name	Total Project Duration (in Months)	Actual Duration required for SOP (in months)	% in months duration only for SOP	Impact on COST when SOP is not freezed				Total No of Clients took extra time to complete SOP
						upto 10%	upto 20%	upto 30%	upto 40%	
1	KPIT Cummins	Client A	10	4	40		Yes			3
		Client B	15	5	33.33	Yes				
		Client C	24	10	41.66667		Yes			
		Client D	12	4	33.33333	Yes				
		Client E	8	4	50			Yes		
						10%	20%	30%	40%	
2	SAP	Client A	18	8	44.44444			Yes		
		Client B	10	5	50.00				Yes	
		Client C	36	15	41.66667			Yes		
		Client D	20	8	40			Yes		
		Client E	12	5	41.66667			Yes		
						upto 10%	upto 20%	upto 30%	upto 40%	

		Client A	**15**	**4**	**26.66667**					
		Client B	**20**	**8**	**40.00**		**Yes**			**3**
3	**Harman**	Client C	**16**	**8**	**50**			**Yes**		
		Client D	**10**	**4**	**40**		**Yes**			
		Client E	**12**	**4**	**33.33333**	**Yes**				
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	10	4	40		**Yes**			
		Client B	**18**	**6**	**33.33**	**Yes**				**3**
4	**Intelzign**	Client C	**24**	**11**	**45.83333**		**Yes**			
		Client D	**14**	**5**	**35.71429**	**Yes**				
		Client E	**12**	**5**	**41.66667**		**Yes**			
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	12	4	33.33333	Yes				
		Client B	**18**	**8**	**44.44**		**Yes**			**3**
	L & T									
5		Client C	**20**	**10**	**50**			**Yes**		
	Infotech									
		Client D	**15**	**6**	**40**		**Yes**			
		Client E	**10**	**3**	**30**	**Yes**				
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	24	10	41.66667		Yes			
		Client B	**15**	**5**	**33.33**	**Yes**				**3**
6	**ATOS**	Client C	**20**	**8**	**40**		**Yes**			
		Client D	**14**	**6**	**42.85714**		**Yes**			
		Client E	**10**	**3**	**30**	**Yes**				
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	**10**	**3**	**30**	Yes				
		Client B	**15**	**5**	**33.33333**	**Yes**				**2**
7	**Davachi**	Client C	**12**	**4**	**33.33333**	**Yes**				
		Client D	**15**	7	**46.66667**			**Yes**		
		Client E	**8**	**4**	**50**				**Yes**	
						upto	**upto**	**upto**	**upto**	
										3
						10%	**20%**	**30%**	**40%**	
		Client A	15	6	40		Yes			
		Client B	**18**	**6**	**33.33**	**Yes**				
8	**CLSA**	Client C	**24**	**12**	**50**				**Yes**	
		Client D	**15**	**4**	**26.66667**	**Yes**				
		Client E	**12**	**5**	**41.66667**			**Yes**		
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	18	6	33.33333	Yes				
		Client B	**12**	**5**	**41.67**		**Yes**			**3**
9	**Zensar**	Client C	**24**	**11**	**45.83333**			**Yes**		
		Client D	**10**	**4**	**40**		**Yes**			
		Client E	**15**	**4**	**26.66667**	**Yes**				
						upto	**upto**	**upto**	**upto**	
						10%	**20%**	**30%**	**40%**	
		Client A	10	4	40		Yes			
		Client B	**15**	**5**	**33.33**	**Yes**				**2**
	TechHigh									
10		Client C	**24**	**10**	**41.66667**		**Yes**			
	way									
		Client D	**12**	**4**	**33.33333**	**Yes**				
		Client E	**8**	**3**	**37.5**	**Yes**				

Table 1. Software Development Life Cycle during Year 2015-2016

Hypothesis 1: If collected needs are not freezed, then it has impact on business. Referring above Table 1. Following hypothesis is proved

H0 -: On an average, clients are taking 35% of time duration for SOP (i.e $\mu = 0.35$)

H1-: On an average, clients are taking more than 35% of time duration for SOP, which has an impact on business (i.e $\mu > 0.35$)

Step IV: Calculation of Z value.

Sample mean = 0.3893

Population mean under H0 is $\mu = 0.35$

Z= diff. / S.E.

where,

diff = 0.3893 – 0.35 = 0.0393

S.E. $= \sigma/\sqrt{n} = 0.009202$

Zcal = 0.0393 / 0.009202 = 4.2738

Table value of Z for one tail test at 5% level of significance is Ztab = 1.64

Step VI: Conclusion:

Since Calculated value of Z (4.2738) > Table value of Z (1.64)Hence, we accept H1 which means that, on an average, clients are taking more than 35% of time duration for SOP, which has an impact on business and hence alternate hypothesis " **If collected needs are not freezed, then it has impact on business"** is accepted.

CONCLUSION :

In this research paper, Researcher has collected data from IT companies and for measuring impact of change in requirements , analysis shows that when there are change in requirements which requires extra time and thus requires extra cost. Thus it has impact on business.

REFERENCES :

1.http://www.enterprisecioforum.com/en/blogs/pearl/it-project-failure-symptoms-and-reasons.

2.http://www.infoq.com/articles/software-failure-reasons

Jones, Capers, *Patterns of Software Systems Failure and Success,*

3.International Thompson Computer Press, Boston, Mass., 1996.

4.https://ankitmathur111. wordpress.com/2012 /06/20/whats-whys-hows- of-software-testing-wwh/ http://istqbexamcertification.com/what-is-fundamental-test-process-in-software-testing/

5.http://www..ibm.com/software/products/en/ratidoorfami

6.https://scitools.com

7.https://en.wikipedia.org/wiki/Software_development_effort_

estimation

REVIEW PAPER ON CONTENT BASED IMAGE RETRIEVAL SYSTEMS USING LOW LEVEL FEATURES

MS.M.Y.PATIL
Computer Science Department,
BaburaojiGholap College, Sangvi, Pune-17.
Manjushap7@gmail.com
MS.M.A.MAHAJAN
Computer Science Department,
BaburaojiGholap College, Sangvi, Pune-17.
Monamhjn2289@gmail.com

Abstract : Content Based Image Retrieval (CBIR) is a technique which uses optical features of image such as color, texture, shape etc. to search user'srequisite image from large image scale database according to user's requirements in the form of a query image. Images are retrieved on the basis of similarity in features where features of the query specification are compared with features from the image database. The image retrieval is interesting and fastest developing methodology in all fields. It is effective and well-organized approach for retrieving the image from large image scale database. Content Based Image Retrieval is a technique to take input as query object and gives output from an image database. In this paper surveys has been conducted on some CBIR features such as color, texture and shape retrieval of images from the large scale database for efficient and accurate image retrieval. After going through exhaustive analysis and combination of CBIR techniques so this paper gives summarization of the different features like Color, Texture and Shape of images with their functionality for content based image retrieval systems.

Keywords : Content Based Image Retrieval (CBIR), feature extraction, Color, Shape, Textures

INTRODUCTION:

Image retrieval is technique concerned with searching and browsing digital images from large image scale database collection. This area of research is very energetic research since the 1970s [1]. More images have been produced in digital form around the world, image retrieval invites interest among researchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, database applications and other related area [1]. The approach of content based image retrieval (CBIR) system is to search image and retrieve relevant images from large image scale database using optical content of an image. CBIR is a technique which uses optical contents called as features. To search images from large scale image database according to user's request in the form of a query image. In today's era, CBIR receives input as a query object and as an output it receives similar objects from alarge scale image database. Generally for image retrieval CBIR used as optical features like color, texture, shapes or any combination of them. In large image scale databases content based image retrieval (CBIR) is searching the problem of digital images. Most of the content based image retrievals (CBIR) use image low level features as color, shape & texture. The database provided by James S. Wang is used for an image database is to store and retrieve an image or image sequences that are relevant to a query.In the WANG database 1,000 and 10,000 images has been used. In this database, the images are divided into 10 classes [2].Each class contain 100 images and 1,000 images has been used. Arrangement of the images in the database into 10 classes makes easy evaluation of the system. Early 1990 CBIR was become a very active research area. Most image retrieval system have been used like Random browsing, Search by example, search by sketch, search by text and navigation with customized image categories [3]. Today, CBIR low level features such as Color, Shape and textures using these content we retrieve images from large image scale database.With the Content Based Image Retrieval there are numbers of advantages over image retrieval techniques compared to other simple retrieval approaches such as text based retrieval techniques.CBIR has two features Low level feature and High Level feature. Low level features including color, shape, texture, spatial information etc. and High level features including human face, text descriptor, keywords. Each of one having different methods to retrieve image from large scale image database which shows in fig 1.

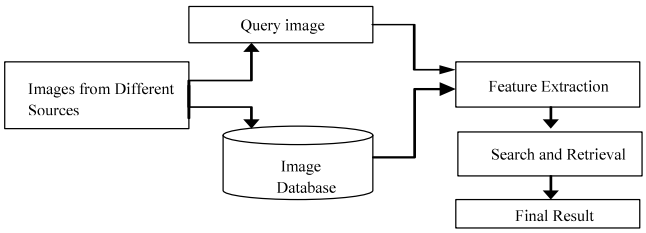


Fig 1: Block diagram of CBIR system

1. Feature Extraction Techniques

Content Based Image Retrieval (CBIR) uses Feature Extraction is a technique toextracts image from large scale image database including optical features can be classified as low level features. Multiple methods have been introduced for each of these optical features and each of them characterizes the feature from a different perception [4].Main three low level features are the Color, Texture and Shape.

A. Color Feature

The color feature is one of the most widely used visual

features in image retrieval .It is basic characteristics for the content of images. With the color feature human can classifies and differentiate between object and images. Colors are used in image retrieval because they are provide powerful information about images. The color feature is one of the most widely used visual features in image retrieval. Images characterized by color features have many advantages like robustness, effectiveness, implementation simplicity, computational simplicity, low storage requirement [5]. To extract the color features from the content of an image, we need to select a color space and use its properties in the extraction. In common, colors are defined in three dimensional color spaces. In digital image purposes, RGB color space is the most prevalent choice [2]. The main drawback of the RGB color space is that it is device dependent system [2]. For the description of color feature many methods can be used likeColor Histogram, Conventional Color Histogram, Invariant Color Histogram, Fuzzy Color Histogram, Geometric Moment, Average RGB, Color Moment, ColorCorrelogram, and Color Coherence Vector.Color moments have been successfully used in content based image retrieval systems. It has been shown [6] that characterizing one dimensional color distributions with the first three moments is more robust and runs faster than the histogram based methods.

B. Texture Feature

Texture is one of the most significant features of normal images. There is no recognised definition for texture, but it can say that it provides the measure of properties. Texture can be represented by Grey Level Co-occurrence [7]**.** To get texture information of image steering features were dragged out. The six image texture properties were coarseness, regularity, directionality, contrast, line likeness and roughness.Texture can be used two major methods Structural and statistical methods. Structural method describe texture by identifying structural primitives. And Statistical method uses statistical distribution of the image intensity [8] to characterize the texture. Many different methods are proposed for computing texture but among those methods, no single method works best with all types of texture. Some common methods are used for texture feature extraction such as Discrete Wavelet Transform, Gabor Wavelet Transform, Haar Discrete Wavelet Transforms, Ranklet Transform, Fourier Transform, discrete cosine transform, Hadamard Transform, Gaussian Pyramid, Laplacian Pyramid, Steerable Pyramid, Gabor Filter.

C. Shape Feature

One more important optical feature is Shape which is most powerful feature used for image classification, indexing and retrievals. Shape is the basic features used to describe image content. Shape properties play asignificant role in content based image database systems invented by computer visualisation researchers [9].Choosing Shape feature for describing an object is for the reason that its inherent properties like identifiability, reliability etc. and retrieval of image can be performed [9].Shape can be characterised based on two classes Boundary Based and Region Based. Boundary Based shape is used for shape representation and region based shape used for describing shape [10]. There are some shape Descriptor methods like Geometric Moments Geometric Moments, Zernike Moments, Fourier Descriptor, Grid Method and Shape Matrix.

2. Literature Review

B. Ramamurthy, K.R. Chandran[5] provide medical image data retrieval from large medical database using image content like shape. Also provide efficient tool for retrieving efficient medical image from large medical image database.

PoojaVerma, Manish Mahajan, annella[11] studied performance comparison of Content Based image retrieval features like shape &color. For finding image there is need to developing effective technique. In this paper an effective image retrieval technique is used which uses color, and shape features of an image [11].

R F Xiang Yuan; Chang-Tsun Li [12] uses CBIR method to retrieve image. Leading to used canny edge detector to extract information from the image. Then, Hough transform is useful to edge map for retrieving image from large scale image database. Then used band-wise matching (BWM) is used to calculate centroid of the Hough peaks. And finally CBIR ranks the image in the database and retrieves number of images with maximum rank.

S. Manoharan, S. Sathappan [13] Applied high level filtering Anisotropic Morphological Filters, hierarchical Kaman filter and particle filter with feature extraction method built on color and grey level feature.

G. Pass [14]proposed a system to describe spatial features. This model is invariant to scaling, rotation and shifting. In this paper segmentations are objects of the images and all images are segmented into several pieces is applied to extract to enhance the user interaction.

Shih and Chen [15] used partition-based color-spatial technique where an image is divided into 100 blocks. For each block, the first three color moments of each color component of each block are extracted and clustered into several classes.

Monika Daga, KamleshLakhwani [16] proposed a new CBIR system using the negative selection algorithm (NSA) and which is used to reduced complexity and efficiency of retrieving image from large image scale database.

Swati Agarwal, A. K. Verma, Preetvanti Singh [17] proposed algorithm for retrieving image based on shape and texture features but not only on the basis of color.In this paper image is decomposed into wavelet coefficients which gives horizontal, vertical and diagonal features in the image. The grouping of DWT and EHD methods increases performance of image retrieval for shape and texture.

Prasad, Biswas, and Gupta [18] recognised for the better result merging color and spatial information. In this paper maximum of three color regions together and obtain Dominant Region segmentation [19].

Xiang-Yang Wang, Hong-Ying Yang, Dong-Ming Li [20]

proposed a new Content Based Image Retrieval technique using color and texture information for higher retrieval efficiency. Firstly the image is transformed from
RGB space to the independence of the color contents of an image by using Zernike chromaticity distribution moments from the chromaticity space. Then the texture attributes are extracted using a rotation-invariant and scale-invariant image descriptor.Finally, the combination of the color and texture information provides a dynamic feature for color image retrieval. And the experimental results proposed color image retrieval is more accurate and efficient in retrieving the image.

Rao, Srihari, and Zhang [21] presented three types of Spatial Color Histogram, which are the Annular, Angular, and Hybrid Color Histogram. All of these three techniques uses color histogram due to the spatial information. These various color-spatial histograms provide a way to explore a particular range of color in spatial domain.

Chan and Chen [22] considered mean value is calculated separately for each *R*, *G*, and *B* color components for each block. The benefit of considering the mean value is that effects from noises in the images and the dissimilarities in sizes of images are expressively reduced.

Yamamoto [23]proposed an account of the spatial information of colors by using multiple histograms.
Researcher has divides two rectangular sub images using a straight line vertically or horizontally. In each sub-image, the division process continues until each region become small.

S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak [24]proposed a technique for image retrieval based on semantic contents. The assemblage of three low level feature extraction methods color, texture, and edge histogram. There is a requirement to contain new features in future for better retrieval effectiveness. Any combination of these techniques, which is more suitable for the application, can be used for retrieval.

Heng Chen and Zhicheng Zhao [25]defined relevance feedback method for image retrieval. This is an efficient method of CBIR. Relevance feedback (RF) is an efficient method for content-based image retrieval (CBIR). SVM-based Relevance feedback (RF) scheme is proposed to improve performance of image retrieval. This scheme is accepted to balance the proportion of positive and negative samples and then synthesis scheme for multiple classifiers based on the final query results.

CONCLUSION :

In this survey paper different types of methods like canny edge detector, Hough transform, band-wise matching, Anisotropic Morphological Filters, hierarchical Kaman filter and particle filter, negative selection algorithm (NSA), Annular, Angular, and Hybrid Color Histogram, Zernike chromaticity distribution moments, Relevance feedback (RF) , SVM-based Relevance feedback (RF) are describe. All these methods given in paper are best to individual. These methods are used by other researchers to bring out different output. There are some combinations are also available forcolor, Shape and texture for retrieving image from large scale image database. In this survey paper discussion on techniques of an image retrieval based on low level features for efficient and accurate image retrieval are provided by literature study. So this paper gives summarization of the different features of images with their functionality for content based image retrieval systems.

REFERENCES :

1. HuiHui Wang, DzulkifliMohamad and N.A. Ismail ―Approaches, Challenges and Future Direction of Image Retrieval journal of computing, volume 2, issue 6, june 2010, issn 2151-9617, pp. 193- 199.

2. Ahmed J. Afifi and Wesam M. Ashour ―Content-Based Image Retrieval Using Invariant Color and Texture Features‖ 978-1-4673- 2181-5/12/$31.00 ©2012 IEEE.

3. Mohd. Danish*, RitikaRawat **, Ratika Sharma- "Comparative Study on CBIR based on Color Feature" Mohd. Danish et al. Int. Journal of Engineering Research and Applications www.ijera.com ISSN: 2248-9622, Vol. 3, Issue 5, Sep-Oct 2013, Pp.839-844

4. AmandeepKhokher, Rajneesh Talwar, ―Content-based Image Retrieval:Feature Extraction Techniques and Applications‖ International Conference on Recent Advances and Future Trends in Information Technology (iRAFIT2012),pp.9-14.

5. B. Ramamurthy, K.R. Chandran "CBMIR: shape based image retrieval using canny edge detection and k means clustering algorithm" International Journal of Computer Application *(0975 – 8887) Volume 17– No.6, March 2011*

6. M. Stricker, and M. Orengo, Similarity of color images, In SPIE Conference on Storage and Retrieval for Image and Video Databases, volume 2420, 1995.

7. M.BabuRao et al., Content Based Image Retrieval using Dominant Color, Texture and Shape, International Journal of Engineering Science and Technology (IJEST), 3(4), 2887-2896, 2011.

8. MeenakshiShruti Pal, Dr.Sushil Kumar Garg, Image Retrieval: A Literature Review, International Journal of Advanced Research in Computer Engineering and Technology, Volume 2, Issue 6, 2013.

9. SaiAnand.C, Tamilarasan.M and Arjun.P ―A Study on Curvature Scale Space ‖ International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Special Issue 3, July 2014,pp.168-174.

10. M. Sifuzzaman, M.R. Islam and M.Z. Ali , Application of Wavelet Transform and its Advantages

Compared to Fourier Transform, Journal of Physical Sciences, Vol. 13, 2009.

11. PoojaVerma and Manish Mahajan, annella, "Content based image retrieval color by comparing color and shape feature" International Journal of Computer Science and Communication Vol. 2, No. 2, July-December 2011, pp. 641-643. J.

12. R F Xiang Yuan; Chang-Tsun Li; "CBIR approach to building image retrieval on linear edge distribution" IEEE International Conference 2006.

13. S. Manoharan, S. Sathappan, "A Novel Approach for Content Based Image Retrieval Using Hybrid Filter Techniques", 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013.Colombo, Sri Lanka.

14. ROI Image Retrieval Based on the Spatial Structure of Objects, Weiwen ZOU1 , Guocan FENG2, 12Mathematics and Computational School, Sun Yat-sen University, Guangzhou, China, 510275 paper: 05170290.

15. J. L. Shih and L. H. Chen, "Color image retrieval based on primitives of color moments", in Proceedings Of the Vision, Image, and Signal Processing, IEE, Vol. 149, No. 6, 2002, pp. 370-376.

16. Monika Daga, KamleshLakhwani, "A Novel Content Based Image Retrieval Implemented By NSA of AIS", International Journal of Scientific & Technology Research Volume 2, Issue 7, July 2013 ISSN 2277-8616.

17. Swati Agarwal, A. K. Verma, Preetvanti Singh, "Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor", International Conference on Information Systems and Computer Networks, proceeding of IEEE xplore-2013.

18. B. G. Prasad, K. K. Biswas, and S. K. Gupta, "Region-based image retrieval using integrated color, shape and location index", Computer Vision and Image Understanding, Vol. 94, 2004, pp. 193-233.

19. K. C. Ravishankar, B. G. Prasad, S. K. Gupta, and K. K. Biswas, "Dominant color region-based indexing Technique for CBIR", in Proceedings of the International Conference on Image Analysis and Processing, ICIAP, Venice, Italy, 1999, pp. 887–892.

20. Xiang-Yang Wang, Hong-Ying Yang, Dong-Ming Li "A new content-based image retrieval technique using color and texture information", Computers & Electrical Engineering, Volume 39, Issue 3, April 2013, Pages 746-761.

21. A. Rao, R. K. Srihari, and Z. Zhang, "Spatial color histograms for content-based image retrieval", in Proceedings of the IEEE International Conference on Tools with Artificial Intelligence, 1999, pp. 183-186.

22. Y. K. Chan and C. Y. Chen, "Image retrieval system based on color-complexity and color-spatialfeatures", Journal of Systems and Software, Vol. 71, 2004, pp. 65-70.

23. H. Yamamoto, H. Iwasa, N. Yokoya, and H. Takemura, "Content- Based Similarity Retrieval of Images Based onSpatial Color Distributions", ICIAP '99 Proceedings of the 10th International Conference on Image Analysis andProcessing.

24. S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak, "A Universal Model for Content-Based Image Retrieval",World Academy of Science, Engineering and Technology, Vol: 2 2008-10-29.

25. Hengchen, zhichengzhao "an effective relevance feedback algorithm for image retrieval" 978-1-4244-6853-9/10/ 2010 IEEE.

REVIEW OF SALT PEPPER NOISE REMOVAL FILTERS

HEMALATA CHAVAN[*], APARNA JOSHI[**]
M.C.A. (Engg.)[*], M.Sc (CS) [**]
Pratibha College of Commerce and Computer Studies, Chinchwad, Pune-33, India
m_hemalata2000@yahoo.com, aajoshi2@gmail.com
9657717307[*], 9420954286

ABSTRACT : To send visual digital images are a major issue in the modern data communication network. The images sent from sender end may not be the same at the receiving end. The image obtained after transmission is often corrupted with noise. Noise is unwanted information present in image that can harm the quality of image. A noise is introduced in the transmission medium due to a noisy channel, errors during the measurement process and during quantization of the data for digital storage. There are different types of noise and noise detection and reduction techniques. This paper deals the performance comparison of the various filters which are used to remove the salt and pepper noise(impulse noise) from the images. The performance criteria are the mean square error (MSE) and peak signal to noise ratio (PSNR).

Keywords: Salt and Pepper, Impulse, MSE, PSNR.

INTRODUCTION:

The basic problem in image processing is the image enhancement and the restoration in the noisy environment. If we want to enhance the quality of images, we can use various filtering techniques which are available in image processing. There are various filters which can remove the noise from images and preserve image details and enhance the quality of image. Salt and pepper noise is an impulse type of noise, which is also referred to as intensity spikes. Impulse noise removal is a mechanism for detection and removal of impulse noise from images.This is caused generally due to dead pixels, analog-to-digital converter errors, errors in data transmission, malfunctioning of pixel elements in the camera sensors, faulty memory locations, or timing errors in the digitization process. It has only two possible values, 'a' and 'b'. The probability of each is typically less than 1. The corrupted pixels are set alternatively to the minimum or to the maximum intensity values, giving the image a “salt and pepper” like appearance. Unaffected pixels remain unchanged. For an 8-bit image, the typical intensity value for pepper noise is 0 and for salt noise 255. [1]

The image which has salt-and-pepper noise present in image will show dark pixels in the bright regions and bright pixels in the dark regions. [1]. The Impulse noise occurs due to the quick transitions such as faulty switching , can be caused by the dead pixels, or due to analog-to-digital conversion errors, or bit errors in the transmission, etc. This all can be eliminated in large amount by using the technique dark frame subtraction and by interpolating around dark/bright pixels.

The Probability Density Function (PDF) of Impulse noise is given by,

$$p(z) = \begin{cases} P_a & \text{for } z=a \\ P_b & \text{for } z=b \\ 0 & \text{otherwise} \end{cases}$$

If b>a, intensity b will appear as a light dot in the image or level a will appear like a dark dot or vice-versa.[3]

TYPES OF FILTER

1) Max. Filter: Max filter is also known as 100th percentile filter. It replaces the value of pixel by the maximum intensity level of the neighborhood of that pixel. This filter finds brightest points in an image.[5]

This filter reduces the pepper noise because it has very low values of intensities. So it can only remove pepper noise.

$$f\hat{}(x,y)=max\{g(s,t)\} \; where \; (s,t)\in Sxy$$

Where, g(s, t) is the sub image area of m*n image.

2) Min Filter: Min filter is also known as 0th percentile filter. It replaces the value of pixel by the minimum intensity level of the neighborhood of that pixel .This filter is used to find the darkest point in an image. It uses the minimum intensity value in a sub image area. It removes salt noise from an image containing salt and pepper noise due to its high intensity value [5][6].

The min filter can be represented by the following equation:

$$f\hat{}(x,y)=min\{g(s,t)\} \; where \; (s,t)\in Sxy$$

3)MINMAX Filter: This filter is the combination of the Min and Max filter. This filter does not remove the impulse noise but it is best suitable for the Gaussian or uniform noise.

4) Mean Filter: It is one of the simplestfilters among the existing spatial filters. It uses a filter window which is usually square. The filter window replaces the center value in the window with the average mean of all the pixels values in the kernel or window. [2]This filter is used to remove the salt & pepper noise both simultaneously from the image. The image details are not preserved in this operation, some details are lost.

5) Median Filter: Median filtering is a nonlinear operation used in image processing to reduce "salt and pepper" noise. Also Mean filter is used to remove the impulse noise. Mean filter replaces the mean of the pixels values but it does not preserve image details. Some details are removes with the mean filter. In the median filter, we do not replace the pixel value with the *mean* of neighboring pixel values, we replaces with the *median* of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the

pixel being considered with the middle pixel value. (If the neighboring pixel which is to be considered contains an even number of pixels, than the average of the two middle pixel values is used.) Fig.1 illustrates an example calculation.

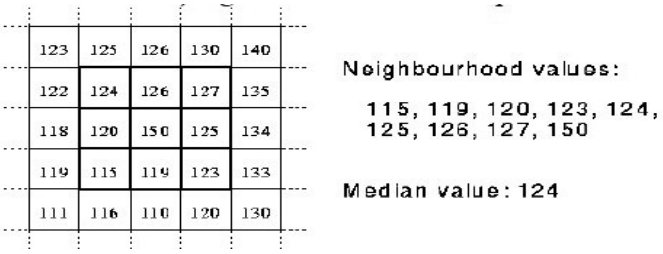


Figure 1: Example of Median Filter

The median filter gives best result when the impulse noise percentage is less than 0.1 %. When the quantity of impulse noise is increased the median filter not gives best result. Now we consider a sub image area of total image

$$f\hat{}(x,y)=median\{g(s,t)\} \qquad where\ (s,t)\in Sxy$$

Where g(s, t) is the sub image area of the m*n image. [7]

6) Weighted Median Filter: In this filter the weights are assigned to the each element in a window. These weights are multiplied to each element in the window. This filter also removes the impulse noise but image details are lost in this filter.

IMAGE QUALITY ASSESSMENT METRICS:

The image quality assessment measures are helpful in detecting the quality of the processed image in comparison with the original image. There are measurements like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) to evaluate the quality of the processed image.

Mean Square Error (MSE): The term MSE (mean square error) is the difference between the original image and the recovered image and it should be as minimum as possible.

Mean square error(MSE)

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\|I(i,j) - u(i,j\|^2$$

Peak Signal-to-Noise Ratio (PSNR): The term peak signal-to-noise ratio, PSNR, is the ratio between the maximum possible Power of a signal and the power of corrupting noise signal. [4]

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$$

CONCLUSION :

In this paper, we have discussed about noise and how it creeps into images while acquiring or sending it using transmission medium. We focused on different filtering techniques which are used to remove Salt and Pepper Noise from image database. Median filters are mostly preferred to remove salt and pepper noise because of their simplicity and less computational complexity. The effectiveness of median filter technique is assessed by image quality assessment metrics such as MSE, PSNR. Experimentally it proven that median filter shows best performance on salt and pepper noise using MSE and PSNR image quality metrics.

REFERENCES :

[1] E. Jebamalar Leavline and D. Asir Antony Gnana Singh, International Journal of Imaging & Robotics, vol. 11, no. 3, (2013).

[2]R. H. Chan, C.-W. Ho, and M. Nikolova, "Convergence of Newton'smethod for a minimization problem in impulse noise removal," J. Comput. Math., vol. 22, pp. 168–177, 2004.

[3] Suk Hwan Lim ; Maurer, R. ; Kisilev, P. "Denoising scheme for realistic digital photos from unknown sources". IEEE International Conference on Acoustics, Speech and Signal Processing, 2009.

[4] J. Harikiran, B. Saichandana and B. Divakar, "Impulse Noise Removal in Digital Image." International Journal of Computer Applications, Vol. 10, no 8, pp. 39-42.

[5] Sitendra Tamrakar and M.R.Aloney, International Journal of Image Processing and Applications, vol.7, no.2, July-December 2016.

[6] Ling Shao, Ruomei Yan and Xuelong Li, "From Heuristic Optimization to Dictionary Learning: A Review and Comprehensive Comparison of Image Denoising Algorithms," IEEE Transactions on Cybernetics, 1-14 August, 2013.

[7] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," second ed., Prentice Hall, Englewood, Englewood, Cliffs, NJ, 2002.

REPORT ON PLANNING AND IMPLEMENTATION OF ICT IN GOVERNMENT SCHOOLS IN MAHARASHTRA

MRS. SUTAPA CHAKRABORTY PAL, MR. RAMDAS S. BOLAGE
Assistant Professors,
Indira College of Commerce and Science,
Tathwade, Pune -33.
sutapa.chakraborty@iccs.ac.in, ramdas.bolage@iics.ac.in

Abstract : Introduction of information and Communication Technology (ICT) in schools of India is a vision which needs to be fulfilled. Government of Maharashtra share the same vision and has set up schemes with proper proposal, under Central Assistance to fulfill this objectives. Our aim of this paper is to give a detailed report on the Recommendations and Proposals of State Education Department on ICT @Government Schools in Maharashtra under financial planning and budget. We have analyzed the report considering the proposed government plan of action and expenditure section. This paper highlights the Governments efforts in the direction of spreading ICT in Government and Government aided schools in Maharashtra.

Keywords: ICT, RMSA, PAB

INTRODUCTION:

Education is the most effective and essential tool for growth of life and society for all ages. The purpose of education is training people for employment and simultaneously helping them to cope with their fast change in life style . Nowadays there is a fast growing interest in application in modern communication technologies in the field of education . The report of Gutterman says , "ICT can be extremely powerful enabler in efforts to bring positive and sustainable development to countries around the globe."

In todays world, 'Education through ICT' is being accepted as a powerful tool to promote social and economic development. ICT gives student and teachers new apparatus to learn and teach . The process of traditional teaching with chalkboard and textbooks can now be supplimented, if not replaced by computer and internet. According to many intellectuals , ICT will allows students to develop their mental human resource by enhancing their knowledge bank. According to **National Curriculum for Education** , " ICT prepares pupils to participate in a rapidly changing world in which work and other activities are increasingly transformed by access to varied and developing technology. Pupils use ICT tools to find , explore , analysis ,exchange and present information responsibly,creatively ."

ICT education in India gained importance as early as 1984-85 with introduction of CLASS (Computer Literacy Studies in Schools) school projects. However , computer-mediated learning in school education system is in it's initial phase for government and government aided schools in India. Concentrating our attention to one of the largest states in India , **Maharashtra**, is yet to implement ICT in most of its Government schools. According to **National University for Education Planning and Administration** (NUEPA) trends , only 50.63% Primary and higher secondary Government schools in Maharashtra had introduced ICT in Computer Education . With one major reason as shortage of fund, the government is allocating fund stagewise to aided secondary and higher secondary schools . Higher secondary schools run by local self government bodies were given computers from **Sarva shiksha Abhiyaan** based on availability of grants.

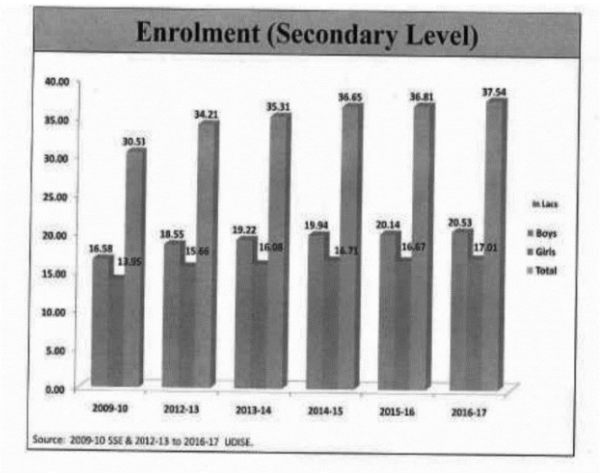
2. Methodology

Our aim of this project is to highlight the initiatives taken by MHRD , Department of School Education andLiteracy , Government of Maharashtra to enhance ICT introduction and implemenatation in government and government aided schools in Maharashtra . Through the **Project Approval Board** (PAB) meeting government has introduced Annual Work Plan and Budget under **Centrally Sponsored Scheme** of the Integrated **Rashtriya Madhyamik Shikasha Abhiyaan** (RMSA).

The following are achievments in education and literacy by Government of maharashtra till date
under **RMSA :-**

1. The overall enrollement increased from 36.81 lakh in 2015-16 to 37.54 lakh in 2016-17.
2. GER increased from 89.65% to 91.74% for the same year.
3. NER increased from 59.94% to 61.36% for the same year.
4. While retention rate has increased to 88.45%, drop out rate has decreased to 11.55% in 2016-17.

Bar graph for enrollment:



Thus overall school enrollment and retention percentage has improved from last few years in Maharashtra.

The following are the other observations :-

- The progerss of GIS mapping of state was found good . Target of completing 100% GIS mapping of schools in near future.
- Under RMSA , total 148 cmputer rooms were sanctioned but no construction started yet. This has led to bottleneck situation with unspend money and no construction . However work for science lab has started .
- Decision was taken by Government to **stop projects which are not yet started** due to the above mentioned bottleneck and **complete the ongoing school projects** .

We will now concentrate on **introduction of ICT** in government schools.

The following are the reports for ICT in Government School:

- Out of 4644 approved schools in 2011-12 , state has not implemented ICT project in 26 schools.

Dropping them 4618 schools stand approved.

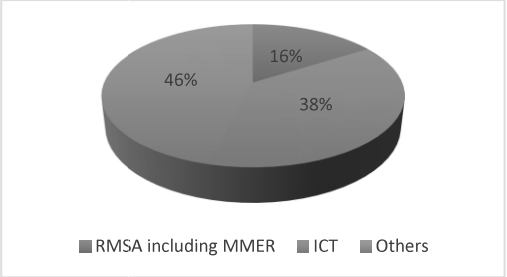
- Rs. 166.40 lakh was cancelled for ICT in the above 26 schools .
- 1500 schools are approved under ICT schools and are scheduled to start very soon.

Proposed teachers salary will be additional Rs. 10,000/- P.M. for ICT only.

- On ICT initiatives it was suggested by Secretary to start implementation of new models like Tablet Based Models.
- In service trainning of teachers for RMSA including ICT is @ Rs. 300/- for 10 days , with total Rs.1248 lakh was approved for 41607 teachers.
- An amount of Rs. 11083.20lakh @ Rs. 2.40 lakh/school , for 4618 schools was approved for recurring Grant for 2016-17 under norms of ICT @ schools scheme on reimbursement basis.
- Funds under above scheme reimbursed after submission of activity wise expenditure report for current year.
- Refresher trainning and MMER of ICT component is included in RMSA.
- ICT resources like JAWA, SAFTA was approved for other school projects.

Financial allocations under the variouas components for school projects :

components	Recurring cost outlay (RS. In Lakh)
RMSA including MMER	4724.33
ICT	11083.2
Others	13601.66
Total	**29409.19**



Government of Maharashtra has vision of extended application of ICT in school administration also . It plans to introduce **e-governance** and below is its proposed budget proposals for IT/e-governance.

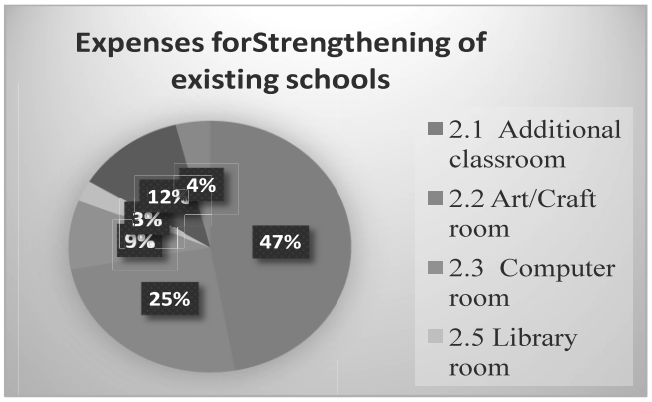
Budget propoals for IT/e-governance:

Sr .No	Activity	Proposal of state		
		Qty.	unit cost	Final cost
1	project - IT/e-governance			
	Complete School Management System- Shaala Darpan	1	70	70
	Total (in lakh)			**70**

Proposals and Recommendations for ICT@School education :

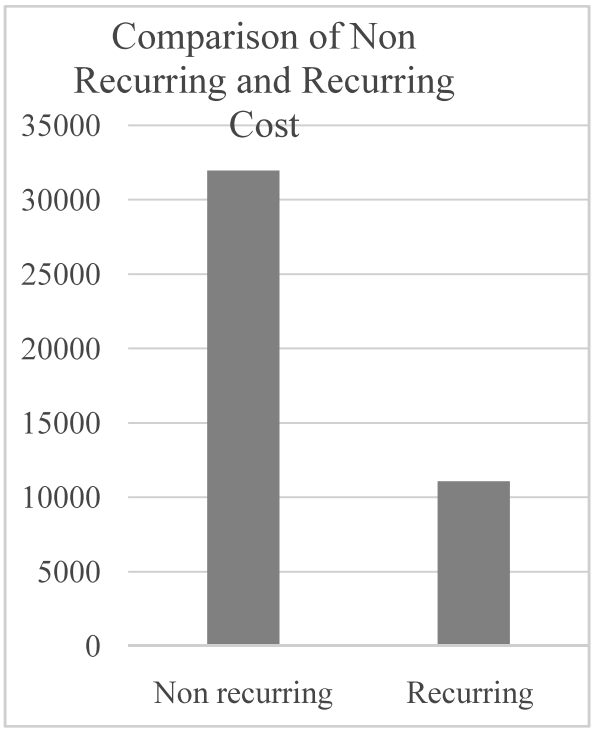
For Strengthening existing and upgrading new schools under RMSA:

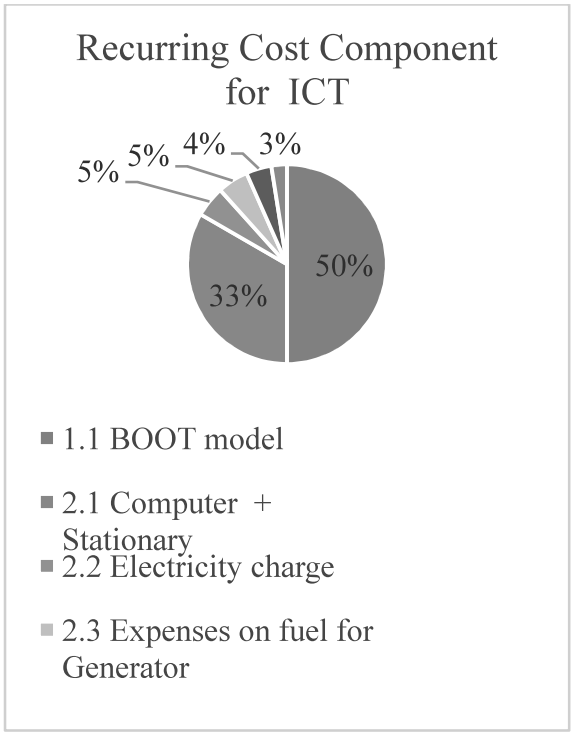
S.No	Activity	Proposal state		
		Qty.	unit cost	Final cost
	RMSA			
	Non recurring			
1	New schools/ Upgraded school			
	1.1 Two section school	26	162.46	4223.96
2	Strengthening of existing schools			
	2.1 Additional classroom	265	18.38	4870.7
	2.2 Art/Craft room	150	17.21	2581.5
	2.3 Computer room	50	18.29	914.5
	2.5 Library room	13	21.06	273.78
	2.6 Science lab	63	19.22	1273.86
	2.7 Other	174	2.344	407.98
	Total (In lakhs)			10322.32



B) ICT

S.No	Activity	Proposal state		
		Qty.	unit cost	Final cost
	Non recurring			
	BOO -model			
	1.1 Hardware and software support	4996	6.4	31974.4
	Total for Non-Recurring (In Lakh)			**31974.4**
	Recurring			
1	Salary of computer teacher			
	1.1 BOOT model	4618	1.2	5541.6
2	Other			
	2.1 Computer + Stationary	4618	0.8	3694.4
	2.2 Electricity charge	4618	0.12	554.16
	2.3 Expenses on fuel for Generator	4618	0.12	554.16
	2.4 Internet / Broad band charges	4618	0.1	461.8
	2.5 Telephone charges	4618	0.06	277.08
	Total for Recurring(In Lakh)			**11083.2**
	TOTAL FOR ICT (IN LAKH)			**43057.6**





CONCLUSION :

Thus we see that the Government of Maharashtra has proposed and recommended grants under Central Assistance for wide spread access of ICT in government schools and government aided schools in Maharashtra. It has plans for every aspect of ICT Education in schools viz classrooms, computer labs,teachers and also e-governence .

Some of the major companies who have been selected for 5000 government schools in key districts of maharashtra are IL&FS Education , Core Education and Technologies and Birla Shloka Edutech . Their contracts have been structured on Build , Own,Operate**(BOOT**) model . It is expected that with these companies, Government's whole hearted plan of action will be fulfilled.

Last but not the least it is worth mentioning that not only government efforts but a great deal of community actions and sincere efforts from segments of population have resulted in **digitalization of many village schools in Maharashtra.** Intact according to NDTV close to 47,000 government primary schools in Maharashtra have been digitally equipped through crowd funding. Thus the real objective of digitalizing young India will be fulfilled by joint cooperation of Government of Maharashtra together with whole hearted efforts of the citizens of the state.

REFERENCES :

1. Website of MHRD- Department of School Education and Literacy
2. Website of Maharashtra School Education Department
3. Report by Project Approval Board , Government of Maharashtra
4. Website of National Education Commission
5. Website of National University for Education Planning and Administration
6. Article on Revised ICT Scheme for Schools from Shodhganga – Inflibnet
7. Reports from India News , NDTV and Bussiness Line

COMPARATIVE STUDY OF BLUETOOTH, ZIGBEE, WI-FI, WI-MAX FOR IOT APPLICATIONS

MRS. NILAMBARI C. KALE.
Department of computer science,
Pratibha College of Com. & Computer Studies, Chinchwad, pune, India
nilambari.mate@gmail.com

Abstract: Internet of things is an emerging technology now a day's. It is becoming an almost basic necessity in general life. Today all aspects of our life, be it health of a person, his location, his movement or communication between different devices etc. can be monitored and analyzed using information captured from various connected devices. IOT fulfills the idea of inter connected devices where the devices are smart enough to share information to each other. So for the sake of communication different IEEE standards used. This paper represents comparision of different standards which are used in IOT applications. Bluetooth (over IEEE 802.15.1), ZigBee (over IEEE 802.15.4), and Wi-Fi (over IEEE 802.11) and WI-MAX(IEEE 802.16) are four protocol standards used for short range wireless communications with low power consumption.

Keywords: wireless standards, Bluetooth, ZigBee, Wi-Fi and WI-MAX

INTRODUCTION:

Internet of Things (IoT) is the computing environment to provide valuable services by interacting with various IoT applications, where diverse devices are connected within the existing internet infrastructure and through intelligent social applications. The IoT enables physical objects to see, hear, think and perform jobs by having them "talk" together, to share information and to coordinate decisions. The IoT transforms these objects from being traditional to smart by exploiting its underlying technologies such as ubiquitous and pervasive computing, embedded devices, communication technologies, sensor networks, Internet protocols and applications. These applications required communication protocols which supports short range wireless communications with low power consumption.

Bluetooth (over IEEE 802.15.1), ZigBee (over IEEE 802.15.4), Wi-Fi (over IEEE 802.11) and Wi-MAX(IEEE 802.16) are four protocol standards at MAC Layer used for IOT applications.

Wireless Protocols :

Bluetooth, also known as the IEEE 802.15.1 standard is based on a wireless radio system designed for short-range and cheap devices to replace cables for computer peripherals, such as mice, keyboards, joysticks, and printers. This range of applications is known as wireless personal area network (WPAN). Two connectivity topologies are defined in Bluetooth: the piconet and scatternet. A piconet is a WPAN formed by a Bluetooth device serving as a master in the piconet and one or more Bluetooth devices serving as slaves. A frequency-hopping channel based on the address of the master defines each piconet. All devices participating in communications in a given piconet are synchronized using the clock of the master. Slaves communicate only with their master in a point-to-point fashion under the control of the master. The master's transmissions may be either point-to-point or point-tomultipoint. Also, besides in an active mode, a slave device can be in the parked or standby modes so as to reduce power consumptions. A scatternet is a collection of operational Bluetooth piconets overlapping in time and space. Two piconets can be connected to form a scatternet. A Bluetooth device may participate in several piconets at the same time, thus allowing for the possibility that information could flow beyond the coverage area of the single piconet. A device in a scatternet could be a slave in several piconets, but master in only one of them.

•

ZigBee is a wireless networking standard that is aimed at remote control and sensor applications which is suitable for operation in harsh radio environments and in isolated locations.ZigBee technology builds on IEEE standard 802.15.4 which defines the physical and MAC layers. The main applications for 802.15.4 are aimed at control and monitoring applications where relatively low levels of data throughput are needed, and with the possibility of remote, battery powered sensors, low power consumption is a key requirement. Sensors, lighting controls, security and many more applications are all candidates for the new technology. Above the physical and MAC layers defined by 802.15.4, the ZigBee standard itself defines the upper layers of the system. This includes many aspects including the messaging, the configurations that can be used, along with security aspects and the application profile layers.

ZigBee supports three different network topologies namely the star, mesh and cluster tree or hybrid networks. Each has its own advantages and can be used to advantage in different situations.

The star network is commonly used, having the advantage of simplicity. As the name suggests it is formed in a star configuration with outlying nodes communicating with a central node.

Mesh or peer to peer networks enable high degrees of reliability to be obtained. They consist of a variety of nodes placed as needed, and nodes within range being able to communicate with each other to form a mesh. Messages may be routed across the network using the different stations as relays. If interference is present on one section of a network, then another can be used instead.

Finally there is what is known as a cluster tree network. This is essentially a combination of star and mesh topologies.

• A Wi-Fi network uses radio waves to wirelessly transmit information across a LAN, the reach of which can be extended by a Wi-Fi range extender. A computer utilizes a

wireless adapter to translate data transmitted by radio waves. These waves are different from those emitted by, for example, FM radios, for which frequency is measured in megahertz (MHz). Wi-Fi's signals are transmitted in frequencies of between 2.5 and 5 gigahertz (GHz). This signal is then transmitted from the adapter through a router, after which it is sent to the internet. Wi-Fi is widely used in businesses, agencies, schools and homes as an alternative to a wired LAN. Many airports, hotels and fast-food facilities offer public access to Wi-Fi networks. These locations are known as hotspots. Many charge a daily or hourly rate for access, but some are free. An interconnected area of hotspots and network access points is known as a hot zone. Modern smart phones and tablets are also able to turn into Wi-Fi hotspots, using their cellular network connections to provide wireless internet connectivity to computers and other devices. To access Wi-Fi hotspots, computers should include wireless adapters. These can be found on laptops and mobile devices, such as tablets or mobile phones. If for some reason your computer doesn't include such an adapter, one can be purchased that can be inserted into the PCI slot or USB port. Your computer should then be able to locate Wi-Fi networks automatically in the area. These can either be open networks or protected networks; the latter can be joined by entering a Wi-Fi password.

• WiMAX(IEEE 802.16): Acronym for **Worldwide Interoperability for Microwave Access**. It is Based on Wireless MAN technology. WiMAX is one of the hottest broadband wireless technologies around today. WiMAX systems are expected to deliver broadband access services to residential and enterprise customers in an economical way.

Loosely, WiMax is a standardized wireless version of Ethernet intended primarily as an alternative to wire technologies (such as Cable Modems, DSL and T1/E1 links) to provide broadband access to customer premises.

WiMAX would operate similar to WiFi, but at higher speeds over greater distances and for a greater number of users. WiMAX has the ability to provide service even in areas that are difficult for wired infrastructure to reach and the ability to overcome the physical limitations of traditional wired infrastructure.

WiMAX was formed in April 2001, in anticipation of the publication of the original 10-66 GHz IEEE 802.16 specifications. WiMAX is to 802.16 as the WiFi Alliance is to 802.11.

• Comparison of Bluetooth, ZigBee, Wi-Fi, Wi-Max

Table described below shows the point of differences among the four protocols. Each protocol is based on an IEEE standard.Wi-Fi and Wi-MAX provides higher data rate, while Bluetooth and ZigBee give a lower one. In general, the Bluetooth, UWB, and ZigBee are intended for WPAN communication (about 10m), while

Wi Fi is oriented to WLAN (about 100m) and Wi-MAX is defines some WMAN(80-90 km) technologies that operate at various frequencies, distances, and speeds to deliver Broadband Wireless Access (BWA).

Standard	Bluetooth	ZigBee	Wi-Fi	Wi-MAX
IEEE spec..	802.15.1	802.15.4	802.11a/b/g	802.16/a/e
Frequency band	2.4GHz	868/915 MHz; 2.4 GHz	2.4 GHz; 5 GHz	2-11 GHz
Max signal rate	1 Mb/s	250kb/s	54Mb/s	134 Mbps
Nominal range	10 m	10 m-100m	100m	80-90 kilometers
Nominal TX power	0 - 10 dBm	(-25) - 0 dBm	15 - 20 dBm	23-43dBm
Number of RF channels	79	1/10;16	14(2.4GHz)	1.25 to 20 MHz
Channel Bandwidth	1MHZ	0.3/0.6 MHz; 2 MHz	22MHz	1.25-28MHZ
Modulation type	GFSK	BPSK (+ ASK), O-QPSK	BPSK, QPSK COFDM, CCK, MQAM	OFDM using QPSK, 16-QAM, 64-QAM, 256-QAM
Spreading	FHSS	DSSS	DSSS, CCK, OFDM	OFDM
Coexistence mechanism	Adaptive freq. hopping	Dynamic freq. selection	Dynamic freq. selection transmit power control	Dynamic freq. selection transmit power control
Basic cell	Piconet	Star	BSS	BSS
Extension of the basic cell	Scatternet	Cluster and mesh	ESS	OSS
Data protection	16-bit CRC	16-bit CRC	32-bit CRC	FCS

CONCLUSION :

Both IEEE standards Bluetooth and ZigBee are used for short range wireless technology but ZigBee has a distinct area upon which it is focused which is sensor networks. Both ensure low power consumption is a key feature. WiMAX is similar to the Wi-Fi, but on a much larger scale and at faster speeds. The bottom line is that Bluetooth, ZigBee, WLANs and WMANs are complementary network architectures, supported by standard technologies that were designed for very different environments and purposes. In this paper we have seen the overview of most popular wireless standards Bluetooth, ZigBee, Wi-Fi and WiMAX.As well as this paper represents some points of difference like Frequency band, Max signal rate, Nominal range, Channel Bandwidth, Data protection and many more. This paper is not to draw any conclusion regarding which one is superior since the suitability of network protocols is greatly influenced by practical applications, of which many other factors such as the network reliability, roaming capability, recovery mechanism, chipset price, and installation cost need to be considered in the future.

REFERENCES :

https://www.tutorialspoint.com/wimax/what_is_wimax.html

http://searchmobilecomputing.techtar get.com/definition/Wi-Fi

"A Comparative Study of Wireless Protocols:Bluetooth, UWB, ZigBee, and Wi-Fi" Karunakar Pothuganti 1 and Anusha Chitneni 2

http://ieeexplore.ieee.org/document/6353314/

NOISE REMOVAL TECHNIQUES IN IMAGE DATABASE FOR QUALITY IMPROVEMENT

MRS. APARNA JOSHI *, DR. VINAY CHAVAN **
Research Student, Sant Gadge Baba Amravati University*,
Research Guide, Sant Gadge Baba Amravati University*,
aajoshi2@gmail.com, drvinaychavan@gmail.com

Abstract : Image denoising technique plays an important role in image processing applications because images are corrupted with various types of noises and so it is very difficult to get useful information from these noisy images. Different techniques of noise removal are used for noise free images and fully recovered by minimum signal distortion.

Keywords : Noise, Spatial Domain, Transform Domain,

INTRODUCTION:

The images which are corrupted by impulse or undesirable random variations in intensity values called noise. Noise is always presents in digital images during image acquisition, coding, transmission, and processing steps which tells unwanted information in digital images. Noise produces undesirable effects such as artifacts, unrealistic edges, unseen line

To reduce these undesirable effects, image denoising techniques are applied for further processing. Different techniques have been proposed for image restoration depending on the type of noise present in image. Noise removal techniques are mainly classified into two categories - spatial domain & Transform domain. And further divided into different categories as shown in fig. 1

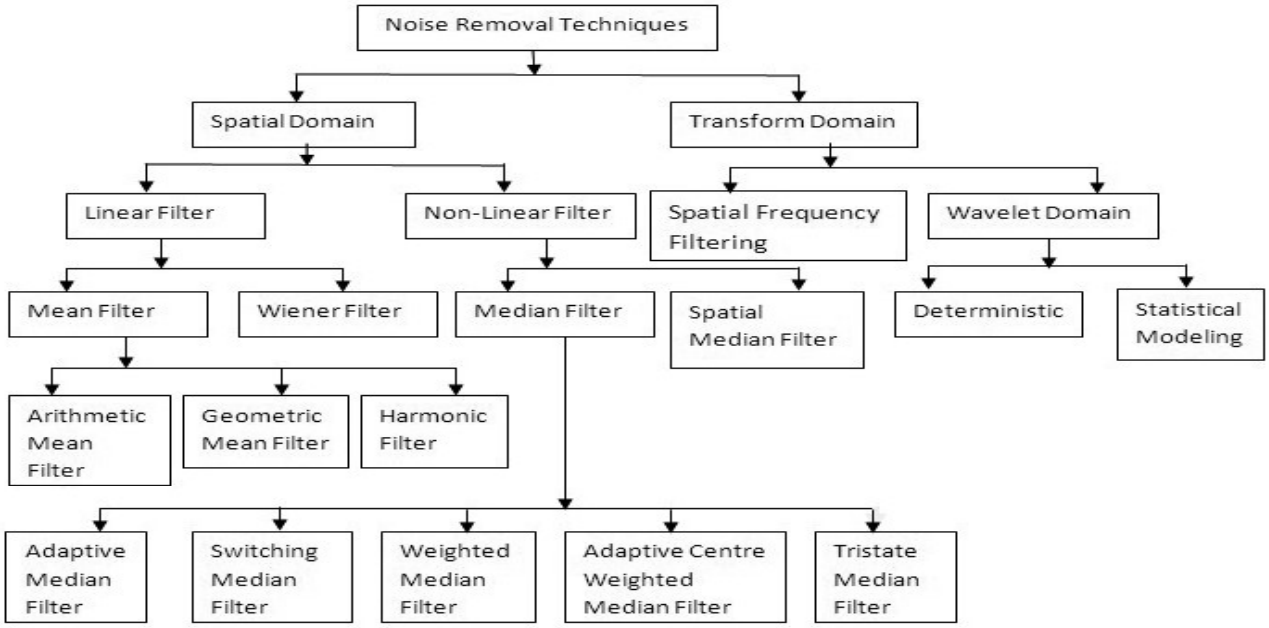


Fig. 1- Noise Removal (Image Denoising) Techniques

1. Spatial Domain :-
This approach of image de-noising directly operates on pixels. One of the biggest advantages of spatial domain is its speed but this technique is unable to preserve edges and fine details in image. This technique categorized as linear and non-linear techniques. [1,2]

A) Linear Filters :-
Linear filters also known as average filters. These filters remove noise from images but blur sharp edges, destroy lines and other fine details in image. This filter is not used in signal dependent noise. It is useful if the signal corruption can be modeled as Gaussian process. Linear filter has two types mean filter & wiener filter. [3,7]

a) Mean Filters :-
Mean filtering is a simple and easy to implement method of smoothing images. It simply replace each pixel from image with mean value of its neighbors, including itself. These types of filters don't preserve fine details in image.[1,4]
Mean filters include –

i) Arithmetic mean filters
ii) Geometric mean filters
iii) Harmonic filters

Arithmetic mean filtering process computes average value of image in the area covered by filter. Then restore value at the

place of corrupted pixel.
In Geometric filtering, each restored pixel is given by product of the pixels in sub image window, raised to power 1/mn. This filter achieves smoothing comparable to arithmetic mean filters but tends to loss less image details in the process.
Harmonic mean filters well for salt noise but fails for pepper noise. It does well also with other types of noise like Gaussian noise.

b) Wiener Filters :-

Wiener is a 2-D adaptive noise removal filter. Wiener filter is a class of optimum linear filter which involves estimation of desired signal sequence from another related sequence. It minimizes overall mean square error in the process of inverse filtering and noise smoothing. To implement this we estimate the power of original image and additive noise. Weiner filter can provide the optimal result and mean square error (MSE) is the accuracy criterion This filter is best suitable to remove Gaussian noise.[12]

B) **Non Linear Filters** :-

In non linear filters noise can be removed without identifying it exclusively. These filtering techniques apply low pass filtering on images on assumption that noise signals always have high frequencies. This method exploits the fact of problem of wavelet transform and maps white noise in the signal domain to white noise in the transform domain. Thus, white signal energy is more concentrated into transform domain, noise energy cannot be accumulated. So, this is the very effective method of noise removal from signal. The method which removes the small coefficients while others are untouched, is known as Hard Thresholding. To cover the demerits of Hard Thresholding, Wavelet transform soft thresholding was also introduced in by Donoho. In this, the coefficients greater than threshold are limited by the absolute value of threshold itself. It has two categories like median filtering and spatial median filter.[3,5,8]

a) **Median Filtering** :-

This method is widely used to preserve edges. This filter works by moving through the image pixel by pixel, replacing the middle pixel in the window with the median value of its neighbouring pixels. The pixel is calculated by first sorting all the pixel values from the pattern of neighbours into numerical order, and then replacing the pixel being considered with median pixel value. It is influenced by the size of filtering window to a great degree, has the conflicts between noise reduction and detail protection. When the filtering window is downsized it maintains the details better while reducing less noise; when the window size is enlarged, it reduces noise in a better way while diminishing the protection of details. Median filter is best suited to remove noise without reducing the sharpness of the image. These filters mostly used to remove Gaussian noise and impulse noise (salt and pepper noise). [3,6,12]

b) Spatial Median Filter :-

In spatial median filter, the spatial median is calculated by calculating the spatial depth between a point and a set of point. This spatial depth is defined as-

$$S_{dept\,h} = 1 - \frac{1}{N-1}\left\|\sum_{i=1}^{N}\frac{X-x_i}{\|X-x_i\|}\right\|$$

In this filter after finding out the spatial depth of each point lying within the filtering mask, this information is used to decide whether the central pixel of window is corrupted or not. If central pixel is uncorrupted then it will not be changed. We then find out the spatial depth of each pixel within the mask and then sort these spatial depths in descending order .The point with largest spatial depth represent spatial median of the set. If central pixel is corrupted with noise then it is replaced by calculated spatial median.

1. Transform Domain :-

Transform domain filtering method includes spatial frequency filtering and wavelet domain. These techniques subdivided on the basis of functions as data adaptive and non data adaptive transforms.

I) Spatial Frequency Filtering :-

This technique refers the use of low pass filters using Fast Fourier Transform(FFT). It removes noise by adapting a frequency domain filter and deciding a cut-off frequency. These methods are time consuming and dependent on cut-off frequency. These methods create artificial frequencies in the processed images.[3,7]

II) Wavelet Domain :-

It focuses on exploring the multi resolution properties of wavelet transform. By observing the signal across multiple resolutions, this technique identifies the close correlation of signal at different resolutions. This method gives the excellent results but is computationally less feasible due to cost and complexity. The Wavelet coefficients can be modeled either in deterministic way or statistical modeling.[3]

i) Deterministic :-

It involves making of tree structure of wavelet coefficients with each level in the tree representing scale of transformation and nodes representing the wavelet coefficients. At particular node, if the wavelet coefficient has the strong presence than the signal, its presence is more pronounced at the parent nodes itself. If there is noisy coefficient, then its consistent presence is missing. [3]

ii) Statistical Modeling :-

This approach explores some properties of Wavelet Transform such as local correlation between neighboring wavelets and multiple and global correlation between the wavelet coefficients etc .It has the inherent goal of perfecting the data of image by using Wavelet Transforms. [3,9]

CONCLUSION :

There are number of techniques developed on the basis of image database and noise present in them. But these techniques are designed for a particular type of noise in image for which it provides good results. Experimentally it is concluded that linear filter removes noise from images but blur sharp edges, destroy lines and other fine detail. Linear filters are good for Gaussian noise. The non-linear type of filter removes noise without identifying it exclusively.

In this, median filter is widely used to preserve edges without reducing the sharpness of image**.** These are mostly used to remove Gaussian noise and impulse noise (salt and pepper noise).

REFERENCES :

1. James C. church, yixinchen,Stephen V rice "A spatial filter for noise removal in digital images",IEEE southeast conference 2008,3-6 april.
2. Sharma A "iMage de-noising using spatial domain filtering:A quantitative study" IEEE 6th international conference 2013.
3. Sarbjit Kaur, Er. Ram Singh, " Image De-noising Techniques : A Review Paper ", International Journal For Technological Research in Engineering, Volume 2, Issue 8, April 2015, ISSN – 2347-4718.
4. Zeinab A. Mustafa BanazierA. Abrahim and Yasser M.kadah, "K11.Modified Hybrid Median Filter For Image Denoising" Biomedical Engineering deprtment, Ciro University,Giza,Egypt
5. T.chan,S.osher and J.Shen "the digital TVfilter and non linear denoising",IEEE transactions on image processing,2001.
6. Rafael C., Gonzalez, and Richard E. Woods "Digital Image Processing" , 2nd ed., Beijing: Publishing House of Electronics Industry,2007
7. LaConte S. et al., "Wavelet Transform-Based Wiener Filtering of Event-Related fMRI Data," MRM,Vol.44, pp.746-757, 2000.
8. D. L. Donoho, "Denoising by soft thresholding," IEEE Trans. Inform. Theory, vol. 41, pp. 613–627, 1995.
9. A.pizurica, w. Philips,I.lemahieu, M.Acheroy"Aversatie wavelet domain noise filtration techniques" IEEE trans. On medical images,2003
10. Abhishek R , Srinivas N," An Advanced Technique for Removal Of Salt & Pepper Noise In Images", International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 2 Issue 9 September 2013 Page No. 2856-2860
11. Ajay Kumar Boyat, Brijendra Kumar Joshi, "A REVIEW PAPER: NOISE MODELS IN DIGITALIMAGE PROCESSING", Signal & Image Processing : An International Journal (SIPIJ) Vol.6, No.2, April 2015
12. Mohd. Tahir, Anas Iqbal, Abdul Samee Khan, "A Review Paper of Various Filters for Noise Removal in MRI Brain Image", International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801 , ISSN (Print): 2320-9798, Vol. 4, Issue 12, December 2016

A REVIEW: BIG DATA & OPEN SOURCE TOOLS

MISS D.P. GITE AND MR. S.N. KAWADE
dipa.gite@gmail.com, sitaram.kawade@gmail.com
S.N. Arts, D.J.M. Commerce & B.N.S. Science College, Sangamner

Abstract: Big Data today influences our lives in the most unexpected ways, and organizations are using it extensively to gain that competitive edge in the market. So let's get acquainted with the open source tools that help us to handle Big Data. Today, on our journey towards a digital India, all government offices are switching to digitization instead of manual record keeping. As we proceed along this path, we see a tremendous increase in the size of data. The term that encapsulates such immense volumes of information is Big Data. Research has focused more on the statistical models, which make soft and probabilistic decisions based on attaching the real-value weights to each input feature. The edge that such models have is that they can express the relative certainty of more than one different possible answer rather than only one, hence producing more reliable results as compared to when such a model is included as one of the components of a larger system.

Keywords: Big Data, Open source, Apache, Hadoop, KNIME

INTRODUCTION:

The term 'Big Data' refers to extremely large data sets, structured or unstructured, that are so complex that they need more sophisticated processing systems than the traditional data processing application software.

It can also refer to the process of using predictive analytics, user behavior analytics or other advanced data analysis technology to extract value from a data set. Big Data is often used in businesses or government agencies to find trends and patterns, which can help them strategic decisions or spot a certain pattern or trend among the masses.

Gone are the days when banks used to store customer information (such as names, photographs and specimen signatures) in individual postcard-like data sheets in thick registers were used in different government offices If an employee had to update any of the registered customer's details, the task could take up the whole day. Hours were wasted searching for that particular customer's details and then creating a new record to replace the old one. The customers, too, had to wait for hours for such minor tasks to be completed. Apart from the tediousness of searching for data from piles of ledgers, such paper files could be lost at any time due to disasters like floods or fire, apart from the degradation of the very paper on which the data was recorded.

Open source tools to handle Big Data

Fig.1 Open Source Tools

Apache Hadoop

a. Apache Hadoop is an open source software framework used for the distributed storage and processing of large data sets using the MapReduce programming model. It consists of computer clusters built using commodity hardware. All the different modules in Hadoop are actually designed with the assumption that different hardware ailures are commonly observed occurrences and they should be automatically handled by the framework.

Features:

· The Hadoop framework is mostly written in Java, with some of its native code in C. Its command line utilities are written as shell scripts.

· Apache Hadoop consists of a large storage part, known as the Hadoop Distributed File System.

· It uses the MapReduce programming model to process large data sets.

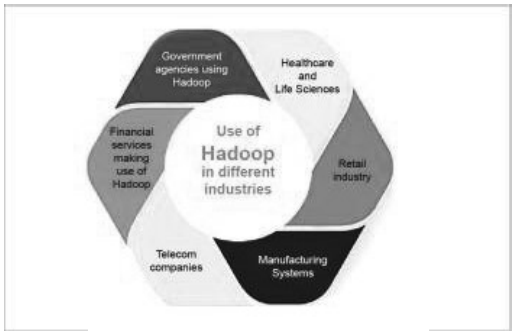


Fig.2 Use of Hadoop

a. Cassandra

This is an open source distributed NoSQL database management system. It's designed to handle large amounts of data across many different commodity servers, hence providing high availability with no single point of failure. It offers strong support for clusters that span various data Centre's,with its asynchronous master less replication allowing low latency operations for all clients.

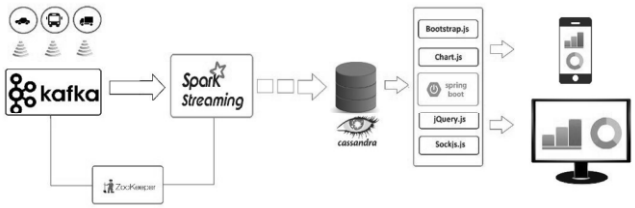


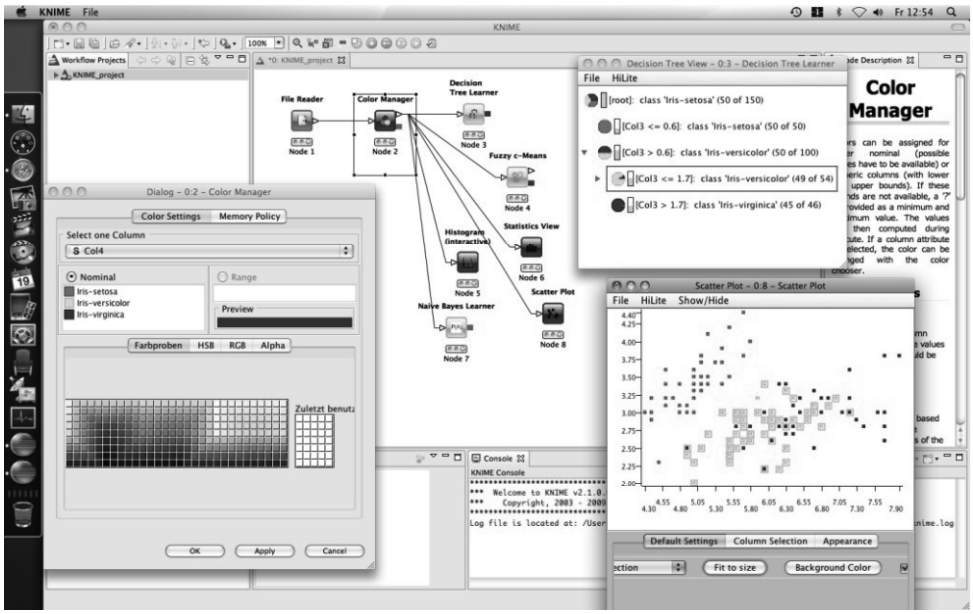
Fig.3 Cassandra

Features:

- It supports replication and multiple data Centre replication.
- It is decentralized.
- It provides MapReduce support.
- It supports Cassandra Query Language (CQL) as an alternative to the Structured Query Language (SQL).

a. KNIME

Also called Konstanz Information Miner, this is an open source data analytics, integration and reporting platform. It integrates different components for data mining and machine learning through its modular data pipelining concept. A graphical user interface allows the assembly of nodes for data pre-processing (which includes extraction, transformation and loading), data modelling, visualization and data analysis.



Features:

- KNIME is written using Java and is based on Eclipse. It makes use of its extension capability to add plugins, hence providing additional functionality.
- The core version of KNIME includes modules for data integration, data transformation as well as the commonly used methods for data visualization and analysis.
- It allows users to create data flows and selectively execute some or all of them.
- It allows us to inspect the models, results and interactive views of the flow.
- KNIME workflows can also be used as data sets to create report templates, which can be exported to different document formats like doc, PPT, etc.

a. Rapid Miner

This is basically a data science software platform. It is used for business and commercial applications as well as for education, research, rapid prototyping, training and application development. It supports all the steps of the machine learning process including data preparation, model validation, results visualization and optimization. It has been developed on an open core model. It provides a graphical user interface to design and execute different analytical workflows.

Features:

- Uses a client or server model with the server offered either on premise, or in private or public cloud infrastructures.
- Provides various machine learning and data mining procedures including data loading and transformation, predictive analytics and statistical modelling, data pre-processing and visualization, evaluation and deployment, etc.
- Is written using the Java programming language
- Provides different learning schemes, algorithms and models, which can be extended using Python and R scripts.

R - Programming

b. R isn't just a software, but also a programming language. Project R is the software that has been designed as a data mining tool, while R programming language is a high-level statistical language that is used for analysis. An open source language and tool, Project R is written is R language and is widely used among data miners for developing statistical software and data analysis. In addition to data mining it provides statistical and graphical techniques, including linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, clustering, and others. You can learn about Project R and R Programming Language

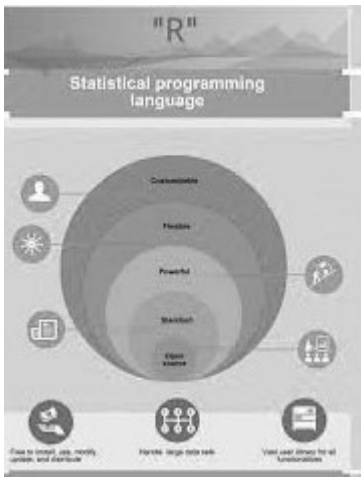


Fig.5 R Programming

f.MongoDB

MongoDB is also a great tool to help store and analyze big data, as well as help make applications. It was originally designed to support humongous databases, with its name MongoDB, actually derived from the word humongous. MongoDB is a no SQL database that is written in C++ with document-oriented storage, full index support, replication and high availability, etc.

g.Lumify

Lumify is a relatively new open source project to create a Big Data fusion and is a great alternative to Hadoop. It has the ability to rapidly sort through numerous quantities of data in different sizes, sources and format. What helps stand out is it's web-based interface allows users to explore relationships between the data via 2D and 3D graph visualizations, full-text faceted search, dynamic histograms, interactive geospatial views, and collaborative workspaces shared in real-time. It also works out of the box on Amazon's AWS environment.

Best practices for handling Big Data

- 1. Always try to bring the huge data set down to its unique set by reducing the amount of data to be managed.
- 2. It's a good practice to leverage the power of virtualization technology. All unique data sets must be virtualized so that multiple applications can reuse the same data footprint. This will also help the smaller data footprint to be stored on any vendor-independent storage device.
- 3. Structure different Big Data environments around analytics, not standard reporting or ad hoc querying.
- 6. Use a metadata-driven codeless development environment as it can increase overall productivity.

CONCLUSION :

Big Data mining and analysis are definitely going to continue to grow in the future, with many companies and agencies spending lots of time and money, for acquiring and analyzing data, making data more powerful. Understand different Big Data technology options and try to go for the best one for handling big data.

REFERENCES :

"Survey on Big Data Using Data Mining"
1. *(PDF). International Journal of Engineering Development and Research. 2015*. Retrieved 14 September 2016.
"Resources on how Topological Data Analysis is used to analyze big data"
2. *Ayasdi.*
3.*CNET News (1 April 2011).* "Storage area networks need not apply".
"How New Analytic Systems will Impact Storage"
4. *September 2011.*
5.Kalil, Tom. "Big Data is a Big Deal". White House. Retrieved 26 September 2012.
"Digital Technology & Social Change"
6.*Canvas.instructure.com*. Retrieved 8 October2017.

VIRAL CONJUNCTIVITIS PROGRESSION DETECTION USING IMAGING METHOD IN MATLAB

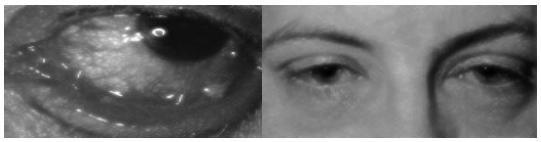
MS.JYOTI PATIL
Jyot.physics@gmail.com , Dr.Sharmila Chaudhari
shar_chaudhari@yahoo.co.in
PDEA'S Baburaoji Gholap Research Institute,Pune ,Maharashtra,India

Abstract: Viral conjunctivitis is the most common cause of infectious conjunctivitis, infected eye then affects the other eye within 24-48 hours. Inflammation or infection of the conjunctiva is known as conjunctivitis and is characterized by dilatation of the conjunctival vessels, resulting in hyperemia and edema of the conjunctiva, typically with associated discharge. which can be so severe that it is difficult to open the eyes. This condition can result in permanent damage to vision, and anyone with this sensation should seek. So to observe disease growth in redness and infection we are using the automated image processing system. In this system algorithm optimized the detection of the vessels and applied a skeletonization transform to allow measurement of vessel diameter and number of branch points. [1]

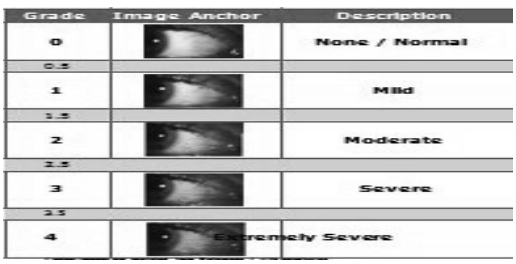
Keywords: *conjunctivitis, Viral, keratitis,*

INTRODUCTION:

The conjunctiva is a thin, translucent, relatively elastic tissue layer with both bulbar and palpebral portions. The bulbar portion of the conjunctiva lines the outer aspect of the globe, while the palpebral portion covers the inside of the eyelids. Underneath the conjunctiva lie the episclera, the sclera and the unveil tissue layers. Hyperemia is viral infection in Conjunctivitis



Ocular redness in the nasal and temporal conjunctiva were assessed separately in both eyes
1) Hyperemia was graded at the following magnifications:
Live in the clinic, at 3x
In 10x images, by 3 separate graders per image
In 25x images, by 1 expert grader
2) Hyperemia at 10x and 3x was graded by using a scale with descriptive anchors and photographic
anchors (shown below), similar to a validated scale1
Hyperemia at 25x was graded on the same scale, but by using the following parameters:
Vessel surface area & Average vessel diameter
Reduction of white surface areas due to emergence of episcleral vasculature and dilation of conjunctival vessels
Injection close to the limbus

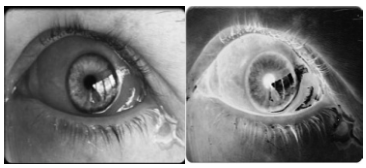


I. METHODS: THE NORMALIZED DIFFERENCE VEGETATION

Subtract the value of the red band from the value of the NIR band and divide by their sum.

ndvi = (NIR - red) ./ (NIR + red);

Apply decorrelation stretch to multichannel decorrstretch applies a decorrelation stretch to a multichannel image and returns the result

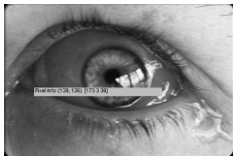


1) Enhancing imagery with a contrast stretch

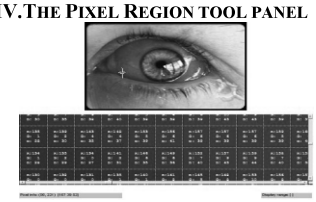
2) Enhancing imagery with a decorrelation stretch

Notice how the array-arithmetic operators in MATLAB make it possible to compute an entire NDVI image in one simple command.Recall that variables red and NIR have class single. This choice uses less storage than class double but unlike an integer class also allows the resulting ratio to assume a smooth gradation of values. Variable ndvi is a 2-D array of class single with a theoretical maximum range of [-1 1].You can specify these theoretical limits when displaying ndvi as a grayscale image.

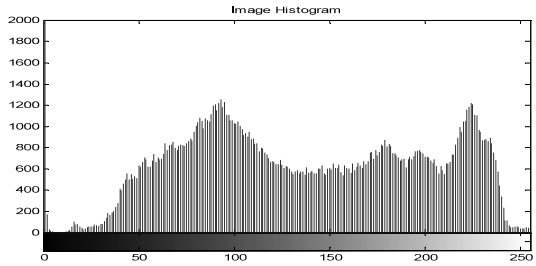
III. RGB VALUES OF IMAGE



IV. THE PIXEL REGION TOOL PANEL



Results: Automatic Image Processing



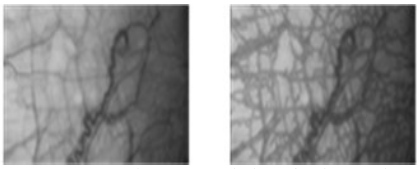
Results: displays Image Histogram

IV. VESSAL DETECTION AND SEGMENTATION

An automatic image processor that detects vessels and computes a set of 50 shape and densitometry measurements, including:

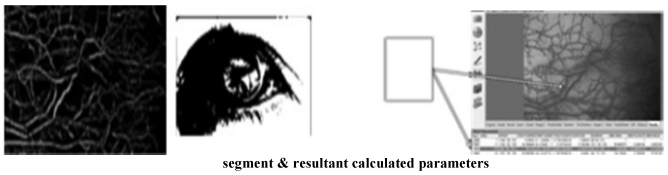
1) Vessel surface area,
2) Maximum vessel diameter
3) Average vessel diameter

An image reviewing interface that allows users to easily compare manual and automatic measurements The automated image processing algorithm optimized the detection of the vessels and applied a skeletonization transform to allow measurement of vessel diameter and number of branch points.



Grade 3, 25x Skeletonized Vessels

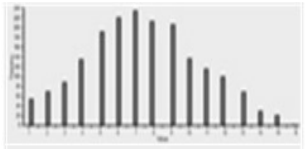
Automatic analysis detected "triple points" (intersections) to quantify vessel ramification (arborization):



segment & resultant calculated parameters

Calculations were returned for each vessel segment:

Automatic measurements yielded a variety of factors that were not evident to clinical observers, including: vessel area,vessel diameter,total vessel length vessel density (vessel area/total area), and other shape factors Measurements could be plotted as histograms, as shown below for vessel segment.



Maximum Vessel Segment Radius, Pixels

Image transformation, vessel detection, and parameter calculation required only a few seconds per image

CONCLUSION :

Thus to calculate the redness of eye the number of red pixels are count. Thus suppose for the segmentation of image is best method in image processing using which the number of pixels can be easily count. The first goal of this study was to find the intensity or the to count number red pixels of infected Of red/pink eye. These tools for the grading of Viral conjunctivitis are fast, reliable, accurate, not prone to human bias, and return information about RGB Intensity of Eye & vessels that was not available with other automated methods.

REFERENCES :

[1] Madhusudhanan Balasubramanian B.E., Bharathiyar University, 1998, " COMPUTER ASSISTED EYE FUNGAL INFECTION DIAGNOSIS"

[2] LIXU GU, SEMINAR SERIES ON ADVANCED MEDICAL IMAGE PROCESSING(2), FUNDAMENTAL CONCEPTS Robarts Research Institute London, Ontario, Canada, July 10, 2002.

[3] Rafael C. Gonzalez and Richard E. Woods. 'Digital Image Processing using MATLAB',2nd edition. Prentice Hall, 2002. ISBN 0-201-18075-8.

[4] Marc M. Schulze, Natalie Hutchings, and Trefford L. Simpson1, The Use of Fractal Analysis and Photometry to Estimate the Accuracy of Bulbar Redness Grading Scales", IOVS, April 2008, Vol. 49, No. 4


INDIRA GROUP
OF INSTITUTES
INDIRA UNIVERSE
CAMPUS
MANAGEMENT
PHARMACY
INDIRA